

Міністерство освіти і науки України
Державний вищий навчальний заклад
«Національний гірничий університет»

Інститут електроенергетики
Факультет інформаційних технологій
Кафедра безпеки інформації та телекомунікацій

ПОЯСНЮВАЛЬНА ЗАПИСКА
дипломної роботи

магістра

(назва освітньо-кваліфікаційного рівня)

галузь знань 17 Електроніка та телекомунікації
(шифр і назва галузі знань)

напрямок підготовки 172 Телекомунікації та радіотехніка
(код і назва спеціальності)

спеціалізація Телекомунікаційні системи та мережі
(код і назва освітньої програми)

освітній рівень магістр
(назва освітнього рівня)

кваліфікація Професіонал в галузі електроніки та телекомунікацій
(код і назва кваліфікації)

на тему: Розробка системи для проведення сумісних конференцій з використанням інтерактивного потокового відео

Виконавець: студент 2 курсу, групи 172м-16

Кошута Тетяна Аделай
(підпис) (прізвище ім'я по-батькові)

Керівники	Прізвище, ініціали	Оцінка	Підпис
роботи	к.ф.-м.н., доцент Магро В.І.		
розділів:			
спеціальний	к.ф.-м.н., доцент Магро В.І.		
економічний	к.е.н., доцент Романюк Н.М.		

Рецензент			
-----------	--	--	--

Нормоконтроль			
---------------	--	--	--

Дніпро
2018

Міністерство освіти і науки України
Державний вищий навчальний заклад
«Національний гірничий університет»

Інститут електроенергетики
Факультет інформаційних технологій
Кафедра безпеки інформації та телекомунікацій

ЗАТВЕРДЖЕНО:
завідувач кафедри
безпеки інформації та телекомунікацій
_____ Корнієнко В.І.

«_____» _____ 2018 року

ЗАВДАННЯ
на виконання кваліфікаційної роботи магістра
спеціальності _____ *172 Телекомунікації та радіотехніка*
(код і назва спеціальності)

студенту _____ *172М-16* _____ *Кошта Твапита Аделай*
(група) (прізвище ім'я по-батькові)

Тема дипломної роботи _____ *Розробка системи для проведення сумісних конференцій з використанням інтерактивного потокового відео*

1 ПІДСТАВИ ДЛЯ ПРОВЕДЕННЯ РОБОТИ

Наказ ректора Державного ВНЗ «НГУ» від _____ **00** № **000**

2 МЕТА ТА ВИХІДНІ ДАНІ ДЛЯ ПРОВЕДЕННЯ РОБІТ

Об'єкт досліджень _____ *засоби розповсюдження інтерактивного потокового відео*

Предмет досліджень _____ *конференцсистема з використанням інтерактивного потокового відео*

Мета НДР зменшення часу очікування при отриманні інтерактивного потокового відео

Вихідні дані для проведення роботи
інтерактивного потокового відео

Система для проведення сумісних конференцій з використанням

3 ОЧІКУВАНІ НАУКОВІ РЕЗУЛЬТАТИ

Наукова новизна комбінування існуючих методів стиснення для з метою зменшення затримки при запиті користувача до сервера додатків або до ігрового сервера

Практична цінність суттєве зменшення часу очікування при запиті потокового відео, в найгіршому випадку час очікування не перевищує 90 мс.

4 ВИМОГИ ДО РЕЗУЛЬТАТІВ ВИКОНАННЯ РОБОТИ

при запиті користувача до сервера додатків або до ігрового сервера з метою отримання потокового відео, в найгіршому випадку час очікування не повинен перевищувати 90 мс.

5 ЕТАПИ ВИКОНАННЯ РОБІТ

Найменування етапів робіт	Строки виконання робіт (початок-кінець)
1. Огляд існуючих систем мультимедіа	15.09.17-15.10.17
2. Розробка системи для проведення сумісних конференцій	16.10.17- 25.12.17
3. Вступ, висновки, реферат	26.12.17-04.01.18
4. Розробка мультимедійної презентації	05.01.18-15.01.18

6 РЕАЛІЗАЦІЯ РЕЗУЛЬТАТІВ ТА ЕФЕКТИВНІСТЬ

Економічний ефект більш ефективно використання існуючих пристроїв користувача, за рахунок зменшення часу затримки потокового відео

Соціальний ефект ширше запровадження систем мультимедіа

до повсякденного життя пересічної людини

7 ДОДАТКОВІ ВИМОГИ

Завдання видав _____
(підпис)

Магро В.І
(прізвище, ініціали)

Завдання прийняв
до виконання _____
(підпис)

Кошта Т.А.
(прізвище, ініціали)

Дата видачі завдання: 14 вересня 2017 року
Термін подання дипломної роботи до ДЕК: 15 січня 2017 року

РЕФЕРАТ

Пояснительная записка: 184с. 39 рис., 1 табл., 10 приложений, 21 источник.

Объект исследования: система для проведения совместных конференций с использованием интерактивного потокового видео.

Цель работы (проекта): разработка системы обработки данных, которая позволит уменьшить время ожидания при получении интерактивного потокового видео.

Методы исследования: системный подход, методы сравнения, индексный метод, структурный анализ, корреляционно-регрессионный анализ.

В специальной части дана характеристика интерактивного потокового видео.

В работе исследовано интерактивное потоковое видео. Проведен анализ компьютерной игры с использованием интерактивного потокового видео проходящей в трех точках: Лиссабон, Луанда, Кабинда.

Данная техническая разработка относится к области систем обработки данных, которые улучшают способность пользователей манипулировать аудио и видеоносителями и подключаться к ним.

В экономическом разделе определены экономические затраты на покупку, установку и настройку необходимого оборудования для проведения компьютерной игры с использованием интерактивного потокового видео.

Практическое назначение работы состоит в уменьшении времени ожидания получения интерактивного потокового сжатого видео в клиентские устройства пользователей по сети Интернет, даже при ограниченной полосе пропускания.

Результаты проведенных в дипломной работе (проекте) исследований могут быть использованы при проведении компьютерных игр с использованием интерактивного потокового видео.

Научная новизна исследований состоит в предложенном способе организации проведения компьютерной игры с использованием интерактивного потокового видео.

Направления дальнейших исследований проектирование систем сотовых связей.

Ключевые слова: ИНТЕРАКТИВНОЕ ПОТОКОВОЕ ВИДЕО, СЕРВЕР, СЛУЖБА ХОСТИНГА, ВИДЕОУСТРОЙСТВО, ВИДЕОИГРА, ПРИЛОЖЕНИЕ, СЖАТИЕ ВИДЕО, СЖАТИЕ ЗВУКА, ВИДЕОПОТОК, СИГНАЛ, ПОЛЬЗОВАТЕЛЬ.

РЕФЕРАТ

Пояснювальна записка: 184 с., 39 мал., 1 таб, 10 додатків, 21 джерел.

Об'єкт дослідження: система для проведення спільних конференцій з використанням інтерактивного потокового відео.

Мета роботи (проекту): розробка системи обробки даних, яка дозволить зменшити час очікування при отриманні інтерактивного потокового відео.

Методи дослідження: системний підхід, методи порівняння, індексний метод, структурний аналіз, кореляційно-регресійний аналіз.

У спеціальній частині дана характеристика інтерактивного потокового відео.

В роботі досліджено інтерактивне потокове відео. Проведено аналіз комп'ютерної гри з використанням інтерактивного потокового відео, яке проходить в трьох точках: Лісабон, Луанда, Кабінда.

Дана технічна розробка відноситься до області систем обробки даних, які покращують здатність користувачів маніпулювати аудіо та відеоніснями і підключатися до них.

В економічному розділі визначені економічні витрати на покупку, установку і налаштування необхідного обладнання для проведення комп'ютерної гри з використанням інтерактивного потокового відео.

Практичне призначення роботи полягає в зменшенні часу очікування отримання інтерактивного потокового стисненого відео у клієнтські пристрої користувачів по мережі Інтернет, навіть при обмеженій смузі пропускання.

Результати проведених в дипломній роботі (проекті) досліджень можуть бути використані при проведенні комп'ютерних ігор з використанням інтерактивного потокового відео.

Наукова новизна досліджень полягає в запропонованому способі організації проведення комп'ютерної гри з використанням інтерактивного потокового відео.

Напрямки подальших досліджень проектування систем стільникових зв'язків.

Ключові слова: ІНТЕРАКТИВНЕ ПОТОКОВЕ ВІДЕО, СЕРВЕР, СЛУЖБА ХОСТИНГУ, ВІДЕОПРИСТРІЙ, ВІДЕОГРА, ДОДАТОК, СТИСК ВІДЕО, СТИСК ЗВУКУ, ВІДЕОПОТІК, СИГНАЛ, КОРИСТУВАЧ.

RESUME

Explanatory note: 184p., 39 images, 1 table, 10 applications, 21 sources.

Object of study: system for joint conferences using interactive streaming video.

The purpose of the work (project): the development of a data processing system that will reduce the waiting time during receiving an interactive streaming video.

Research methods: system approach, comparison methods, index method, structural analysis, correlation-regression analysis.

In a special section, an interactive streaming video is described.

In this paper, an interactive streaming video is explored. The analysis of a computer game with the use of interactive streaming video passing in three points is carried out: Lisbon, Luanda, Cabinda.

This technical development refers to the field of data processing systems that improve the ability of users to manipulate audio and video media and connect to them.

In the economic section, economic costs are determined for the purchase, installation and configuration of the necessary equipment for computer games using interactive video streaming.

The practical purpose of the job is to reduce the waiting time for receiving interactive streamed compressed video to client devices of users over the Internet, even with limited bandwidth.

The results of research carried out in the thesis (project) can be used in computer games using interactive video streaming.

The scientific novelty of the research consists in the proposed method of organizing a computer game using interactive video streaming.

Directions for further research are the design of cellular communication systems.

Keywords: interactive video streaming, servers, hosting services, video devices, video games, APPENDIX, compressed video, compressed audio, the video signal to the user.

СОДЕРЖАНИЕ

ВВЕДЕНИЕ.....	9
РАЗДЕЛ 1. СОСТОЯНИЕ ВОПРОСА. ПОСТАНОВКА ЗАДАЧИ.....	11
1.1 История развития домашних видеоустройств.....	11
1.2 Возникновение современных видеоустройств.	13
1.3 Недостатки современных видеоустройств.....	33
1.4 Постановка задачи.....	41
РАЗДЕЛ 2. СПЕЦИАЛЬНАЯ ЧАСТЬ.....	44
2.1 Краткое описание предлагаемого технического решения.....	44
2.2 Подробное описание предлагаемого технического решения.....	68
2.2.1 Характеристики сжатого видео.....	72
2.2.2 Архитектура службы хостинга.....	74
2.2.3 Анализ времени передачи сигнала туда и обратно.....	84
2.2.4 Сжатие видео с малым временем ожидания.....	90
2.2.5 Сброшенные или задержанные пакеты.....	116
2.2.6 Реализация устройства сжатия/устройства восстановления сжатого видео и аудио.....	122
2.2.7 Распределение серверного центра службы хостинга.....	125
РАЗДЕЛ 3. ЭКОНОМИЧЕСКАЯ ЧАСТЬ.....	149
3.1 Экономические затраты на приобретение оборудования.....	149
3.2 Экономические затраты на установку и настройку оборудования.....	150
ЗАКЛЮЧЕНИЕ.....	154
СПИСОК ИСПОЛЬЗОВАННОЙ ЛИТЕРАТУРЫ.....	156
ПРИЛОЖЕНИЕ.....	157

ВВЕДЕНИЕ

Основная составная часть интернета – это информация. Она имеет несколько общепонятных форм: текст, звук, изображение, видео. Качественная информация вбирает в себя все четыре аспекта. И конечно, самой понятной, самой качественно и полной информацией является именно видео, поскольку в нем могут присутствовать все составные: изображения, звук, текст. Именно поэтому, потоковое видео является самым востребованным типом передачи информации в интернете.

Само видео - это последовательный набор изображений(кадров), с параллельной записью звуков. Потоковое видео не исключение и является абсолютно тем же самым, что и обычное видео. Основное различие заключается в принципе распространения и потребления. Если обычное видео, будь то фильм, видео урок или любительская запись концерта, мы сначала скачиваем из сети интернет, и только после этого смотрим, то суть потокового видео заключена в объединении двух этих действий. Мы сразу можем приступить к просмотру интересующего нас видеоматериала, и одновременно, в то время пока мы смотрим, видео подгружается в кэш память нашего браузера, тем самым позволяя нам сразу приступить к просмотру, без предварительной загрузки и скачки видео файла. Таким образом, потоковое видео, это то видео, которое можно смотреть еще до того, как оно полностью загрузится.

Потоковое видео (Streaming Video) – это технология буферизации и сжатия данных, позволяющая вести трансляцию мультимедийного контента (видео) через Интернет в режиме реального времени. Иными словами, потоковое видео – это процесс преобразования видео и аудио контента в сжатый цифровой формат с его последующим распространением через компьютерные сети. Сжатые данные легко могут быть доставлены с использованием компьютерных сетей в силу их небольших размеров. Видео и аудио может быть постоянно потоковым, либо доставляться по требованию. Для просмотра потокового видео пользователю обычно нужен специальный проигрыватель, который

декодирует данные и выводит расшифрованное изображение на экран. Этот плеер может интегрироваться в оболочку браузера или же работать как самостоятельная программа.

Потоковое видео обладает еще одним весьма важным достоинством: оно не занимает места в памяти устройства, и после просмотра, автоматически стирается, давая возможность использовать накопительные свойства устройства по нашему усмотрению.

Интернет потокового видео схож со стандартом ТВ-вещания. Программное обеспечение используется для конвертации медиафайлов в подходящий формат, который подходит для доставки через Интернет.

Потоковое видео является достаточно модным направлением в потребительских технологиях на сегодняшний день. Благодаря гораздо меньшим размерам, в отличие от исходного видео файла и повсеместному распространению в мировой паутине интернета, потоковое видео стало самым простым и доступным способом просмотра видеоматериалов в интернете.

Объектом исследования является система для проведения совместных конференций с использованием интерактивного потокового видео.

Целью дипломной работы является разработка системы обработки данных, которая позволит уменьшить время ожидания при получении интерактивного потокового видео.

РАЗДЕЛ 1. СОСТОЯНИЕ ВОПРОСА. ПОСТАНОВКА ЗАДАЧИ.

1.1 История развития домашних видеоустройств

Записываемые аудионосители и киноносители являются одной стороной общественной жизни со времен Томаса Эдисона. В начале 20-го века были широко распространены записываемые аудионосители (валики и грампластинки) и киноносители (синематограф и фильмы), но обе технологии находились, однако, в своей ранней стадии развития. В конце 1920-х годов кинофильмы были объединены со звуковым сопровождением на основе рынка товаров широкого потребления, вслед за этим появились цветные кинофильмы со звуковым сопровождением. Радиовещание постепенно приняло форму ширококвещательных средств распространения аудиоинформации рынка товаров широкого потребления, пользующихся поддержкой рекламы. Когда в середине 1940-х годов был установлен стандарт телевизионного (TV) вещания, телевидение присоединилось к радио как форма ширококвещательных средств распространения рынка товаров широкого потребления, и ранее записанные кинофильмы или киноизображения в прямом эфире пришли в дом. К середине 20-го века в большинстве американских домов были граммофоны для воспроизведения записанных аудионосителей, радиоприемник для приема прямых аудиопередач и телевизор для воспроизведения аудио/видео (A/V) носителей прямой передачи. Очень часто эти 3 "медиаплеера" (проигрыватель, радиоприемник и телевизор) были объединены в одном корпусе с совместным использованием обычных динамиков, который стал "медиацентром" для дома. Несмотря на то что выбор носителей информации был ограничен для потребителя, "экосистема" носителей информации была довольно устойчивой. Большинство потребителей знали, как пользоваться "медиаплеерами", и могли в полной мере использовать их функциональные возможности. Вместе с тем, издатели мультимедиа (в значительной степени киностудии и телестудии и музыкальные компании) могли поставлять свои носители информации и в театры и для домашнего использования и не страдали от широко

распространенного пиратства или ”вторичных продаж”, т.е. перепродаж используемого носителя информации. Как правило, издатели не получают доход от вторичных продаж”, и, по существу, это уменьшает доход, который издатели иначе могли бы получить от покупателя используемого носителя информации вследствие новых продаж. Несмотря на то что в середине 20-го века, конечно, использовались проданные грампластинки, такие продажи не оказывали большого влияния на издателей грампластинок, потому что, в отличие от кино или видеопрограммы которые взрослый человек, как правило, смотрит один раз или только несколько раз музыкальная фонограмма может прослушиваться сотни или даже тысячи раз.

Поэтому носитель музыкальной информации является гораздо менее ”преходящим” (т.е. он имеет непреходящую ценность для взрослого потребителя), чем кино/видеоносители. После покупки грампластинки если потребителю понравилась музыка, то он, вероятно, будет хранить ее в течение длительного времени. С середины 20-го века до настоящего времени экосистема носителей информации подверглась ряду радикальных перемен, как с выгодой для потребителей и издателей, так и в ущерб им. При широком распространении введения магнитофонов, особенно кассетных магнитных лент с высококачественным стереозвуком, конечно, потребителям стало намного удобнее. Но это также отметило начало того, что является теперь широко распространенной практикой по отношению к потребительским носителям информации, пиратство. Конечно, многие потребители использовали кассетные магнитные ленты для записи на нее своих собственных грампластинок просто для удобства, но все возрастающее количество потребителей например, студенты в студенческом общежитии при свободном доступе к коллекциям пластинок друг друга) могли делать пиратские копии. Кроме того, потребители

могли записывать на магнитную ленту музыку, передаваемую по радио, вместо покупки грампластинки или магнитной ленты у издателя.

1.2 Возникновение современных видеоустройств

Появление потребительского VCR (кассетного видеомаягнитофона) привело к еще большему удобству потребителей, так как теперь кассетный видеомаягнитофон мог быть установлен для записи телепрограммы, которую можно было посмотреть в более позднее время, и это также привело к созданию видеопроката, где доступ к фильмам, а также телепрограммам мог предоставляться по требованию. Быстрое развитие домашних устройств хранения данных рынка товаров широкого потребления с середины 1980-х годов привело к беспрецедентному уровню выбора и удобства для потребителя, а также привело к быстрому расширению издательского рынка мультимедиа. В настоящее время потребители сталкиваются с большим выбором мультимедиа, а также с множеством устройств хранения данных, многие из которых привязаны к конкретным видам мультимедиа или конкретным издателям. Страстный потребитель мультимедиа может иметь стек устройств, подключенных к телевизорам и компьютерам, находящимся в различных помещениях дома, что в результате приводит к паутине кабелей, ведущих к одному или нескольким телевизорам и/или персональным компьютерам (PC), а также к группе пультов дистанционного управления. Эти устройства могут включать в себя видеоигровую консоль, кассетный видеомаягнитофон, DVD-плеер, звуковой процессор/усилитель объемного звука, спутниковую телевизионную абонентскую приставку, телевизионную абонентскую приставку кабельного телевидения и т.д. И для страстного потребителя может существовать множество устройств с аналогичными функциями из-за вопросов совместимости. Например, потребителю могут принадлежать как HD-DVD, так и Blu-ray DVD плеер или как Microsoft Xbox®, так и Sony Playstation® игровая видеосистема. На самом деле из-за несовместимости некоторых игр по всем версиям игровых консолей потребителю могут принадлежать как XBox, так и более поздняя версия, например Xbox 360®. Часто потребителей сбивает с толку

то, какой видеовход и какой пульт дистанционного управления использовать. Даже после того как диск вставлен в соответствующий плеер, выбраны видеовход и аудиовход для этого устройства и найден соответствующий пульт дистанционного управления, потребитель, тем не менее, сталкивается с техническими проблемами. Например, в случае широкоэкранный DVD пользователю может потребоваться сначала определить и затем установить соответствующий формат изображения на своем экране монитора или телевизора (например, 4:3, Full, Zoom, Wide Zoom, Cinema Wide и т.д.). Аналогично, пользователю может потребоваться сначала определить и затем установить соответствующий аудиоформат системы объемного звучания (например, AC3, Dolby Digital, DTS и т.д.). Часто потребитель не осознает, что он, возможно, не использует мультимедийное содержимое в полном объеме функциональных возможностей своей телевизионной или аудиосистемы. Все больше и больше устройств хранения данных, основанных на Internet- технологиях, добавляются к стеку устройств. Звуковые устройства, подобные системе Sonos® Digital Music, передают потоком аудио непосредственно из Internet.

Аналогично, устройства, подобные телевизионной приставке Slingbox™, записывают видео и передают его потоком через домашнюю сеть или через Internet, где его можно смотреть удаленно на PC. И службы IP-телевидения (IPTV) предлагают услуги, подобные кабельному телевидению, через цифровую абонентскую линию (DSL) или другие соединения с Internet в доме. Недавно также предпринимались попытки для интеграции множества мультимедиафункций в одном устройстве, например Moxi® Media Center и PC под управлением операционной системы Windows XP Media Center Edition. Несмотря на то что каждое из этих устройств предлагает элемент механизма для функций, которые оно выполняет, в каждом отсутствует повсеместно распространенный и простой доступ к большинству носителей информации. Кроме того, производство таких устройств часто стоит сотни долларов, часто из-за необходимости дорогостоящей обработки и/или локального запоминающего устройства. Кроме того, эти современные потребительские

электронные устройства, как правило, потребляют много мощности, даже в режиме ожидания, что означает, что они требуют больших затрат со временем и нерационально используют энергоресурсы. Например, устройство может продолжать работать, если потребитель забыл выключить его или переключился на другой видеовход. И так как ни одно из устройств не является законченным решением, то оно должно быть интегрировано с другим стеклом устройств в доме, что по-прежнему оставляет пользователя с паутиной проводов и большим количеством пультов дистанционного управления. Кроме того, когда многие новейшие устройства, основанные на Internet-технологиях, действительно работают должным образом, они, как правило, предлагают мультимедиа в более общей форме, чем оно могло иначе быть доступным. Например, устройства, которые передают поток видео через Internet, часто передают поток только видеоматериал, а не интерактивные "дополнительные материалы", которые часто сопровождают DVD, подобные "мнению о" видео, играм или комментариям кинорежиссера. Это происходит вследствие того, что часто интерактивный материал выводится в конкретном формате, предназначенном для конкретного устройства, которое обрабатывает интерактивность локально. Например, каждый из дисков DVD, HD-DVD и Blu-ray имеет свой собственный конкретный интерактивный формат. Любое домашнее устройство хранения данных или локальный компьютер, который мог быть разработан для поддержки всех популярных форматов, могут потребовать некоторого уровня усложнения и гибкости, что, вероятно, сделало бы их функционирование недопустимо дорогим и сложным для потребителя. Кроме указанной проблемы, если позже будет введен новый формат, то в локальном устройстве может не существовать аппаратного обеспечения для поддержки этого нового формата, что будет означать то, что потребитель должен будет купить модернизированное локальное устройство хранения данных. Например, если позже введено видео с высоким разрешением или стереоскопическое видео, то локальное устройство может не иметь вычислительной мощности для

декодирования видео, или оно может не иметь аппаратного обеспечения для вывода видео в новом формате.

Вопрос сложности и морального износа устройства хранения данных является серьезной проблемой, когда дело доходит до усложненного интерактивного мультимедиа, особенно видеоигр. Современные видеоигровые приложения в основном разделены на четыре главные машинозависимые аппаратные платформы, Sony PlayStation® 1, 2 и 3 (PS1, PS2 и PS3), Microsoft Xbox® и Xbox 360® и Nintendo Gamecube® и Wii™, а также основанные на PC игры. Каждая из этих платформ отличается от других так, что игры, написанные для исполнения на одной платформе, обычно не исполняются на другой платформе. Могут также существовать проблемы совместимости устройства от поколения к поколению. Несмотря на то что большинство разработчиков игровых программ создают игровые программы, независимые от конкретной платформы, для исполнения конкретной игры на конкретной платформе, требуется собственный слой программного обеспечения (часто называемый "механизмом разработки игровых программ") для адаптации игры для использования на конкретной платформе. Каждая платформа продается потребителю как "консоль" (т.е. автономный блок, подключаемый к телевизору или монитору/динамикам), или она сама является PC. Как правило, видеоигры продаются на оптическом носителе информации, например DVD Blu-ray, DVD-ROM или CD-ROM, который содержит видеоигру, воплощенную как усложненное приложение реального времени. Так как скорости домашнего широкополосного доступа увеличились, видеоигры становятся все более и более доступными для загрузки. Требования к специфике для достижения совместимости платформы с программным обеспечением видеоигры являются чрезвычайно высокими из-за характера режима реального времени и высоких вычислительных требований усовершенствованных видеоигр. Например, можно было бы ожидать полную совместимость видеоигр от поколения к поколению аналогично тому, как существует общая совместимость рабочих приложений от одного PC к другому с более быстрым процессором или ядром.

Однако с видеоиграми это не так. Так как производители видеоигр при выпуске поколения видеоигр, как правило, пытаются добиться максимально возможной производительности для данной ценовой точки, то часто существенно изменяют архитектуру системы так, что многие игры, написанные для системы предшествующего поколения, не исполняются на системе более позднего поколения. Например, XBox выполнен на основе семейства процессоров x86, тогда как XBox 360 выполнен на основе семейства PowerPC. Могут быть использованы способы эмуляции предшествующей архитектуры, но с учетом того, что видеоигры являются приложениями реального времени, часто невозможно достичь идентичного поведения при эмуляции. Это на носит ущерб потребителю, производителю видеоигровой консоли и издателю программного обеспечения видеоигры. Для потребителя это означает необходимость наличия видеоигровых консолей, как старого, так и нового поколения, подключенных к телевизору для возможности ведения всех игр. Для производителя консоли это означает затраты, связанные с эмуляцией, и более медленное внедрение новых консолей. И для издателя это означает необходимость выпуска множества версий новых игр для охвата всех возможных потребителей - не только выпуск версии для каждого бренда видеоигры (например, XBox, Playstation), но часто версии для каждой версии данного бренда (например, PS2 и PS3).

Портативные устройства, например сотовые ("cell") телефоны и портативные медиаплееры, также представляют проблемы для разработчиков игр. Все большее количество таких устройств подключаются к беспроводным сетям передачи данных и могут загружать видеоигры. Но на рынке существует широкий выбор сотовых телефонов и устройств хранения данных с широким диапазоном разных разрешающих способностей дисплея и вычислительных возможностей. Кроме того, так как такие устройства, как правило, имеют ограничения по потреблению мощности, стоимости и весу, у них, как правило, отсутствует усовершенствованное аппаратное обеспечение ускорения выполнения графических операций, подобное графическому процессору ("GPU"), например, устройства изготавливаемые корпорацией NVIDIA, Santa

Clara, CA (Санта-Клара, Калифорния). Следовательно, разработчики игровых программ, как правило, разрабатывают данную компьютерную игру одновременно для множества разных типов портативных устройств. Пользователь может обнаружить, что данная компьютерная игра не доступна конкретно для его сотового телефона или портативного медиаплеера. В случае домашних игровых консолей производители аппаратных платформ, как правило, назначают лицензионный платеж для разработчиков игровых программ для возможности издания игры на своей платформе. Операторы сотовой связи также, как правило, назначают лицензионный платеж для издателя игр для загрузки игры в сотовый телефон. В случае компьютерных игр не существует лицензионного платежа, выплачиваемого для издания игр, но разработчики игр, как правило, сталкиваются с высокими затратами из-за более высокого уровня косвенных затрат на обслуживание клиента для поддержки широкого диапазона конфигураций PC и из-за вопросов инсталляции, которые могут возникнуть. Кроме того, PC, как правило, представляют меньшие препятствия для пиратства игровых программ, так как опытный в техническом отношении пользователь может легко их перепрограммировать, и игры можно легче нелегально переиздавать и легче распространять (например, через Internet). Соответственно, для разработчика игровых программ существуют затраты и неблагоприятные условия при публикации на игровых консолях, сотовых телефонах и PC. Для издателей игровой консольной и компьютерной программы затраты не ограничиваются этим. Для распространения игр через розничные каналы издатели назначают оптовую цену ниже продажной цены для розничных торговцев для получения чистой прибыли. Издатель также, как правило, должен оплачивать затраты на производство и распространение физических носителей информации, содержащих игру. Розничный торговец также часто возлагает на издателя расход по "оплате ценовой защиты" для покрытия возможных непредвиденных расходов, например когда игра не раскупается, или если цена игры снижается, или если розничный торговец должен возместить часть или всю оптовую цену и/или принять обратно игрот

покупателя. Кроме того, розничные торговцы также, как правило, возлагают на издателей расходы по оплате поддержки сбыта игр в рекламных листовках. Кроме того, розничные торговцы все чаще выкупают игры у пользователей, которые перестали играть в них, и затем продают их как подержанные игры, как правило, не разделяя дохода от их продажи с издателем игр. Дополнением к бремени расходов, возложенных на издателей игр, является тот факт, что игры часто нелегально переиздают и распространяют через Internet для загрузки пользователями и изготовления ими бесплатных копий. Так как скорости широкополосного доступа в Internet увеличиваются и широкополосные соединения стали более широко распространены в США и во всем мире, в частности к дому и к Internet-кафе, где арендуются PC, подключенные к Internet, игры все чаще распространяются посредством загрузки в PC или консоли. Кроме того, широкополосные соединения все чаще используют для ведения многопользовательских онлайн-игр и массовых многопользовательских онлайн-игр. Эти изменения уменьшают некоторые затраты и вопросы, связанные с распространением в розницу. Загрузка онлайн-игр направлена на преодоление некоторых неблагоприятных условий для издателей игр за счет того, что затраты на распространение, как правило, уменьшаются, и затраты из-за непроданного носителя информации являются небольшими или отсутствуют. Но загруженные игры, по-прежнему, подвержены пиратству, и из-за своего размера они могут загружаться очень долго. Кроме того, множество игр могут заполнить дисководы небольшой емкости, как те, которые продаются с портативными компьютерами или с видеоигровыми консолями. Однако для игр большого размера или MMOG требуется, чтобы можно было вести эту игру по онлайн-соединению, проблема пиратства уменьшается, так как обычно требуется, чтобы пользователь имел действительную учетную запись пользователя. В отличие от линейных мультимедиа (например, видео и музыка), которые можно скопировать камерой при съемке видео с экрана дисплея или микрофоном при записи аудио с динамиков, каждый опыт использования видеоигры уникален и не может быть

скопирован с использованием простой видео/аудиозаписи. Соответственно, даже в регионах, где не обеспечивается строгое выполнение законов об авторском праве и пиратство сильно распространено, ММОГ могут быть ограждены от пиратства, и, следовательно, может поддерживаться предпринимательская деятельность. Несмотря на то что пиратство во многих случаях может быть уменьшено из-за характера онлайн-игр или ММОГ, оператор интерактивных игр, тем не менее, сталкивается с остальными проблемами. Многие игры для исполнения их должным образом требуют значительных локальных (т.е. домашних) ресурсов обработки для онлайн-игр или ММОГ. Если производительность локального компьютера пользователя является низкой, то он, возможно, не сможет вести игру. Кроме того, так как игровые консоли устаревают, то они также отстают от современного уровня развития и, возможно, не могут обрабатывать усовершенствованные игры. Даже если предположить, что локальный РС пользователя может удовлетворять вычислительным требованиям игры, то часто существует сложность инсталляции. Может существовать несовместимость драйверов. Консоль может исчерпать пространство на локальном диске, когда загружается большее количество игр. Сложные игры, как правило, принимают загружаемые заплатки от разработчика игр с течением времени, так как обнаружены и исправлены ошибки или если в игре выполнены модификации. Эти заплатки требуются снова загружать. Но иногда не все пользователи выполняют загрузку всех заплат. В других случаях загруженные заплатки вводят другие вопросы совместимости или расхода пространства на диске. Кроме того, во время ведения игры может потребоваться загрузка большого количества данных для обеспечения графической или поведенческой информации в локальный РС или консоль. Соответственно, онлайн-игры или игры ММОГ часто ограничивают свои требования к вычислительной сложности и/или по памяти. Кроме того, они часто ограничивают количество передач данных во время игры. И онлайн-игры или игры ММОГ также могут сузить рынок пользователей, которые могут вести эти игры. Кроме того, опытные в техническом отношении

пользователи все чаще декомпилируют локальные копии игр и модифицируют эти игры так, чтобы они могли обманывать. Обманы могут быть такими простыми, как то, чтобы сделать повторное нажатие клавиши более быстрым, чем это может человек. В играх, которые поддерживают внутриигровые транзакции ресурсов, обман может достигать уровня фальсификации, которая в результате приводит к мошенническим транзакциям, включающим в себя ресурсы фактической экономической ценности. Когда экономическая модель онлайн-игр или игр MMOG основана на таких транзакциях ресурсов, это может в результате привести к последствиям, причиняющим значительный ущерб операторам игр. Стоимость разработки новой игры выросла, когда РС и консоли стали в состоянии выводить все более усложненные игры. На заре видеоигровой промышленности разработка видеоигр была процессом, очень похожим на разработку прикладного программного обеспечения, то есть большую часть стоимости разработки составляла разработка программного обеспечения, в отличие от разработки таких графических, звуковых и поведенческих элементов или "ресурсов", как те, которые могут быть разработаны для кинофильма с пространственными спецэффектами. В настоящее время многие программы работ по разработке усложненных видеоигр больше напоминают разработку кинофильмов со спецэффектами, чем разработку программного обеспечения. Например, многие видеоигры обеспечивают имитации 3D (трехмерных) миров и формируют все более фотореалистичные (т.е. компьютерную графику, которая кажется такой реалистичной, как кадр с изображением игры актеров, фотографически) персонажи, реквизит и окружающую обстановку. Одним из самых перспективных аспектов разработки фотореалистичных игр является создание сформированного компьютером человеческого лица, которое нельзя отличить от человеческого лица при игре актеров. Технологии захвата лица, такие как Contour™ Reality Capture, разработанный Mova, San Francisco, CA (Сан-Франциско, Калифорния), захватывают и отслеживают точную геометрию лица актера с высоким разрешением, когда оно находится в движении. Эта технология

обеспечивает возможность визуализации лица 3D на PC или игровой консоли, которое практически нельзя отличить от захваченного лица при игре актеров.

Наблюдение за медленным перемещением индикатора выполнения по экрану при выводе на экран PC или консоли сообщения наподобие "Loading..." ("Загрузка...") принимается современными пользователями как неотъемлемый недостаток сложных видеоигр. Задержка при загрузке следующей сцены с диска может занимать несколько секунд или даже несколько минут. Это является пустой тратой времени и может совершенно разочаровать игрока. Как обсуждалось ранее, большая часть или вся задержка может происходить из-за времени загрузки многоугольника, текстуры или других данных с диска, но также может случиться так, что часть времени загрузки тратится на подготовку процессором и/или GPU в PC или консоли данных для сцены. Соответственно, существующие PC и консольные системы, как правило, являются ограниченными как по сложности, так и по продолжительности воспроизведения данных сцен и страдают от долгого времени загрузки для сложных сцен. Другим существенным ограничением систем прикладного программного обеспечения и игровых видеосистем на предшествующем уровне техники является то, что они все чаще используют большие базы данных, например, таких 3D объектов, как многоугольники и текстуры, которые требуется загружать в PC или игровую консоль для обработки. Как обсуждалось выше, загрузка таких баз данных может занимать много времени при хранении на локальном диске. Время загрузки, однако, обычно намного больше, если база данных хранится удаленно и доступ к ней осуществляется через Internet. В таком случае могут потребоваться минуты, часы или даже дни для загрузки большой базы данных. Кроме того, создание таких баз данных часто связано с большими расходами, и они предназначены для продажи локальному конечному пользователю. Однако база данных подвергается риску несанкционированного использования после ее загрузки локальным пользователем. Во многих случаях пользователю требуется загрузка базы данных просто для ее оценки, чтобы проверить, удовлетворяет ли она его

потребностям. Длительное время загрузки может быть сдерживающим фактором для пользователя, оценивающего 3D базу данных до того, как принять решение о покупке. Аналогичные вопросы имеют место в ММОГ, в частности, как например, в играх, которые обеспечивают возможность пользователям использовать все более настраиваемых самостоятельно персонажей. Чтобы РС или игровая консоль выводили на экран персонажа, у них должен быть доступ к базе данных 3D геометрии (многоугольники, текстуры и т.д.), а также линий поведения (например, имеет ли персонаж щит, достаточно ли прочный этот щит, чтобы отклонить копьё или нет) для этого персонажа. Как правило, когда пользователь впервые ведет ММОГ, большое количество баз данных для персонажа уже поставляется с исходной копией игры, которая доступна локально на оптическом диске игры или загружается на диск. Но, по мере продвижения игры, если пользователь сталкивается с персонажем или объектом, база данных которого не доступна локально, то до того, как этот персонаж или объект могут быть выведены на экран, их база данных должна быть загружена. Это в результате может привести к значительной задержке игры.

С учетом уровня усложненности и сложности видеоигр другой проблемой для разработчиков видеоигр и издателей, связанной с видеоигровыми консолями предшествующего уровня техники, является то, что разработка видеоигры часто занимает от двух до трех лет при стоимости десятки миллионов долларов. С учетом того, что новые платформы видеоигровой консоли вводят со скоростью примерно по одной через каждые пять лет, разработчики игр должны начинать разработку этих игр за годы до выпуска новой игровой консоли, чтобы видеоигры появились в продаже одновременно с выпуском новой платформы. Иногда выпускают несколько консолей от конкурирующих производителей примерно в одно время (например, с интервалом один или два года), но еще неизвестна популярность каждой консоли, например какая консоль сделает самые большие объемы продаж программного обеспечения видеоигры. С учетом роста уровня инвестиций,

требуемых для видеоигр, производство игр все больше становится похожим на производство кинофильмов, и компании по производству игр обычно отдают свои производственные ресурсы исходя из своей оценки будущего успеха конкретной видеоигры. Но в отличие от кинокомпаний, эта ставка основана не просто на успехе самого производства, скорее она основана на успехе игровой консоли, на которой эта игра должна исполняться. Выпуск игры на многих консолях одновременно может уменьшить риск, но из-за этой дополнительной программы работ возрастает стоимость, и часто происходит задержка фактического выпуска игры. Условия работы пользователя и условия разработки прикладного программного обеспечения на PC требуют все большего объема вычислений, становятся динамичными и интерактивными, не только чтобы сделать их более привлекательными визуально для пользователей, но также и сделать их более полезными и интуитивно понятными. Усовершенствованные графические средства, например Maya™ от Autodesk, Inc., обеспечивают способность к воспроизведению динамических изображений и самой усложненной 3D визуализации, которые расширяют границы современных CPU и GPU. Однако вычислительные требования этих новых инструментальных средств вызывают несколько практических вопросов у пользователей и разработчиков программного обеспечения таких продуктов. Так как визуальное отображение операционной системы (OS) должно работать на широком диапазоне классов компьютеров - в том числе на компьютерах предшествующего поколения, которые больше не продаются, но на которых, тем не менее, можно заменить операционную систему (OS) на новую, - графические требования OS ограничены в значительной степени "наименьшим общим знаменателем" компьютеров, для которых предназначена эта OS, который, как правило, включает в себя компьютеры, которые не включают в себя GPU.

Это очень ограничивает функциональные возможности графических средств OS. Кроме того, портативные компьютеры с батарейным питанием (например, ноутбуки) ограничивают возможности визуального отображения, так как высокий уровень вычислительной активности CPU или GPU, как

правило, в результате приводит к более высокому уровню потребляемой мощности и меньшему времени работы аккумулятора. Портативные компьютеры, как правило, включают в себя программное обеспечение, которое автоматически понижает активность процессора для уменьшения потребляемой мощности, когда процессор не используется. В некоторых моделях компьютера пользователь может понизить активность процессора вручную. OS, под управлением которой работает портативный компьютер, должна быть пригодной к использованию даже в случае, если компьютер работает с производительностью, равной доле от своей максимальной производительности. Соответственно, производительность графической подсистемы OS часто остается гораздо ниже доступной на современном уровне техники вычислительной мощности. Приложения с широкими функциональными возможностями, требующие большого объема вычислений, подобные Maya, часто продают с ожиданием того, что они будут использоваться на высокопроизводительных PC. Этим, как правило, определяется требование к "наименьшему общему знаменателю" более дорогому и менее портативному с гораздо большей производительностью. Как следствие, у таких приложений гораздо более ограниченная целевая аудитория, чем у универсальной OS, и, как правило, их продают в гораздо меньшем объеме, чем программное обеспечение универсальной OS или универсальное прикладное программное обеспечение.

Потенциальная аудитория также ограничена, потому что потенциальному пользователю часто трудно испытать такие приложения, требующие большого объема вычислений, заранее. Это существенно ограничивает аудиторию для таких приложений с широкими функциональными возможностями. Это также способствует увеличению продажной цены, так как стоимость разработки обычно погашается гораздо меньшим количеством покупок, чем в случае универсального приложения. Дорогостоящие приложения также побуждают людей и компании к использованию "пиратских" копий прикладного программного обеспечения. В результате высокопроизводительное прикладное программное обеспечение страдает от сильно распространенного

пиратства, несмотря на существенные усилия издателей такого программного обеспечения уменьшить такое пиратство различными способами. Однако даже при использовании "пиратских" приложений с широкими функциональными возможностями пользователи не могут избежать необходимости вкладывать деньги в дорогостоящие современные РС для исполнения пиратских копий. Следовательно, несмотря на то что пользователи пиратского программного обеспечения могут использовать приложения за цену, равную доле от его фактической розничной цены, они, тем не менее, должны покупать или приобретать дорогостоящий РС для полного использования этого приложения. Это относится и к пользователям пиратских видеоигр с высокими характеристиками. Несмотря на то что пираты могут получить игры за цену, равную доле от их фактической цены, они, тем не менее, должны покупать дорогостоящую вычислительную технику, необходимую для ведения игры должным образом. С учетом того, что видеоигры, как правило, являются времяпрепровождением для потребителей, дополнительные затраты на высокопроизводительную игровую видеосистему могут являться препятствием. Эта ситуация ухудшается в странах (например, в Китае), где среднегодовой доход рабочих в настоящее время является довольно низким относительно среднегодового дохода рабочих в США. В результате гораздо меньший процент людей владеет высокопроизводительной игровой видеосистемой или высокопроизводительным РС. В таких странах весьма распространены Internet-кафе, в которых пользователи оплачивают использование компьютера, подключенного к Internet. Часто такие Internet-кафе имеют старую модель или низкопроизводительный РС без высокопроизводительных признаков, например GPU, которые могли бы иначе обеспечить возможность игрокам вести видеоигры, требующие большого объема вычислений. Это является решающим фактором успеха игр, которые исполняются на низкопроизводительных РС, например "World of Warcraft" медиаконгломерата Vivendi, которая является очень успешной в Китае, и там в нее обычно играют в Internet кафе. Напротив, в игру, требующую большого объема вычислений, подобную "Second Life", с гораздо

меньшей вероятностью можно играть на PC, установленном в китайском Internet-кафе. Такие игры фактически недоступны пользователям, которые имеют доступ только к низкопроизводительным PC в Internet-кафе. Препятствия также существуют для пользователей, которые рассматривают покупку видеоигры и хотели бы сначала испытать демонстрационную версию игры с загрузкой ее через Internet в свой домашний компьютер. Демонстрационная версия видеоигры часто является полнофункциональной версией игры с некоторыми недоступными признаками или с ограничениями, наложенными на количество проведенных игр. Может подразумеваться длительный процесс загрузки гигабайтов данных до того, как игра будет инсталлирована и исполняться на PC или на консоли. В случае PC может также подразумеваться выяснение того, какие специальные драйверы необходимы для этой игры, загрузка соответствующей версии, инсталляция их, и затем определение того, можно ли на PC вести эту игру. Этот последний этап может включать в себя определение того, существуют ли у PC достаточные возможности для обработки (CPU и GPU), достаточная RAM и совместимая OS. Соответственно, после длительного процесса при попытке запуска демонстрационной версии видеоигры пользователь может обнаружить, что демонстрационная версия видеоигры, возможно, не может быть воспроизведена, с учетом конфигурации PC пользователя. Хуже, если после загрузки новых драйверов пользователем для испытания демонстрационной версии эти версии драйверов могут оказаться несовместимыми с другими играми или приложениями, которые пользователь постоянно использует на PC, соответственно, инсталляция демонстрационной версии может привести ранее работающие игры или приложения в неработающее состояние. Эти препятствия не только разочаровывают пользователя, но они создают препятствия издателям программного обеспечения видеоигры и разработчикам видеоигр для сбыта своих игр.

Другая проблема, которая в результате приводит к экономической неэффективности, имеет отношение к тому, что данные PC или игровая консоль обычно предназначены для обеспечения определенного уровня требования к

производительности для приложений и/или игр. Например, некоторые PC имеют большую или меньшую RAM, более медленные или более быстрые CPU и более медленные или более быстрые GPU, если они вообще имеют GPU. Некоторые игры или приложения используют преимущества всей вычислительной мощности данного PC или консоли, в то время как многие игры или приложения не используют их. Если игре или приложению, которые выбирает пользователь, требуется производительность меньше максимальной производительности локального PC или консоли, то пользователь может зря потратить деньги на неиспользуемые признаки PC или консоли. В случае консоли производитель консоли может заплатить больше, чем было необходимо для финансирования затрат на консоль. Другая проблема, которая существует в сбыте и использовании видеоигр, включает в себя учет того, что пользователь просматривает другие игры до того, как он совершает покупку этой игры. В данное время существует несколько подходов для записи видеоигр для повторного воспроизведения в дальнейшем. Например, в некоторых технических решениях сообщается о записи информации о состоянии игры, включающей в себя операции игрового контроллера, в течение "геймплей" (процесс игры) в компьютере клиента видеоигры (принадлежащего идентичному или другому пользователю). Эта информация о состоянии может быть использована в дальнейшем для повторного воспроизведения некоторых или всех операций игры на компьютере клиента видеоигры (например, PC или консоли). Существенным недостатком этого подхода является то, что для того, чтобы пользователю просмотреть записанную игру, он должен иметь компьютер клиента видеоигры, который может воспроизводить игру, и должен иметь видеоигровое приложение, исполняющееся на этом компьютере так, что геймплей является идентичным при повторном воспроизведении состояния записанной игры. Кроме того, видеоигровое приложение должно быть написано таким способом, что не существует возможности различия исполнения между записанной игрой и повторно воспроизводимой игрой. Например, графическое представление игры обычно вычисляется на покaдровой основе. Для многих игр

игровая логика иногда может занимать меньше или больше времени, чем один период кадра, для вычисления графического представления, выводимого на экран для следующего кадра, в зависимости от того, является ли сцена особенно сложной или существуют ли другие задержки, которые замедляют исполнение. В такой игре может со временем встретиться "пороговый" кадр, который вычисляется за несколько меньшее время, чем один период кадра (скажем, на несколько тактов CPU меньше). Когда снова вычисляется идентичная сцена с использованием идентичной информации о состоянии игры, она может легко занять на несколько тактов CPU больше, чем один период кадра. Следовательно, когда игра повторно воспроизводится, кадр вычисляется в течение двух периодов кадра, а не в течение одного периода кадра. Некоторые линии поведения основаны на том, как часто в игре вычисляется новый кадр (например, когда игра производит выборку входных данных из игровых контроллеров). Когда ведется игра, это несоответствие в привязке ко времени для разных линий поведения не влияет на ведение игры, но это может в результате привести к тому, что при повторном воспроизведении игры выведется другой результат.

Соответственно, использование состояния игры для записи видеоигр требует очень осторожного проектирования игровой программы для обеспечения того, что повторное воспроизведение с использованием идентичной информации о состоянии игры выведет идентичный исход. Другим подходом предшествующего уровня техники для записи видеоигры является простая запись видеовыхода РС или игровой видеосистемы. Видео после этого может быть перемотано и повторно воспроизведено, или в качестве альтернативы, записанное видео может быть выгружено в Internet, как правило, после сжатия. Недостатком этого подхода является то, что когда последовательность кадров 3D игры повторно воспроизводится, пользователь ограничен просмотром этой последовательности кадров только с точки зрения, с которой эта последовательность кадров была записана. Другими словами, пользователь не может изменить точку обзора сцены.

Кроме того, когда сжатое видео записанной последовательности кадров игры, воспроизводимое на домашнем PC или игровой консоли, делают доступным другим пользователям через Internet, даже если это видео сжато в реальном времени, может оказаться невозможной выгрузка этого сжатого видео в реальном времени в Internet. Причиной, по которой это происходит, является то, что много домов в мире, которые подключены к Internet, имеют в высокой степени асимметричные широкополосные соединения. Сжатые видеопоследовательности с высоким разрешением часто имеют большие полосы пропускания, чем пропускная способность восходящего потока данных сети, что делает невозможной выгрузку их в реальном времени.

Соответственно, может иметь место существенная задержка после воспроизведения последовательности кадров игры (возможно, минуты или даже часы) до того, как другой пользователь в Internet сможет просмотреть игру. Несмотря на то что эта задержка допустима в определенных ситуациях (например, для просмотра выполнений игрока, которые происходили раньше), она делает невозможным просмотр игры в прямом эфире (например, соревнование по баскетболу с участием победителей) или с возможностью "повтора", когда игра воспроизводится в прямом эфире. Другой подход предшествующего уровня техники обеспечивает возможность зрителю, имеющему телевизор, просматривать видеоигры в прямом эфире, но только под управлением телевизионной производственной группы. Некоторые телевизионные каналы, как в США, так и в других странах, обеспечивают каналы для просмотра видеоигр, по которым телевизионная аудитория может смотреть определенных пользователей видеоигры (например, игроков с самым высоким рейтингом, участвующих в соревнованиях).

Это выполняется при наличии видеовыхода игровых видеосистем (PC и/или консолей), подаваемого в оборудование для обработки и распространения видео для телевизионного канала. После этого телевизионный канал может использовать свое оборудование для спецэффектов и обработки видео/аудио для манипуляции выходом из различных игровых видеосистем. Например,

телевизионный канал может накладывать текст поверх видео из видеоигры, в которой указывается статус разных игроков, и телевизионный канал может накладывать дополнительное аудио от комментатора, который может обсуждать действия, происходящие во время игр. Кроме того, видеоигровой выход может быть комбинирован с камерами, записывающими видео фактических игроков игр. Одной проблемой, связанной с этим подходом, является то, что такие линии передачи видео в прямом эфире должны быть доступны для оборудования телевизионного канала для обработки и распространения видео в реальном времени для того, чтобы у него было возбуждение прямой передачи. Однако, как обсуждалось ранее, это часто бывает невозможно осуществить, когда игровой видеосистемой управляют из дома, особенно если часть трансляции включает в себя видео в прямом эфире из камеры, которая захватывает реальное видео игрока. Кроме того, во время соревнования существует беспокойство, что игрок, находящийся дома, может модифицировать игру и обманывать, как описано выше. По этим причинам такие трансляции видеоигры по телевизионным каналам часто компонуют так, что игроки и игровые видеосистемы сосредоточены в общественном месте (например, в телевизионной студии или на арене), где телевизионное производственное оборудование может принимать линии передачи видео из множества игровых видеосистем и, возможно, камер для прямой передачи.

Несмотря на то, что такие видеоигровые телевизионные каналы предшествующего уровня техники могут обеспечить очень захватывающее представление для телевизионной аудитории, которое является практическим опытом, близким к реальному спортивному событию, например с видеоигроками, представленными как "спортсмены", и в смысле их действий в мире видеоигры, и в смысле их действий в реальном мире, эти игровые видеосистемы часто ограничены положениями, когда игроки находятся в непосредственной физической близости друг с другом. И так как телевизионные каналы транслируют, то каждый вещательный канал может показывать только один видеопоток, который выбирается производственной группой телевизионного

канала. Из-за этих ограничений и высокой стоимости эфирного времени, производственного оборудования и производственных групп по таким телевизионным каналам, как правило, показывают только игроков с самым высоким рейтингом, участвующих в самых важных соревнованиях. Кроме того, данный телевизионный канал, транслирующий полноэкранное изображение видеоигры всей телевизионной аудитории, показывает только одну видеоигру в данный момент времени. Это очень ограничивает возможности выбора телезрителя. Например, телезритель, возможно, не интересуется игрой(ами), показываемой(ыми) в данное время. Другой зритель может интересоваться только просмотром игры конкретного игрока, которую не показывают по телевизионному каналу в данное время. В других случаях зритель может интересоваться только просмотром того, как опытный игрок управляет конкретным уровнем в игре. Однако другие зрители могут захотеть управлять точкой обзора, с которой смотрят видеоигру, которая отличается от точки обзора, выбранной производственной группой и т.д. Одним словом, телезритель может иметь множество предпочтений при просмотре видеоигр, которые не обеспечиваются конкретной трансляцией по телевизионной сети, даже если доступны несколько разных телевизионных каналов.

1.3 Недостатки современных видеоустройств

По всем вышеупомянутым причинам существующие сейчас видеоигровые телевизионные каналы имеют существенные ограничения в представлении видеоигр телезрителям. Другой недостаток игровых видеосистем предшествующего уровня техники и систем прикладного программного обеспечения состоит в том, что они являются сложными и обычно страдают от ошибок, аварийных отказов и/или непреднамеренных и нежелательных линий поведения (обобщенно "ошибки"). Несмотря на то что игры и приложения, как правило, проходят через процесс настройки и отладки (часто называемый "гарантии качества программного обеспечения" или SQA) перед выпуском, почти неизменно, после выпуска игры или приложения для широкой аудитории, в условиях эксплуатации неожиданно возникают ошибки. К сожалению,

разработчику программного обеспечения трудно идентифицировать и обнаружить многие из ошибок после выпуска. Разработчикам программного обеспечения может быть трудно узнать об ошибках.

Еще одной проблемой, связанной с РС и игровыми консолями, является то, что они подлежат обслуживанию, что причиняет большое неудобство потребителю. Вопросы обслуживания также влияют на производителя РС или игровой консоли, так как они, как правило, обязаны отправлять специальный ящик для безопасной перевозки сломанных РС или консоли и затем нести расходы по ремонту, если РС или консоль находятся на гарантии. На издателя прикладного программного обеспечения или игры может также влиять потеря продаж (или использование онлайн-обслуживания) вследствие того, что РС и/или консоли находятся в состоянии ремонта.

На рис.1.1 изображена игровая видеосистема предшествующего уровня техники, например Sony Playstation® 3, Microsoft Xbox 360®, Nintendo Wii™, персональный компьютер на базе Windows или Macintosh корпорации Apple. Каждая из этих систем включает в себя центральный процессор (CPU) для исполнения управляющей программы, как правило, графический процессор (GPU) для выполнения усовершенствованных графических операций и разнообразные виды ввода/вывода (I/O) для обмена информацией с внешними устройствами и пользователями. Для простоты эти компоненты изображены объединенными вместе как один блок. Также изображено, что игровая видеосистема предшествующего уровня техники по рис.1.1 включает в себя накопитель 5 на оптических носителях информации (например, накопитель высокой емкости для дисков), накопитель 4 на жестких дисках для хранения данных и управляющей программы видеоигры, сетевое соединение 6 для ведения игр с несколькими участниками, для загрузки игр, заплат, демонстрационных версий или другого мультимедиа, оперативное запоминающее устройство (RAM) 2 для хранения управляющей программы, в настоящее время исполняемой CPU/GPU 1, игровой контроллер 7 для приема входных команд от

пользователя во время геймплей и дисплей 3 (например, SDTV/HDTV или компьютерный монитор).

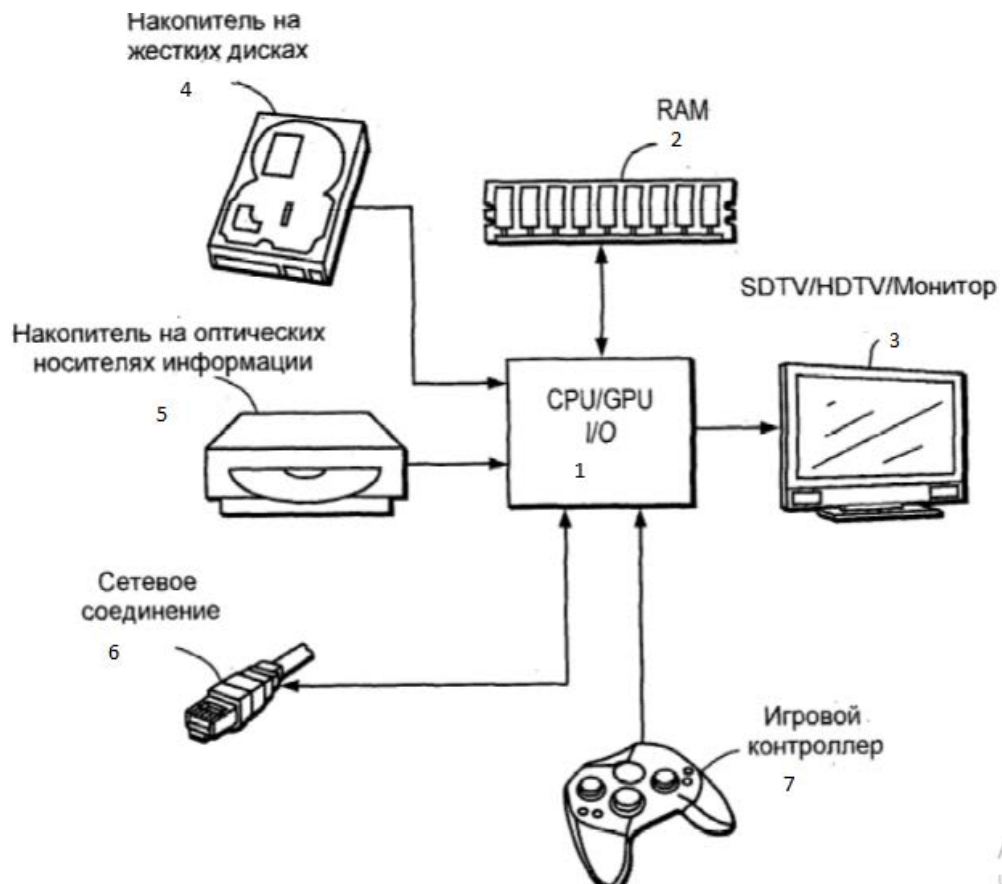


Рисунок 1.1 – Схема архитектуры игровой видеосистемы предшествующего уровня техники

Система предшествующего уровня техники, изображенная на рис.1.1, имеет несколько ограничений. Во-первых, накопители 4 на оптических дисках и накопители 3 на жестких дисках, как правило, имеют гораздо более медленные скорости доступа по сравнению со скоростью доступа RAM 1. При работе непосредственно через RAM 1 CPU/GPU 100 может на практике обрабатывать гораздо больше многоугольников в секунду, чем это возможно, когда управляющая программа и данные считываются непосредственно с накопителя 3 на жестких дисках или накопителя 4 на оптических дисках вследствие того, что RAM 1 обычно имеет гораздо большую полосу пропускания и не страдает от задержек дисковых механизмов относительно долгого поиска. Но в этих

системах предшествующего уровня техники обеспечен только ограниченный объем RAM (например, 256-512 МБайт). Следовательно, часто требуется последовательность кадров "Загрузка...", во время которой RAM 1 периодически заполняется данными для следующей сцены видеоигры.

В некоторых системах делается попытка совмещать загрузку управляющей программы одновременно с геймплей, но это можно сделать, только когда известна последовательность событий (например, если автомобиль едет по дороге, то может быть загружена геометрия для приближающиеся зданий, находящихся на краю дороги, в то время как автомобиль едет). Для сложной и/или быстрой смены сцен этот тип совмещения обычно не работает. Например, в случае когда пользователь находится в середине сражения и RAM 1 полностью заполнена данными, представляющими объекты, находящиеся в поле зрения в этот момент, если пользователь быстро перемещает изображение влево для просмотра объектов, которые в настоящее время не загружены в RAM 1, то в результате произойдет нарушение непрерывности в действии, так как не будет достаточно времени, чтобы загрузить новые объекты из накопителя 3 на жестких дисках или оптического носителя 4 информации в RAM 1.

Другая проблема, связанная с системой по рис.1.1, возникает из-за ограничений емкости памяти накопителей 3 на жестких дисках и оптического носителя информации 4. Несмотря на то что дисковые запоминающие устройства могут быть изготовлены с относительно большой емкостью памяти, они, тем не менее, не обеспечивают достаточное количество емкости памяти для определенных сценариев, встречающихся в современных видеоиграх.

Другая проблема, связанная с такими подходами предшествующего уровня техники, как тот, который изображен на рис.1.1, состоит в том, что со временем видеоигры становятся более усовершенствованными и требуют большей вычислительной мощности CPU/GPU.

Соответственно, даже при предположении неограниченного объема RAM требования к аппаратному обеспечению видеоигры превышают

максимальный уровень вычислительной мощности, имеющейся в этих системах. В результате пользователи должны модернизировать игровую аппаратуру через каждые несколько лет, чтобы не отставать (или вести более новые игры при более низких уровнях качества). Одним последствием тенденции еще большего усовершенствования видеоигр является то, что домашние видеоигровые вычислительные машины, как правило, являются экономически неэффективными, потому что их стоимость обычно определяется в соответствии с требованиями игры самой высокой производительности, которую они могут поддерживать.

Видеоигровые вычислительные машины, как правило, выключены в течение большего количества времени в неделю. Согласно исследованию Nielsen Entertainment в июле 2006 г. активных игроков в возрасте 13 лет и старше, в среднем, активные игроки проводят четырнадцать часов в неделю за игрой в консольные видеоигры, или только 12% всего времени в неделю. Это означает, что средняя видеоигровая консоль не используется 88% времени, что является неэффективным использованием дорогостоящего ресурса. Это особенно важно с учетом того, что видеоигровые консоли часто финансируются производителем для снижения покупной цены. Видеоигровые консоли также несут расходы, связанные с почти любым потребительским электронным устройством. Например, электроника и механизмы систем должны быть смонтированы в корпусе. Производитель должен предоставлять гарантию на сервисное обслуживание. Розничный торговец, который продает систему, должен получать прибыль от любой продажи системы и/или от продажи программного обеспечения видеоигры. Все эти факторы добавляют к стоимости видеоигровой консоли, которая должна финансироваться или производителем, с передачей потребителю, или обоими. Кроме того, основной проблемой видеоигровой промышленности является пиратство. Механизмы безопасности, используемые фактически на каждой основной игровой видеосистеме, "взламывают" со временем, что в результате приводит к несанкционированному копированию видеоигр. Игры, которые можно

загружать (например, игры для PC или Mac), являются особенно уязвимыми для пиратства. В определенных регионах мира, где пиратство слабо контролируется, по существу, не существует жизнеспособного рынка для автономной программного обеспечения видеоигры, потому что пользователи могут покупать пиратские копии так же легко, как и легальные копии за цену, равную очень незначительной доле от цены. Кроме того, во многих частях мира цена игровой консоли составляет такой высокий процент от дохода, что даже если бы пиратство контролировали, то мало людей могли бы позволить себе современную игровую систему.

Кроме того, рынок подержанных игр уменьшает доход видеоигровой промышленности. Когда пользователь теряет интерес к игре, он может продать игру розничному торговцу, который перепродает эту игру другим пользователям. Это несанкционированная, но обычная практика значительно уменьшает доходы издателей игр. Аналогично, обычно происходит сокращение продаж приблизительно на 50% при переходе на другую платформу каждые несколько лет. Это происходит потому, что пользователи перестают покупать игры для старых платформ, когда они узнают, что скоро должна быть выпущена новая версия платформы (например, когда собираются выпустить Playstation 3, пользователи перестают покупать игры для Playstation 2). Вместе взятые, уменьшение продаж и увеличение затрат на разработку, связанные с новыми платформами, могут оказать очень существенное неблагоприятное влияние на рентабельность разработчиков игр.

Новые игровые консоли также являются очень дорогими. Xbox 360, Nintendo Wii и Sony Playstation 3 - все в розницу продаются за сотни долларов. Мощные игровые системы персональных компьютеров могут стоить до 8000 долларов. Это представляет для пользователей существенное вложение денег, особенно с учетом того, что аппаратное обеспечение устаревает после нескольких лет, и того, что многие системы покупают для детей. Одним подходом к решению вышеизложенных проблем являются онлайн-игры, в которых программа, управляющая игрой, и данные размещены на сервере и

доставляются в клиентские машины по требованию как сжатое видео и аудио, которые передаются потоком по широкополосной цифровой сети. Основным недостатком этих систем является проблема времени ожидания, т.е. время, которое требуется сигналу для прохождения до и от игрового сервера, который обычно расположен в "центральной станции" оператора.

Динамичные видеоигры-боевики (также известные как видеоигры "twitch" ("подергивания")) требуют очень малого времени ожидания между временем, когда пользователь выполняет операцию с игровым контроллером, и временем, когда экран дисплея обновляется и показывает результат операции пользователя. Малое время ожидания требуется для того, чтобы у пользователя было ощущение, что игра отвечает "немедленно". Пользователи могут быть удовлетворены разными интервалами времени ожидания, в зависимости от типа игры и уровня мастерства пользователя.

Когда игровой сервер или сервер приложений установлены в близлежащей управляемой сетевой среде или там, где сетевой тракт до пользователя является предсказуемым и/или он может выдерживать пики полосы пропускания, намного проще управлять временем ожидания, и с точки зрения максимального времени ожидания, и с точки зрения постоянства времени ожидания. Такого уровня управления можно достичь между распределительным устройством сети кабельного телевидения до дома абонента кабельного телевидения, или от центральной АТС DSL до дома абонента DSL, или в среде коммерческой учрежденческой локальной сети (LAN) от сервера или пользователя. Также можно получить специально разделенные частные соединения точка-точка между компаниями, которые имеют гарантированную полосу пропускания и время ожидания. Но в игровой или прикладной системе, которая размещает игры в серверном центре, подключенном к общему Internet, и затем передает потоком сжатое видео пользователю через широкополосное соединение, многие факторы воздействуют на время ожидания, что в результате приводит к серьезным ограничениям в применении систем предшествующего уровня техники. В типичном доме,

подключенном к широкополосному каналу, пользователь может иметь DSL или кабельный модем для услуг широкополосного доступа. Для таких услуг широкополосного доступа время передачи сигнала туда и обратно между домом пользователя и общим Internet обычно составляет 25 мс (а иногда и больше). Кроме того, существуют величины времени ожидания передачи сигнала туда и обратно, которые являются следствием маршрутизации данных через Internet в серверный центр. Время ожидания передачи через Internet меняется в зависимости от маршрута, который задается данным, и задержек, которые происходят вследствие маршрутизации по нему. В дополнение к задержкам маршрутизации, время передачи сигнала туда и обратно также является следствием скорости света, проходящего через светопровод, который соединяет большую часть Internet. Например, на каждые 1000 миль (1600 км) время передачи сигнала туда и обратно приблизительно равно 22 мс вследствие скорости 30 света через светопровод и других потерь. Дополнительное время ожидания может произойти из-за скорости передачи данных, передаваемых потоком через Internet. Например, если пользователю предоставляется услуга DSL, которую продают как "услугу DSL 6 мегабит в секунду", то на практике пользователь, вероятно, получит меньше 5 мегабит в секунду пропускной способности нисходящего потока данных в лучшем случае и, вероятно, будет наблюдать, что соединение периодически ухудшается из-за различных факторов, например перегрузки во время периодов пиковой нагрузки в мультиплексоре доступа к цифровой абонентской линии (DSLAM). Аналогичный вопрос может возникнуть с уменьшением скорости передачи данных кабельного модема, используемого для соединения, проданного как "услуга кабельной модемной связи 6 мегабит в секунду", до скорости, гораздо меньшей, чем указанная, если произойдет перегрузка в локальном общем коаксиальном кабеле, образующем линию связи через окрестности, или в другом месте в сети системы кабельных модемов. Если пакеты данных с устойчивой скоростью 4 Мбит/с будут передаваться потоком в одном направлении в формате протокола передачи дейтаграмм пользователя (UDP) от серверного центра через такие соединения,

если все будет работать исправно, то пакеты данных будут передаваться без дополнительного времени ожидания, но в случае перегрузки (или других препятствий) и если потоковым данным к пользователю доступно только 3,5 Мбит/с, то в типичной ситуации или пакеты будут сброшены, что в результате приведет к потере данных, или пакеты будут стоять в очереди в точке перегрузки, пока они не будут отправлены, следовательно, с введением дополнительного времени ожидания. В разных точках перегрузки существует разная вместимость очереди для хранения задерживаемых пакетов, соответственно, в некоторых случаях пакеты, которые не могут пройти через точку перегрузки, немедленно сбрасываются. В других случаях в очереди стоят несколько мегабитов данных и в итоге отправляются. Но почти во всех случаях очереди в точках перегрузки ограничены по вместимости, и после превышения этих ограничений очереди переполняются, и пакеты сбрасываются.

1.4 Постановка задачи

Соответственно, для устранения дополнительного времени ожидания (или, что еще хуже, потери пакетов) необходимо устранение превышения возможностей скорости передачи данных из игрового сервера или сервера приложений пользователю. Время ожидания также является следствием времени, требуемого для сжатия видео в сервере и восстановления сжатого видео в клиентском устройстве. Время ожидания также имеет место, когда в видеоигре, исполняющейся на сервере, вычисляется следующий кадр, который должен быть выведен на экран. Имеющиеся в настоящее время доступные алгоритмы сжатия видео страдают или из-за высоких скоростей передачи данных, или из-за большого времени ожидания. Например, алгоритм сжатия движущихся изображений MJPEG является только-внутрикадровым алгоритмом сжатия с потерями, который отличается малым временем ожидания. Каждый кадр видео сжимается независимо от каждого другого кадра видео. Когда клиентское устройство принимает кадр видео, сжатого по алгоритму сжатия движущихся изображений MJPEG, то оно может немедленно восстановить сжатый кадр и вывести его на экран с получающимся в результате очень малым временем

ожидания. Но из-за того, что каждый кадр сжимается отдельно, упомянутый алгоритм не может использовать сходные черты между последовательными кадрами, и в результате только внутрикадровые алгоритмы сжатия видео страдают от очень высоких скоростей передачи данных. Например, видео 640×480 , 60 кадр/с (кадров в секунду), сжатое по алгоритму сжатия движущихся изображений MJPEG, может потребоваться 40 Мбит/с (мегабит в секунду) или большая (скорость передачи) данных. Такие высокие скорости передачи данных для видеокна с таким низким разрешением требуют чрезмерно больших затрат во многих широкополосных приложениях (и конечно для большинства потребительских приложений, основанных на Internet-технологиях). Кроме того, из-за того, что каждый кадр сжимается независимо, артефакты в кадрах, которые могут получиться в результате сжатия с потерями, вероятно, будут появляться в разных местах в последовательных кадрах. Это может в результате привести к тому, что для зрителя выглядит как движущиеся визуальные артефакты, когда восстанавливается сжатое видео. Другие алгоритмы сжатия, например MPEG2, H.264 или VC9 от корпорации "Майкрософт", как они используются в конфигурациях предшествующего уровня техники, могут достигать высоких коэффициентов сжатия, но за счет большого времени ожидания. Такие алгоритмы используют межкадровое, а также внутрикадровое сжатие. Периодически такие алгоритмы выполняют только внутрикадровое сжатие кадра. Такой кадр известен как ключевой кадр (как правило, называемый I-кадр). После этого эти алгоритмы, как правило, сравнивают I-кадр с предшествующими кадрами и с последующими кадрами. Вместо того чтобы сжимать предшествующие кадры и последующие кадры независимо, упомянутый алгоритм определяет то, что изменилось в изображении предшествующих и последующих кадров по сравнению с I-кадром, и после этого сохраняет эти изменения как так называемые B-кадры для изменений, предшествующих I-кадру, и P-кадры для изменений, следующих за I-кадром. Это в результате приводит к гораздо меньшим скоростям передачи данных, чем только-внутрикадровое сжатие. Но это, как правило, достигается за счет

большого времени ожидания. I-кадр обычно намного больше, чем В-кадр или Р-кадр (часто в 10 раз больше), и в результате его передача занимает пропорционально больше времени при данной скорости передачи данных.

РАЗДЕЛ 2. Специальная часть

Рассмотрим, например, ситуацию, когда I-кадры в 10 раз больше В-кадров и Р-кадров и существует 29 В-кадров+30 Р-кадров=59 промежуточных кадров для каждого отдельного I-интракадра, или всего 60 кадров для каждой "Группы Кадров" (GOP). Соответственно, при 60 кадр/с каждую секунду существует 1 60-кадровая GOP. Предположим, что максимальная скорость передачи данных на канале передачи равна 2 Мбит/с. Для получения видео высшего качества в этом канале алгоритм сжатия может выводить поток данных 2 Мбит/с, и с учетом вышеупомянутых коэффициентов в результате этого получается 2 мегабита (Мбит)/(59+10)=30 394 бита в каждом интракадре и 303 935 битов в каждом I-кадре. Когда алгоритм восстановления сжатых данных принимает сжатый видеопоток, для устойчивого воспроизведения видео, требуется восстанавливать каждый сжатый кадр и выводить их на экран с постоянным интервалом (например, 60 кадр/с). Для получения этого результата если для какого-либо кадра существует время ожидания передачи, то все кадры должны быть задержаны, по меньшей мере, на время этого времени ожидания, соответственно, самое большое время ожидания кадра определяет время ожидания для каждого видеокадра. I-кадры вводят самые большие величины времени ожидания передачи, так как они являются наибольшими, и весь I-кадр должен быть принят до того, как может быть восстановлен сжатый I-кадр и выведен на экран (или любой промежуточный кадр, зависящий от I-кадра). С учетом того, что скорость передачи данных по каналу равна 2 Мбит/с, передача I-кадра займет $303\,935/2$ Мбит = 145 мс.

2.1 Краткое описание предлагаемого технического решения

В системе межкадрового сжатия видео, как описано выше, с использованием большого процента полосы пропускания канала передачи существуют большие величины времени ожидания из-за большого размера I-кадра относительно среднего размера кадра. Или, другими словами, несмотря на то что при алгоритмах межкадрового сжатия предшествующего уровня техники достигается меньшая средняя покадровая скорость передачи данных,

чем при только-внутрикадровых алгоритмах сжатия (например, 2 Мбит/с по сравнению с 40 Мбит/с), они, тем не менее, страдают от высокой пиковой покадровой скорости передачи данных (например, $303935 \times 60 = 18,2$ Мбит/с) из-за больших I-кадров. Однако следует принимать во внимание то, что вышеупомянутый анализ предполагает, что все P-кадры и B-кадры намного меньше I-кадров. Несмотря на то что это в общем справедливо, это не справедливо для кадров с высоким уровнем сложности изображения, некоррелированных с предшествующим кадром, с большим перемещением или со сменами сцен. В таких ситуациях P-кадры или B-кадры могут становиться такими большими, как I-кадры (если P-кадр или B-кадр становится больше I-кадра, то усложненный алгоритм сжатия, как правило, "обеспечивает" I-кадр и заменяет P-кадр или B-кадр I-кадром). Соответственно, в цифровом видеопотоке в любой момент могут иметь место пики скорости передачи данных, имеющие размер I-кадра. Соответственно, в случае со сжатым видео, когда средняя скорость передачи данных видео приближается к пропускной способности скорости передачи данных каналов передачи (как это часто происходит, с учетом требований высокой скорости передачи данных для видео), высокие пиковые скорости передачи данных из-за I-кадров или больших P-кадров или B-кадров в результате приводят к большому времени ожидания кадра.

Несомненно, в вышеизложенном обсуждении описывается только время ожидания алгоритма сжатия, создаваемое большими B-кадрами, P-кадрами или I-кадрами в GOP. Если используются B-кадры, то время ожидания будет еще больше. Причиной, по которой это происходит, является то, что до того, как B-кадр может быть выведен на экран, все B-кадры, следующие за B-кадром и I-кадром, должны быть приняты. Соответственно, в последовательности группы изображений (GOP), например, BBBBBIPPPPBVVVBIPPPPP, в которой существуют 5 B-кадров перед каждым I-кадром, первый B-кадр не может быть выведен на экран устройством восстановления сжатого видео до тех пор, пока не будут приняты последующие B-кадры и I-кадр.

Соответственно, если видео передают потоком со скоростью 60 кадр/с (т.е. 16,67 мс/кадр), то до того, как может быть восстановлен сжатый первый В-кадр, потребуется $16,67 \times 6 = 100$ мс для приема пяти В-кадров и I-кадра, независимо от скорости полосы пропускания канала, и это только с 5 В-кадрами. Сжатые видеопоследовательности с 30 В-кадрами являются довольно распространенными. И при небольшой полосе пропускания канала, как 2 Мбит/с, влияние времени ожидания, вызванного размером I-кадра, является большим дополнением к влиянию времени ожидания из-за ожидания поступления В-кадров. Соответственно, на канале 2 Мбит/с с большим количеством В-кадров довольно просто превысить время ожидания 500 мс или больше с использованием технологии сжатия видео предшествующего уровня техники. Если В-кадры не используются (за счет меньшего коэффициента сжатия для данного уровня качества), то В-кадр не вызывает время ожидания, но по-прежнему время ожидания вызывают пиковые размеры кадра, описанные выше. Проблема усугубляется самим характером многих видеоигр. Алгоритмы сжатия видео, использующие структуру GOP, описанную выше, в значительной степени оптимизированы для использования с видео в прямом эфире или материалом кинофильма, предназначенного для пассивного просмотра. Как правило, камера (либо реальная камера или виртуальная камера в случае создаваемой компьютером анимации) и сцена являются относительно устойчивыми, просто потому, что если камера или сцена перемещаются слишком отрывисто, то материал фильма или видео (а) как правило, не приятно смотреть, и (б) если его смотрят, то обычно зритель не внимательно следит за действием, когда камера внезапно резко поворачивается (например, если камера ударяется при съемке, когда ребенок задувает свечи на торте ко дню рождения, и внезапно резко поворачивается в сторону от торта и возвращается обратно, то зрители, как правило, сосредоточены на ребенке и торте и не обращают внимания на краткое прерывание, когда камера внезапно перемещается). В случае видеointервью или видеотелеконференции камера может удерживаться в фиксированном положении и совсем не двигаться, что в результате приводит к

совсем очень маленькому количеству пиков данных. Но 3D видеоигра с высокой активностью отличается постоянным движением (например, рассмотрим 3D гонки, где весь кадр находится в быстром движении в течение гонки, или рассмотрим игры-боевики от первого лица, где виртуальная камера постоянно перемещается рывками). Такие видеоигры могут в результате привести к последовательностям кадров с большими и частыми пиками, где пользователю может быть необходимо ясно видеть то, что происходит в течение этих внезапных движений. Как таковые артефакты сжатия гораздо менее приемлемы в 3D видеоиграх с высокой активностью.

Соответственно, видеовыход многих видеоигр, вследствие их характера, выводит сжатый видеопоток с очень высокими и частыми пиками. С учетом того, что пользователи динамичных видеоигр-боевиков не терпимы к большому времени ожидания, и с учетом всех вышеупомянутых причин времени ожидания до настоящего времени существовали ограничения для размещенных на сервере видеоигр, которые передавали потоком видео в Internet. Кроме того, пользователи приложений, которые требуют высокой степени интерактивности, страдают от аналогичных ограничений, если эти приложения размещены в общем Internet и передают видео потоком. Такие услуги требуют конфигурации сети, в которой хостинговые серверы устанавливаются непосредственно в распределительном устройстве (в случае кабельного широкополосного канала), или в центральной АТС (в случае цифровых абонентских линий (DSL)), или в пределах LAN (или на специально разделенном частном соединении) в коммерческой обстановке так, что маршрутом и расстоянием от клиентского устройства до сервера управляют для минимизации времени ожидания, и пики могут быть адаптированы так, что они не будут вызывать время ожидания. Сети LAN (как правило, имеющие скорость в пределах 100 Мбит/с-1 Гбит/с) и арендуемые линии связи с соответствующей полосой пропускания, как правило, могут поддерживать требования по пиковой полосе пропускания (например, пиковая полоса пропускания 18 Мбит/с равна незначительной доле от пропускной способности LAN 100 Мбит/с).

Требования по пиковой полосе пропускания могут также быть обеспечены инфраструктурой широкополосной сети, связанной с жилыми домами, если выполнены специальные адаптации. Например, в системе кабельного телевидения цифровому видеотрафику может быть предоставлена выделенная полоса пропускания, которая может обрабатывать пики, например большие I-кадры. И в системе DSL может быть обеспечен модем DSL с более высокой скоростью, с учетом больших пиков, или может быть обеспечено специально разделенное соединение, которое может обрабатывать данные при более высоких скоростях передачи. Но общепринятый кабельный модем и инфраструктура DSL, подсоединенные к общему Internet, имеют гораздо меньший допуск для требований по пиковой полосе пропускания для сжатого видео.

Соответственно, онлайн-службы, которые размещают видеоигры или приложения в серверных центрах, находящихся на большом расстоянии от клиентских устройств, и затем передают потоком сжатый видеовыход по Internet через обычные широкополосные соединения, связанные с жилыми домами, страдают от существенного времени ожидания и ограничений по пиковой полосе пропускания - особенно в отношении игр и приложений, которые требуют очень малого времени ожидания (например, игры-боевики от первого лица и другие многопользовательские, интерактивные игры-боевики или приложения, требующие малого времени отклика).

Таким образом, в данной технической разработке необходимо обеспечить малое время ожидания для удаленного воспроизведения и/или использования пользователями игр и приложений, исполняемых на серверах службы хостинга, посредством передачи потокового сжатого интерактивного видео в клиентские устройства пользователей по сети, такой как сеть Интернет, даже при ограниченной полосе пропускания.

Для этого обеспечивается служба хостинга, содержащая множество распределенных географически серверных центров службы хостинга, каждый из которых содержит серверы, которые, при исполнении игры или приложения,

формируют поток сжатого с малым временем ожидания интерактивного потокового видео и/или аудио. Также обеспечивается средство для групповой передачи упомянутого потока во множество клиентских устройств с адаптацией к разным видео, аудио и/или сетевым характеристикам отдельного клиентского устройства и средство для обеспечения возможности пользовательского ввода из одного или нескольких клиентских устройств в службу хостинга, для управления интерактивным потоковым видео.

На примере рисунков рассмотрим краткое описание предлагаемого технического решения.

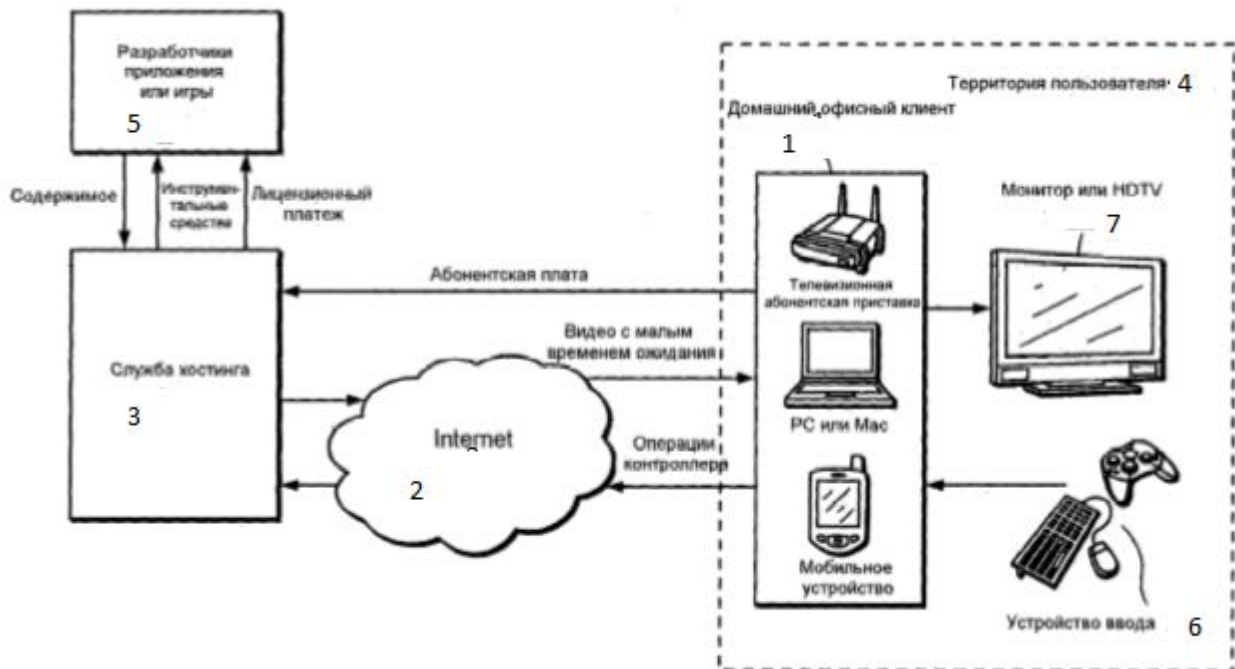


Рисунок 2.1 – Первый вариант схемы архитектуры системы высокого уровня

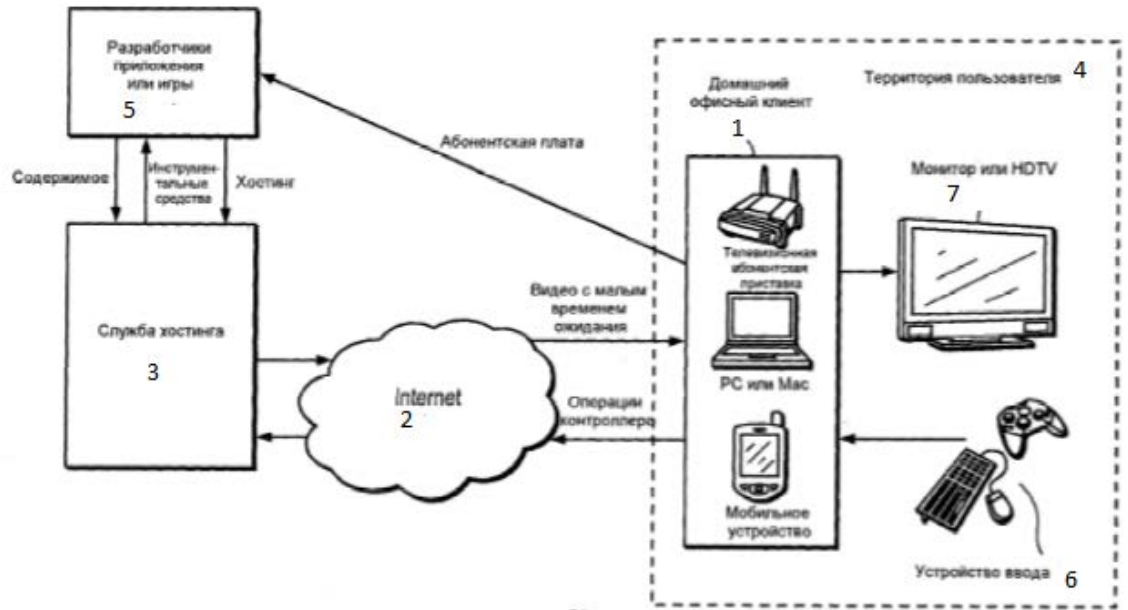


Рисунок 2.2 – Второй вариант схемы архитектуры системы высокого уровня

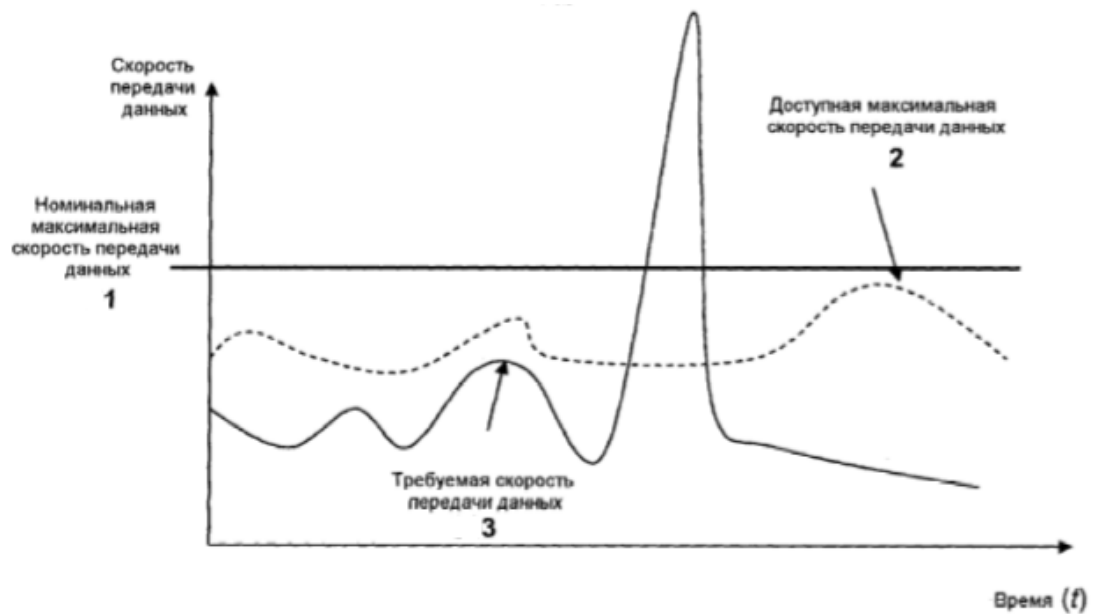


Рисунок 2.3 – Схема фактических, оцененных и требуемых скоростей передачи данных для передачи информации между клиентом и сервером

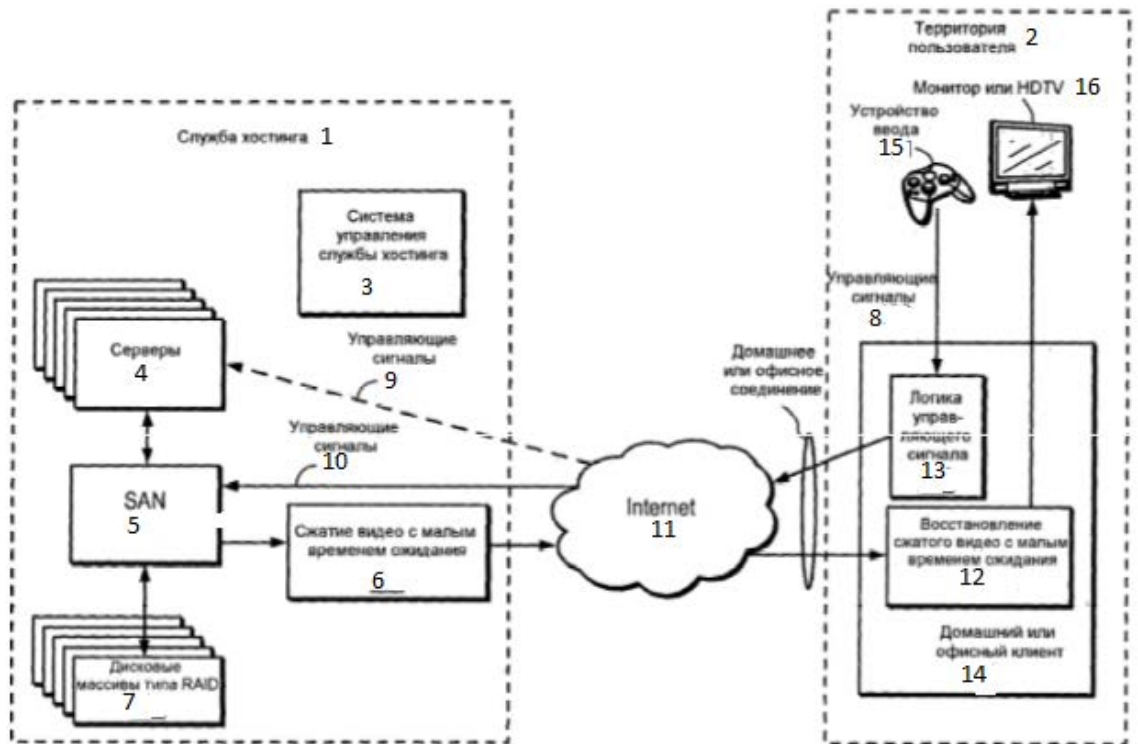


Рисунок 2.4 – Схема службы хостинга и клиента, применяемые согласно одному варианту осуществления

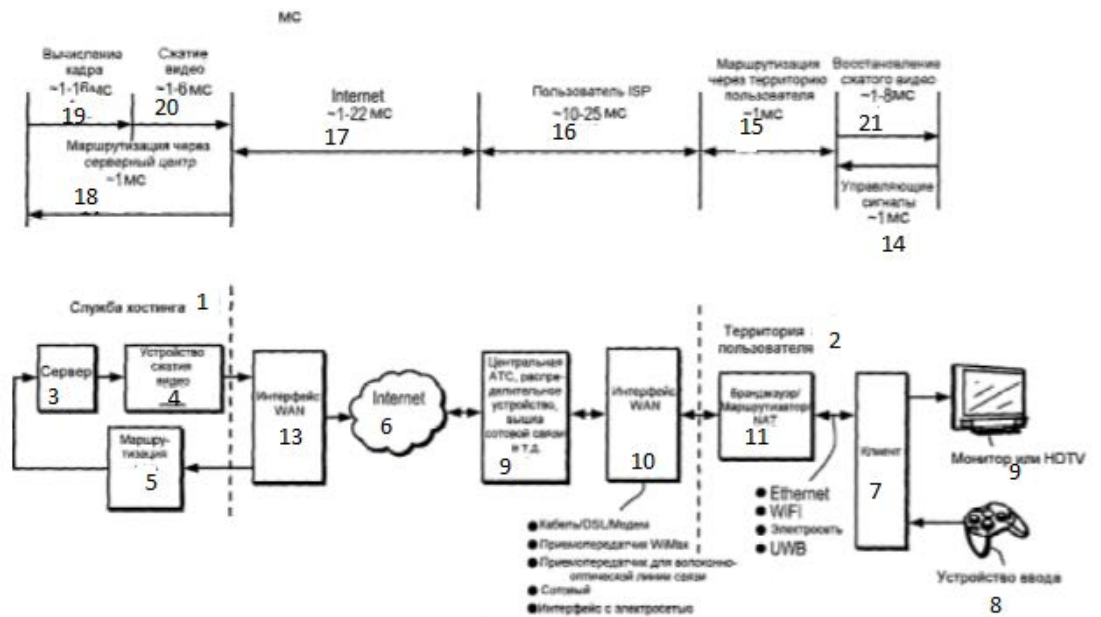


Рисунок 2.5 – Схема иллюстративных величин времени ожидания, связанных с передачей информации между клиентом и службой хостинга

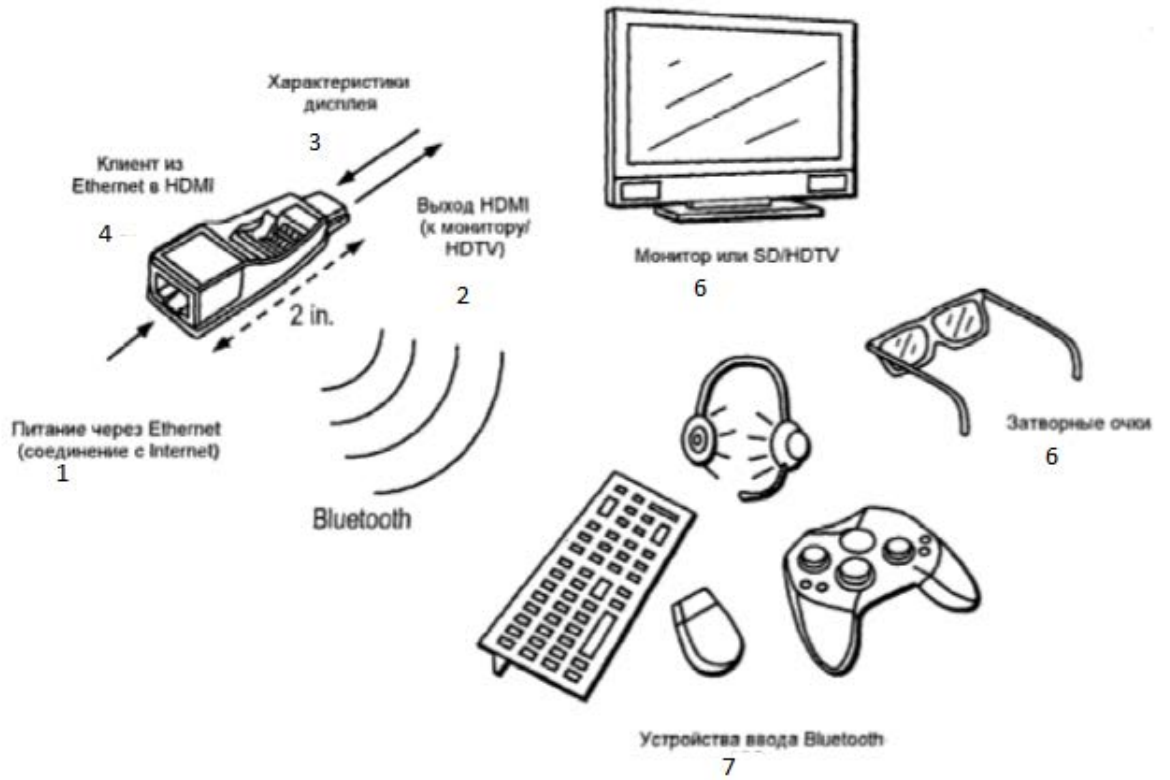


Рисунок 2.6 - Схема клиентского устройства согласно одному варианту осуществления

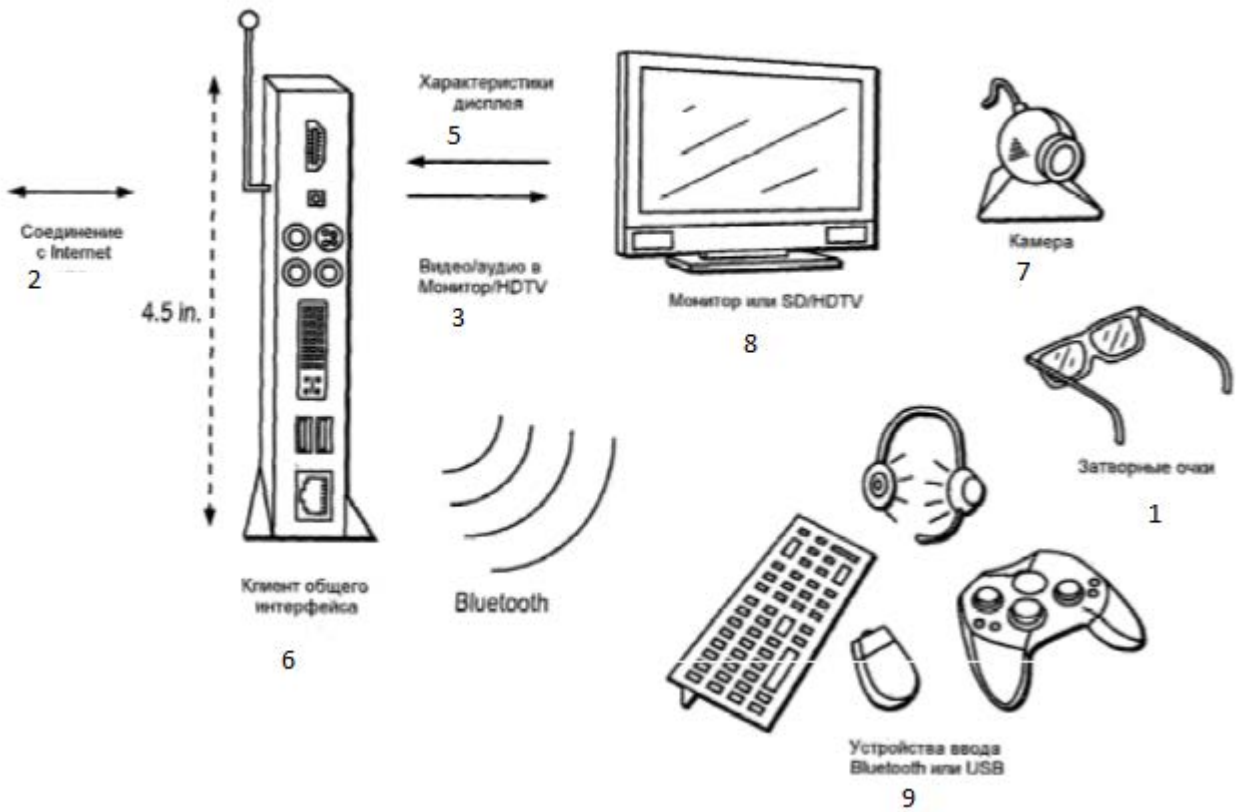


Рисунок 2.7 - Схема клиентского устройства согласно другому варианту осуществления

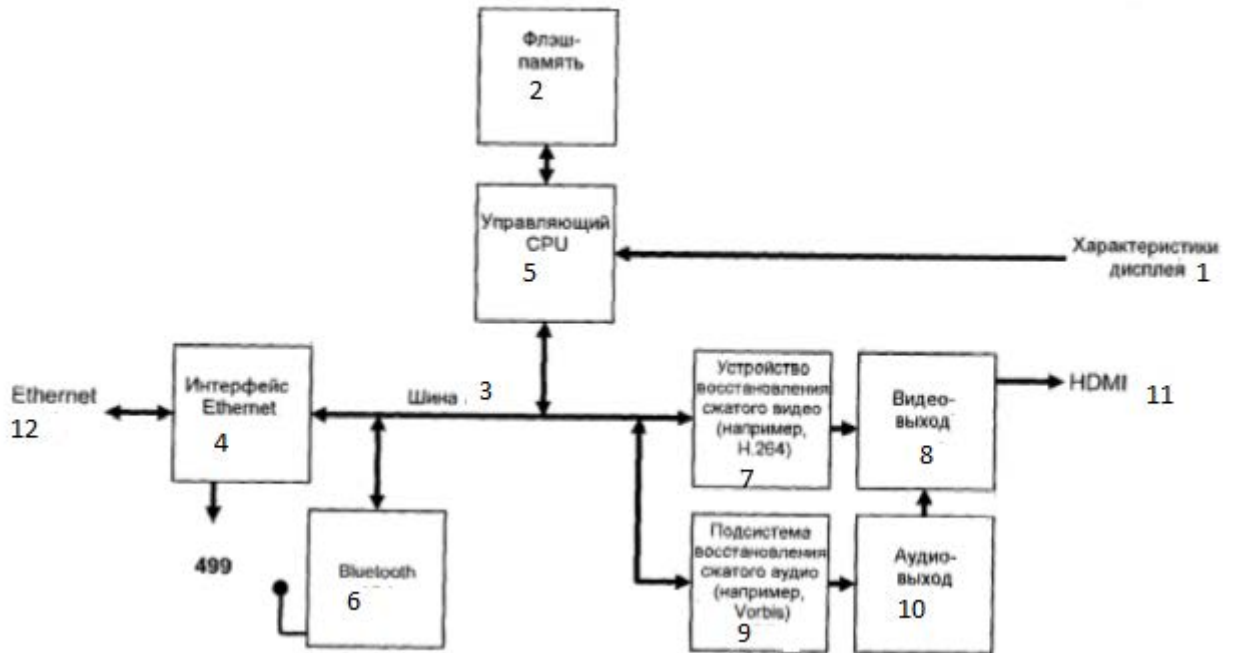


Рисунок 2.8 - Схема иллюстративной блок-схемы клиентского устройства по рис. 2.6

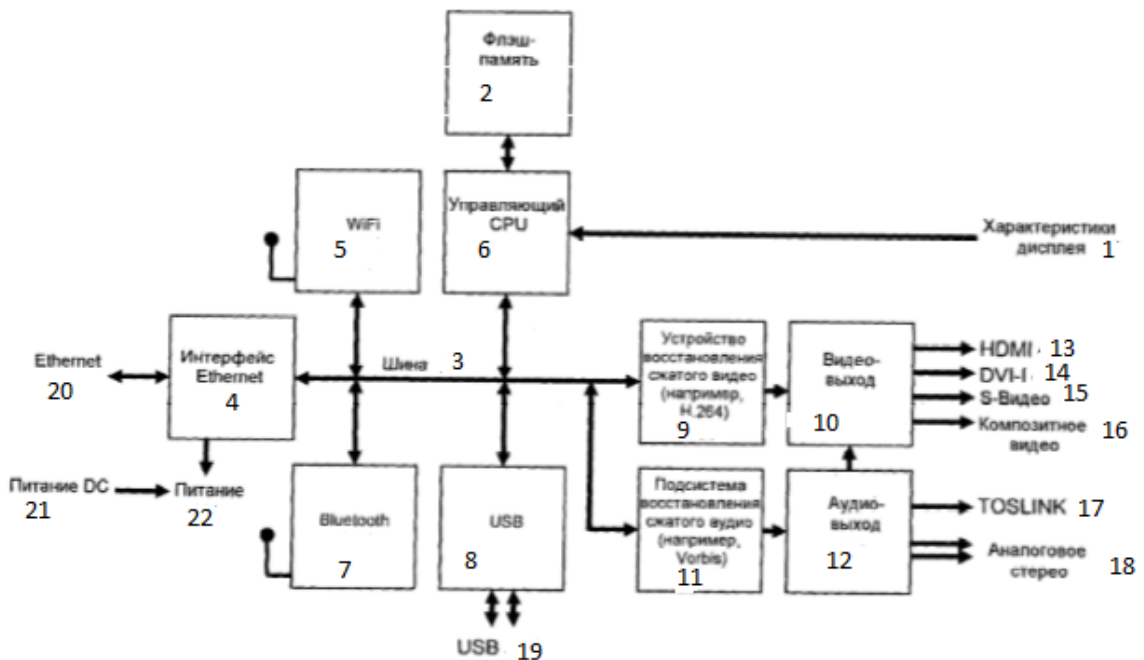


Рисунок 2.9 - Схема иллюстративной блок-схемы клиентского устройства по рис.2.7

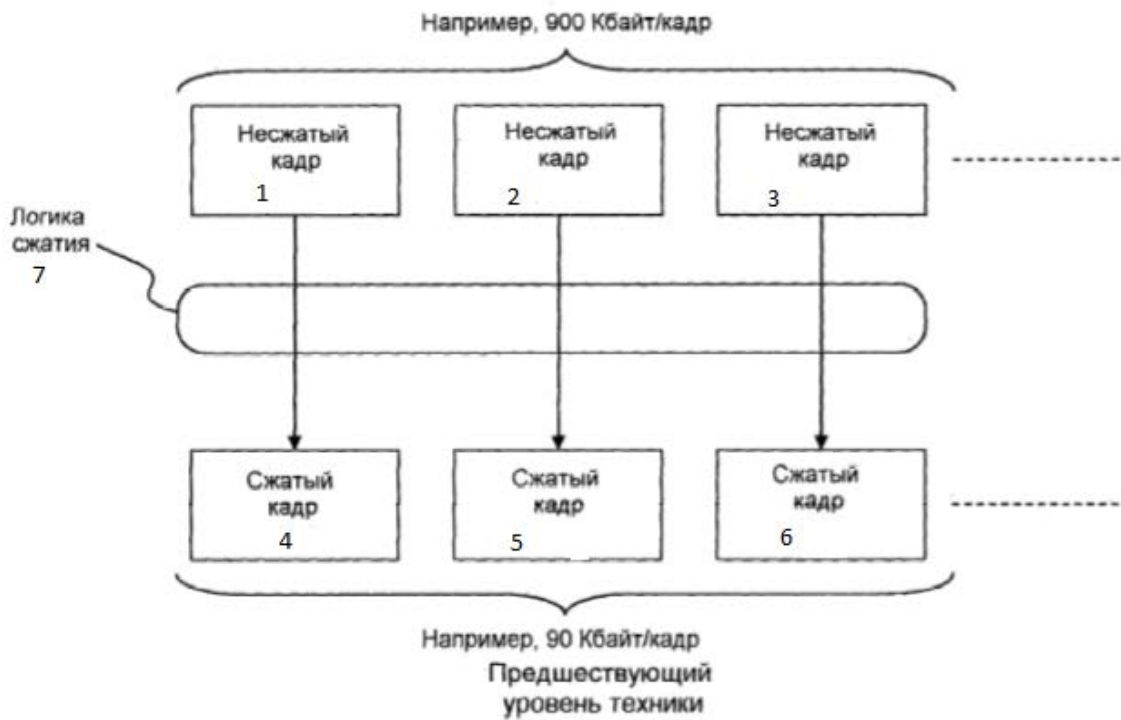


Рисунок 2.10 - Схема иллюстративной формы сжатия видео, которая может быть применена согласно одному варианту осуществления

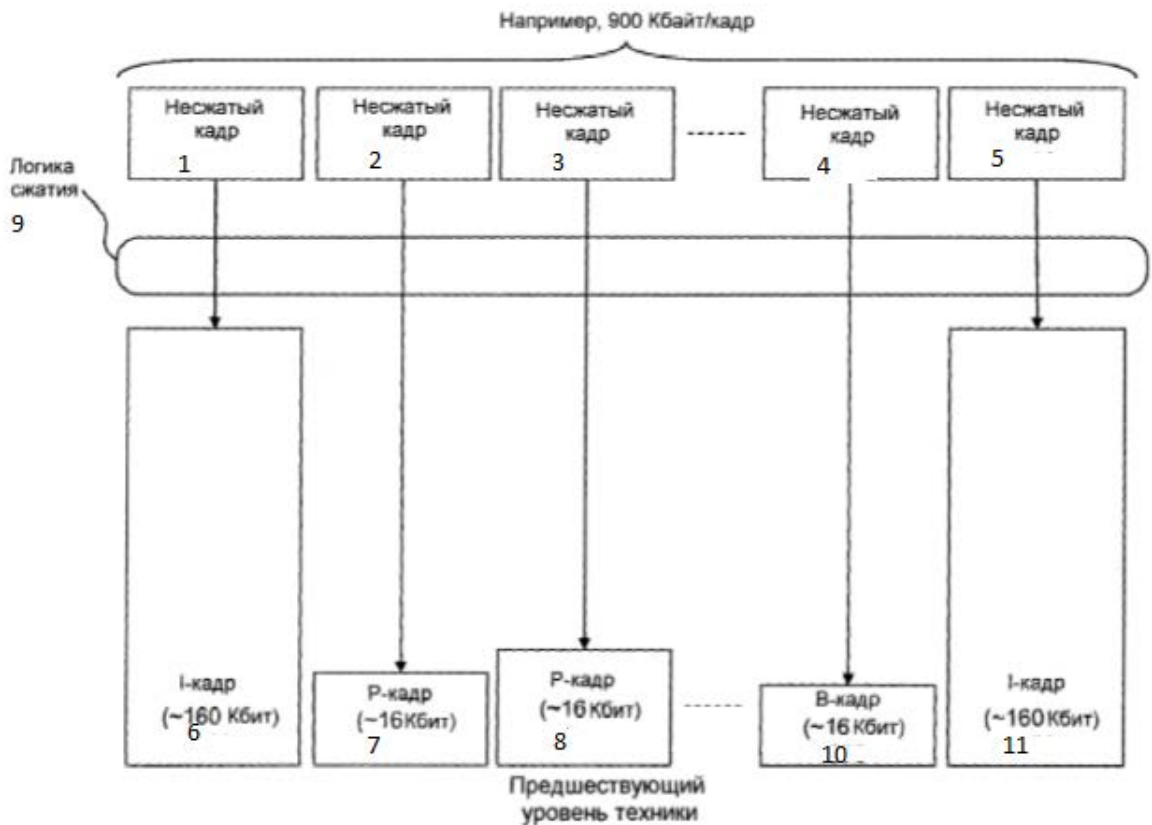


Рисунок 2.11 - Схема иллюстративной формы сжатия видео, которая может быть применена в другом варианте осуществления

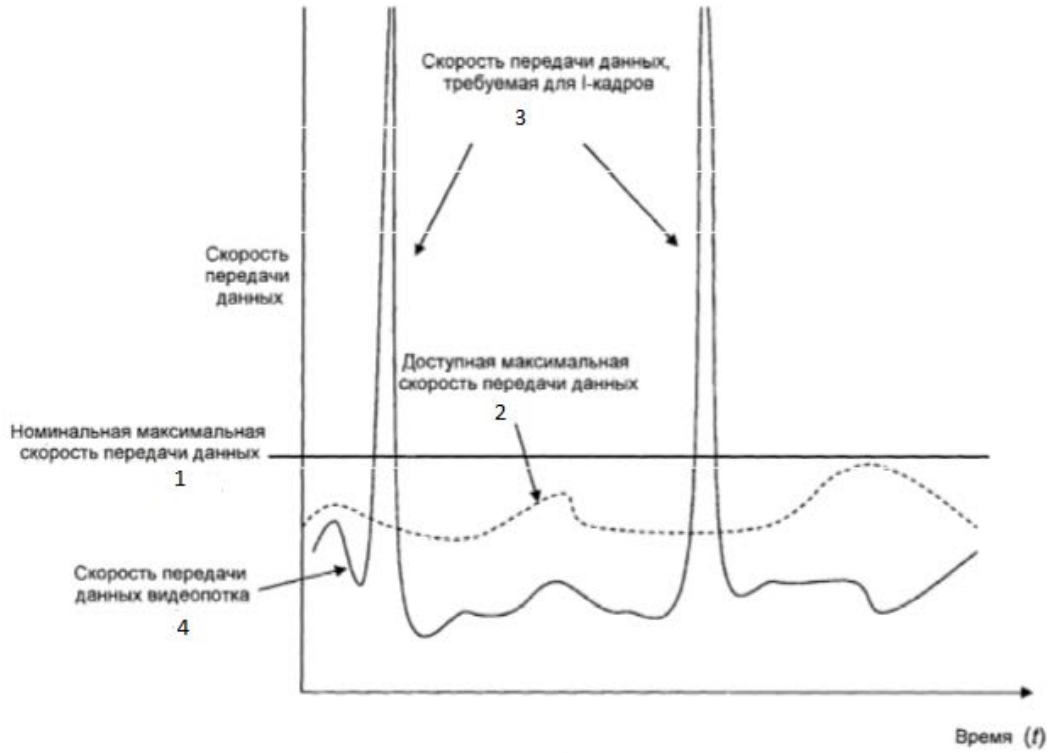


Рисунок 2.12 - Схема пиков скорости передачи данных, связанной с передачей видеопоследовательности с низкой активностью, низким уровнем сложности

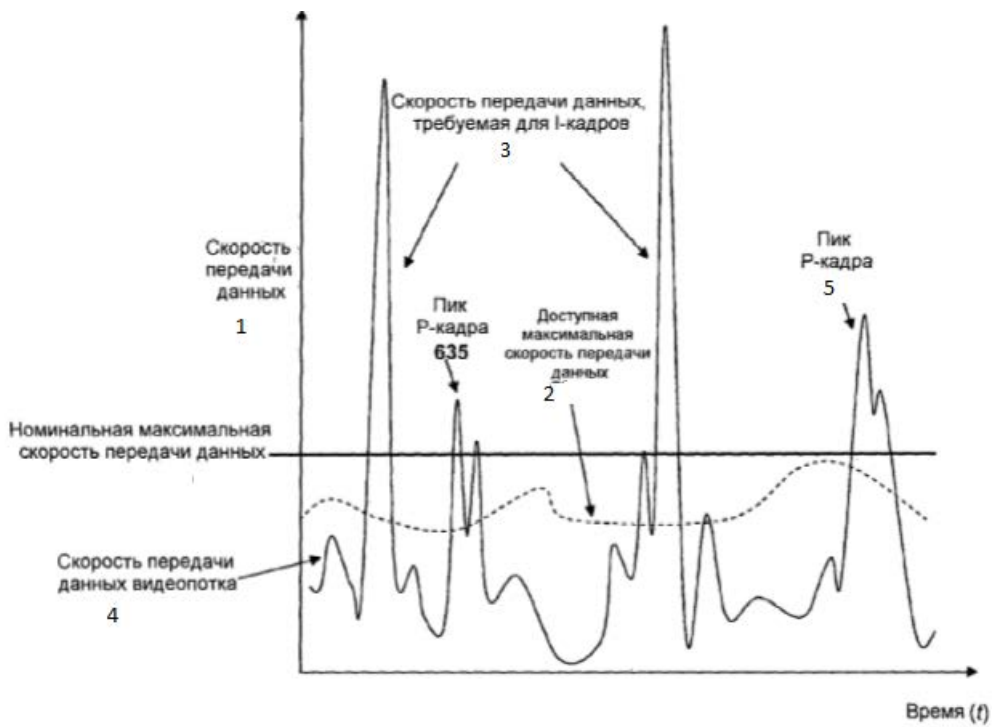


Рисунок 2.13-Схематика скорости передачи данных, связанной с передачей видеопоследовательности с высокой активностью, высоким уровнем СЛОЖНОСТИ

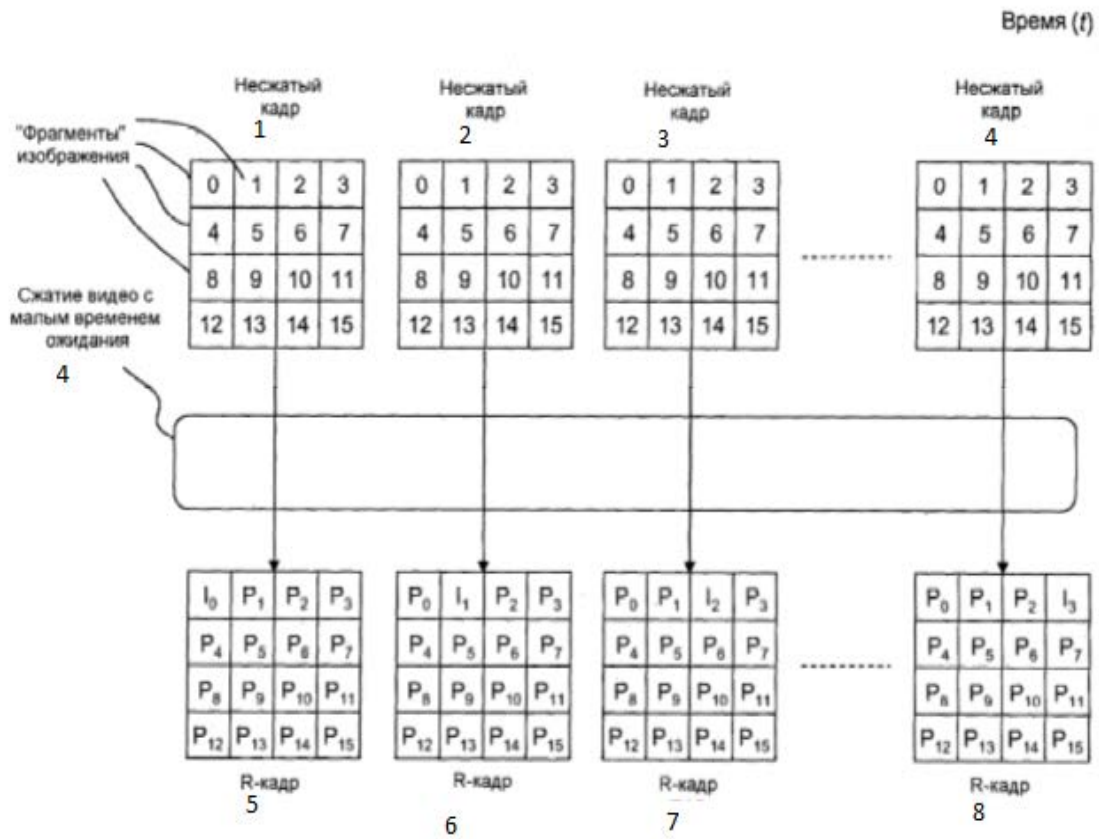


Рисунок 2.14 - Схема иллюстративных способов сжатия видео, применяемых в одном варианте осуществления

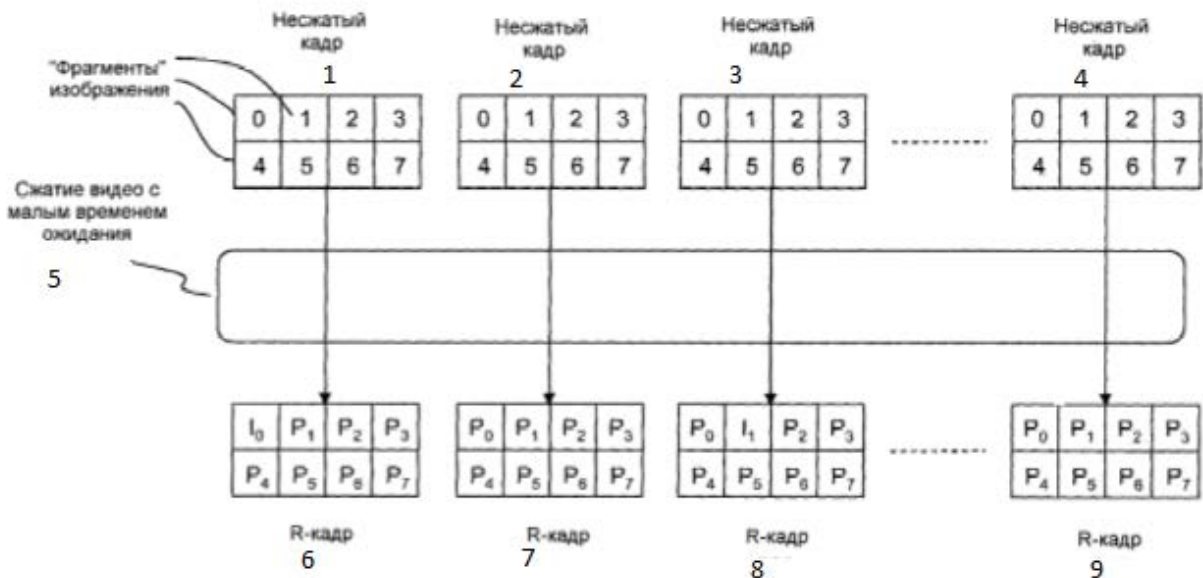


Рисунок 2.15 - Схема иллюстративных способов сжатия видео, применяемых в одном варианте осуществления

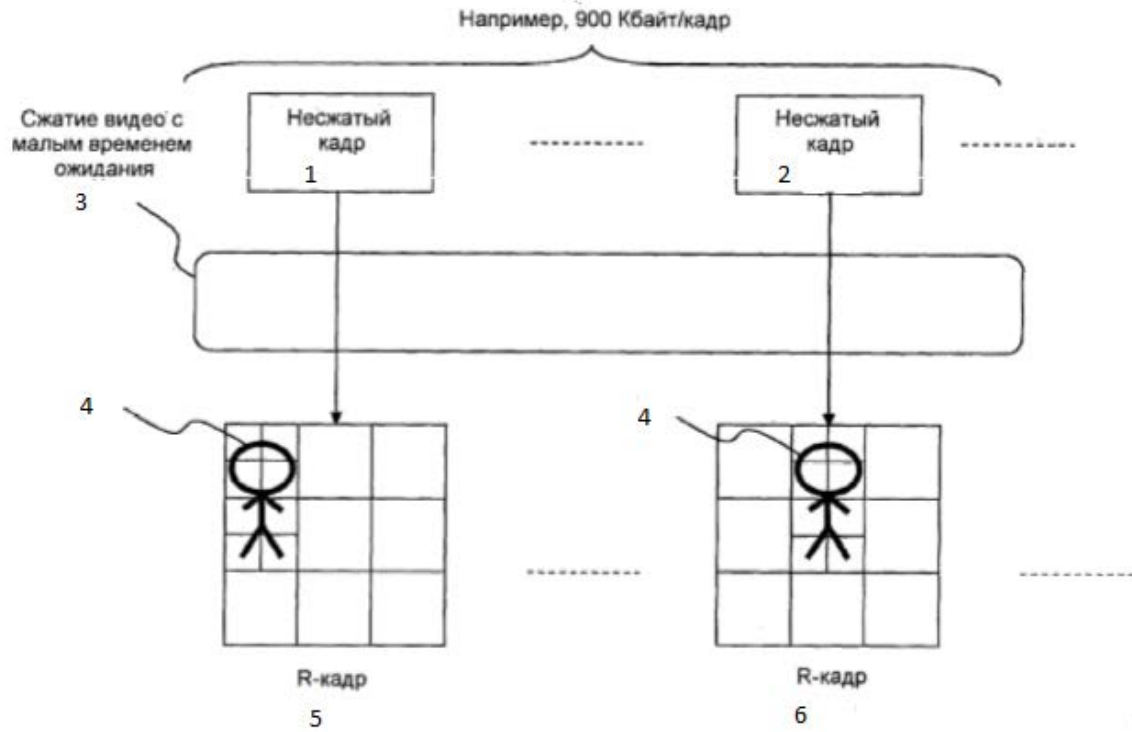


Рисунок 2.16 - Схема дополнительных иллюстративных способов сжатия видео, применяемых в одном варианте осуществления

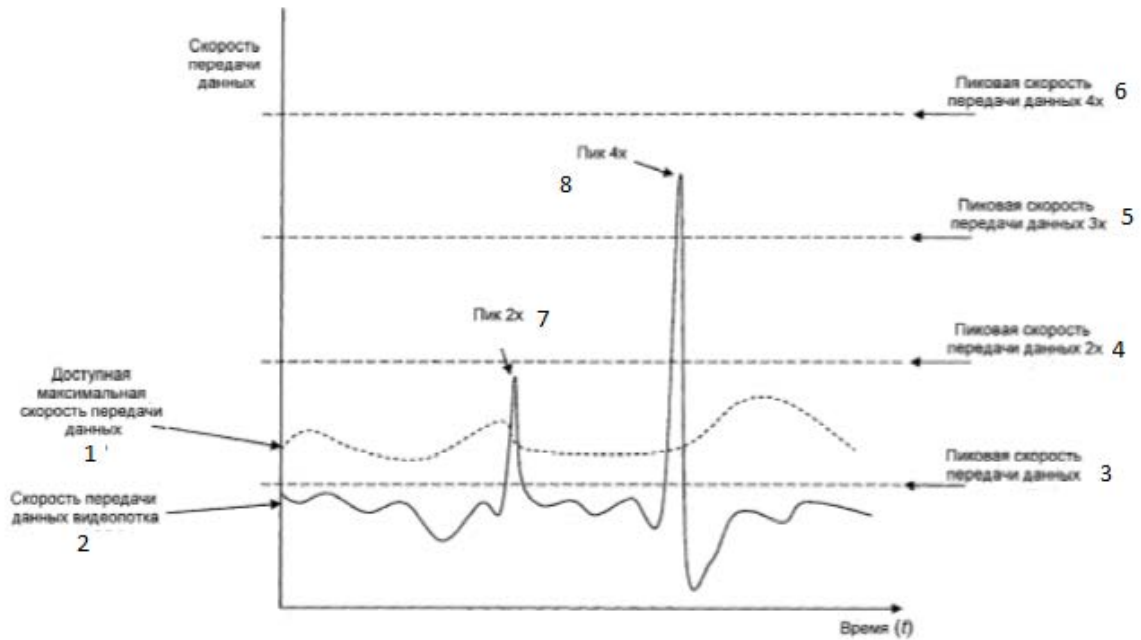


Рисунок 2.17 - Схема иллюстративных способов, применяемых для уменьшения пиков скорости передачи данных

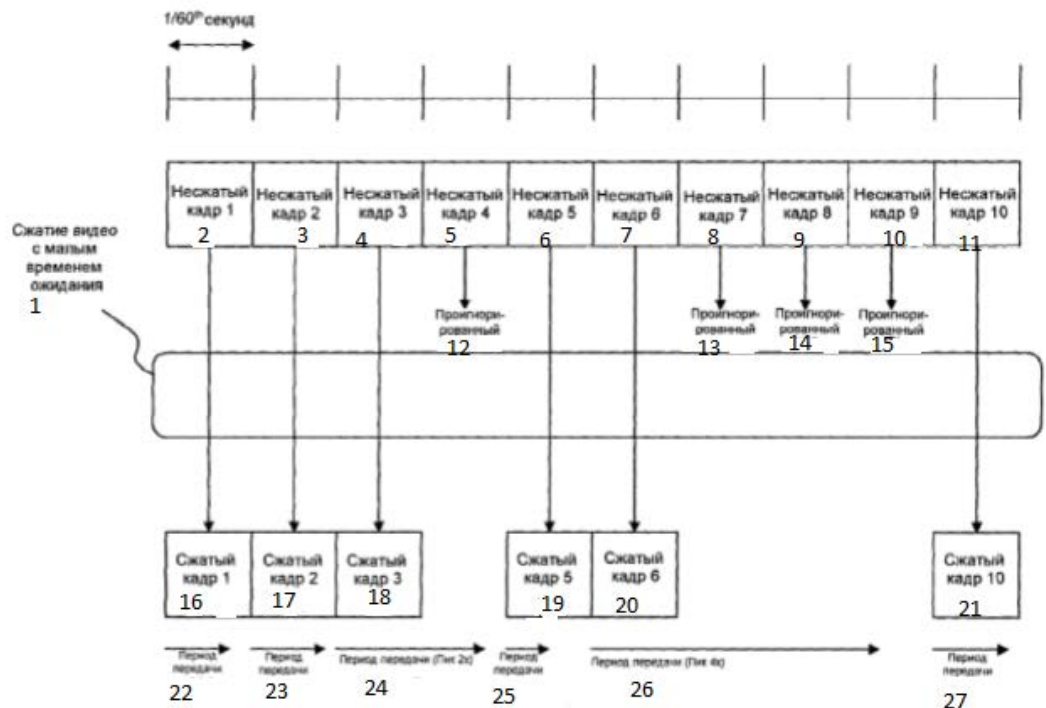


Рисунок 2.18 - Схема иллюстративных способов, применяемых в одном варианте осуществления для уменьшения пиков скорости передачи данных

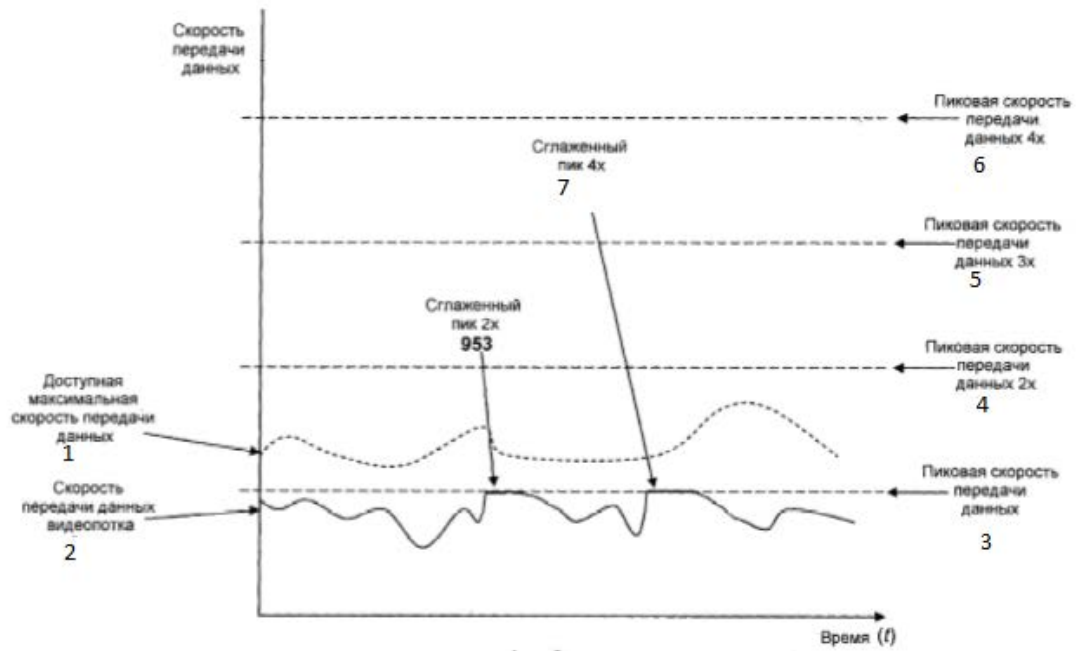


Рисунок 2.19 - Схема иллюстративных способов, применяемых в одном варианте осуществления для уменьшения пиков скорости передачи данных

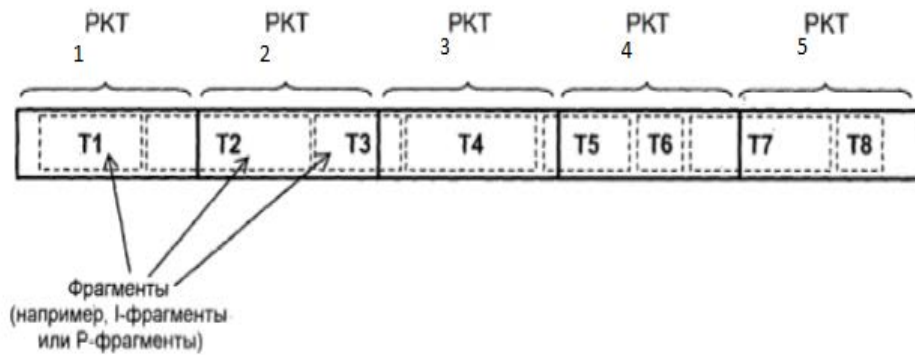


Рисунок 2.20 - Схема одного варианта осуществления, который эффективно упаковывает фрагменты внутри пакетов

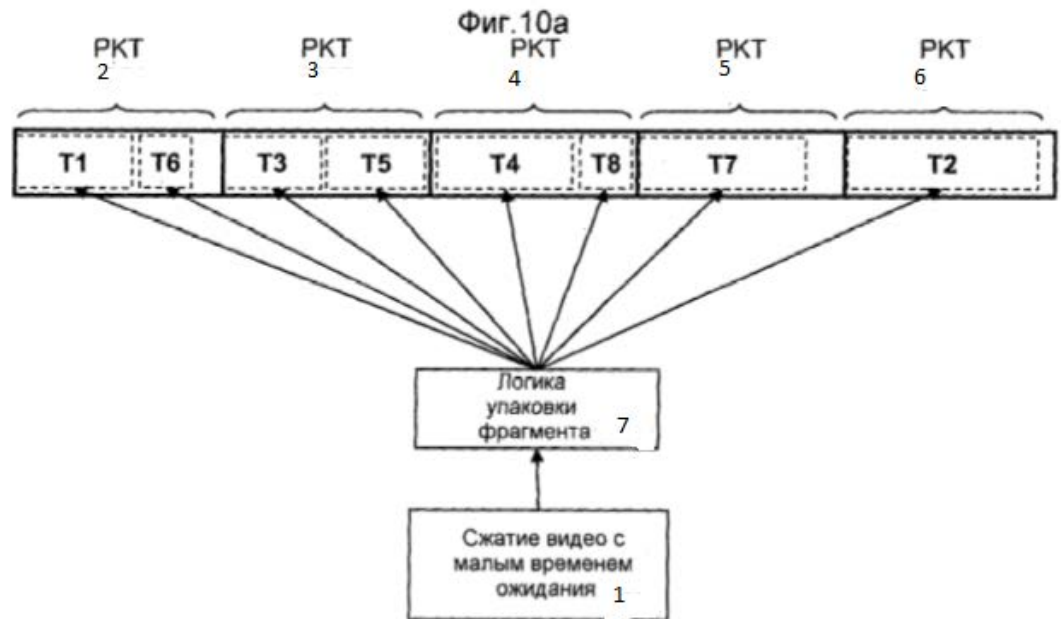


Рисунок 2.21 - Схема одного варианта осуществления, который эффективно упаковывает фрагменты внутри пакетов

RU 2 504 908 C2

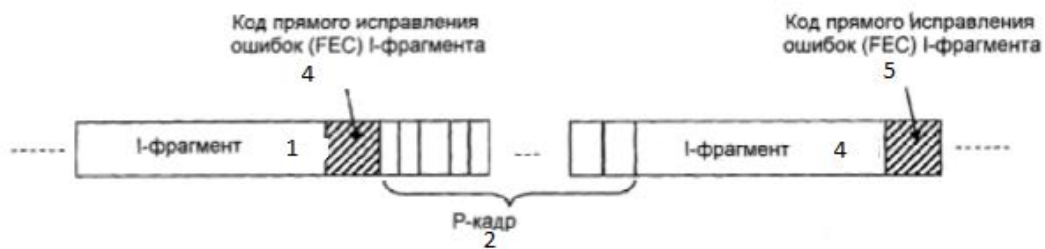


Рисунок 2.22 - Схема вариантов осуществления, которые применяют способы прямого исправления ошибок

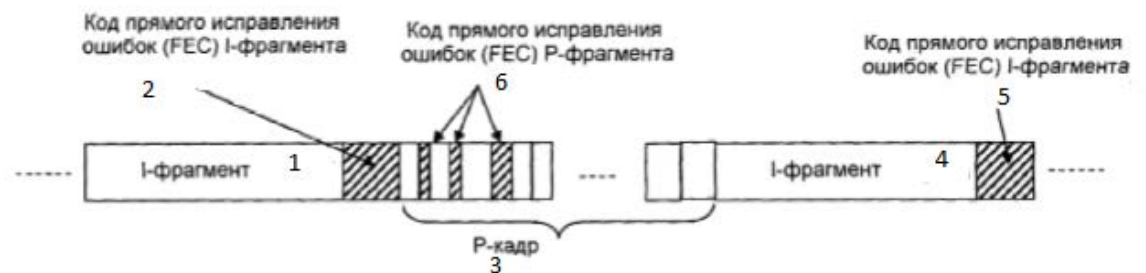


Рисунок 2.23 - Схема вариантов осуществления, которые применяют способы прямого исправления ошибок

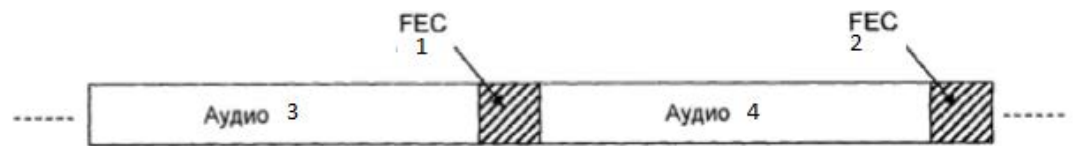


Рис 2.24 - Схема вариантов осуществления, которые применяют способы прямого исправления ошибок



Рисунок 2.25 - Схема вариантов осуществления, которые применяют способы прямого исправления ошибок

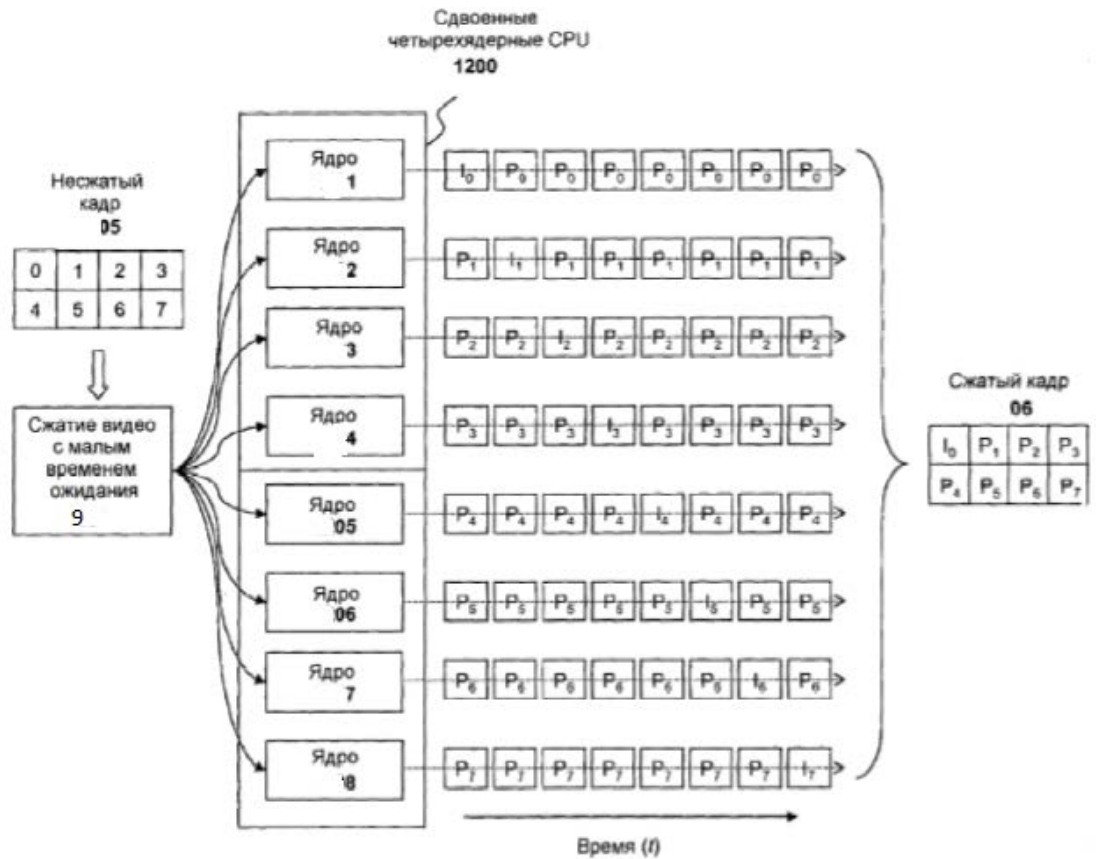


Рисунок 2.26 - Схема одного варианта осуществления, который использует многоядерные процессоры для сжатия

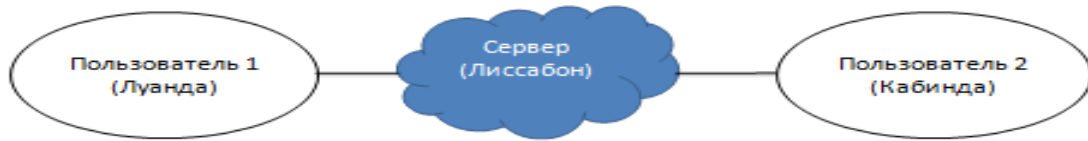


Рисунок 2.27 - Схема географического положения и связи между службами хостинга

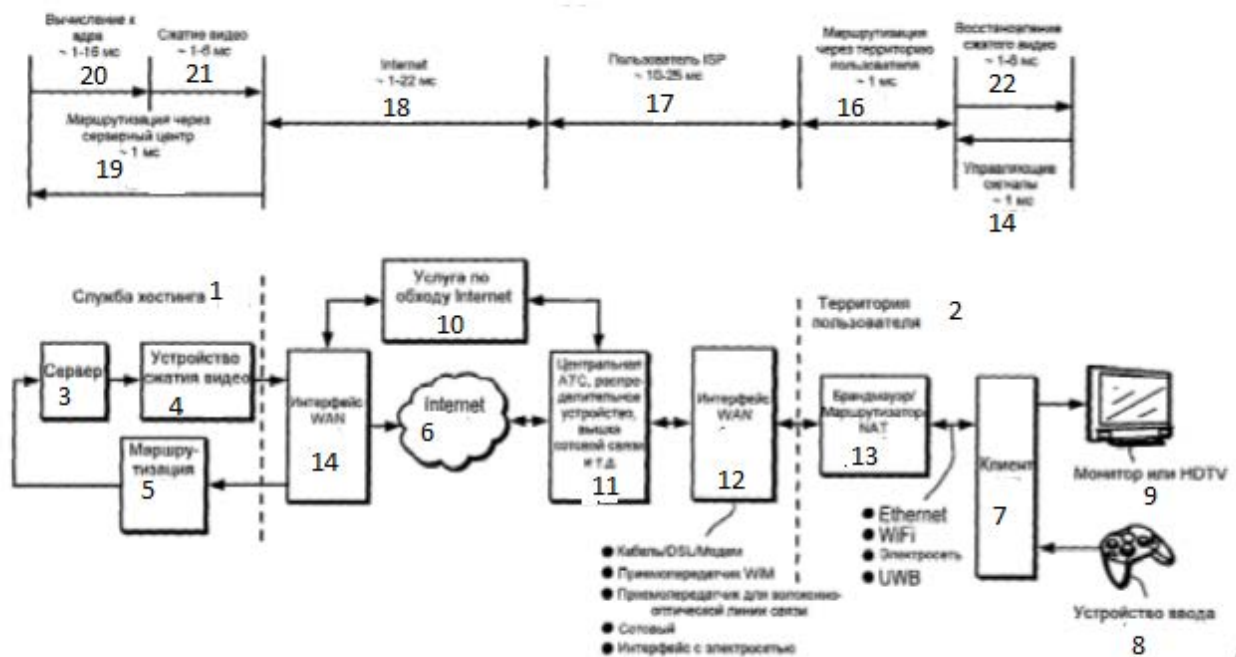


Рисунок 2.28 - Схема иллюстративных величин времени ожидания, связанных с передачей информации между клиентом и службой хостинга

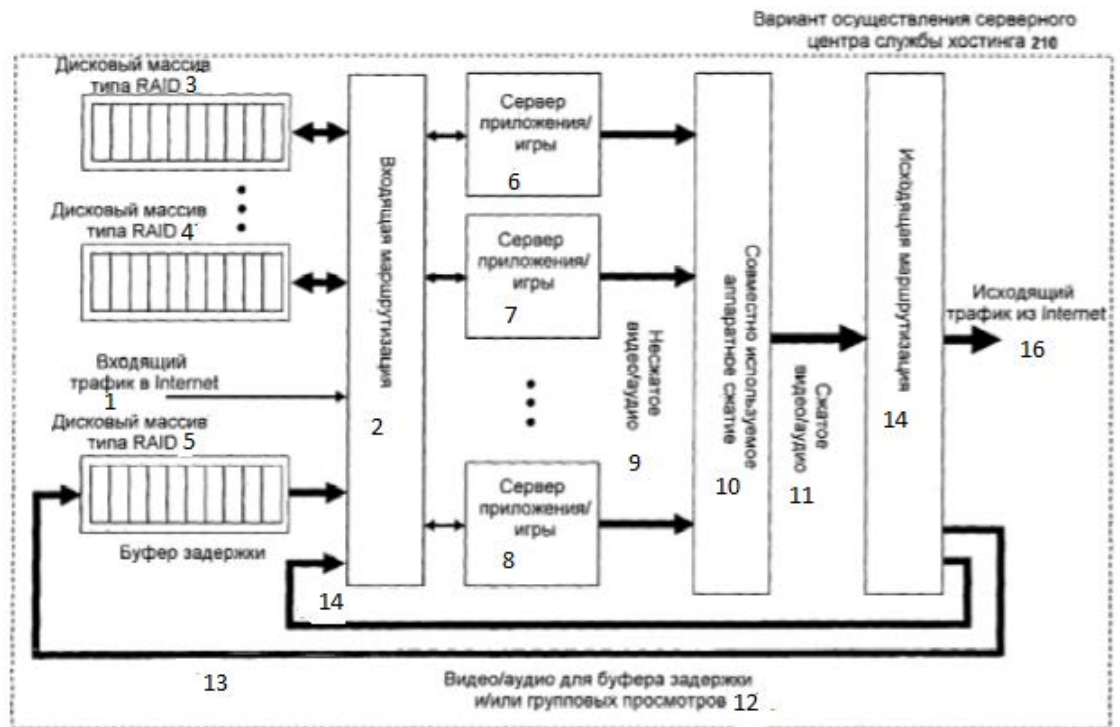


Рисунок 2.29 - Схема архитектуры серверного центра службы хостинга

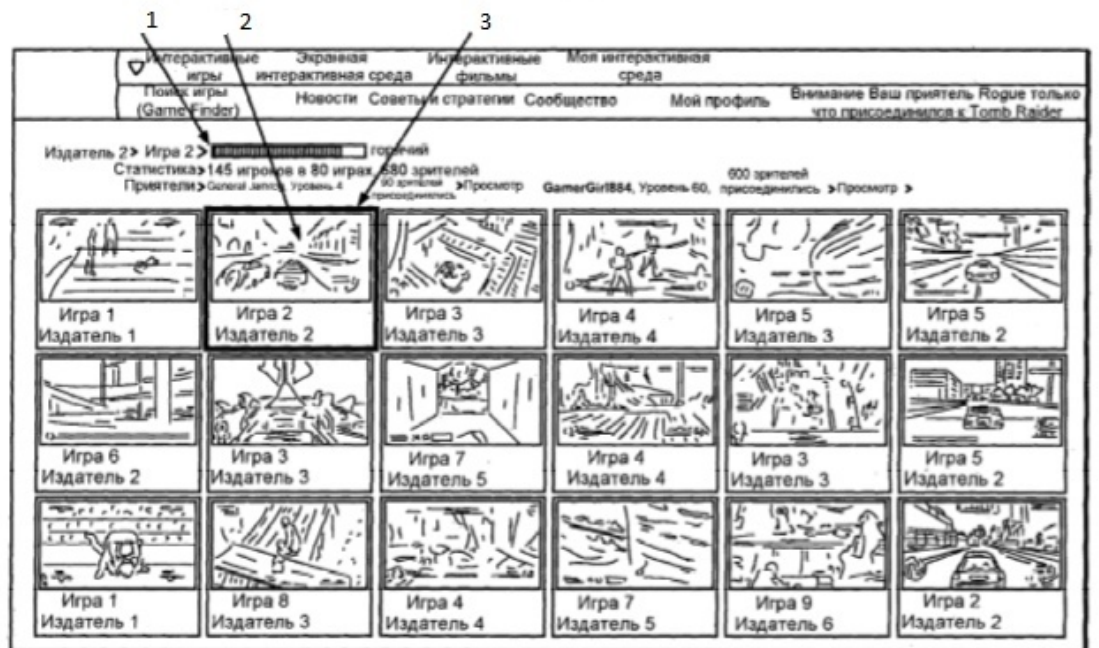


Рисунок 2.30 - Схема иллюстративного моментального снимка экрана одного варианта осуществления интерфейса пользователя, который включает в себя множество видеоокон в реальном масштабе времени

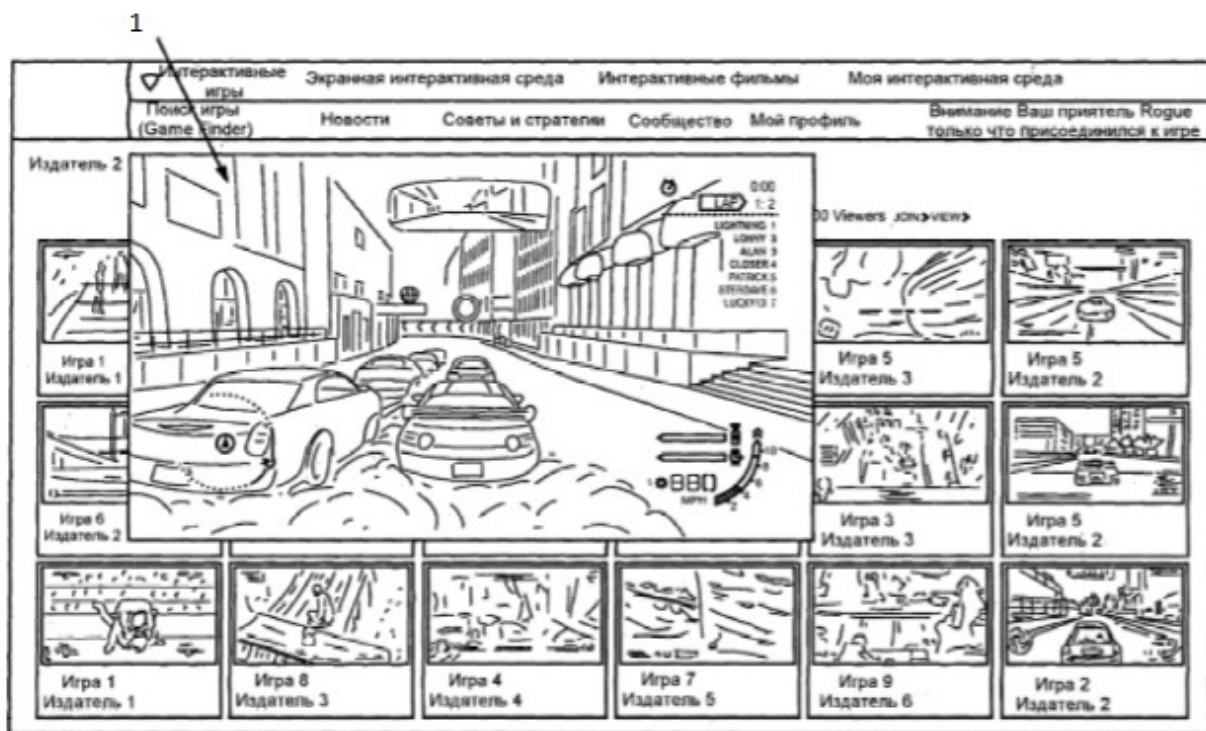


Рисунок 2.31 - Схема интерфейса пользователя по рис.2.30 после выбора конкретного видеоокна

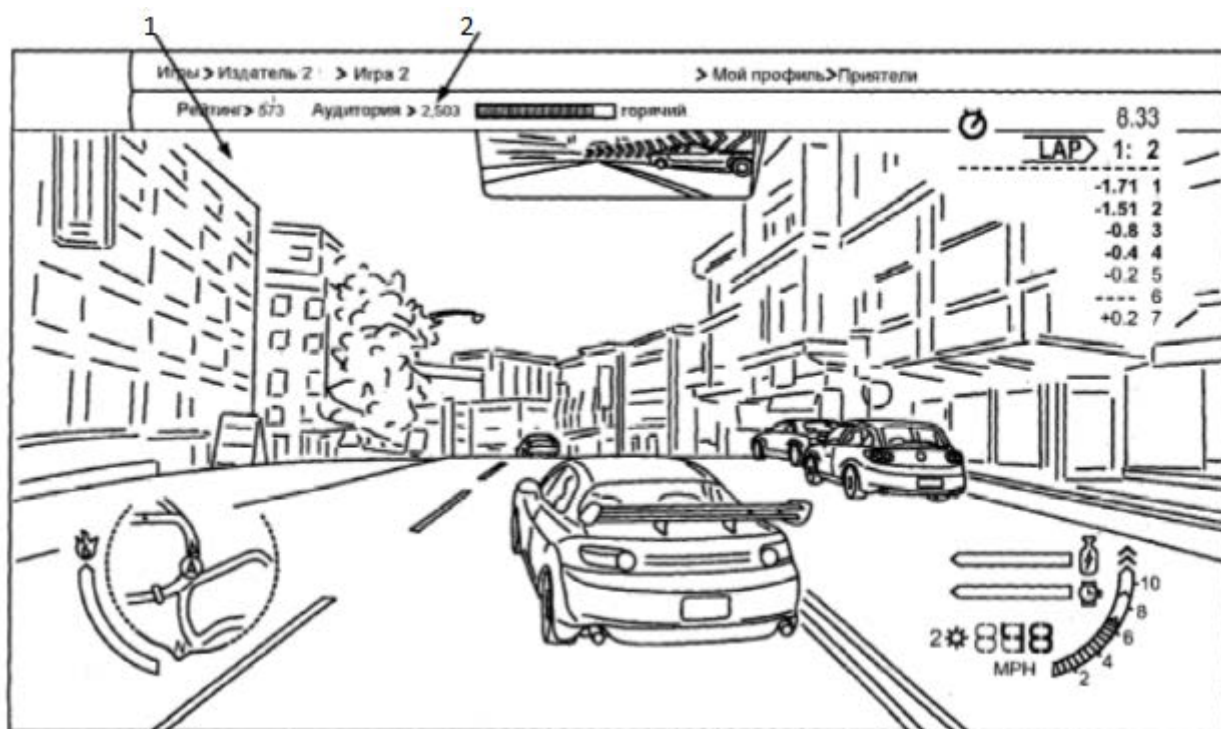


Рисунок 2.32 - Схема интерфейса пользователя по рис. 2.31 после распаковки упомянутого конкретного видеоокна во весь экран

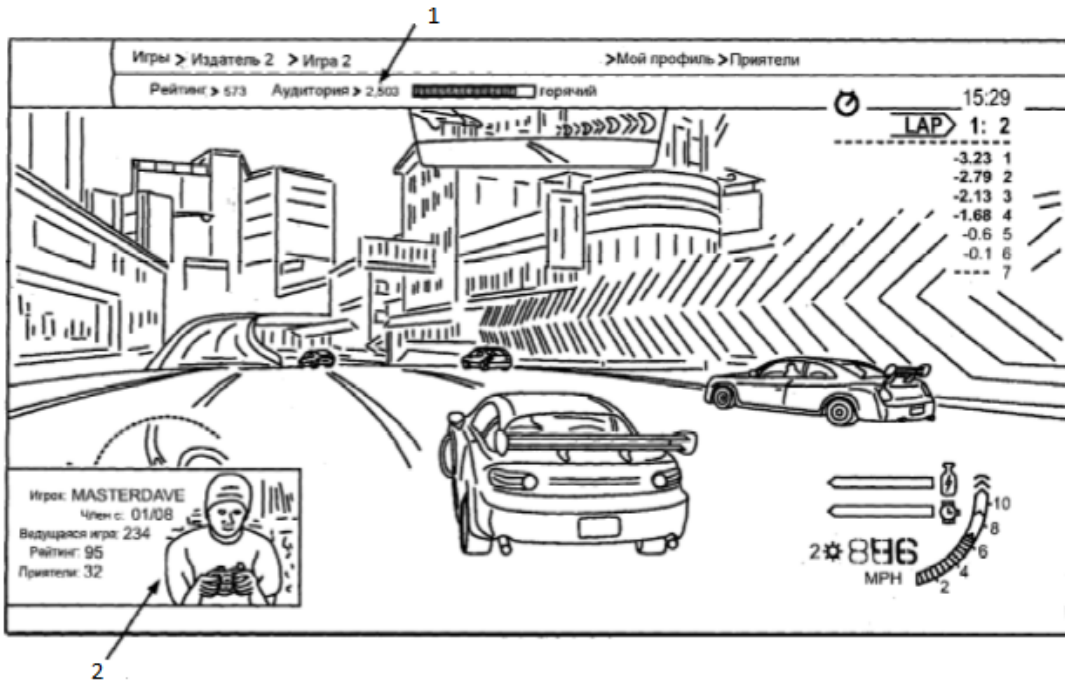


Рисунок 2.33 - Схема иллюстративных совместных пользовательских видеоданных, совмещенных на экране игры с несколькими участниками

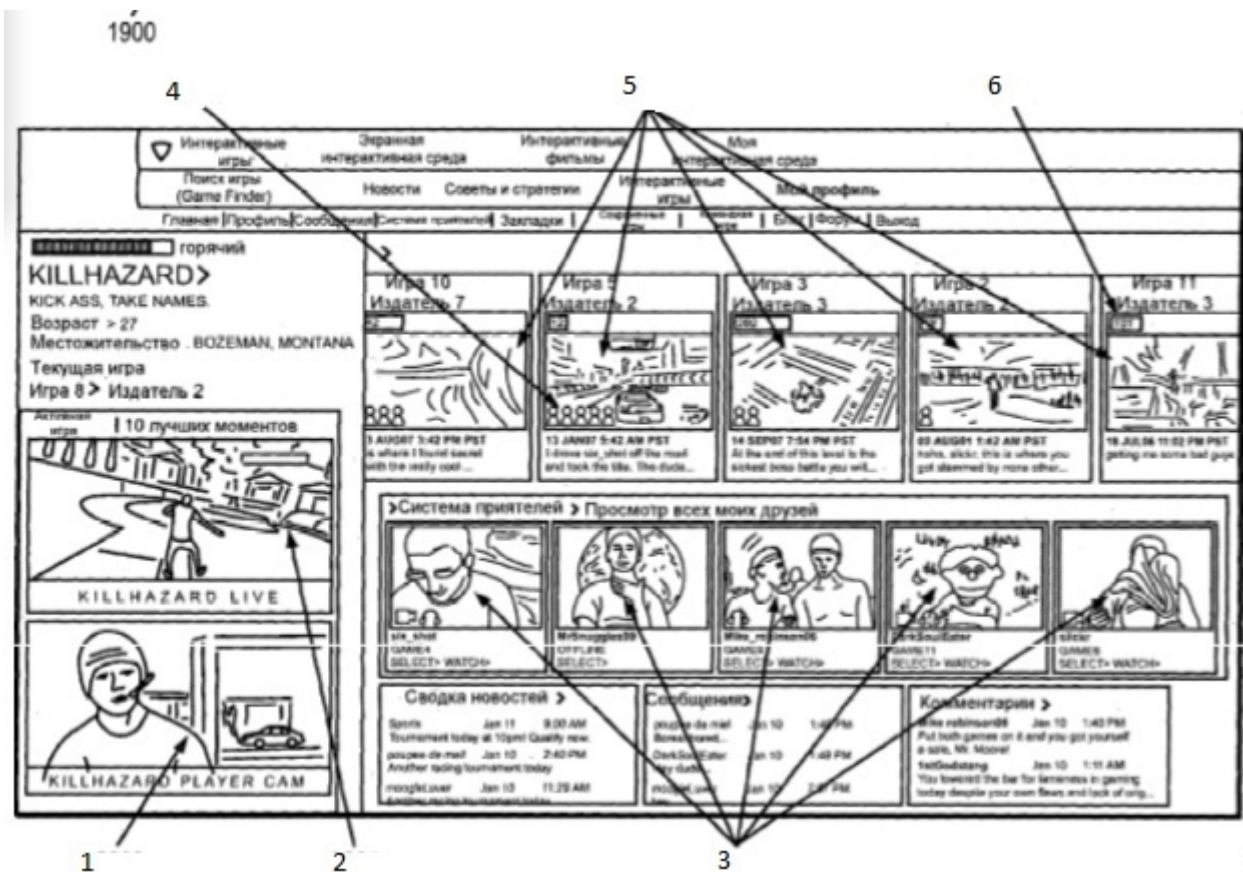


Рисунок 2.34-Схема иллюстративной страницы пользователя для игрока в службе хостинга

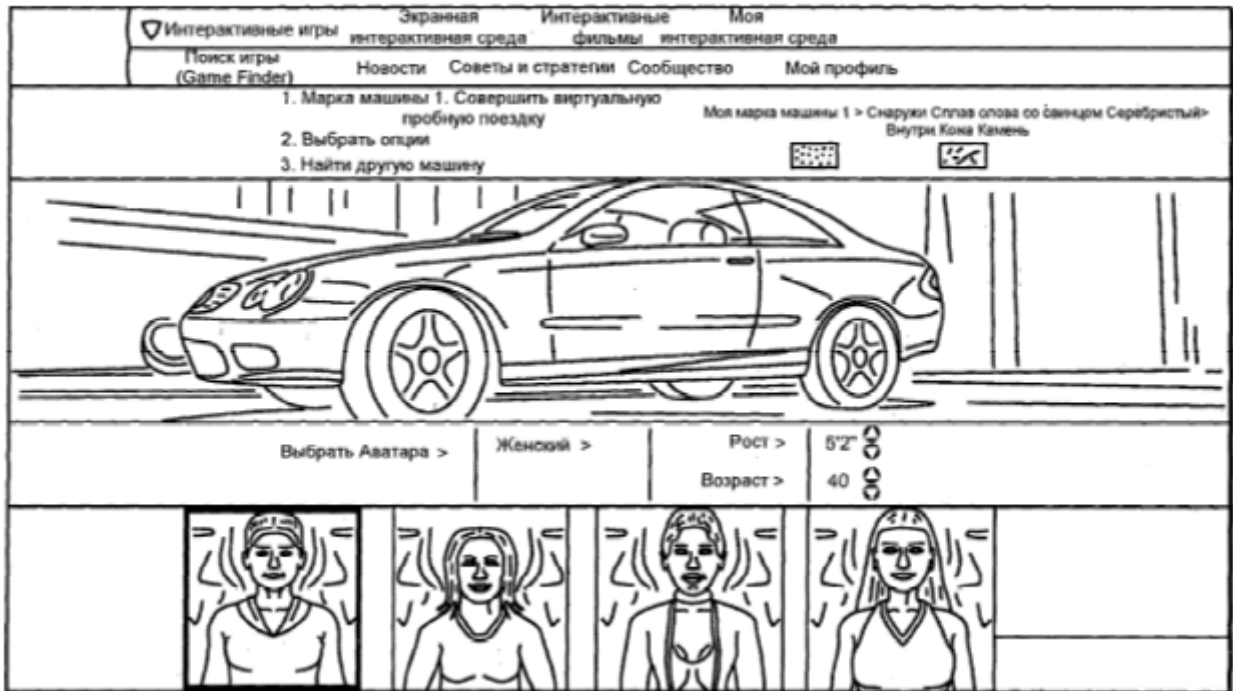


Рисунок 2.35 - Схема иллюстративной 3D интерактивной рекламы

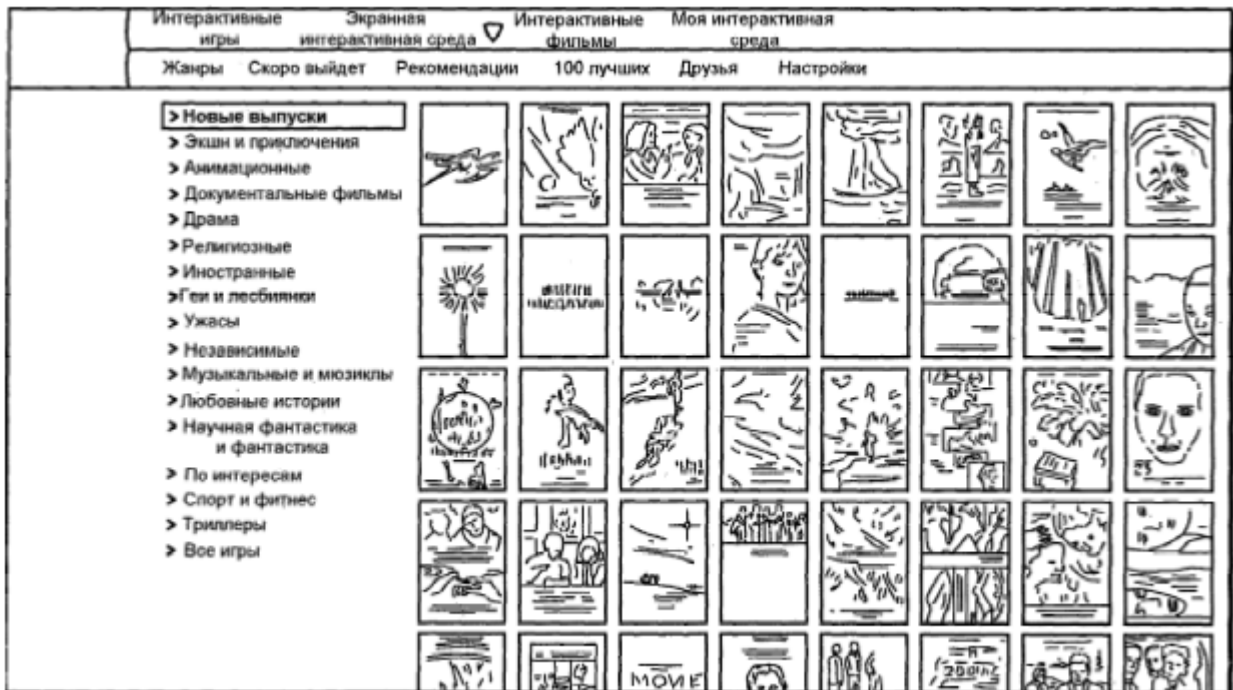


Рисунок 2.36— Схема иллюстративной последовательности этапов для вывода фотореального изображения, имеющего текстурированную поверхность, исходя из захвата поверхности живого исполнения

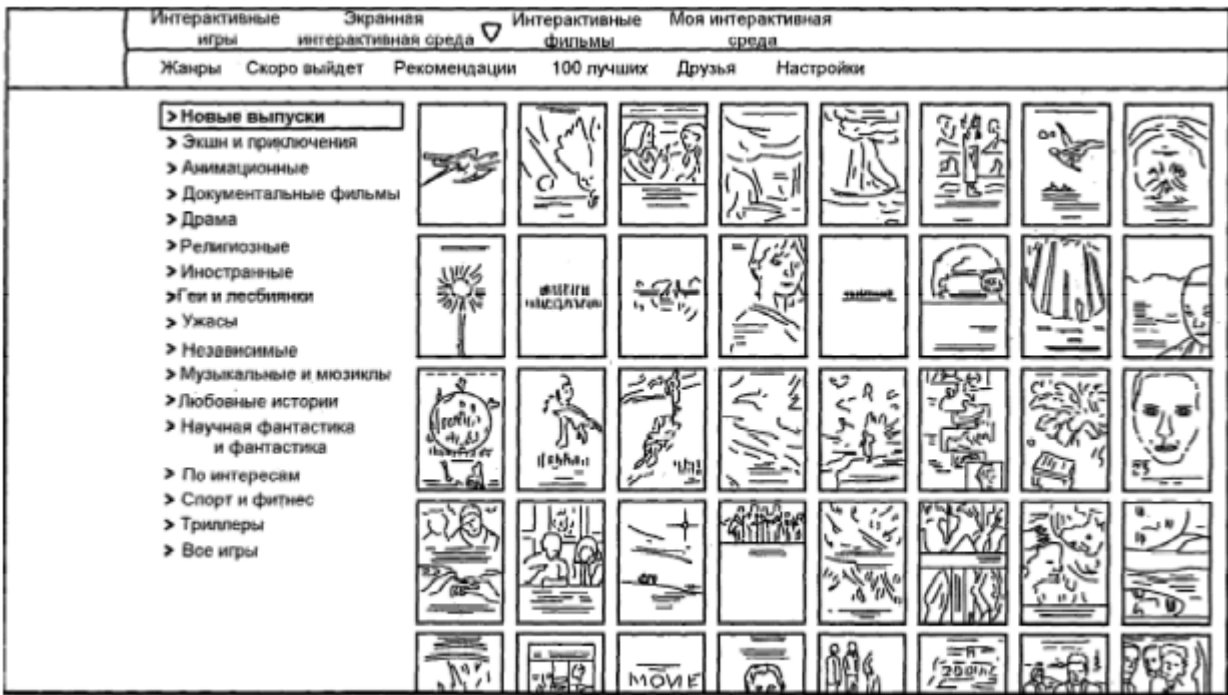


Рисунок 2.37 - Схема иллюстративной страницы интерфейса пользователя, которая обеспечивает возможность выбора линейного мультимедийного содержимого

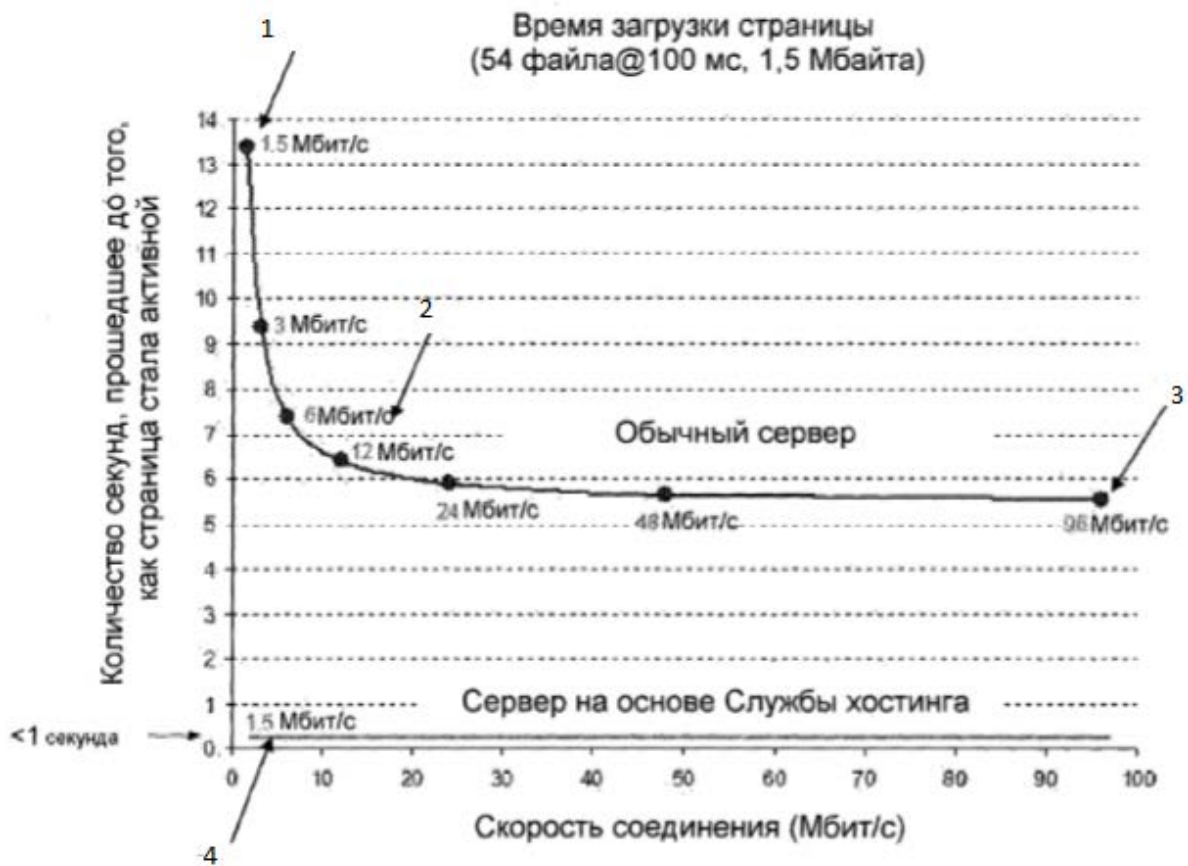


Рисунок 2.38 - Схема, которая иллюстрирует время, которое пройдет до того, как web-страница станет активной, по сравнению со скоростью соединения. Описание иллюстративных вариантов осуществления

2.2 Подробное описание предлагаемого технического решения

На рис. 2.1-рис. 2.2 обеспечена архитектура высокого уровня двух вариантов осуществления, в которых видеоигры и приложения размещаются службой 3 хостинга и к ним получают доступ клиентские устройства 1, находящиеся на территории 4 пользователя (отметим, что "территория пользователя" означает любое место, где находится пользователь, в том числе на открытом воздухе, если он использует мобильное устройство), через Internet 2 (или другую общедоступную или частную сеть) согласно абонентскому обслуживанию. Клиентские устройства 1 могут быть универсальными компьютерами, например PC Windows Microsoft, или PC Linux, или компьютерами Macintosh корпорации Apple, Inc. с проводным или беспроводным соединением с Internet, с встроенным или внешним дисплеем 7, или они могут быть специализированными клиентскими устройствами, например телевизионной абонентской приставкой (с проводным или беспроводным соединением с Internet), которая выводит видео и аудио на монитор или телевизор 7, или они могут быть мобильными устройствами, предположительно с беспроводным соединением с Internet. Любое из этих устройств может иметь свои собственные устройства ввода пользователя (например, клавиатуры, кнопки, сенсорные экраны, трекпад или световое перо, камеры видеозахвата и/или камеры, отслеживающие движение, и т.д.) или они могут использовать внешние устройства 6 ввода (например, клавиатуры, мыши, игровые контроллеры, световое перо, камеры видеозахвата и/или камеры, отслеживающие движение, и т.д.), подключенные посредством проводов или беспроводным способом. Как описано более подробно ниже, служба 3 хостинга включает в себя серверы различных уровней производительности, в том числе серверы с мощными средствами обработки CPU/GPU. Во время ведения

игры или использования приложения в службе 3 хостинга, домашнее или офисное клиентское устройство 1 принимает входные данные контроллера и/или клавиатуры от пользователя, и затем оно отправляет эти входные данные контроллера через Internet 2 в службу 3 хостинга, который в ответ исполняет программу, управляющую игрой, и формирует последовательные кадры видеовыхода (последовательность видеоизображений) для игры или прикладного программного обеспечения (например, если пользователь нажимает на кнопку, которая может перемещать персонажа на экране вправо, тогда игровая программа создает последовательность видеоизображений, на которых изображен персонаж, перемещающийся вправо). Эта последовательность видеоизображений после этого сжимается с использованием устройства сжатия видео с малым временем ожидания, и после этого служба 3 хостинга передает видеопоток с малым временем ожидания через Internet 2. Домашнее или офисное клиентское устройство после этого декодирует сжатый видеопоток и визуализирует восстановленные видеоизображения на мониторе или телевизоре. Следовательно, требования к вычислительной и графической технике клиентского устройства 1 значительно уменьшаются. Требуется, чтобы клиент 1 имел вычислительную мощность только для пересылки входных данных клавиатуры/контроллера в Internet 2 и декодирования и восстановления сжатого видеопотока, принятого из Internet 2, что в настоящее время фактически любой персональный компьютер может выполнять в программном обеспечении на своем центральном процессоре (например, двухядерный процессор корпорации Intel (Intel Corporation Core Duo CPU), работающий с тактовой частотой приблизительно 2 ГГц, может восстанавливать сжатое HDTV 1280×720, закодированное с использованием таких устройств сжатия, как H.264 и Windows Media VC9). И, в случае любых клиентских устройств, также специализированные интегральные схемы могут выполнять восстановление сжатого видео для таких стандартов в реальном времени с гораздо меньшими затратами и с гораздо меньшим потреблением мощности, чем универсальный CPU, например, который может потребоваться для современного PC. А именно,

для выполнения функции пересылки входных данных контроллера и восстановления сжатого видео, в домашних клиентских устройствах 1 не требуются никакие специализированные графические процессоры (GPU), накопитель на оптических дисках или накопители на жестких дисках, например игровая видеосистема предшествующего уровня техники, изображенная на рис. 1.1.

По мере того как игры и прикладное программное обеспечение становится более сложным и более фотореалистичным, они требуют GPU, CPU более высокой производительности, большего количества RAM и более быстрых накопителей на магнитных дисках большей емкости, и вычислительная мощность службы 3 хостинга может непрерывно модернизироваться, но конечному пользователю не потребуется обновлять домашнюю или офисную клиентскую платформу 1, так как требования к ее обработке будут оставаться постоянными для разрешающей способности дисплея и частоты кадров при данном алгоритме восстановления сжатого видео. Соответственно, аппаратные ограничения и вопросы совместимости, имеющиеся в настоящее время, не существуют в системе, изображенной на рис. 2.1-рис. 2.2. Кроме того, так как игра и прикладное программное обеспечение исполняются только в серверах в службе 3 хостинга, не существует копии игры или прикладного программного обеспечения (ни в виде оптического носителя информации, ни как загружаемого программного обеспечения) в офисе ("офис", как используется в этом описании, если не оговорено иное, включает в себя любое нежилое окружение, не связанное с постоянным проживанием, в том числе классные комнаты, например) или в доме пользователя. Это значительно уменьшает вероятность того, что игра или прикладное программное обеспечение будут незаконно скопированы (незаконно использоваться), а также уменьшает вероятность того, что ценная база данных, которая может быть (использована) игрой или прикладным программным обеспечением, будет незаконно использоваться. На самом деле, если для ведения игры требуются специализированные серверы (например, с требованием очень

дорогого, большого или шумного оборудования) или прикладное программное обеспечение, которое на практике нельзя использовать дома или в офисе, то, даже если будет получена пиратская копия игры или прикладного программного обеспечения, то ее нельзя будет использовать дома или в офисе. В одном варианте осуществления, служба 3 хостинга обеспечивает инструментальные средства разработки программного обеспечения разработчикам 5 прикладного программного обеспечения или игры (что, как правило, относится к компаниям по разработке программного обеспечения, игровым или телевизионным студиям или издателям прикладного программного обеспечения или игры), которые проектируют видеоигры, для того, чтобы они могли проектировать игры, которые можно будет исполнять в службе 3 хостинга. Такие инструментальные средства обеспечивают возможность разработчикам использовать признаки службы хостинга, которые обычно не доступны в автономном РС или игровой консоли (например, быстрый доступ к очень большим базам данных сложной геометрии ("геометрия", если не оговорено иное, используется в этом описании для ссылки на многоугольники, текстуры, снаряжение, освещение, линии поведения и другие компоненты и параметры, которые определяют базы данных 3D)).

Разные бизнес-модели возможны по этой архитектуре. По одной модели, служба 3 хостинга получает абонентскую плату от конечного пользователя и платит лицензионный платеж разработчикам 5, как изображено на рис. 2.1. В альтернативной реализации, изображенной на рис. 2.2, разработчики 5 получают абонентскую плату непосредственно от пользователя и платят службе 3 хостинг за хостинг содержимого приложения или игры. Эти лежащие в основе принципы не ограничены какой-либо конкретной бизнес-моделью для обеспечения хостинга приложения или онлайн-игр.

2.2.1 Характеристики сжатого видео

Как обсуждалось ранее, одной существенной проблемой, связанной с оказанием услуг по видеоиграм или услуг по прикладному программному обеспечению в режиме онлайн, является проблема времени ожидания. Время

ожидания 70-80 мс (от момента, когда устройство ввода приводится в действие пользователем, до момента, когда ответ выводится на экран дисплея) находится на верхней границе для игр и приложений, требующих малого времени отклика. Однако достичь такой задержки очень трудно в контексте архитектуры, изображенной на рис. 2.1 и рис. 2.2, из-за нескольких практических и физических ограничений. Как указано на рис. 2.3, когда пользователь подписывается на Internet-услугу, соединение обычно оценивается номинальной максимальной скоростью 1 передачи данных до офиса или дома пользователя. В зависимости от политики провайдеров и характеристик оборудования для маршрутизации, может быть более или менее обеспечено строгое выполнение этой максимальной скорости передачи данных, но, как правило, фактическая доступная скорость передачи данных является более низкой по одной или нескольким разным причинам. Например, в центральной АТС DSL или в локальной линии кабельной модемной связи может находиться слишком много сетевого трафика, или в кабельной сети может присутствовать шум, вызывающий сброс пакетов, или провайдер может установить максимальное количество битов в месяц для каждого пользователя. В настоящее время, максимальная скорость передачи нисходящего потока данных для услуг DSL и кабельной связи, как правило, колеблется от нескольких сотен килобит в секунду (Кбит/с) до 30 Мбит/с. Услуги сотовой связи обычно ограничены до сотен Кбит/с нисходящего потока данных.

Однако скорость услуг широкополосного доступа и количество пользователей, которые подписываются на услуги широкополосного доступа, существенно увеличатся со временем. В настоящее время, некоторые аналитики оценивают, что 33% американских абонентов широкополосного доступа имеют скорость передачи нисходящего потока данных 2 Мбит/с или больше. Например, некоторые аналитики предсказывают, что к 2010 г. свыше 85% американских абонентов широкополосного доступа будут иметь скорость передачи данных 2 Мбит/с или больше. Как указано на рис. 2.3, фактическая доступная максимальная скорость 2 передачи данных может колебаться с течением

времени. Соответственно, в контексте прикладного программного обеспечения или онлайн-игр с малым временем ожидания иногда трудно предсказать фактическую доступную скорость передачи данных для конкретного видеопотока. Если скорость 3 передачи данных, требуемая для поддержки данного уровня качества при данном количестве кадров в секунду (кадр/с) с данной разрешающей способностью (например, 640×480, 60 кадр/с) для определенной величины сложности сцены и перемещения, поднимается выше фактической доступной максимальной скорости 2 передачи данных (как указано посредством пика на рис. 2.3), то может возникнуть несколько проблем. Например, в некоторых Internet-услугах будут просто сбрасываться пакеты, что в результате приведет к потере данных и искаженным/потерянными изображениям на видеозэкране пользователя. В других услугах дополнительные пакеты будут временно помещены в буфер (т.е. поставлены в очередь), и эти пакеты будут обеспечены клиенту с имеющейся скоростью передачи данных, что в результате приведет к увеличению времени ожидания - неприемлемый результат для многих видеоигр и приложений.

Наконец, некоторые провайдеры Internet-услуг будут рассматривать увеличение скорости передачи данных как злонамеренную атаку, например как атаку типа отказ в обслуживании (известный способ, используемый хакерами для блокирования сетевых соединений), и прервут соединение пользователя с Internet на определенный период времени. Соответственно, в вариантах осуществления, описанных в этом документе, предпринимаются шаги для обеспечения того, что скорость передачи данных, требуемая для видеоигры, не будет превышать максимальную имеющуюся скорость передачи данных.

2.2.2 Архитектура службы хостинга

На рис. 2.4 изображена архитектура 1 службы хостинга. Служба 1 хостинга или может быть расположена в одном серверном центре, или может быть распределена по всему множеству серверных центров (для обеспечения соединений с пользователями с меньшим временем ожидания, которые имеют маршруты с меньшим временем ожидания до определенных серверных центров,

чем другие, для обеспечения выравнивания нагрузки среди пользователей и обеспечения избыточности на случай отказа одного или нескольких серверных центров). Служба 1 хостинга может со временем включать в себя сотни тысяч или даже миллионов серверов 4, обслуживающих очень большое количество пользователей. Система 3 управления службы хостинга обеспечивает централизованное управление для службы 1 хостинга и управляет маршрутизаторами, серверами, системами сжатия видео, системами учета и формирования счетов и т.д. В одном варианте осуществления система 3 управления службы хостинга реализована в распределенной системе обработки данных на основе операционной системы Linux, соединенной с дисковыми массивами типа RAID, используемыми для хранения баз данных для пользовательской информации, информации для сервера и статистических данных системы. В вышеизложенных описаниях различные операции, реализованные службой 1 хостинга, если они не отнесены к другим специальным системам, инициируются и управляются системой 3 управления службы хостинга.

Служба хостинга(рис. 2.2) включает в себя несколько серверов, например таких серверов, которые в настоящее время имеются в продаже от Intel, IBM, Hewlett Packard и других. В качестве альтернативы, серверы могут быть собраны в заказной конфигурации из компонентов, или, в конце концов, они могут быть интегрированы так, что весь сервер будет реализован как однокристалльная интегральная схема.

Несмотря на то что на этой схеме для примера изображено небольшое количество серверов 4, при фактическом применении может быть только один сервер 4, или миллионы серверов 4, или больше. Все серверы 4 могут быть сконфигурированы одинаково (в качестве примера некоторых параметров конфигурации, с одинаковыми производительностью и типом CPU, с или без GPU, и, если с GPU, то с одинаковыми производительностью и типом GPU, с одинаковым количеством CPU и GPU, с одинаковыми объемом и типом/скоростью RAM и с одинаковой конфигурацией RAM), или различные подмножества серверов 4 могут иметь одинаковую конфигурацию (например,

25% серверов могут быть выполнены определенным способом, 50% другим способом и 25% еще одним способом), или все серверы 4 могут быть разными.

В одном варианте осуществления серверы 4 без диска, т.е. вместо того чтобы иметь свое собственное локальное массовое запоминающее устройство (независимо от того, является ли оно оптическим или магнитным запоминающим устройством, или запоминающим устройством на полупроводниках, например флэш-памятью, или другим средством массовой памяти, выполняющим аналогичную функцию), каждый сервер имеет доступ к общей массовой памяти через быстродействующую магистраль или ускоренное сетевое соединение. В одном варианте осуществления этим высокоскоростным соединением является сеть 5 хранения данных (SAN), соединенная с набором матриц 7 независимых дисковых накопителей с избыточностью (RAID) с соединениями между устройствами, реализованными с использованием технологии Gigabit Ethernet. Как известно специалистам в данной области техники, SAN 5 можно использовать для объединения множества дисковых массивов 7 типа RAID вместе, что в результате приводит к чрезвычайно широкой полосе пропускания - приближающейся или, возможно, превышающей полосу пропускания, доступную из RAM, используемой в современных игровых консолях и PC. И, в то время как дисковые массивы типа RAID на основе вращающихся носителей информации, например магнитных носителей информации, часто имеют существенное время ожидания доступа во время поиска, дисковые массивы типа RAID на основе полупроводниковой памяти могут быть реализованы с гораздо меньшим временем ожидания доступа. В другой конфигурации некоторые или все серверы 4 предоставляют некоторые или все свои собственные массовые запоминающие устройства локально. Например, на сервере 4 может храниться информация, к которой часто обращаются, например ее операционная система и копия видеоигры или приложения, на локальном запоминающем устройстве на основе флэш-памяти с малым временем ожидания, но он может использовать SAN для доступа к дисковым массивам 405 типа RAID, на основе вращающихся носителей информации с большим временем поиска, для получения доступа к

большим базам данных геометрии или информации о состоянии игры с меньшей частотой. Кроме того, в одном варианте осуществления служба 1 хостинга применяет логику б сжатия видео с малым временем ожидания, описанную подробно ниже. Логика б сжатия видео может быть реализована в программном обеспечении, аппаратном обеспечении или их любой комбинации (определенные варианты осуществления которой описаны ниже). Логика б сжатия видео включает в себя логику для сжатия аудио- и визуального материала.

При операции во время ведения видеоигры или использования приложения на территории 2 пользователя посредством клавиатуры, мыши, игрового контроллера или другого устройства 15 ввода, логика 13 управляющего сигнала на клиенте 14 передает управляющие сигналы 9-10 (как правило, в виде пакетов UDP), представляющие нажатия клавиши (и другие виды вводов пользователя), приводимые в действие пользователем, в службу 1 хостинга. Управляющие сигналы от данного пользователя маршрутизируются в соответствующий сервер 4 (или серверы, если на устройство ввода пользователя реагируют множество серверов). Как изображено на рис. 2.4, управляющие сигналы 9 могут маршрутизироваться в серверы 4 через SAN. В качестве альтернативы или в дополнение, управляющие сигналы 10 могут маршрутизироваться непосредственно в серверы 4 по сети службы хостинга (например, по локальной сети Ethernet). Независимо от того как они были переданы, сервер или серверы исполняют игру или прикладное программное обеспечение в ответ на управляющие сигналы 9-10. Несмотря на то что на рис. 2.4 не изображены различные сетевые компоненты, например брандмауэр(ы) и/или шлюз(ы), они могут обрабатывать входящий и исходящий трафик на границе службы 1 хостинга (например, между службой 1 хостинга и Internet 11) и/или на границе территории 2 пользователя между Internet 11 и домашним или офисным клиентом 14. Графический выход и аудиовыход исполняемой игры или прикладного программного обеспечения - т.е. новые последовательности видеоизображений - обеспечиваются в логику б сжатия видео с малым временем

ожидания, которая сжимает последовательности видеоизображений согласно способам сжатия видео с малым временем ожидания, например, согласно способам, описанным в этом документе, и передает сжатый видеопоток, обычно со сжатым или несжатым аудио, назад в клиента 14 через Internet 11 (или, как описано ниже, через оптимизированную высокоскоростную сетевую службу, которая преобходит в обход общего Internet). После этого логика 12 восстановления сжатого видео с малым временем ожидания в клиенте 14 восстанавливает сжатые видео- и аудиопотоки и визуализирует восстановленный видеопоток, и, как правило, воспроизводит восстановленный аудиопоток на дисплее 16. В качестве альтернативы, аудио может быть воспроизведено в динамиках, отделенных от дисплея 16, или совсем не воспроизводиться. Отметим, что, несмотря на то что устройство 15 ввода и дисплей 16 изображены на рис. 2.1 и рис. 2.2 как автономные устройства, они могут быть интегрированы внутри клиентских устройств, например портативных компьютеров или мобильных устройств.

Домашний или офисный клиент 14 (описанный ранее на рис. 2.1 и рис. 2.2 как домашний или офисный клиент 1) может быть очень экономичным устройством с малым энергопотреблением, с очень ограниченной вычислительной производительностью или производительностью графической подсистемы, и может иметь очень ограниченную локальную массовую память, или она совсем может отсутствовать. Напротив, каждый сервер 4, соединенный с SAN 5 и множеством RAID 7, может быть исключительно высокопроизводительной вычислительной системой, и, безусловно, если совместно используются множество серверов в конфигурации параллельной обработки, то почти не существует ограничений по вычислительной производительности системы и производительности графической подсистемы, которые могут быть использованы. И, из-за сжатия 6 видео с малым временем ожидания и сжатия 12 видео с малым временем ожидания, пользователя создается впечатление, что упомянутая вычислительная мощность серверов 4 предоставляется пользователю. Когда пользователь нажимает клавишу на

устройстве 15 ввода, изображение на дисплее 16 обновляется в ответ на нажатие этой клавиши без значимой задержки с точки зрения восприятия так, как в случае локального исполнения игры или прикладного программного обеспечения. Соответственно, с домашним или офисным клиентом 14, который является компьютером с очень низкой производительностью или только экономичной интегральной схемой, в которой реализовано восстановление сжатого видео с малым временем ожидания и логика 13 управляющего сигнала, пользователь обеспечен фактически произвольно выбираемой вычислительной мощностью из удаленного местоположения, причем создается впечатление, что они находятся локально. Это дает пользователям возможность запускать наиболее усовершенствованные, требующие интенсивной работы процессора (обычно новые) видеоигры и самые высокопроизводительные приложения.

На рис. 2.6 изображено очень элементарное и экономичное домашнее или офисное клиентское устройство 4. Это устройство является вариантом осуществления домашнего или офисного клиента 14 по рис. 2.4 и рис. 2.5. Длина его приблизительно равна 5 сантиметрам. У него есть гнездо 462 Ethernet, которое обеспечивает интерфейс с кабелем Ethernet с питанием через Ethernet (PoE), через которое оно получает питание и получает соединение с Internet. Оно может исполнять преобразование сетевых адресов (NAT) внутри сети, которая поддерживает NAT. В офисной обстановке многие новые коммутаторы Ethernet имеют PoE и доводят PoE непосредственно до гнезда Ethernet в офисе. В таком случае все, что требуется, - это кабель Ethernet от настенного гнезда до клиента 4. Если имеющееся соединение Ethernet не переносит энергию (например, в доме с DSL или кабельным модемом, но без PoE), то в продаже имеются экономичные настенные "модули" (т.е. источники питания), которые принимают кабель Ethernet без питания и выводят Ethernet с PoE. Клиент 4 содержит логику 13 управляющего сигнала (по рис. 2.4), которая соединена с беспроводным интерфейсом Bluetooth, который обеспечивает интерфейс с устройствами 7 ввода Bluetooth, например клавиатурой, мышью, игровым контроллером и/или микрофоном, и/или наушниками. Кроме того, один вариант осуществления

клиента 4 может выводить видео со скоростью 120 кадр/с, при соединении с дисплеем 6 может поддерживать видео 120 кадр/с и передавать сигналы (как правило, через инфракрасное излучение) в затворные очки 5 для попеременного закрытия то одного глаза, то другого с каждым последовательным кадром. Результатом, воспринимаемым пользователем, является стереоскопическое 3D изображение, которое ”выскакивает” из экрана дисплея. Одним таким дисплеем 6, который поддерживает такую операцию, является Samsung HL-T5076S. Так как видеопоток для каждого глаза является отдельным, в одном варианте осуществления два независимых видеопотока сжимаются службой 1 хостинга, кадры чередуются во времени, и сжатые кадры восстанавливаются как два независимых процесса восстановления сжатых данных внутри клиента 4. Клиент 4 также содержит логику 12 восстановления сжатого видео с малым временем ожидания, которая восстанавливает сжатое входящее видео и аудио и выводит через HDMI (мультимедийный интерфейс высокой четкости), соединительный кабель 2, который подключают к SDTV (телевидение стандартной четкости) или HDTV (телевидение высокой четкости) 6, в случае если телевизор с видео и аудио, или к монитору 6, который поддерживает HDMI. Если монитор 6 пользователя не поддерживает HDMI, то может быть использован HDMI-DVI (цифровой видеоинтерфейс), но аудио будет утрачено. Согласно стандарту HDMI, характеристики 3 дисплея (например, поддерживаемые разрешающие способности, частоты кадров) передаются из дисплея 6, и эта информация затем возвращается через подключение 1 к Internet обратно в службу 1 хостинга, поэтому она может передавать потоком сжатое видео в формате, подходящем для дисплея.

На рис. 2.7 изображено домашнее или офисное клиентское устройство 5, которое является идентичным домашнему или офисному клиентскому устройству 4, изображенному на рис. 2.6, за исключением того, что у него больше внешних интерфейсов. Кроме того, клиент 5 может принимать любой PoE для питания, или он может работать от внешнего адаптера источника питания (не изображен), который вставляется в контактное гнездо на стене. С использованием входа

USB клиента 5, видеокамера 6 обеспечивает сжатое видео в клиент 5, которое выгружается клиентом 5 в службу 1 хостинга для использования, описанного ниже. Устройство сжатия с малым временем ожидания, использующее способы сжатия, описанные ниже, является встроенным в камеру 6. В дополнение к разъему Ethernet для подключения к Internet клиент 5 также имеет беспроводной интерфейс 802.11g с Internet. Оба интерфейса могут использовать NAT внутри сети, которая поддерживает NAT. Кроме того, в дополнение к разъему HDMI для вывода видео и аудио, клиент 5 также имеет разъем Dual Link DVI-I (двухканальный DVI-I), который включает в себя аналоговый выход (и с эталонным кабелем адаптера обеспечивает выход VGA). Он также имеет аналоговые выходы для композитного видео и S-видео. Для аудио, клиент 5 имеет левое/правое гнезда RCA аналогового стерео, и для вывода цифрового аудио он имеет выход TOSLINK (оптический выход).

В дополнение к беспроводному интерфейсу Bluetooth с устройствами 9 ввода, он также имеет гнезда USB для обеспечения интерфейса с устройствами ввода. На рис. 2.8 изображен один вариант осуществления внутренней архитектуры клиента 4. Или все, или некоторые из устройств, изображенные на схеме, могут быть реализованы в программируемой пользователем логической матрице, заказной схеме ASIC или в нескольких отдельных устройствах, или заказных, или имеющих в готовом виде. Ethernet с PoE 12 подсоединяется к интерфейсу 4 Ethernet. Питание 13 подается из Ethernet с PoE 12 и соединяется с остальными устройствами в клиенте 5. Шина 3 является общей шиной для обмена информацией между устройствами. Управляющий CPU 5 (почти любой небольшой CPU, например, CPU серии R4000 с быстродействием миллион команд в секунду (MIPS) и тактовой частотой 100 МГц со встроенной RAM отвечает требованиям), исполняющий небольшое клиентское управляющее приложение из флэш-памяти 2, обеспечивает выполнение стека протоколов для сети (т.е. интерфейс Ethernet), и также осуществляет связь со службой 1 хостинга, и конфигурирует все устройства в клиенте 5. Он также управляет интерфейсами с устройствами 7 ввода и отправляет пакеты обратно в службу 1

хостинга с пользовательскими данными контроллера, защищенными посредством прямого исправления ошибок, в случае необходимости. Кроме того, управляющий CPU 5 осуществляет текущий контроль пакетного трафика (например, если пакеты утеряны или задержаны, а также делает отметку времени их поступления). Эта информация отправляется обратно в службу 1 хостинга так, чтобы она могла постоянно осуществлять текущий контроль сетевого соединения и корректировать то, что она отправляет соответственно. Во флэш-память 2 первоначально во время изготовления загружают управляющую программу для управляющего CPU 5, а также серийный номер, который является уникальным для конкретной клиентской 5 единицы оборудования. Этот серийный номер обеспечивает возможность службе 1 хостинга однозначно идентифицировать клиентскую 5 единицу оборудования. Интерфейс 6 Bluetooth осуществляет связь с устройствами 7 ввода беспроводным способом через свою антенну, находящуюся внутри клиента 5. Устройство 7 восстановления сжатого видео является устройством восстановления сжатого видео с малым временем ожидания, сконфигурированным для реализации восстановления сжатого видео, описанного в этом документе. Существует большое количество устройств восстановления сжатого видео, или имеющих в готовом виде, или в виде интеллектуальной собственности (IP) конструкции, которая может быть интегрирована в FPGA или заказную схему ASIC. Одной компанией, предлагающей IP для декодера H.264, является Ocean Logic, Manly, NSW Australia. Преимуществом использования IP является то, что способы сжатия, используемые в этом описании, не соответствуют стандартам сжатия. Некоторые стандартные устройства восстановления сжатых данных являются достаточно гибкими для обеспечения способов сжатия, описанных в этом документе, но некоторые не могут их обеспечивать. Но, посредством IP, существует полная гибкость для перепроектирования устройства восстановления сжатых данных так, как требуется. Выход устройства восстановления сжатого видео соединен с подсистемой 8 вывода видео, которая соединяет видео с видеовыходом интерфейса 11 HDMI.

Подсистема 11 восстановления сжатого аудио реализуется или с использованием стандартного устройства восстановления сжатого аудио, которое имеется в продаже, или она может быть реализована как IP, или восстановление сжатого аудио может быть реализовано внутри управляющего процессора 6, который может, например, реализовывать устройство восстановления сжатого аудио Vorbis. Устройство, которое осуществляет восстановление сжатого аудио, соединено с подсистемой 12 вывода аудио, которая соединяет аудио с аудиовыходом интерфейса 490 HDMI.

На рис. 2.9 изображен один вариант осуществления внутренней архитектуры клиента 5. Как можно видеть, эта архитектура является идентичной архитектуре клиента 4 за исключением дополнительных интерфейсов и необязательного внешнего питания DC от адаптера источника питания, который вставляют в контактное гнездо на стене, и, если так используется, заменяет питание, которое может поступать из PoE 20 Ethernet. Функциональные возможности, которые присутствуют у клиента 5 не будут повторяться ниже, но дополнительные функциональные возможности описываются следующим образом.

CPU 6 осуществляет связь с дополнительными устройствами и конфигурирует их. Подсистема 5 WiFi обеспечивает беспроводной доступ к Internet, в качестве альтернативы, к Ethernet 10 через свою антенну. В продаже имеются подсистемы WiFi от широкого круга производителей, включающих в себя Atheros Communications, Santa Clara, CA (Санта-Клара, Калифорния).

Подсистема 8 USB обеспечивает альтернативный вариант по отношению к связи Bluetooth для проводных устройств 22 ввода USB. Подсистемы USB являются довольно стандартными и общедоступными для матриц FPGA и схем ASIC, а также они часто являются встроенными в имеющиеся в готовом виде устройства, выполняющие другие функции, подобные восстановлению сжатого видео. Подсистема 20 вывода видео выводит более широкий диапазон видеовыходов, чем внутри клиента 5. В дополнение к обеспечению видеовыхода HDMI 13, она обеспечивает DVI-I 491, S-видео 15 и композитное видео 16. Кроме того, когда для цифрового видео используется интерфейс DVI-I 491,

характеристики 1 дисплея передаются обратно из дисплея в управляющий CPU 6 для того, чтобы он мог уведомлять службу 1 хостинга о характеристиках дисплея 7. Все интерфейсы, обеспечиваемые подсистемой 10 вывода видео, являются довольно стандартными интерфейсами и общедоступными во многих видах.

Подсистема 12 вывода аудио выводит аудио в цифровой форме через цифровой интерфейс 17 (S/PDIF и/или TOSLINK) и аудио в аналоговой форме через стереофонический аналоговый интерфейс 18.

2.2.3 Анализ времени передачи сигнала туда и обратно

Несомненно, для реализации преимуществ, изложенных в предыдущем разделе, время передачи сигнала туда и обратно между действием пользователя с использованием устройства 8 ввода и появлением результата этого действия на дисплее 420 не должно превышать 70-80 мс. Это время ожидания должно учитывать все факторы на пути от устройства 8 ввода на территории 4 пользователя до службы 1 хостинга и обратно до территории 4 пользователя, до дисплея 9. На рис. 2.5 изображены различные компоненты и сети, через которые должны проходить сигналы, и выше этих компонентов и сетей находится временная шкала, на которой выведены в упорядоченном виде иллюстративные величины времени ожидания, которые можно ожидать в практической реализации. Отметим, что рис. 2.5 является упрощенной, и изображена только маршрутизация по критическому пути. Другая маршрутизация данных, используемая для других признаков системы, описана ниже. Двухнаправленные стрелки (например, стрелка 16) указывают время передачи сигнала туда и обратно, и однонаправленные стрелки (например, стрелка 20) указывают время передачи сигнала в одну сторону, и "≈" обозначают приблизительное измерение. Следует указать, что будут реальные ситуации, в которых нельзя будет получить перечисленные величины времени ожидания, но в большинстве случаев в США, с использованием соединений посредством кабельного модема и DSL с территорией 4 пользователя, эти величины времени ожидания могут быть получены в случаях, описанных в следующем разделе. Также отметим, что,

несмотря на то что сотовая беспроводная связь с Internet безусловно будет функционировать в изображенной системе, в большинстве современных американских сотовых систем передачи данных (например EVDO) существуют очень большие величины времени ожидания, и не могут быть получены величины времени ожидания, изображенные на рис. 2.5. Однако эти лежащие в основе принципы могут быть реализованы в будущих сотовых технологиях, в которых может быть реализован этот уровень времени ожидания.

Начиная с устройства 15 ввода на территории 4 пользователя, после того как пользователь приводит в действие устройство 15 ввода, управляющий сигнал пользователя отправляется в клиент 14 (который может быть автономным устройством, например, телевизионной абонентской приставкой, или он может быть программным обеспечением, или аппаратным обеспечением, функционирующими в другом устройстве, например PC или мобильном устройстве) и разбивается на пакеты (в формате UDP в одном варианте осуществления), и пакету дается адрес назначения для передачи в службу 1 хостинга. Пакет будет также содержать информацию для указания того, от какого пользователя поступают управляющие сигналы. После этого пакет(ы) управляющего сигнала пересылается(ются) через устройство 11 брандмауэра/маршрутизатора/NAT (преобразование сетевых адресов) в интерфейс 10 WAN. Интерфейс 10 WAN является интерфейсным устройством, обеспечиваемым для территории 2 пользователя провайдером ISP пользователя (провайдер Internet- услуг). Интерфейс 10 WAN может быть кабелем или модемом DSL, приемопередатчиком WiMax, приемопередатчиком для волоконно-оптической линии связи, интерфейсом сотовой системы передачи данных, интерфейсом передачи данных по протоколу IP через электросеть (Internet Protocol-over-powerline) или любым другим из множества интерфейсов с Internet. Кроме того, устройство 11 брандмауэра/маршрутизатора/NAT (и, возможно, интерфейс 442 WAN) может быть интегрировано в клиента 14. Примером этого может быть мобильный телефон, который включает в себя программное обеспечение для реализации функциональных возможностей домашнего или

офисного клиента 14, а также средство для маршрутизации и подключения к Internet беспроводным способом через некоторый стандарт (например, 802.11g). После этого интерфейс 10 WAN маршрутизирует управляющие сигналы в то, что называется в этом описании "точка входа в Internet" 12 для провайдера Internet-услуг (ISP) пользователя, которая является средством, которое обеспечивает интерфейс между средством передачи данных WAN, соединенным с территорией 2 пользователя, и общим Internet или частной сетью. Характеристики точки входа в Internet изменяются в зависимости от характера обеспечиваемой Internet-услуги. Для DSL это, как правило, будет центральная АТС телефонной компании, где расположен DSLAM. Для кабельных модемов это, как правило, будет центральная станция мультисистемного оператора (MSO) кабельной линии. Для систем сотовой связи это, как правило, будет диспетчерская, связанная с антенной мачтой сотовой связи. Но какова бы ни была природа точки входа в Internet, она далее маршрутизирует пакет(ы) управляющего сигнала в общий Internet 6. Пакет(ы) управляющего сигнала далее маршрутизируются в интерфейс 12 WAN со службой 1 хостинга, через то, что наиболее вероятно будет интерфейсом приемопередатчика для волоконно-оптической линии связи. WAN 12 далее маршрутизирует пакеты управляющего сигнала в логику 5 маршрутизации (которая может быть реализована многими различными способами, включающими в себя коммутатор Ethernet и серверы маршрутизации), которая определяет адрес пользователя и маршрутизирует управляющий(ие) сигнал(ы) в соответствующий сервер 3 для данного пользователя.

После этого сервер 3 принимает управляющие сигналы как входные сигналы для игры или прикладного программного обеспечения, которые исполняются на сервере 4, и использует их для обработки следующего кадра игры или приложения. После формирования следующего кадра, видео и аудио выводятся из сервера 4 в устройство сжатия видео. Видео и аудио могут быть выведены из сервера 4 в устройство сжатия посредством различных средств. Во-первых, устройство сжатия может быть встроено в сервер 4, соответственно, сжатие

может быть реализовано локально внутри сервера 4. Или видео и/или аудио могут выводиться в виде пакетов через сетевое соединение, например соединение Ethernet, в сеть, которая является любой частной сетью между сервером 4 и устройством 6 сжатия видео, или через совместно используемую сеть, например SAN 5. Или видео может выводиться через разъем видеовыхода из сервера 5, например разъем VGA или DVI, и затем захватываться устройством сжатия 6 видео. Кроме того, аудио может выводиться из сервера 4 или как цифровое аудио (например, через разъем S/PDIF или TOSLINK), или как аналоговое аудио, которое оцифровано и закодировано логикой сжатия аудио внутри устройства 6 сжатия видео. После захвата устройством 6 сжатия видео видеокadra и аудио, формируемого в течение этого периода кадра, из сервера 4, устройство сжатия видео сжимает видео и аудио с использованием способов, описанных ниже. После сжатия видео и аудио, формируются пакеты с адресом для отправки обратно в клиент 14 пользователя, и их маршрутизируют в интерфейс 12 WAN, который далее маршрутизирует видео- и аудиопакеты через общий Internet 6, который далее маршрутизирует видео- и аудиопакеты в точку 12 входа в Internet ISP пользователя, которая маршрутизирует видео- и аудиопакеты в интерфейс 10 WAN на территории пользователя, который маршрутизирует видео- и аудиопакеты в устройство 11 брандмауэра/маршрутизатора/NAT, которое далее маршрутизирует видео- и аудиопакеты в клиента 4. Клиент 4 восстанавливает сжатые видео и аудио и после этого выводит видео на экран дисплея 9 (или на встроенный дисплей клиента) и отправляет аудио на дисплей 9, или в отдельные усилитель/динамики, или в усилитель/динамики, встроенные в клиента.

Для того чтобы пользователь воспринимал весь только описанный процесс без запаздывания, двухсторонняя задержка должна быть меньше 70 или 80 мс. Некоторые из задержек времени ожидания в описанном пути туда и обратно контролируются службой 1 хостинга и/или пользователем, а другие не контролируются. Однако следующие измерения, основанные на анализе и тестировании большого количества реальных сценариев, являются

приблизительными. Время передачи в одном направлении для отправки управляющих сигналов 14 обычно меньше 1 мс, маршрутизация туда и обратно через территорию 15 пользователя обычно выполняется с использованием общедоступных коммутаторов потребительского уровня брандмауэра/маршрутизатора/NAT по локальной сети Ethernet приблизительно за 1 мс. I (Соединения) пользователя ISP значительно отличаются по их задержкам 16 передачи сигнала туда и обратно, но с провайдерами кабельной модемной связи и DSL, как правило, наблюдается задержка между 10 и 25мс. Время передачи сигнала туда и обратно по общему Internet 6 может очень отличаться в зависимости от того, как маршрутизируется трафик, и существуют ли какие-нибудь отказы на маршруте (и эти вопросы обсуждаются ниже), но, как правило, общий Internet обеспечивает довольно оптимальные маршруты, и время ожидания в значительной степени определяется скоростью света через светопровод с учетом расстояния до пункта назначения. Как также обсуждается ниже, авторы установили службу 1 хостинга самое большее на расстоянии примерно 1000 миль (1600 км), на котором, как ожидается, она будет располагаться от территории 21 пользователя. Фактическое время передачи сигнала через Internet на расстояние 1000 миль (1600 км) (2000 миль (3200 км) туда и обратно) приблизительно равно 22 мс. Интерфейс 441 WAN со службой 1 хостинга является обычно интерфейсом волоконно-оптической линии связи с высокой скоростью коммерческого уровня с незначительным временем ожидания. Соответственно, время 17 ожидания общего Internet обычно находится между 1 и 10 мс. Может быть достигнуто время ожидания маршрутизации 18 в одном направлении через службу 1 хостинга меньше 1 мс. Сервер 4, как правило, вычисляет новый кадр для игры или приложения за время меньшее, чем один период кадра (который при 60 кадр/с равен 16,7 мс), соответственно, 16 мс является приемлемым для использования максимальным временем 19 передачи сигнала в одну сторону. В оптимизированной аппаратной реализации алгоритмов сжатия аудио и сжатия видео, описанных в этом документе, сжатие 20 может быть выполнено за 1 мс. В менее

оптимизированных версиях, сжатие может занимать 6 мс (несомненно, еще менее оптимизированные версии могут занимать больше времени, но такие реализации могут влиять на общее время ожидания передачи туда и обратно, и может потребоваться, чтобы другие величины времени ожидания были меньше (например, может быть уменьшено допустимое расстояние через общий Internet) для поддержания целевого времени ожидания 70-80 мс). Величины времени передачи сигнала туда и обратно для Internet 18, (соединения) 16 пользователя ISP, и маршрутизации 19 по территории пользователя уже были рассмотрены, поэтому остается рассмотреть время ожидания восстановления 18 сжатого видео, которое, в зависимости от того, реализовано ли восстановление 18 сжатого видео в специализированном аппаратном обеспечении, или реализовано ли оно в программном обеспечении на клиентском устройстве 7 (например, PC или мобильном устройстве), может изменяться в зависимости от размера дисплея и производительности CPU для восстановления сжатых данных. Как правило, восстановление 21 сжатых данных занимает между 1 и 8 мс.

Соответственно, посредством суммирования всех величин времени ожидания в самом плохом случае, наблюдаемом на практике, можно определить время передачи сигнала туда и обратно в самом плохом случае, которое, как можно ожидать, будет испытывать пользователь системы, изображенной на рис. 2.4. Они равны $1+1+25+22+1+16+6+8=80$ мс. И, безусловно, на практике (с оговорками, обсуждаемыми ниже), это примерно равно времени передачи сигнала туда и обратно, наблюдаемому при использовании версии пилотной модели, изображенной на рис. 2.4 с использованием имеющихся в готовом виде PC Windows в качестве клиентских устройств и домашних соединений кабельной модемной связи и DSL в пределах Европы. Несомненно, сценарии, которые лучше самого плохого случая, могут в результате привести к гораздо меньшим величинам времени ожидания, но на них нельзя основываться при разработке коммерческой услуги, которая широко используется.

Для получения величин времени ожидания, выведенных в упорядоченном виде на рис. 2.5, при передаче сигнала через общий Internet, требуется, чтобы

устройство 4 сжатия видео и устройство 12 восстановления сжатого видео по рис. 2.4 в клиенте 7 формировали поток пакетов с очень конкретными характеристиками, например, последовательность пакетов, формируемая на всем пути от службы 1 хостинга до дисплея 9, не подвергалась задержкам или чрезмерной потере пакетов и, в частности, постоянно удовлетворяла ограничениям полосы пропускания, доступной пользователю по подключению к Internet пользователя через интерфейс 13 WAN и брандмауэр/маршрутизатор/NAT 11. Кроме того, устройство сжатия видео должно создавать поток пакетов, который достаточно устойчив, чтобы он мог допускать неизбежную потерю пакетов и изменение порядка следования пакетов, которое происходит в обычном Internet и сетевых передачах.

2.2.4 Сжатие видео с малым временем ожидания

Для достижения вышеизложенных целей, в одном варианте осуществления используется новый подход к сжатию видео, посредством которого уменьшают время ожидания и требования по пиковой полосе пропускания для передачи видео. До описания этих вариантов осуществления будет обеспечен анализ современных способов сжатия видео согласно рис. 2.10 и рис. 2.11-рис. 2.12. Несомненно, эти способы могут быть применены согласно лежащим в основе принципам, если пользователь будет обеспечен полосой пропускания, достаточной для обработки данных со скоростью передачи, требуемой этими способами. Отметим, что в этом описании не рассматривается сжатие аудио, за исключением констатации того, что оно реализуется одновременно и синхронно со сжатием видео. Существуют способы сжатия аудио предшествующего уровня техники, которые удовлетворяют требованиям для этой системы.

На рис. 2.10 изображен один конкретный способ предшествующего уровня техники для сжатия видео, в котором каждый отдельный видеокادر 1-3 сжимается посредством логики 7 сжатия с использованием конкретного алгоритма сжатия для формирования последовательности сжатых кадров 4-6. Одним вариантом осуществления этого способа является ”алгоритм сжатия

движущихся изображений MJPEG”, в котором каждый кадр сжимается согласно алгоритму сжатия неподвижного изображения, разработанный Объединенной группой экспертов по машинной обработке фотографических изображений (JPEG) на основе дискретного косинусного преобразования (DCT). Однако могут быть применены различные другие типы алгоритмов сжатия, хотя, по-прежнему, с выполнением этих лежащих в основе принципов (например, алгоритмы сжатия, основанные на вейвлете, например, JPEG- 2000).

Одна проблема, связанная с этим типом сжатия, состоит в том, что оно уменьшает скорость передачи данных каждого кадра, но оно не использует сходные черты между последовательными кадрами для уменьшения скорости передачи данных всего видеопотока. Например, как изображено на рис. 2.10 предположим, что частота кадров $640 \times 480 \times 24$ бит/пиксел = $640 \times 480 \times 24 / 8 / 1024 = 900$ Килобайт/кадр (Кбайт/кадр), для данного качества изображения, посредством алгоритма сжатия движущихся изображений MJPEG можно сжать поток с коэффициентом 10, что в результате приводит к потоку данных 90 Кбайт/кадр. При 60 кадров/сек потребовалась бы полоса пропускания канала $90 \text{ Кбайт} \times 8 \text{ бит} \times 60 \text{ кадров/сек} = 42,2 \text{ Мбит/с}$, которая была бы слишком широкой полосой пропускания почти для всех домашних подключений к Internet в США в настоящее время и слишком широкой полосой пропускания для многих офисных подключений к Internet. Безусловно, с учетом того, что при этом потребуется постоянный поток данных с такой широкой полосой пропускания, и он будет обслуживать только одного пользователя, даже в офисной среде LAN, он будет потреблять большую часть полосы пропускания LAN Ethernet 100 Мбит/с и сильно нагружать коммутаторы Ethernet, поддерживающие LAN. Соответственно, сжатие для движущегося видео неэффективно при сравнении с другими способами сжатия (например, способами сжатия, описанными ниже). Кроме того, алгоритмы сжатия одного кадра, подобные JPEG и JPEG-2000, которые используют алгоритмы сжатия с потерями, выводят артефакты сжатия, которые могут быть не заметны на неподвижных изображениях (например, артефакт внутри густой листвы в сцене может не казаться артефактом, так как

глаз не знает точно, как густая листва должна выглядеть). Но, если сцена движется, то артефакт может быть заметным, потому что глаз замечает, что артефакт (изменяется) от кадра к кадру, несмотря на то что артефакт находится в области сцены, где он мог быть незаметным в неподвижном изображении. Это в результате приводит к восприятию "фоновых помех" в последовательности кадров, сходных по виду с помехами "снег", видимых во время граничного аналогового телевизионного приема. Несомненно, этот тип сжатия может, тем не менее, использоваться в определенных вариантах осуществления, описанных в этом документе, но, вообще говоря, для данного качества восприятия требуется устранение фоновых помех в сцене, высокая скорость передачи данных (т.е. низкий коэффициент сжатия).

Другие типы сжатия, например H.264 или Windows Media VC9, MPEG2 и MPEG4, все являются более эффективными при сжатии видеопотока, потому что они используют сходные черты между последовательными кадрами. Все эти способы основаны на одинаковых общих способах сжатия видео. Соответственно, несмотря на то что будет описан стандарт H.264, однако идентичные общие принципы относятся к различным другим алгоритмам сжатия. Большое количество устройств сжатия H.264 и устройство восстановления сжатых данных имеются в наличии, в том числе библиотека программного обеспечения с открытым исходным текстом x264 для сжатия H.264 и библиотека программного обеспечения с открытым исходным текстом FFmpeg для восстановления сжатых данных H.264.

На рис. 2.11 и рис. 2.12 изображен иллюстративный способ сжатия предшествующего уровня техники, в котором последовательность несжатых видеок кадров 1-3, 559-561 сжимают посредством логики 9 сжатия в последовательность "I-кадров" 611, 671, "P-кадров" 612-613 и "B-кадров" 670. Вертикальная ось на рис. 2.11 в общем обозначает получающийся в результате размер каждого из закодированных кадров (несмотря на то что кадры не вычерчены в масштабе). Как описано выше, кодирование видео с использованием I-кадров, B-кадров и P-кадров понятно специалистам в данной

области техники. В нескольких словах, I-кадр 611 является сжатием на основе DCT полного несжатого кадра 11 (аналогичным сжатому изображению JPEG, как описано выше). P-кадры 5-6, в общем, значительно меньше по размеру, чем I-кадры 611, потому что они используют данные предыдущего I-кадра или P-кадра, то есть они содержат данные, указывающие на изменения между предыдущим I-кадром или P-кадром. В-кадры 9 являются аналогичными P-кадрам за исключением того, что В- кадры используют кадр в следующем опорном кадре, а также, возможно, кадр в предыдущем опорном кадре.

В нижеследующем обсуждении будет предполагаться, что требуемая частота кадров равна 60 кадров/секунда, что каждый I-кадр равен приблизительно 160 Кбит, средний P-кадр и В-кадр равен 16 Кбит, и что новый I-кадр формируется каждую секунду. С этим набором параметров средняя скорость передачи данных равна $160 \text{ Кбит} + 16 \text{ Кбит} \times 59 = 1,1 \text{ Мбит/с}$. Эта скорость передачи данных вполне находится в пределах максимальной скорости передачи данных для многих современных высокоскоростных подключений к Internet в домах и офисах. Посредством этого способа также можно решать проблему фоновых помех при только-внутрикадровом кодировании, потому что P- и В-кадры отслеживают различия между кадрами, поэтому артефакты сжатия не должны появляться и исчезать от кадра к кадру, следовательно, уменьшается проблема фоновых помех, описанная выше.

Одна проблема, связанная с вышеизложенными видами сжатия, состоит в том, что, несмотря на то что средняя скорость передачи данных является относительно низкой (например, 1,1 Мбит/с), передача одного I-кадра может занимать несколько периодов кадра. Например, с использованием способов предшествующего уровня техники, сетевого соединения 2,2 Мбит/с (например, DSL или кабельный модем с пиком 2,2 Мбит/с максимальной доступной скорости передачи данных 2 по рис. 2.3) обычно достаточно для передачи потоком видео со скоростью 1,1 Мбит/с с I-кадром 160 Кбит/с каждые 60 кадров. Это может быть выполнено, если устройство восстановления сжатых

данных будет ставить в очередь 1 секунду видео до восстановления сжатого видео. За 1 секунду передается 1,1 Мбит данных, что может быть легко обеспечено при максимальной доступной скорости передачи данных 2,2 Мбит/с, даже с предположением, что доступная скорость передачи данных может периодически снижаться на 50%. К сожалению, этот подход предшествующего уровня техники в результате приводит ко времени ожидания в 1 секунду для видео из-за 1-секундного видеобуфера в приемнике. Такая задержка отвечает требованиям многих приложений предшествующего уровня техники (например, воспроизведение линейного видео), но является слишком долгим временем ожидания для динамичных видеоигр-боевиков, в которых не допускается время ожидания более 70-80 мс.

Если попытаться исключить 1-секундный видеобуфер, то это по-прежнему не приведет к сокращению времени ожидания, достаточному для динамичных видеоигр-боевиков. Например, использование В-кадров, как описано ранее, неизбежно влечет за собой прием всех В-кадров, предшествующих I-кадру, а также I-кадра. Если предположить, что 59 кадров не-I-кадров примерно разделены между Р- и В-кадрами, то принимается, по меньшей мере, 29 В-кадров и I-кадр до того, как любой В-кадр может быть выведен на экран. Соответственно, независимо от доступной полосы пропускания канала, это неизбежно влечет за собой задержку $29+1=30$ кадров длительностью $1/60$ секунды каждый, или 500 мс времени ожидания. Очевидно, что это слишком долго.

Соответственно, другой подход должен исключать В-кадры и использовать только I- и Р-кадры. Этот подход исключает неизбежное время ожидания, вводимое В-кадрами, так как декодирование каждого Р-кадра зависит только от предыдущего принятого кадра. Проблема, которая остается при этом подходе, состоит в том, что I-кадр настолько больше среднего Р-кадра, что на канале с небольшой полосой пропускания, как обычно бывает в большинстве домов и во многих офисах, передача I-кадра добавляет значительное время ожидания. Это изображено на рис. 2.12. Скорость 4 передачи данных видеопотока ниже

доступной максимальной скорости 1 передачи данных за исключением I-кадров, где пиковая скорость передачи данных, требуемая для I-кадров 3 намного превышает доступную максимальную скорость 2 передачи данных (и даже номинальную максимальную скорость 1 передачи данных). Скорость передачи данных, требуемая P-кадрами, меньше доступной максимальной скорости передачи данных. Даже если пики в 2,2 Мбит/с доступной максимальной скоростью передачи данных постоянно остаются на своей пиковой скорости 2,2 Мбит/с, то передача I-кадра займет $160 \text{ Кбит} / 2,2 \text{ Мбит} = 71 \text{ мс}$, и если доступная максимальная скорость 2 передачи данных снизится на 50% (1,1 Мбит/с), то передача I-кадра займет 142 мс. Соответственно, время ожидания при передаче I-кадра находится приблизительно между 71-142 мс. Это время ожидания является дополнением к величинам времени ожидания, идентифицированным на фиг.4b, которые в самом плохом случае составляют в сумме 70 мс, поэтому в результате общее время передачи сигнала туда и обратно, от момента, когда пользователь приводит в действие устройство 8 ввода до того, как изображение появляется на дисплее 9, равно 141-222 мс, которое является слишком большим. И если доступная максимальная скорость передачи данных становится меньше 2,2 Мбит/с, то время ожидания увеличивается дополнительно.

Отметим также, что, в общем, существуют серьезные последствия "создания пробки" ISP посредством пиковой скорости 3 передачи данных, которая намного превышает доступную скорость 2 передачи данных. Оборудование у разных ISP ведет себя по-разному, но следующие линии поведения являются довольно распространенными среди ISP кабельной модемной связи и DSL при приеме пакетов с гораздо большей скоростью передачи данных, чем доступная скорость 2 передачи данных, (а) задержка пакетов посредством постановки их в очередь (введение времени ожидания), (б) сброс некоторых или всех пакетов, (с) прерывание соединения на некоторый период времени (скорее всего потому, что ISP будет обеспокоен тем, что это является злонамеренной атакой, например, атакой "отказ в обслуживании").

Соответственно, передача потока пакетов с полной скоростью передачи данных с такими характеристиками, как характеристики, изображенные на рис. 2.12, не является практически осуществимым вариантом. Пики 3 могут стоять в очереди в службе 1 хостинга и отправляться со скоростью передачи данных ниже доступной максимальной скорости передачи данных, что вводит неприемлемое время ожидания, описанное в предыдущем абзаце. Кроме того, последовательность 4 скорости передачи данных видеопотока, изображенная на рис. 2.12, является очень "нормальной" последовательностью скорости передачи данных видеопотока и может быть видом последовательности скорости передачи данных, которая, как можно ожидать, получается в результате сжатия видео из видеопоследовательности, которая не изменяется очень сильно и обнаруживает очень мало движения. Последовательность 4 скорости передачи данных видеопотока, изображенная на рис.6с, является последовательностью, которую обычно можно ожидать наблюдать из видео с гораздо большим количеством действий, например которая может формироваться в кинофильме или видеоигре, или в некотором прикладном программном обеспечении. Отметим, что наряду с пиками 3 I-кадра, существуют также пики Р-кадра, например 4 и 5, которые являются довольно большими и превышают доступную максимальную скорость передачи данных во многих случаях. Несмотря на то что эти пики Р-кадра не совсем такие большие, как пики I-кадра, они, тем не менее, являются слишком большими для передачи их по каналу с полной скоростью передачи данных, и как и в случае с пиками I-кадра, они, пики Р-кадра, должны передаваться медленно (следовательно, с возрастающим временем ожидания).

На канале с широкой полосой пропускания (например, LAN 100 Мбит/с или частное соединение с широкой полосой пропускания 100 Мбит/с) в сети могут допускаться большие пики, например пики 3 I-кадра или пики 5 Р-кадра, и, в принципе, может поддерживаться малое время ожидания. Но такие сети часто совместно используются многими пользователями (например, в офисной обстановке), и такие данные "с пиками" влияют на производительность LAN, в

частности, если сетевой трафик маршрутизируется к частному, совместно используемому соединению (например, из удаленного центра обработки и хранения данных в офис). Прежде всего следует учитывать, что в этом примере видеопоток с относительно низким разрешением 640×480 пикселей с 60 кадр/с. Потоки HDTV 1920×1080 с 60 кадр/с легко обрабатываются современными компьютерами и дисплеями, и все больше имеется в продаже дисплеев с разрешающей способностью 2560×1440 с 60 кадр/с (например, дисплей 30" корпорации Apple, Inc). Видеопоследовательности с высокой активностью с 1920×1080 с 60 кадр/с может потребоваться 4,5 Мбит/с с использованием сжатия H.264 для приемлемого уровня качества. Если предположить, что I-кадры достигают пика при номинальной скорости передачи данных 10X, то в результате получаются пики 45 Мбит/с, а также меньший, но, тем не менее, значительный, пик P- кадра. Если несколько пользователей принимают видеопотоки по идентичной сети 100 Мбит/с (например, частное сетевое соединение между офисом и центром обработки и хранения данных), то легко увидеть, как пики видеопотока нескольких пользователей могут объединяться с переполнением полосы пропускания сети и с возможным переполнением полосы пропускания магистралей коммутаторов, поддерживающих пользователей в сети. Даже в случае сети по технологии Gigabit Ethernet, если достаточное количество пиков достаточного количества пользователей объединяются одновременно, то это может переполнить сеть или сетевые коммутаторы. И когда видео с разрешающей способностью 2560×1440 станет более обычным явлением, средняя скорость передачи данных видеопотока может быть равна 9,5 Мбит/с, что в результате, возможно, приведет к пиковой скорости передачи данных 95 Мбит/с. Не вызывает сомнения, что соединение 100 Мбит/с между центром обработки и хранения данных и офисом (которое в настоящее время является исключительно высокоскоростным соединением) будет полностью переполнено пиковым трафиком отдельного пользователя. Соответственно, несмотря на то что LAN и частные сетевые соединения могут допускать потоковое видео с пиками, потоковое видео с высокими пиками нежелательно,

и может потребоваться специальное планирование и адаптация отделом IT офиса. Несомненно, для стандартных приложений линейного видео, эти вопросы не являются проблемой, потому что скорость передачи данных ”сглаживается” в момент передачи, и данные для каждого кадра ниже максимальной доступной скорости 622 передачи данных, и в буфере в клиенте сохраняется последовательность I-, P- и B-кадров до того, как они будут восстановлены. Соответственно, скорость передачи данных по сети остается близкой к средней скорости передачи данных видеопотока. К сожалению, это вводит время ожидания, даже если B-кадры не используются, что является неприемлемым для приложений с малым временем ожидания, например видеоигры и приложения требуют малого времени отклика.

Ранее для уменьшения видеопотоков, которые имеют высокие пики, является использовались способ, часто называемый ”Постоянный битрейт” (постоянная скорость передачи битов, CBR). Несмотря на то что представляется, что термин CBR подразумевает, что все кадры сжимаются так, чтобы у них была одинаковая скорость передачи битов (т.е. размер), он обычно относится к схеме сжатия, в которой допускается максимальная скорость передачи битов по определенному количеству кадров (в нашем случае, 1 кадр). Например, в случае рис.2.13, если ограничение CBR применяется к кодированию, которое ограничивает скорость передачи битов, например до 70% номинальной максимальной скорости 1 передачи данных, тогда алгоритм сжатия ограничивает сжатие каждого из кадров так, чтобы любой кадр, который обычно сжимается с использованием больше 70% номинальной максимальной скорости 1 передачи данных, сжимается посредством меньшего количества битов. Результатом этого является то, что кадры, которым обычно требуется больше битов для поддержания данного уровня качества, ”испытывают недостаток” в битах, и качество изображения этих кадров хуже, чем качество изображения других кадров, которым не требуется большее количество битов, чем 70% (скорости) максимальной скорости 1 передачи данных. Этот подход может привести к приемлемым результатам для определенных типов сжатого видео, когда (а)

ождается мало движения или изменений сцены и (b) пользователи могут считать приемлемым периодическое ухудшение качества. Хорошим примером применения, подходящего для CBR, является проведение видеотелеконференции, так как в ней мало пиков, и если качество ухудшается на короткий период (например, если камерой выполняют панорамирование, что в результате приводит к существенному движению сцены и большим пиками, то во время панорамирования может быть не достаточно битов для сжатия высококачественного изображения, что в результате может привести к ухудшенному качеству изображения), это является приемлемым для большинства пользователей. К сожалению, CBR плохо подходит для многих других применений, в которых существуют сцены высокого уровня сложности или с большим движением, и/или в которых требуется достаточно постоянный уровень качества.

Логика 4 сжатия с малым временем ожидания, применяемая в одном варианте осуществления, использует несколько разных способов для решения круга проблем, связанных с передачей потоком сжатого видео с малым временем ожидания, при поддержании высокого уровня качества. Во-первых, логика 4 сжатия с малым временем ожидания формирует только I-кадры и P-кадры, следовательно, уменьшается необходимость ожидания в течение нескольких периодов кадра для декодирования каждого B-кадра. Кроме того, как изображено на рис. 2.14, в одном варианте осуществления, логика 4 сжатия с малым временем ожидания разбивает каждый несжатый кадр 1-4 в последовательность "фрагментов" ("tile", "тайл") и отдельно кодирует каждый фрагмент или как I-кадр, или как P-кадр. Группа сжатых I- кадров и P-кадров называется в этом описании "R-кадры" 6-4. В конкретном примере, изображенном на рис. 2.14, каждый несжатый кадр разбивается в матрицу 4×4 из 16 фрагментов. Однако эти лежащие в основе принципы не ограничены никакой конкретной схемой подразделения. В одном варианте осуществления логика 5 сжатия с малым временем ожидания разделяет видеокادر на несколько фрагментов и кодирует (т.е. сжимает) один фрагмент из каждого кадра как I-

кадр (т.е. этот фрагмент сжимается так, как если бы он был отдельным видеокадром размером 1/16 полного изображения, и сжатие, используемое для этого "мини"-кадра, является сжатием I-кадра), а остальные фрагменты изображения как P-кадры (т.е. сжатие, используемое для каждой "мини" 1/16 (части) кадра, является сжатием P-кадра). Фрагменты, сжатые как I-кадры и как P-кадры, должны называться "I-фрагментами" и "P-фрагментами" соответственно. С каждым последующим видеокадром, фрагмент, который должен быть закодирован как I-фрагмент, меняется. Соответственно, в данный период кадра, только один фрагмент из фрагментов в видеокадре является I-фрагментом, а остальные фрагменты являются P-фрагментами. Например, на рис. 2.14, фрагмент 0 несжатого кадра 1 закодирован как I-фрагмент I_0 , а остальные 1-15 фрагменты закодированы как P-фрагменты с P_1 по P_{15} для вывода R-кадра 711. В следующем несжатом видеокадре 2, фрагмент 1 несжатого кадра 1 закодирован как I-фрагмент I_1 , а остальные фрагменты 0 и с 2 по 15 закодированы как P-фрагменты, P_0 и с P_2 по P_{15} , для вывода R-кадра 7.

Соответственно, I-фрагменты и P-фрагменты для фрагментов постепенно чередуются во времени по последовательным кадрам. Процесс продолжается до тех пор, пока не будет сформирован R-(кадр) 4 с последним фрагментом в матрице, закодированным как I-фрагмент (т. е. I_{15}). После этого процесс повторяется с формированием еще одного R-кадра, например кадра 6 (т.е. кодирование I-фрагмента для фрагмента 0) и т.д. Несмотря на то что на рис.2.4 не 30 изображено, в одном варианте осуществления первый R-кадр видеопоследовательности из R-кадров содержит только I-фрагменты (т.е. последующие (R)-кадры содержат данные опорного изображения, исходя из которого вычисляется движение). В качестве альтернативы, в одном варианте осуществления последовательность в начальный момент использует схему I-фрагмента, идентичную обычной, но не включает в себя P-фрагменты для тех фрагментов, которые еще не закодированы посредством I-фрагмента. Другими словами, определенные фрагменты не кодируются никакими данными до тех пор, пока не поступит первый I-фрагмент, следовательно, устраняются начальные

пики в скорости 2 передачи данных видеопотока по рис. 2.17, что более подробно объясняется ниже. Кроме того, как описано ниже, различные другие размеры и формы могут использоваться для фрагментов, хотя, по-прежнему, с выполнением этих лежащих в основе принципов. Логика 12 восстановления сжатого видео, исполняющаяся в клиенте 7, восстанавливает каждый сжатый фрагмент, как если бы это была отдельная видеопоследовательность маленьких I- и P-кадров, и затем передает каждый фрагмент в буфер кадров, управляющий дисплеем 14. Например, I0 и P0 из R-кадров 6-4 используются для восстановления сжатого фрагмента 0 видеоизображения и его визуализации. Точно так же I1 и P1 из R-кадров 6-4 используются для восстановления фрагмента 1 и так далее. Как упоминалось выше, восстановление сжатых I-кадров и P-кадров известно в данной области техники, а восстановление сжатых I-фрагментов и P-фрагментов может выполняться посредством наличия множества устройств восстановления сжатого видео, функционирующих в клиенте 7. Несмотря на то что представляется, что увеличение количества процессов увеличивает затраты вычислительных ресурсов в клиенте 7, это фактически не так, потому что сам фрагмент пропорционально меньше относительно количества дополнительных процессов, поэтому количество выводимых на экран пикселей является идентичным, как если бы существовал один процесс и использовались обычные I- и P-кадры полного размера.

Этот способ R-кадра значительно уменьшает пики полосы пропускания, обычно связанные с I-кадрами, изображенными на рис. 2.12 и рис. 2.13, потому что любой данный кадр обычно составлен из P-кадров, которые обычно меньше I-кадров. Например, снова предположим, что обычный I-кадр равен 160 Кбит, тогда I-фрагменты каждого из кадров, изображенных на рис. 2.14, примерно равны 1/16 этой величины или 10 Кбит. Аналогично, предположим, что обычный P-кадр равен 16 Кбит, тогда P-кадры для каждого из фрагментов, изображенные на рис. 2.14, могут быть примерно равны 1 Кбит. Конечным результатом является R-кадр приблизительно $10 \text{ Кбит} + 15 \times 1 \text{ Кбит} = 25 \text{ Кбит}$. Соответственно, каждая последовательность из 60 кадров равна $25 \text{ Кбит} \times 60 = 1,5 \text{ Мбит/с}$.

Соответственно, при 60 кадров/секунда потребуется канал с возможностью поддержки полосы пропускания 1,5 Мбит/с, но с гораздо меньшими пиками, появляющимися из-за I-фрагментов, распределенных по всему интервалу из 60 кадров.

Отметим, что в предыдущих примерах с идентичным предположением о скоростях передачи данных для I-кадров и P-кадров средняя скорость передачи данных была равна 1,1 Мбит/с. Причиной этого является то, что в предыдущих примерах новый I- кадр вводится только один раз каждые 60 периодов кадра, тогда как в этом примере 16 фрагментов, которые составляют I-кадр, циклически повторяются через 16 периодов кадра, и, по существу, эквивалент I-кадра вводится каждые 16 периодов кадра, что в результате приводит к несколько более высокой средней скорости передачи данных. На практике, тем не менее, более частое введение I-кадров не увеличивает скорость передачи данных линейно. Это происходит вследствие того, что P-кадр (или P-фрагмент) первоначально кодирует отличие последующего кадра от предшествующего. Соответственно, если предшествующий кадр является довольно похожим на следующий кадр, то P-кадр является очень маленьким, если предшествующий кадр весьма отличается от следующего кадра, то P-кадр является очень большим. Но так как P-кадр в значительной степени получается из предыдущего кадра, а не из фактического кадра, то получающийся в результате закодированный кадр может содержать больше ошибок (например, визуальные артефакты), чем I-кадр с соответствующим количеством битов. И, когда один P-кадр следует за другим P-кадром, то может произойти накопление ошибок, которое ухудшается при наличии длинной последовательности P-кадров. Далее, усложненное устройство сжатия видео обнаруживает, что качество изображения ухудшается после последовательности P-кадров, и, в случае необходимости, оно выделяет больше битов последующим P-кадрам для повышения качества или, если это будет наиболее эффективным образом действия, заменяет P-кадр I-кадром. Соответственно, когда используются длинные последовательности P-кадров (например, 59 P-кадров, как в предыдущих примерах выше), в частности,

когда сцена очень сложная и/или в ней много движения, обычно требуется больше битов для Р-кадров по мере их удаления от I-кадра.

Другими словами, если посмотреть на Р-кадры с противоположной точки зрения, Р-кадры, которые непосредственно следуют за I-кадром, как правило, требуют меньшего количества битов, чем Р-кадры, которые являются более удаленными от I-кадра. Так, в примере, изображенном на рис. 2.14, все Р-кадры удалены от предшествующего им I-кадра не более чем на 15 кадров, тогда как, например, в предыдущем примере, Р-кадр может находиться на расстоянии 59 кадров от I-кадра. Соответственно, при большей частоте I-кадров Р-кадры меньше. Несомненно, точные относительные размеры меняются в зависимости от характера видеопотока, но в примере по рис. 2.14, если I-фрагмент равен 10 Кбит, то размер Р-фрагментов в среднем может быть равен только 0,75 Кбит, что в результате приводит к $10 \text{ Кбит} + 15 \times 0,75 \text{ Кбит} = 21,25 \text{ Кбит}$, или при (скорости) 60 кадров в секунду скорость передачи данных равна $21,25 \text{ Кбит} \times 60 = 1,3 \text{ Мбит/с}$, или приблизительно на 16% больше, чем скорость передачи данных потока с I-кадром, за которым следуют 59 Р-кадров, в 1,1 Мбит/с. И опять же, относительные результаты этих двух подходов к сжатию видео меняются в зависимости от видеопоследовательности, но, как правило, опыт показывает, что использование Р-кадров требует приблизительно на 20% больше битов для данного уровня качества, чем использование последовательностей I/P-кадров. Но, несомненно, Р-кадры существенно уменьшают пики, что намного уменьшает время ожидания при использовании упомянутых видеопоследовательностей, по сравнению со временем ожидания для последовательности I/P-кадров. Р-кадры могут быть сконфигурированы множеством различных способов, в зависимости от характера видеопоследовательности, надежности канала и доступной скорости передачи данных. В альтернативном варианте осуществления используется количество фрагментов, отличное от 16 в конфигурации 4×4. Например, может использоваться 2 фрагмента в конфигурации 2×1 или 1×2, может использоваться 4 фрагмента в конфигурации 2×2, 4×1 или 1×4, может использоваться 6 фрагментов в конфигурации 3×2, 2×3, 6×1 или 1×6, или может использоваться 8

фрагментов в конфигурации 4×2 (как изображено на рис. 2.15), 2×4 , 8×1 или 1×8 . Отметим, что фрагменты не обязательно должны быть квадратными, также видеокادر не обязательно должен быть квадратным или даже прямоугольным. Фрагменты могут принимать любую форму, которая лучше всего подходит для используемого приложения или видеопотока.

В другом варианте осуществления циклическое повторение I- и P-фрагментов не фиксировано количеством фрагментов. Например, в конфигурации 4×2 с 8 фрагментами может, однако, использоваться последовательность с периодическим повторением через 16 элементов, как изображено на рис.2.15. Последовательные несжатые кадры 1, 2, 4, каждый, разделены на 8 фрагментов, 0-7, и каждый фрагмент сжимают отдельно. В R-кадре 6 только фрагмент 0 сжат как I-фрагмент, а остальные фрагменты сжаты как P-фрагменты. В следующем R-кадре 7 все 8 фрагментов сжаты как P-фрагменты, и далее в следующем R-кадре 8 фрагмент 1 сжат как I-фрагмент, а остальные фрагменты все сжаты как P-фрагменты. И такое установление последовательности продолжается для 16 кадров, при этом I-фрагмент формируется только через кадр, соответственно, последний I-фрагмент будет сформирован для фрагмента 7 в течение периода 15-го кадра (на рис. 2.15 не изображен), и в течение периода 16-го кадра сжимают R-кадр 780 с использованием всех P-фрагментов. Далее последовательность снова начинается с фрагмента 0, сжатого как I-фрагмент, при этом другие фрагменты сжаты как P-фрагменты. Как в предшествующем варианте осуществления, самый первый кадр всей видеопоследовательности обычно состоит из всех I-фрагментов для обеспечения опорных элементов (reference) для P-фрагментов, начиная с этого момента. Циклическое повторение I-фрагментов и P-фрагментов даже не должно быть четным, кратным числом количества фрагментов. Например, с 8 фрагментами, за каждым кадром с I-фрагментом может следовать 2 кадра со всеми P-фрагментами, прежде чем будет использован другой I-фрагмент. В еще одном варианте осуществления, для определенных фрагментов может устанавливаться последовательность I-фрагментов чаще, чем для других фрагментов, если, например, известно, что в

определенных областях экрана больше движения, что требует частых I-фрагментов, в то время как другие являются более статическими (например, с изображением счета в игре), что требует менее частых I-фрагментов. Кроме того, несмотря на то что каждый кадр на рис. 2.14-рис. 2.15 изображен с одним I-фрагментом, в одном кадре может быть закодировано множество I-фрагментов (в зависимости от полосы пропускания канала передачи). И наоборот, определенные кадры или последовательности кадров могут быть переданы без I-фрагментов (т.е. только P-фрагменты). Причина, по которой подходы, описанные в предыдущем абзаце, работают хорошо, состоит в том, что, если не распределить I-фрагменты по каждому отдельному кадру, то, как представляется, это в результате приведет к большим пикам, поведение системы не является таким простым. Так как каждый фрагмент сжимают отдельно от других фрагментов, то, по мере того как фрагменты становятся меньше, кодирование каждого фрагмента может стать менее эффективным, потому что устройство сжатия данного фрагмента не может использовать сходные признаки изображения и сходное движение из остальных фрагментов. Соответственно, деление экрана на 16 фрагментов в общем в результате приведет к менее эффективному кодированию, чем деление экрана на 8 фрагментов. Но, если экран разделен на 8 фрагментов, и это вызывает ввод данных полного I-кадра каждые 8 кадров, вместо каждых 16 кадров, то это в результате приводит к гораздо более высокой скорости передачи данных в целом. Соответственно, при введении полного I-кадра каждые 16 кадров, вместо того чтобы вводить его каждые 8 кадров, общая скорость передачи данных уменьшается. Кроме того, при использовании 8 больших фрагментов, вместо 16 меньших фрагментов, уменьшается общая скорость передачи данных, что также уменьшает до некоторой степени пики данных, вызываемые большими фрагментами.

В другом варианте осуществления логика 4 сжатия видео с малым временем ожидания по рис. 2.14 и рис. 2.15 управляет выделением битов различным фрагментам в R- кадрах или с предварительным конфигурированием

посредством установок, на основе известных характеристик видеопоследовательности, которая должна быть сжата, или автоматически, на основе непрерывного анализа качества изображения в каждом фрагменте. Например, в некоторых видеоиграх-гонках, передняя часть автомобиля игрока (которая является относительно неподвижной в сцене) занимает большую часть нижней половины экрана, тогда как верхняя половина экрана полностью заполнена надвигающимся шоссе, зданиями и пейзажем, которые почти всегда находятся в движении. Если логика 4 сжатия выделяет равное количество битов каждому фрагменту, то фрагменты в нижней половине экрана (фрагменты 4-7) в несжатом кадре 1 по рис.2.15, будут в общем, сжиматься с более высоким качеством, чем фрагменты в верхней половине экрана (фрагменты 0-3) в несжатом кадре 1 по рис.2.15. Если известно, что эта конкретная игра или эта конкретная сцена игры имеют такие характеристики, то операторы 1 службы хостинга могут сконфигурировать логику 4 сжатия для выделения большего количества битов фрагментам, находящимся в верхней части экрана, чем фрагментам, находящимся в нижней части экрана. Или, логика 4 сжатия может оценивать качество сжатия фрагментов после сжатия кадров (с использованием одного или нескольких из множества показателей качества сжатия, например, пиковое отношение сигнал/шум (PSNR)), и если она определяет, что за определенное временное окно качество определенных фрагментов постоянно улучшается, тогда она постепенно выделяет большее количество битов фрагментам, качество которых ухудшается до тех, пока уровень качества различных фрагментов не станет примерно одинаковым. В альтернативном варианте осуществления логика 4 устройства сжатия выделяет биты для получения более высокого качества в конкретном фрагменте или группе фрагментов. Например, с более высоким качеством в центре экрана, чем на краях, она может обеспечивать лучшее общее восприятие.

В одном варианте осуществления, для улучшения разрешающей способности определенных участков видеопотока, логика 4 сжатия видео использует меньшие фрагменты для кодирования областей видеопотока с

относительно большей сложностью сцены и/или большим движением, чем для областей видеопотока с относительно меньшей сложностью сцены и/или меньшим движением. Например, как изображено на рис. 2.16, меньшие фрагменты применяются вокруг движущегося персонажа 4 в одной области одного R-кадра 5 (за которым, возможно, следует последовательность R-кадров с идентичными размерами фрагментов (не изображены)). Далее, когда персонаж 4 перемещается в новую область изображения, меньшие фрагменты используются вокруг этой новой области внутри другого R-кадра 6, как изображено. Как упоминалось выше, различные другие размеры и формы могут быть применены как "фрагменты", хотя, по-прежнему, с выполнением этих лежащих в основе принципов. Несмотря на то что циклические I/P-фрагменты, описанные выше, существенно уменьшают пики в скорости передачи данных видеопотока, они не устраняют пики полностью, в частности в случае таких быстро изменяющихся или очень сложных видеоизображений, которые встречаются в кинофильмах, видеоиграх и некотором прикладном программном обеспечении. Например, во время внезапного монтажного перехода, за сложным кадром может следовать другой сложный кадр, который полностью отличается от него. Несмотря на то что несколько I-фрагментов могут находиться перед монтажным переходом только на расстоянии нескольких периодов кадра, они не помогут в этой ситуации, потому что материал нового кадра не имеет связи с предыдущими I-фрагментами. В такой ситуации (и в других ситуациях, когда, несмотря на то что не все изменяется, большая часть изображения изменяется) устройство 4 сжатия видео определяет, что многие, если не все, P-фрагменты более эффективно кодировать как I-фрагменты, и в результате получается очень большой пик в скорости передачи данных для этого кадра.

Как обсуждалось ранее, это просто имеет место, что с большинством подключений к Internet потребительского уровня (и многими офисными соединениями) просто не выполнимо для данных, создающих "пробку", которые превышают доступную максимальную скорость передачи данных, изображенную позицией 2 на рис. 2.13, вместе с номинальной максимальной

скоростью 1 передачи данных. Отметим, что номинальная максимальная скорость 1 передачи данных (например, "DSL 6 Мбит/с") является, по существу, цифрой сбыта для пользователей, рассматривающих покупку подключения к Internet, но, в общем, она не гарантирует уровень производительности. Для целей этой заявки это не имеет отношение к делу, так как единственной проблемой является доступная максимальная скорость 2 передачи данных во время передачи потоком видео через соединение. Следовательно, на рис. 2.17 и рис. 2.19, так как мы описываем решение проблемы пиков, номинальная максимальная скорость передачи данных не изображена на графике, и изображена только доступная максимальная скорость 1 передачи данных. Скорость передачи данных видеопотока не должна превышать доступную максимальную скорость 1 передачи данных. Для решения этой проблемы, первым, что делает устройство 4 сжатия видео, является определение пиковой скорости 3 передачи данных, которая является скоростью передачи данных, которую канал в состоянии обрабатывать устойчиво.

Эта скорость может быть определена несколькими способами. Один такой способ состоит в постепенной отправке тестового потока с все более высокой скоростью передачи данных из службы 1 по размещению информации на сервере в клиента 7 по рис. 2.4 и рис. 2.5 и обеспечении клиентом обратной связи в службу хостинга относительно уровня потери пакетов и времени ожидания. Когда потеря пакетов и/или время ожидания начинают резко увеличиваться, это указывает на приближение к доступной максимальной скорости 1 передачи данных. После этого служба 1 хостинга может постепенно уменьшать скорость передачи данных тестового потока до тех пор, пока клиент 7 не сообщит, что за достаточный период времени тестовый поток был принят с приемлемым уровнем потери пакетов и время ожидания является почти минимальным. Это устанавливает пиковую максимальную скорость 3 передачи данных, которая далее используется как пиковая скорость передачи данных для потокового видео. С течением времени пиковая скорость 3 передачи данных будет колебаться (например, если другой пользователь, живущий в этом доме, начнет интенсивно

использовать подключение к Internet), и клиент 7 должен будет постоянно осуществлять текущий контроль над увеличением времени ожидания или потери пакетов с указанием того, что доступная максимальная скорость 1 передачи данных падает ниже ранее установленной пиковой скорости 3 передачи данных, и если это так, то пиковую скорость 3 передачи данных. Аналогично, если с течением времени клиент 7 находит, что потеря пакетов и время ожидания остаются на оптимальных уровнях, то он может отправлять запрос, чтобы устройство сжатия видео медленно увеличивало скорость передачи данных для проверки того, увеличилась ли доступная максимальная скорость передачи данных (например, прекратил ли другой пользователь, живущий в доме, интенсивно использовать подключение к Internet), и снова ожидать до тех пор, пока потеря пакетов и/или более высокое время ожидания не будут указывать на то, что доступная максимальная скорость 1 передачи данных была превышена, и снова может быть найден более низкий уровень для пиковой скорости 3 передачи данных, но тот, который может быть выше, чем уровень, который был до тестирования на увеличение скорости передачи данных. Соответственно, с использованием этого способа (и других способов, подобных этому) может быть найдена пиковая скорость 3 передачи данных и периодически корректироваться, по мере необходимости. Пиковая скорость 3 передачи данных устанавливает максимальную скорость передачи данных, которая может использоваться устройством 4 сжатия видео для передачи потоком видео пользователю. Логика для определения пиковой скорости передачи данных может быть реализована на территории 2 пользователя и/или в службе 1 хостинга. На территории 2 пользователя клиентское устройство 7 выполняет вычисления для определения пиковой скорости передачи данных и передает эту информацию обратно в службу 1 хостинга, в службе 1 хостинга сервер 3 выполняет вычисления для определения пиковой скорости передачи данных на основе статистики, принятой из клиента 7 (например, потеря пакетов, время ожидания, максимальная скорость передачи данных и т.д.).

На рис. 2.17 изображена иллюстративная скорость 2 передачи данных видеопотока, который содержит значительную сложность сцены и/или значительное движение, сформированного с использованием способа сжатия циклического I/P-фрагмента, описанного ранее и изображенного на рис. 2.14, рис.2.15 и рис. 2.16. Устройство 4 сжатия видео было сконфигурировано для вывода сжатого видео со средней скоростью передачи данных, которая ниже пиковой скорости 3 передачи данных, и отметим, что почти все время скорость передачи данных видеопотока остается ниже пиковой скорости 3 передачи данных. При сравнении скорости 2 передачи данных со скоростью 4 передачи данных видеопотока, изображенной на рис. 2.13, создаваемой с использованием I/P/B- или I/P-кадров, видно, что на выходе сжатия циклического I/P- фрагмента получается гораздо более гладкая скорость передачи данных. Однако в пике 7 2× кадра (который приближается к пиковой скорости 4 передачи данных 2×) и пике 8 4× кадра (который приближается к пиковой скорости 6 передачи данных 4×) скорость передачи данных превышает пиковую скорость 3 передачи данных, что является неприемлемым. На практике даже с видео с высокой активностью в видеоиграх с быстрым изменением пики, превышающие пиковую скорость 3 передачи данных, встречаются меньше чем в 2% кадров, пики, превышающие пиковую скорость 4 передачи данных 2×, встречаются редко, а пики, превышающие пиковую скорость 5 передачи данных 3×, не встречаются почти никогда. Но, когда они действительно встречаются (например, во время монтажного перехода), требующаяся им скорость передачи данных является необходимой для вывода видеоизображения хорошего качества.

Один способ решения этой проблемы состоит в том, чтобы просто сконфигурировать устройство 4 сжатия видео так, что выход его максимальной скорости передачи данных является пиковой скоростью 3 передачи данных. К сожалению, получающееся в результате качество видео на выходе во время пиковых кадров является плохим, так как алгоритм сжатия ”испытывает недостаток” в битах. В результате появляются артефакты сжатия, когда осуществляются внезапные переходы или быстрое движение, и со временем

пользователь понимает, что артефакты всегда неожиданно возникают, когда происходят внезапные изменения или быстрое движение, и это может его весьма раздражать. Несмотря на то что зрительная система человека довольно чувствительна к визуальным артефактам, которые появляются во время внезапных изменений или быстрого движения, она не является очень чувствительной, чтобы заметить уменьшение частоты кадров в таких ситуациях.

Фактически, когда происходят такие внезапные изменения, оказывается, что зрительная система человека занята отслеживанием изменений, и она не замечает, если частота кадров на короткое время падает с 60 кадр/с до 30 кадр/с и затем сразу возвращается к 60 кадр/с. И в случае очень значительного перехода, подобного внезапному изменению сцены, зрительная система человека не замечает, если частота кадров падает до 20 кадр/с или даже 15 кадр/с и затем сразу возвращается к 60 кадр/с. До тех пор пока уменьшение частоты кадров происходит редко, человеку-наблюдателю кажется, что видео непрерывно передается со скоростью 60 кадр/с. Это свойство зрительной системы человека используется в способах, изображенных на рис.2.18. Сервер 4 (по рис.2.4 и рис.2.5) выводит несжатый выходной видеопоток с устойчивой частотой кадров (с 60 кадр/с в одном варианте осуществления). На временной шкале изображено, что каждый кадр 2-11 выводится каждую 1/60-ую долю секунды. Каждый несжатый видеокадр, начиная с кадра 2, выводится в устройство 4 сжатия видео с малым временем ожидания, которое сжимает кадр за время, меньшее, чем период кадра, с выводом для первого кадра сжатого кадра 16. Данные, выводимые для сжатого кадра 16, могут быть большего размера или меньшего размера, в зависимости от многих факторов, как описано ранее. Если размер данных такой маленький, что они могут быть переданы клиенту 7 в течение периода кадра (1/60-ая доля секунды) или меньше с пиковой скоростью 3 передачи данных, то они передаются в течение периода 991 передачи (xmit time) (длина стрелки указывает на продолжительность периода передачи). В течение следующего периода кадра сервер 8 выводит несжатый кадр 962 2, он сжимается в сжатый кадр 982 2, и он передается в клиент 7 в течение периода 23 передачи, который

меньше, чем период кадра, с пиковой скоростью 3 передачи данных. Далее в течение следующего периода кадра сервер 3 выводит несжатый кадр 4. Когда он сжимается устройством 4 сжатия видео, получающийся в результате сжатый кадр 983 3 является большим количеством данных, чем может быть передано с пиковой скоростью 3 передачи данных в течение одного периода кадра. Соответственно, он передается в течение периода 24 передачи (пик 2×), который занимает весь период кадра и часть следующего периода кадра. Далее в течение следующего периода кадра сервер 4 выводит другой несжатый кадр 964 4 и выводит его в устройство 6 сжатия видео, но данные игнорируются, и они изображены посредством позиции 12. Это происходит потому, что устройство 6 сжатия видео сконфигурировано для игнорирования последующих несжатых видеок кадров, которые поступают, когда оно все еще передает предыдущий сжатый кадр. Несомненно, устройство восстановления сжатого видео клиента 7 не примет кадр 4, но оно просто будет продолжать выводить на экран дисплея 16 кадр 3 в течение двух периодов кадра (т.е. на короткое время уменьшит частоту кадров с 60 кадр/с до 30 кадр/с).

В течение следующего кадра 5 сервер 4 выводит несжатый кадр 965 5, который сжимается в сжатый кадр 985 5 и передается внутри 1 кадра в течение периода 25 передачи. Устройство восстановления сжатого видео клиента 415 восстанавливает сжатый кадр 5 и выводит его на экран дисплея 9. Далее сервер 4 выводит несжатый кадр 966 6, устройство 6 сжатия видео сжимает его в сжатый кадр 986 6, но на этот раз размер получившихся в результате данных является очень большим. Сжатый кадр передается в течение периода 26 передачи (пик 4×) с пиковой скоростью 3 передачи данных, но передача кадра занимает почти 4 периода кадра. В течение следующих 3 периодов кадра устройство 6 сжатия видео игнорирует 3 кадра из сервера 4, и устройство восстановления сжатых данных клиента 7 постоянно выводит кадр 6 на дисплей 16 в течение 4 периодов кадра (т.е. на короткое время уменьшает частоту кадров с 60 кадр/с до 15 кадр/с). И наконец, сервер 4 выводит кадр 970 10, устройство 6 сжатия видео сжимает его в сжатый кадр 987 10, и он передается в течение периода 997 передачи, и

устройство восстановления сжатых данных клиента 7 восстанавливает сжатый кадр 10 и выводит его на экран дисплея 16, и снова вывод видео возобновляется со скоростью 60 кадр/с.

Отметим, что несмотря на то, что устройство 6 сжатия видео сбрасывает видеокadres видеопотока, формируемого сервером 4, оно не сбрасывает аудиоданные независимо от того, в какой форме поступает аудио, и оно продолжает сжимать аудиоданные, когда видеокadres сбрасываются, и передает их в клиент 7, который продолжает восстанавливать сжатые аудиоданные и обеспечивать аудио в устройство, которое использует пользователь для воспроизведения аудио, каково бы оно ни было. Соответственно, аудио продолжает воспроизводиться полностью в течение периодов, когда кадры сбрасываются. Сжатое аудио использует относительно маленький процент от полосы пропускания по сравнению со сжатым видео и в результате не оказывает значительного влияния на общую скорость передачи данных.

Несмотря на то что это не изображено ни в одной из схем скорости передачи данных, всегда существует пропускная способность скорости передачи данных, зарезервированная для сжатого аудиопотока в пределах пиковой скорости 3 передачи данных.

Пример, только что описанный по рис. 2.18, был выбран для иллюстрации того, как частота кадров падает во время пиков скорости передачи данных, но он не иллюстрирует то, как используются способы с циклическим I/P-фрагментом, описанные ранее, такие пики скорости передачи данных и последующие сброшенные кадры являются редкими даже во время последовательностей с высокой активностью/с высокой сложностью сцены, например последовательностей, которые встречаются в видеоиграх, кинофильмах и некотором прикладном программном обеспечении. Следовательно, уменьшенные частоты кадров являются нечастыми и кратковременными и зрительная система человека не замечает их.

Если только что описанный механизм уменьшения частоты кадров применяется к скорости передачи данных видеопотока, изображенной на рис.

2.17, то получающаяся в результате скорость передачи данных видеопотока изображена на рис. 2.19. В этом примере пик 952×2 был уменьшен до сглаженного пика 953×2 и пик 955×4 был уменьшен до сглаженного пика 955×4 , и вся скорость 2 передачи данных видеопотока остается на уровне пиковой скорости 3 передачи данных или ниже его.

Соответственно, с использованием способов, описанных выше, видеопоток с высокой активностью может быть передан с малым временем ожидания через общий Internet и через подключение к Internet потребительского уровня. Кроме того, в офисной обстановке в LAN (например, Ethernet 100 Мбит или беспроводная связь 802.11g) или в частной сети (например, соединение 100 Мбит/с между центром обработки и хранения данных и офисами) видеопоток с высокой активностью может передаваться без пиков так, что множество пользователей (например, передающих 1920×1080 с 60 кадр/с с 4,5 Мбит/с) могут использовать LAN или совместно используемое частное информационное соединение без накладывающихся пиков, переполняющих сеть или магистрали сетевого коммутатора.

Корректировка скорости передачи данных

Возможно что служба 1 хостинга сначала оценивает доступную максимальную скорость 2 передачи данных и время ожидания канала для определения соответствующей скорости передачи данных для видеопотока и затем динамически корректирует скорость передачи данных в ответе. Для корректировки скорости передачи данных служба 1 хостинга может, например, модифицировать разрешающую способность изображения и/или количество кадров/секунда видеопотока, который должен быть отправлен в клиент 7. Кроме того, служба хостинга может корректировать уровень качества сжатого видео. При изменении разрешающей способности видеопотока, например с разрешающей способности 1280×720 на 640×360 , логика 12 восстановления сжатого видео в клиенте 7 может приводить к масштабу изображение для поддержания идентичного размера изображения на экране дисплея.

В одном варианте осуществления в ситуации, когда в канале пропадает сигнал, служба 1 хостинга приостанавливает игру. В случае игры с несколькими участниками служба хостинга сообщает другим пользователям, что упомянутый пользователь выбыл из игры и/или приостанавливает игру с другими пользователями.

2.2.5 Сброшенные или задержанные пакеты

В одном варианте осуществления, если данные потеряны из-за потери пакетов между устройством сжатия видео и клиентом 7 по рис. 2.4 или рис. 2.5 или из-за пакета, принимаемого не по порядку, который поступает слишком поздно, для восстановления сжатого кадра и удовлетворения требований ко времени ожидания восстановленного кадра, логика 12 восстановления сжатого видео может уменьшать визуальные артефакты. При реализации потокового I/P-кадра если существует потерянный/задержанный пакет, то это влияет на весь экран, возможно, вызывая полное зависание экрана на некоторый период времени или показ других визуальных артефактов по всему экрану. Например, если потерянный/задержанный пакет вызывает потерю I-кадра, то в устройстве восстановления сжатых данных будет отсутствовать опорный элемент для всех последующих P-кадров до тех пор, пока не будет принят новый I-кадр. Если P-кадр будет потерян, то это повлияет на P-кадры 20 для всего экрана, которые за ним следуют. В зависимости от того, сколько времени пройдет до того, как появится I-кадр, это окажет более длительное или более короткое визуальное воздействие. С использованием чередующихся I/P-фрагментов, как изображено на рис. 2.14 и рис. 2.15, гораздо менее вероятно, что потерянный/задержанный пакет повлияет на весь экран, так как он только повлияет на фрагменты, содержащиеся в затронутом пакете. Если данные каждого фрагмента отправляются в отдельном пакете, то, если пакет потерян, это повлияет только на один фрагмент. Несомненно, продолжительность визуального артефакта будет зависеть от того, потерян ли пакет I-фрагмента и, если P-фрагмент потерян, то, через какое количество кадров появится I-фрагмент. Но с учетом того, что разные фрагменты на экране обновляются посредством I-кадров очень часто

(возможно, каждый кадр), даже если затрагивается один фрагмент на экране, то другие фрагменты могут не затрагиваться. Кроме того, если некоторое событие вызывает потерю нескольких пакетов одновременно (например, всплеск напряжения электропитания рядом с линией DSL, который на короткое время прерывает поток данных), тогда некоторые из фрагментов будут затронуты больше, чем другие, но вследствие того что некоторые фрагменты будут быстро обновлены посредством нового I-фрагмента, то их это затронет только на короткое время. Кроме того, с реализацией потокового I/P-кадра I-кадры являются не только наиболее важным кадром, но и I-кадры являются чрезвычайно большими, поэтому если существует событие, которое вызывает сброс/задержку пакета, то существует большая вероятность того, что будет затронут I- кадр (т.е. если какая-либо часть I-кадра потеряна, то маловероятно, что I-кадр может быть вообще восстановлен), чем намного меньший I-фрагмент. Вследствие всех этих причин использование I/P-фрагментов в результате приводит к гораздо меньшим визуальным артефактам, когда пакеты сбрасываются/задерживаются, чем с I/P- кадрами.

Один вариант осуществления направлен на уменьшение влияния потерянных пакетов посредством рациональной упаковки сжатых фрагментов внутри пакетов TCP (протокол управления передачей) или пакетов UDP (протокол передачи дейтаграмм пользователя). Например, в одном варианте осуществления фрагменты выравниваются по границам пакета всегда, когда это возможно. На рис.2.20 изображено то, как фрагменты могут быть упакованы внутри последовательности пакетов 1001-1005 без реализации этого признака. А именно, на рис. 2.20 фрагменты пересекают границы пакета и упакованы неэффективно так, что потеря одного пакета в результате приводит к потере множества кадров. Например, если потеряны пакеты 1003 или 1004, то потеряны три фрагмента, что в результате приводит к визуальным артефактам.

В отличие от этого на рис. 2.21 изображена логика 7 упаковки фрагмента для рациональной упаковки фрагментов внутри пакетов для уменьшения влияния потери пакета. Сначала логика 7 упаковки фрагмента выравнивает фрагменты

по границам пакета. Соответственно, фрагменты T1, T3, T4, T7 и T2 выровнены по границе пакетов 2-6 соответственно. Логика упаковки фрагмента также направлена на размещение фрагментов внутри пакетов наиболее эффективным возможным способом без пересечения границ пакета. На основе размера каждого из фрагментов фрагменты T1 и T6 объединены в один пакет 2, T3 и T5 объединены в один пакет 3, фрагменты T4 и T8 объединены в один пакет 4, фрагмент T7 добавлен в пакет 5 и фрагмент T2 добавлен в пакет 6. Соответственно, согласно этой схеме, потеря одного пакета в результате приведет к потере не больше чем 2 фрагментов (а не 3 фрагментов, как изображено на рис.2.20).

Одним дополнительным преимуществом в отношении варианта осуществления, изображенного на рис. 2.21, является то, что фрагменты передаются в порядке, отличном от того, в котором они выводятся на экран внутри изображения. Соответственно, если соседние пакеты потеряны из-за идентичного события, препятствующего передаче, то это затронет области, которые не находятся рядом друг с другом на экране, что создаст менее заметные артефакты на дисплее.

В одном варианте осуществления применяются способы прямого исправления ошибок (FEC) для защиты определенных частей видеопотока от канальных ошибок. Как известно в данной области техники, способы FEC, например Рида-Соломона и Витерби, формируют и добавляют информацию данных для исправления ошибок к данным, передаваемым по каналу связи. Если происходит ошибка в основных данных (например, I-кадр), то можно использовать FEC для исправления этой ошибки.

Коды FEC увеличивают скорость передачи данных передачи, поэтому предпочтительно они используются только там, где они больше всего необходимы. Если отправляются данные, которые не могут в результате привести к очень заметному визуальному артефакту, то может оказаться предпочтительным не использовать коды FEC для защиты данных. Например, потерянный P-фрагмент, который непосредственно идет перед I-фрагментом,

создаст на экране визуальный артефакт (т.е. на фрагменте на экране не будет обновлен) только на 1/60-ую долю секунды. Человеческий глаз не замечает такой визуальный артефакт. Чем дальше Р- фрагменты отстоят от следующего I-фрагмента, тем все более заметной становится потеря Р-фрагмента. Например, если циклической схемой фрагмента является I-фрагмент, за которым следуют 15 Р-фрагментов до того, как I-фрагмент будет снова доступен, то, если потерян Р-фрагмент, непосредственно следующий за I-фрагментом, это в результате приведет к тому, что этот фрагмент будет показывать неправильное изображение в течение 15 периодов кадра (при 60 кадр/с это равно 250 мс).

Человеческий глаз легко замечает разрыв в потоке продолжительностью 250 мс. Соответственно, чем дальше отстоит Р-фрагмент от следующего нового I-фрагмента (т.е. чем ближе Р-фрагменты расположены к предыдущему I-фрагменту), тем более заметным становится артефакт. Как обсуждалось ранее, несмотря на то что, в общем, чем ближе Р-фрагмент расположен к предыдущему I-фрагменту, тем меньше данные для этого Р-фрагмента. Соответственно, Р-фрагменты, которые следуют за I-фрагментами, не только более важно защищать от потери, но и они меньше по размеру. И, в общем, чем меньше данных, которые необходимо защищать, тем меньший код FEC нужен для их защиты. Соответственно, как изображено на рис. 2.22, в одном варианте осуществления из-за важности I-фрагментов в видеопотоке только I-фрагменты обеспечиваются кодами FEC. Соответственно, FEC 2 содержит код исправления ошибок для I-фрагмента 1, и FEC 5 содержит код исправления ошибок для I-фрагмент 2. В этом варианте осуществления FEC не формируется для Р-фрагментов.

В одном варианте осуществления, изображенном на рис. 2.23, коды FEC также формируются для Р-фрагментов, которые, наиболее вероятно, вызовут визуальные артефакты, если будут потеряны. В этом варианте осуществления коды FEC 1105 обеспечивают коды исправления ошибок для первых 3 Р-фрагментов, но не для последующих Р-фрагментов. В другом варианте осуществления коды FEC формируются для Р-фрагментов, размер данных

которых является наименьшим (которые будут, как правило, самостоятельно выбирать Р-фрагменты, раньше всего встречающиеся после I-фрагмента, защита которых наиболее важна). В другом варианте осуществления вместо того чтобы отправлять код FEC с фрагментом, фрагмент передается два раза, причем каждый раз в другом пакете. Если потерян/задержан один пакет, то используется другой пакет.

В одном варианте осуществления, изображенном на рис. 2.24, коды 1 и 1113 FEC формируются для аудиопакетов, 3 и 4, соответственно, передаются из службы хостинга одновременно с видео. Особенно важно поддерживать целостность аудио в видеопотоке, потому что искаженное аудио (например, щелканье или шипенье) в результате приводит к особенно неприятному практическому опыту пользователя. Коды FEC способствуют обеспечению представления содержимого аудио в компьютере 7 клиента без искажения.

В другом варианте осуществления вместо того чтобы отправлять код FEC с аудиоданными, аудиоданные передаются два раза, причем каждый раз в другом пакете. Если потерян/задержан один пакет, то используется другой пакет. Кроме того, в одном варианте осуществления, изображенном на рис. 2.25, коды 2 и 4 FEC используются для команд 1 и 3 ввода пользователя, соответственно (например, нажатие клавиши) передаваемых в восходящем направлении из клиента 7 в службу 1 хостинга. Это важно, потому что потеря нажатия клавиши или движения мыши в видеоигре или приложении могут в результате привести к неприятному практическому опыту пользователя.

В другом варианте осуществления вместо того чтобы отправлять код FEC с данными команды ввода пользователя, данные команды ввода пользователя передаются два раза, причем каждый раз в другом пакете. Если потерян/задержан один пакет, то используется другой пакет. В одном варианте осуществления служба 1 хостинга оценивает качество канала связи с клиентом 7 для определения того, использовать ли FEC и если да, то какие части видео, аудио и команд пользователя с какими FEC следует применять. Оценка "качества" канала может включать в себя такие функции, как оценка потери пакетов,

времени ожидания и т.д., как описано выше. Если канал является особенно ненадежным, то служба 1 хостинга может применять FEC ко всем I-фрагментам, P-фрагментам, аудио и командам пользователя. Напротив, если канал является надежным, то служба 1 хостинга может применять FEC только к аудио и командам пользователя или может не применять FEC ни к аудио ни к видео, или может совсем не использовать FEC. Могут применяться различные другие перестановки применения FEC, хотя, по-прежнему, с выполнением этих лежащих в основе принципов. В одном варианте осуществления, служба 1 хостинга непрерывно осуществляет текущий контроль условий на канале и соответственно изменяет политику FEC.

В другом варианте осуществления согласно рис. 2.4 и рис. 2.5, когда пакет потерян/задержан, что в результате приводит к потере данных фрагмента, или если, возможно, из-за особенно большой потери пакетов, FEC не может исправить потерянные данные фрагмента, то клиент 7 оценивает, сколько кадров остается до того, как новый I-фрагмент будет принят, и сравнивает это количество со временем передачи сигнала туда и обратно, от клиента 415 до службы 1 хостинга. Если время передачи сигнала туда и обратно меньше, чем количество кадров, оставшееся до ожидаемого поступления нового I-фрагмента, то клиент 7 отправляет сообщение в службу по размещению 1 информации на сервере с запросом нового I-фрагмента. Это сообщение маршрутизируется в устройство 4 сжатия видео, и вместо формирования P-фрагмента для фрагмента, данные которого были потеряны, оно формирует I-фрагмент. С учетом того что система, изображенная на рис. 2.4 и рис. 2.5, предназначена для обеспечения времени передачи сигнала туда и обратно, которое, как правило, меньше 80 мс, в результате получим, что фрагмент исправляется в пределах 80 мс (при 60 кадр/с продолжительность кадров равна 16,67 мс, соответственно, в полных периодах кадра время ожидания 80 мс в результате приводит к исправлению фрагмента в пределах 83,33 мс, которые равны 5 периодам кадра - заметный разрыв, но намного менее заметный, чем, например, разрыв 250 мс для 15 кадров). Когда устройство 4 сжатия формирует такой I-фрагмент вне его обычного

циклического порядка, если I-фрагмент вызывает то, что полоса пропускания этого кадра превышает доступную полосу пропускания, то устройство 4 сжатия задерживает циклы других фрагментов так, что другие фрагменты принимают P-фрагменты в течение этого периода кадра (даже если один фрагмент обычно должен принимать I-фрагмент в течение этого кадра), и далее, начиная со следующего кадра, продолжается обычное циклическое повторение, и фрагмент, который обычно принимал I-фрагмент в предыдущем кадре, принимает I-фрагмент. Несмотря на то что это действие на короткое время задерживает фазу циклического повторения R-кадра, визуально это обычно не заметно.

2.2.6 Реализация устройства сжатия/устройства восстановления сжатого видео и аудио

На рис. 2.26 изображен один конкретный вариант осуществления, в котором используется многоядерный процессор и/или мультипроцессор 1200 для сжатия 8 фрагментов параллельно. В одном варианте осуществления используется компьютерная система с четырехядерным CPU Xeon, вдвоенный процессор, функционирующая с тактовой частотой 2,66 ГГц или выше, причем каждое ядро реализует устройство сжатия H.264 \times 264 с открытым исходным текстом как независимый процесс. Однако могут использоваться различные другие конфигурации аппаратного/программного обеспечения, хотя по-прежнему с выполнением этих лежащих в основе принципов. Например, каждое из ядер CPU может быть заменено устройством сжатия H.264, реализованным в FPGA. В примере, изображенном на рис. 2.26, ядра 1-8 используются для одновременной обработки I-фрагментов и P-фрагментов как восемь независимых потоков. Как известно в данной области техники, современные многоядерные и мультипроцессорные компьютерные системы по своей природе могут осуществлять многопоточность при интеграции с многопоточными операционными системами, например, Microsoft Windows XP Professional Edition (или 64-разрядная или 32-разрядная версия) и Linux.

В варианте осуществления, изображенном на рис. 2.26 так как каждое из этих 8 ядер отвечает только за один фрагмент, то оно функционирует в значительной

степени независимо от других ядер, причем каждое исполняет отдельный экземпляр $\times 264$. Карта захвата DVI на основе PCI Express $\times 1$, например Sendero Video Imaging IP.

Development Board от Microtronix, Oosterhout, Нидерланды, используется для захвата несжатого видео с разрешающей способностью 640×480 , 800×600 или 1280×720 , и FPGA на карте использует прямой доступ к памяти (DMA) для передачи захваченного видео через DVI в системную RAM. Фрагменты компоуются в компоновку $1205 \ 4 \times 2$ (несмотря на то что они изображены как квадратные фрагменты, в этом варианте осуществления они имеют разрешающую способность 160×240). Каждый экземпляр $\times 264$ сконфигурирован для сжатия одного из 8 фрагментов 160×240 , и они синхронизированы так, что после начального сжатия I-фрагмента каждое ядро входит в цикл, причем каждый кадр не совпадает по фазе с другим, для сжатия одного I-фрагмента, за которым следуют семь P-фрагментов, и изображенный на рис. 2.26.

Каждый период кадра, получающиеся в результате сжатые фрагменты объединяются в поток пакетов с использованием способов, описанных ранее, и после этого сжатые фрагменты передаются в целевой клиент 7.

Несмотря на то что не изображено на рис. 2.26, если скорость передачи данных объединенных 8 фрагментов превышает заданную пиковую скорость 3 передачи данных, то все 8 процессов $\times 264$ откладываются на такое количество периодов кадра, которое необходимо до тех пор, пока все данные упомянутых объединенных 8 фрагментов не смогут быть переданы. В одном варианте осуществления клиент 7 реализован в виде программного обеспечения на PC, на котором исполняются 8 экземпляров FFmpeg. Принимающий процесс принимает 8 фрагментов, и каждый фрагмент маршрутизируется в экземпляр FFmpeg, который восстанавливает сжатый фрагмент и визуализирует его в соответствующем месте расположения фрагмента на дисплее 9.

Клиент 7 принимает ввод клавиатуры, мыши или игрового контроллера из драйверов устройства ввода PC и передает его в сервер 3. Сервер 3 далее

применяет принятые данные устройства ввода и применяет их к игре или приложению, исполняющемуся на сервере 3, который является PC, работающим под управлением Windows с использованием двухядерного (Core Duo) CPU Intel 2,16 ГГц. Сервер 3 далее выводит новый кадр и выводит его через свой выход DVI или из расположенной на системной плате графической системы, или через выход DVI платы PCI 8800GTX NVIDIA. Одновременно сервер 3 выводит аудио, выводимое игрой или приложениями, через свой цифровой аудиовыход (например, S/PDIF), который соединен с цифровым аудиовходом на сдвоенном четырехядерном PC на основе Xeon, который реализует сжатие видео. Устройство сжатия аудио с открытым исходным текстом Vorbis используется для сжатия аудио одновременно с видео с использованием любого ядра, доступного для потока обработки. В одном варианте осуществления ядро, которое выполняет сжатие своего фрагмента, сначала исполняет сжатие аудио. Сжатое аудио далее передается вместе со сжатым видео и восстанавливается в клиенте 7 с использованием устройства восстановления сжатого аудио Vorbis.

2.2.7 Распределение серверного центра службы хостинга

Свет через стекло, например светопровод, проходит со скоростью, равной некоторой доле от скорости света в вакууме, и, соответственно, может быть определена точная скорость распространения света в светопроводе. Но на практике с учетом времени из-за задержек маршрутизации, неэффективности передач и других потерь авторы наблюдали, что оптимальные величины времени ожидания в Internet отражают скорости передачи ближе к 50% от скорости света. Соответственно, оптимальное время передачи сигнала туда и обратно на расстояние 1000 миль (1600 км) равно приблизительно 22 мс, и оптимальное время передачи сигнала туда и обратно на расстояние 3000 миль (4800 км) равно приблизительно 64 мс. Соответственно, один сервер, находящийся в столице Португалии, находится слишком далеко для обслуживания клиентов

Анголе (который может находиться на расстоянии 3000 миль (4800 км)) с требуемым временем ожидания.

Однако, как изображено на рис. 2.27, если серверный центр 1300 службы 1 хостинга расположен в Португалии (например, в Лиссабоне и т.д.) так, что расстояние до любой точки в Анголе равно приблизительно 1500 миль (2400 км) или меньше, то время передачи сигнала туда и обратно по Internet может быть равно только 32 мс. Согласно рис. 2.4 отметим, что, несмотря на то что величины времени ожидания в самом плохом случае, допустимые для (соединения) 16 пользователя ISP, равны 25 мс, как правило, авторы наблюдали величины времени ожидания ближе к 10-15 мс с системами кабельной модемной связи и DSL. Кроме того, согласно рис. 2.5 предполагается, что максимальное расстояние от территории 21 пользователя до хостинг-центра 1 равно 1000 миль (1600км). Соответственно, при типичном используемом времени передачи сигнала туда и обратно (по соединению) пользователя ISP 15 мс и максимальном расстоянии через Internet 1500 миль (2400 км) для времени передачи сигнала туда и обратно 32 мс общее время передачи сигнала туда и обратно от момента, когда пользователь приводит в действие устройство 15 ввода и наблюдает ответ на дисплее 9, равно $1+1+15+32+1+16+6+8=80$ мс. Соответственно, при передаче по Internet на расстояние 1500 миль (2400 км), как правило, может получаться время ответа 80 мс. Это может обеспечить любую территорию пользователя достаточно коротким временем 16 ожидания (для соединения) пользователя ISP в Анголе для доступа к одному серверному центру, который расположен в Португалии.

Если конкретный серверный центр службы хостинга перегружен и/или для приложения или игры пользователя время ожидания для соединения с другим, менее загруженным серверным центром службы хостинга является допустимым, то клиент 7 может быть перенаправлен в другой серверный центр службы хостинга. В такой ситуации игра или приложение, исполняемые пользователем, приостанавливаются на сервере 3 в перегруженном серверном центре пользователя и данные о состоянии приложения или игры передаются в

сервер 3, находящийся в другом серверном центре службы хостинга. После этого игра или приложение возобновляются. В одном варианте осуществления служба 1 хостинга ожидает до тех пор, пока игра или приложение не достигнут естественной точки приостановки (например, между уровнями в игре или после того, как пользователь инициирует операцию "сохранение" в приложении) для упомянутой передачи. В еще одном варианте осуществления служба 1 хостинга ожидает до тех пор, пока активность пользователя не прекратится на заданный период времени (например, на 1 минуту), и тогда инициирует упомянутую передачу в это время. Как описано выше, в одном варианте осуществления служба 1 хостинга абонирует услугу 440 по обходу Internet по рис. 2.28 с целью обеспечения своим клиентам гарантируемого времени ожидания. Услуги по обходу Internet, как используется в этом описании, являются услугами, которые обеспечивают маршруты через частную сеть от одной точки до другой (точки) в Internet с гарантируемыми характеристиками (например, время ожидания, скорость передачи данных и т.д.). Например, если служба 1 хостинга принимает большой объем трафика от пользователей, использующих услугу DSL AT&T, предлагаемую в Анголе, то вместо маршрутизации в находящиеся в Португалии центральные АТС AT&T служба 1 хостинга может арендовать частное информационное соединение с большой пропускной способностью у провайдера услуг между находящимися в Анголе центральными АТС и одним или несколькими серверными центрами для службы 1 хостинга. Тогда, если маршруты от всех серверных центров HS1-HS6 службы хостинга через общий Internet до пользователя, находящегося в Анголе, с использованием DSL AT&T в результате приводят к слишком большому времени ожидания, то вместо них может быть использовано частное информационное соединение.

Несмотря на то что частные информационные соединения, как правило, являются более дорогими, чем маршруты через общий Internet, до тех пор пока они составляют небольшой процент от соединений службы 1 хостинга с пользователями, общее влияние на стоимость будет оставаться на низком уровне и пользователи будут испытывать более постоянный практический опыт

обслуживания. Серверные центры часто имеют два уровня резервного питания в случае отказа в системе питания. Первый уровень, как правило, является резервным питанием от батарей (или от альтернативного, находящегося в готовности к немедленному использованию источника энергии, такого как маховик, который поддерживается в рабочем состоянии и подсоединен к генератору), которое немедленно обеспечивает питание, когда происходит отказ в питающей линии, и поддерживает в рабочем состоянии серверный центр. Если отказ в системе питания является непродолжительным и происходит быстрый возврат питающей линии (например, в течение одной минуты), то батареи - это все, что необходимо для поддержания серверного центра в рабочем состоянии. Но если отказ в системе питания происходит в течение более длительного периода времени, то, как правило, запускают генераторы (например, дизельные), которые принимают на себя функции батарей и могут работать, пока у них есть топливо. Такие генераторы являются чрезвычайно дорогими, так как они должны быть способны выводить такую мощность, какую обычно получает серверный центр из питающей линии.

В одном варианте осуществления каждая из служб HS1-HS5 хостинга совместно использует пользовательские данные так, что, если в одном серверном центре происходит отказ в системе питания, то он может приостановить игры и приложения, которые находятся в процессе (исполнения), и затем передает данные о состоянии приложения или игры из каждого сервера 3 в серверы 3, находящиеся в других серверных центрах, и затем уведомят клиента 415 каждого пользователя и предписывает, чтобы он передавал информацию в новый сервер 3. С учетом того, что такие ситуации происходят редко, может оказаться приемлемой передача пользователя серверному центру службы хостинга, который не может обеспечивать оптимальное время ожидания (т.е. пользователь просто должен будет мириться с более высоким временем ожидания в течение отказа в системе питания), что обеспечивает гораздо более широкие возможности для передачи пользователей. Например, с учетом разных часовых поясов в США пользователи на Восточном побережье могут ложиться

спать в 23,30, в то время как пользователи на Западном Побережье в 2030 начинают максимально использовать видеоигры. Если в это время происходит отказ в системе питания в серверном центре службы хостинга, находящемся на Западном Побережье, то, возможно, нет достаточного количества серверов 3, находящихся на Западном Побережье в других серверных центрах службы хостинга для управления всеми пользователями. В такой ситуации некоторые пользователи могут быть переданы в серверные центры службы хостинга, находящиеся на Восточном побережье, в которых существуют доступные серверы 3, и единственным последствием для пользователей будет большее время ожидания. После передачи пользователей из серверного центра, который остался без электричества, серверный центр может начать в должном порядке отключение своих серверов и оборудования так, чтобы все оборудование было отключено до разрядки батарей (или другого немедленно подключаемого резервного питания). Следовательно, серверный центр может избежать затрат на генератор.

В одном варианте осуществления во время периодов интенсивной нагрузки служб 1 хостинга (или вследствие пиковой пользовательской нагрузки, или вследствие отказа одного или нескольких серверных центров) пользователей передают в другие серверные центры на основе требований к времени ожидания игры или приложения, которые они используют. Соответственно, пользователям, использующим игры или приложения, которые требуют малого времени ожидания, отдается предпочтение (при предоставлении) доступных соединений с серверами с малым временем ожидания при ограниченном питании.

Признаки службы хостинга

На рис. 2.29 изображен вариант осуществления компонентов серверного центра для службы 1 хостинга, используемый в нижеследующем описании признаков. Как и в случае со службой 1 хостинга, изображенной на рис. 2.1, компонентами этого серверного центра управляет и их работу координирует система 3 управления службы 1 хостинга, если не оговорено иное.

Входящий трафик 1 в Internet из клиентов 7 пользователя направляется во входящую маршрутизацию 2. Как правило, входящий трафик 1 в Internet входит в серверный центр через высокоскоростное оптоволоконное соединение с Internet, но будет достаточно любых средств сетевого соединения с отвечающими требованиям полосой пропускания, надежностью и малым временем ожидания. Входящая маршрутизация 2 является системой сетевых (сеть может быть реализована как сеть Ethernet, сеть волоконно-оптических каналов или посредством любых других транспортных средств) коммутаторов и серверов маршрутизации, поддерживающих коммутаторы, которая принимает поступающие пакеты и маршрутизирует каждый пакет в соответствующий сервер 6-8 приложения/игры ("app/game"). В одном варианте осуществления пакет, который доставляется на конкретный сервер приложения/игры, представляет подмножество данных, принимаемых из клиента, и/или может быть преобразован/изменен другими сетевыми компонентами (например, такими сетевыми компонентами, как шлюзы и маршрутизаторы) в центре обработки и хранения данных. В некоторых случаях пакеты маршрутизируются в несколько серверов 6-8 одновременно, например если игра или приложение исполняются параллельно, одновременно на множестве серверов. Дисковый массив 3-4 типа RAID соединен с сетью 2 с входящей маршрутизацией так, что серверы 6-8 приложения/игры могут считывать с дисковых массивов 3-4 типа RAID и записывать на них. Кроме того, дисковый массив 5 типа RAID (который может быть реализован как множество дисковых массивов типа RAID) также соединен с входящей маршрутизацией 2, и данные из дискового массива 5 типа RAID могут быть считаны из серверов 6-8 приложения/игры. Входящая маршрутизация 2 может быть реализована в широком диапазоне архитектур сети предшествующего уровня техники, включающих в себя древовидную структуру коммутаторов, с входящим трафиком 1 в Internet в его корне, в ячеистой структуре, соединяющей все различные устройства, или как набор взаимосвязанных подсетей, с трафиком, сосредоточенным между устройствами внутренней связи, отделенным от трафика, сосредоточенного среди других устройств. Одним типом

конфигурации сети является SAN, которая, несмотря на то что она, как правило, используется для запоминающих устройств, также может использоваться для обычной высокоскоростной передачи данных между устройствами. Кроме того, серверы 6-8 приложения/игра каждый может иметь множество сетевых соединений во входящую маршрутизацию 2. Например, сервер 6-8 может иметь сетевое соединение с подсетью, подсоединенной к дисковым массивам 3-4 типа RAID, и другое сетевое соединение с подсетью, подсоединенной к другим устройствам.

Серверы 6-8 приложения/игры могут все быть сконфигурированы одинаково, некоторые - по-разному, или все - по-разному, как описано ранее в отношении серверов 4 в варианте осуществления, изображенном на рис. 2.4. В одном варианте осуществления каждый пользователь при использовании службы хостинга обычно (использует), по меньшей мере, один сервер 6-8 приложения/игры. Для простоты объяснения предположим, что данный пользователь использует сервер 6 приложения/игры, но множество серверов могут использоваться одним пользователем и множество пользователей могут совместно использовать один сервер 6-8 приложения/игры. Ввод сигнала управления пользователя, отправленный из клиента 7, как описано ранее, принимается как входящий трафик 81 в Internet и маршрутизируется через входящую маршрутизацию 2 в сервер 6 приложения/игры. Сервер 6 приложения/игры использует ввод сигнала управления пользователя как ввод сигнала управления в игру или приложение, исполняющиеся на сервере, и вычисляет следующий кадр видео и аудио, связанного с ним. Сервер 6 приложения/игры после этого выводит несжатое видео/аудио 1529 в совместно используемое сжатие 10 видео. Сервер приложения/игры может выводить несжатое видео посредством любых средств, включающих в себя одно или несколько соединений по технологии Gigabit Ethernet, но в одном варианте осуществления видео выводится через соединение DVI, а аудио и другие сжатые данные и информация о состоянии канала связи выводятся через соединение универсальной последовательной шины (USB).

Совместно используемое сжатие 1530 видео сжимает несжатое видео и аудио из серверов 6-8 приложения/игры. Сжатие может быть реализовано полностью в аппаратном обеспечении или в аппаратном обеспечении, исполняющем программное обеспечение. Для каждого сервера 6-8 приложения/игры может (существовать) выделенное устройство сжатия, или если устройства сжатия являются достаточно быстрыми, то данное устройство сжатия можно использовать для сжатия видео/аудио из нескольких серверов 6-8 приложения/игры. Например, при 60 кадр/с время видеокadra равно 16,67 мс. Если устройство сжатия может сжимать один кадр в 1 мс, то это устройство сжатия можно использовать для сжатия видео/аудио из 16 серверов 6-8 приложения/игры посредством приема ввода из одного сервера за другим, причем устройство сжатия сохраняет состояние каждого процесса сжатия видео/аудио и переключает контекст по мере того, как оно циклически переключается между потоками видео/аудио из серверов. Это в результате приводит к существенной экономии затрат на аппаратное обеспечение для сжатия. Так как разные серверы будут завершать кадры в разное время, в одном варианте осуществления ресурсы устройства сжатия находятся в совместно используемом пуле 10 с совместно используемым средством хранения (например, RAM, флэш-память) для хранения состояния каждого процесса сжатия, и когда кадр сервера 6-8 закончен и готов для сжатия, средство управления определяет, какой ресурс сжатия является доступным в этот момент, обеспечивает ресурс сжатия состоянием процесса сжатия сервера и кадром несжатого видео/аудио для сжатия.

Отметим, что часть состояния процесса сжатия каждого сервера включает в себя информацию о самом сжатии, например восстановленные данные буфера кадра предыдущего кадра, которые могут использоваться как опорный элемент для IP-фрагментов, разрешающая способность видеовыхода, качество сжатия, мозаичная структура, выделение битов для каждого фрагмента, качество сжатия, аудиоформат (например, стерео, объемный звук, AC Dolby® 3). Но состояние процесса сжатия также включает в себя информацию о состоянии канала связи

относительно пиковой скорости 3 передачи данных и того, выводится ли в настоящее время предыдущий кадр (как изображено может выводить данный видео/аудио поток во множество пунктов назначения одновременно. В одном варианте осуществления это реализовано с использованием групповой передачи по протоколу Internet (IP), в которой передается данный поток UDP, предназначенный для передачи потоком в множество пунктов назначения одновременно, и это широковещание повторяется коммутаторами и серверами маршрутизации в исходящей маршрутизации 1540. Множеством пунктов назначения широковещания могут быть клиенты 415 множества пользователей, (доступные) через Internet, множество серверов 6-8 приложения/игры, (доступные) через входящую маршрутизацию 2, и/или один или несколько буферов 5 задержки.

Соответственно, выход данного сервера 6-7 сжимают в один или множество форматов, и каждый сжатый поток направляется в один или множество пунктов назначения. Кроме того, в другом варианте осуществления если множество серверов 6-8 приложения/игры одновременно используются одним пользователем (например, в конфигурации параллельной обработки для создания 3D вывода сложной сцены) и каждый сервер выводит часть получающегося в результате изображения, то видеовыход из множества серверов 6-8 может объединяться совместно используемым аппаратным сжатием 10 в комбинированный 10, и с этого момента и далее он обрабатывается так, как описано выше, как если бы он выходил из одного сервера 1521-1525 приложения/игры.

Отметим, что в одном варианте осуществления копия (по меньшей мере, с разрешающей способностью видео, просматриваемого пользователем, или выше) всего видео, формируемого серверами 6-8 приложения/игры, записывается в буфер 10 задержки, по меньшей мере, на некоторое количество минут (15 минут в одном варианте осуществления). Это обеспечивает возможность каждому пользователю "перематывать" видео из каждого сеанса для просмотра предыдущих действия или деяний (в случае игры). Соответственно, в одном

варианте осуществления в отношении каждого выходного 11 потока сжатого видео/аудио, маршрутизируемого в клиент 7 пользователя, также осуществляется групповая передача в буфер 10 задержки. Когда видео/аудио сохраняется в буфере 10 задержки, каталог в буфере 10 задержки обеспечивает перекрестную ссылку между сетевым адресом сервера 6-8 приложения/игры, который является источником задержанного видео/аудио, и местом в буфере 10 задержки, где задержанно видео/аудио может быть найдено.

По мере того как пиковая скорость 3 передачи данных или другие характеристики канала изменяются со временем, как определяется сервером 6-8 приложения/игры, поддерживающим данные текущего контроля каждого пользователя, отправляемые из клиента 7, сервер 6-8 приложения/игры отправляет соответствующую информацию в совместно используемое аппаратное сжатие 10. Совместно используемое аппаратное сжатие 10 также разбивает на пакеты сжатое видео/аудио с использованием средств, например средств, описанных ранее, и если требуется, то с применением кодов FEC, дублированием определенных данных или принятием других мер для соответствующего обеспечения возможности принятия клиентом 7 потока видео/аудио данных и восстановления сжатых данных с максимально возможным качеством и надежностью. Некоторые приложения, например описанные ниже, требуют, чтобы видео/аудиовыход данного сервера 6-8 приложения/игры был доступным со многими разрешающими способностями (или в других разнообразных форматах) одновременно. Если сервер 6-8 приложения/игры так уведомляет ресурс совместно используемого аппаратного сжатия 1530, то несжатые видео/аудио 9 этого сервера 6-8 приложения/иг 6-8 одновременно сжаты в разных форматах, с разными разрешающими способностями и/или с разными структурами исправления ошибок/пакетов. В некоторых случаях некоторые ресурсы сжатия могут совместно использоваться множеством процессов сжатия, сжимающих идентичное видео/аудио (например, во многих алгоритмах сжатия, существует этап, на котором изображение масштабируется во множество размеров перед применением сжатия. Если требуется вывод

изображений разных размеров, то этот этап может использоваться для обслуживания нескольких процессов сжатия одновременно.

В других случаях для каждого формата требуются отдельные ресурсы сжатия. В любом случае сжатое видео/аудио 11 со всеми требуемыми различными разрешающими способностями и во всех требуемых различных форматах для данного сервера 6-8 приложения/игр (будь то один или много) выводится одновременно в исходящую маршрутизацию 12. В одном варианте осуществления вывод сжатого видео/аудио 11 осуществляется в формате UDP, соответственно, он является однонаправленным потоком пакетов. Сеть 12 исходящей маршрутизации содержит набор коммутаторов и серверов маршрутизации, которые направляют каждый поток сжатого видео/аудио пользователю(ям), которому(ым) он предназначен, или в другие пункты назначения через интерфейс исходящего трафика 16 из Internet (которые, как правило, подключены к волоконно-оптическому интерфейсу с Internet), и/или обратно в буфер 5 задержки, и/или обратно во входящую маршрутизацию 2, и/или через частную сеть (не изображена) для распределения видео. Отметим, что (как описывается ниже) исходящая маршрутизация 12 может выводить данный видео/аудио поток во множество пунктов назначения одновременно. В одном варианте осуществления это реализовано с использованием групповой передачи по протоколу Internet (IP), в которой передается данный поток UDP, предназначенный для передачи потоком в множество пунктов назначения одновременно, и это ширококочание повторяется коммутаторами и серверами маршрутизации в исходящей маршрутизации 12. Множеством пунктов назначения ширококочания могут быть клиенты 7 множества пользователей, (доступные) через Internet, множество серверов 6-8 приложения/игры, (доступные) через входящую маршрутизацию 2, ли один или несколько буферов 5 задержки.

Соответственно, выход данного сервера 6-8 сжимают в один или множество форматов, и каждый сжатый поток направляется в один или множество пунктов назначения. Кроме того, в другом варианте осуществления если множество серверов 6-8 приложения/игры одновременно используются

одним пользователем (например, в конфигурации параллельной обработки для создания 3D вывода сложной сцены) и каждый сервер выводит часть получающегося в результате изображения, то видеовыход из множества серверов 1 6-8 может объединяться совместно используемым аппаратным сжатием 10 в комбинированный 10, и с этого момента и далее он обрабатывается так, как описано выше, как если бы он выходил из одного сервера 6-8 приложения/игры.

Отметим, что в одном варианте осуществления копия (по меньшей мере, с разрешающей способностью видео, просматриваемого пользователем, или выше) всего видео, формируемого серверами 6-8 приложения/игры, записывается в буфер 1515 задержки, по меньшей мере, на некоторое количество минут (15 минут в одном варианте осуществления). Это обеспечивает возможность каждому пользователю "перематывать" видео из каждого сеанса для просмотра предыдущих действия или деяний (в случае игры). Соответственно, в одном варианте осуществления в отношении каждого выходного 11 потока сжатого видео/аудио, маршрутизируемого в клиент 7 пользователя, также осуществляется групповая передача в буфер 5 задержки. Когда видео/аудио сохраняется в буфере 5 задержки, каталог в буфере 5 задержки обеспечивает перекрестную ссылку между сетевым адресом сервера 6-8 приложения/игры, который является источником задержанного видео/аудио, и местом в буфере 1515 задержки где задержанно видео/аудио может быть найдено.

Игры в реальном масштабе времени, немедленно выводимые на экран с возможностью немедленного ведения игры

Серверы 6-8 приложения/игры могут использоваться не только для исполнения данного приложения или видеоигры для пользователя, но они также могут использоваться для создания интерфейсных приложений пользователя для службы 1 хостинга, которые поддерживают навигацию в службе 1 хостинга и другие признаки. Моментальный снимок экрана одного такого интерфейсного приложения пользователя изображен на рис. 2.30, экран "Game Finder" ("средство поиска игры"). Этот конкретный экран интерфейса пользователя обеспечивает

возможность пользователю смотреть 15 игр, которые ведутся в реальном масштабе времени (или отложены) другими пользователями. Каждое из "миниатюрных" видеоокон, например 1, является окном с видео в реальном масштабе времени в движении, показывающим одно видео из игры одного пользователя. Вид, показанный в миниатюре, может быть видом, идентичным тому, который видит пользователь, или он может быть задержанным видом (например, если пользователь ведет боевую игру, то пользователю может потребоваться, чтобы другие пользователи не видели, где он скрывается, и он может решить задержать какой-либо вид его геймплей на некоторый период времени, скажем, на 10 минут). Вид также может быть полем зрения камеры игры, который отличается от вида любого пользователя. Посредством выбора пунктов меню (не изображено в этом примере) пользователь может отобрать выборку игр для одновременного просмотра исходя из различных критериев. Как небольшое осуществление выборки иллюстративных отборов пользователь может выбрать случайную выборку игр (например, изображенную на рис. 2.30), все игры одного вида (все, которые ведут разные игроки), только высоко котирующиеся игроки игры, игроки на данном уровне в игре, или низко котирующиеся игроки (например, если игрок изучает основы), игроки, которые являются "приятелями" (или соперниками), игры, у которых наибольшее количество зрителей и т.д. Отметим, что, в общем, каждый пользователь принимает решение, могут ли другие смотреть видео его игры или приложения и если это так, то кто из них и когда может его смотреть, можно ли его смотреть только с задержкой. Сервер 6-8 приложения/игры, который формирует экран интерфейса пользователя, изображенный на рис. 2.30, запрашивает 15 линий передач видео/аудио посредством отправки сообщения в сервер 6-8 приложения/игры для каждого пользователя, у которого он запрашивает игру. Сообщение отправляется через входящую маршрутизацию 2 или другую сеть. Сообщение включает в себя размер и формат запрашиваемого видео/аудио и идентифицирует пользователя, просматривающего экран интерфейса пользователя. Данный пользователь может решить выбрать режим "privacy"

(”конфиденциальность”) и не обеспечить возможность каким-либо другим пользователям просматривать видео/аудио его игры (или с его точки обзора или с другой точки обзора) или, как описано в предыдущем абзаце, пользователь может решить обеспечить возможность просмотра видео/аудио из его игры, но задержать просматриваемое видео/аудио. Сервер 6-8 приложения/игры пользователя, принимающий запрос и соглашающийся обеспечить возможность просмотра его видео/аудио, отправляет такое подтверждение в запрашивающий сервер, и он также уведомляет совместно используемое аппаратное сжатие 10 о необходимости формирования дополнительного потока сжатого видео в запрашиваемом формате или с запрашиваемым размером экрана (при предположении, что формат и размер экрана отличаются от тех, которые уже формируются), и он также указывает пункт назначения для сжатого видео (т.е. запрашивающий сервер). Если запрашиваемое видео/аудио только задержано, то запрашивающий сервер 6-8 приложения/игры уведомляют об этом, и он запрашивает задержанное видео/аудио из буфера 5 задержки посредством поиска места расположения видео/аудио в каталоге, находящемся в буфере 5 задержки, и сетевого адреса сервера 6-8 приложения/игры, который является источником задержанного видео/аудио. После формирования и обработки всех этих запросов до 15 потоков видео миниатюрного размера в масштабе реального времени маршрутизируются из исходящей маршрутизации 12 во входящую маршрутизацию 2 в сервер 6-85 приложения/игры, формирующий экран интерфейса пользователя, и восстанавливаются и выводятся на экран сервером. Задержанные видео/аудио потоки могут быть со слишком большим размером экрана, и если это так, то сервер 6-8 приложения/игры восстановит сжатые потоки и уменьшит масштаб видеопотоков до размера миниатюры. В одном варианте осуществления запросы на аудио/видео отправляют (и она ими управляет) в центральную службу ”управления”, подобную системе управления службы хостинга по рис. 2.4 (не изображена на рис. 2.29), которая далее перенаправляет эти запросы в соответствующий сервер 6-8 приложения/игры. Кроме того, в одном варианте осуществления может не требоваться запроса,

потому что миниатюры "проталкиваются" в клиенты тех пользователей, которые дают разрешение на него. Аудио из 15 игр, все смикшированное одновременно, может создать какофонию звуков. Пользователь может решить смикшировать все звуки вместе таким способом (возможно только для получения ощущения "шума", создаваемого всем просматриваемым действием), или пользователь может решить прослушивать только аудио из одной игры в данный момент времени. Выбор одной игры выполняется посредством перемещения желтой рамки 2 выбора фрагмента в данную игру (перемещение желтой рамки может выполняться при использовании клавиш курсора на клавиатуре, при перемещении мыши, при перемещении джойстика или при нажатии на кнопки направления на другом устройстве, например, на мобильном телефоне). После выбора одной игры из этой игры воспроизводится только аудио. Кроме того, показывается информация 3 об игре. В случае этой игры, например, логотип издателя ("EA") и логотип игры "Need for Speed Carbon" ("Жажда скорости"), и оранжевая горизонтальная полоса указывает в относительном выражении количество людей, ведущих эту игру или наблюдающих за ней в этот конкретный момент (много в этом случае, поэтому игра является "горячей"). Кроме того, обеспечены "Статистические данные" ("Stats") с указанием того, что существуют 145 игроков, активно ведущих 80 разных экземпляров Need for Speed Game (т.е. она может вестись или как игра отдельного игрока или как игра с несколькими участниками), и существует 680 зрителей (одним из которых является этот пользователь). Отметим, что эти статистические данные (и другие статистические данные) собираются системой 401 управления службы хостинга и сохраняются в дисковых массивах 3-4 типа RAID для ведения журналов регистрации работы службы 1 хостинга и для надлежащего биллинга пользователей и выплат издателям, которые обеспечивают содержимое. Некоторые статистические данные записываются вследствие действий системы 3 управления службы, и некоторые сообщаются в систему 3 управления службы отдельным сервером 6-8 приложения/игры. Например, сервер 6-8 приложения/игры, исполняющий это приложение Game

Finder (средство поиска игры), отправляет сообщения в систему 3 управления службы хостинга, когда игры просматриваются (и когда их прекращают для просмотра), чтобы она могла обновить статистические данные о том, сколько игр отображается на экране). Некоторые статистические данные доступны для интерфейсных приложений пользователя, например, для этого приложения Game Finder (средство поиска игры).

Если пользователь щелкает по кнопке активации на своем устройстве ввода, то он видит, что миниатюрное видео в желтой раме увеличивается, в то время как оно остается в режиме реального времени, до размера во весь экран. Это воздействие изображено в процессе на рис. 2.31. Отметим, что видеоокно 1 увеличилось в размере. Для реализации этого воздействия сервер 6-8 приложения/игры запрашивает сервер 6-8 приложения/игры, исполняющий выбранную игру, для получения копии видеопотока для размера во весь экран (с разрешающей способностью дисплея 9 пользователя) игры, маршрутизируемой в него. Сервер 6-8 приложения/игры, исполняющий игру, уведомляет совместно используемое аппаратное устройство 1530 сжатия, 10 копия миниатюрного размера игры больше не требуется (если другой сервер 6-8 приложения/игры не требует такую 10 миниатюру), и после этого он предписывает, чтобы он отправлял копию видео размера во весь экран в сервер 6-8 приложения/игры с распакованием видео. Пользователь, ведущий игру, может иметь или может не иметь дисплей 9 с разрешающей способностью, идентичной разрешающей способности дисплея пользователя, увеличивающего изображение игры. Кроме того, другие зрители игры могут иметь или могут не иметь дисплеи 9 с разрешающей способностью, идентичной разрешающей способности дисплея пользователя, увеличивающего изображение игры (и могут иметь другие средства воспроизведения аудио, например стерео или объемный звук). Соответственно, совместно используемое аппаратное устройство 10 сжатия определяет то, формируется ли уже подходящий поток сжатого видео/аудио, который отвечает требованиям пользователя, запрашивающего упомянутый видео/аудио поток, и если он уже существует, то оно уведомляет исходящую

маршрутизацию 12, чтобы она маршрутизировала копию потока в сервер 6-8 приложения/игры с изменением масштаба изображения видео, и если (оно) не сжимает другую копию видео, которая подходит для этого пользователя, (то) выдает команду в исходящую маршрутизацию отправить упомянутый поток обратно во входящую маршрутизацию 2 и сервер 6-8 приложения/игры, изменяющий масштаб изображения видео. Этот сервер, теперь принимающий полноэкранную версию выбранного видео, восстанавливает его и постепенно масштабирует его до полного размера.

На рис. 2.32 изображено то, как экран выглядит после окончательного увеличения масштаба изображения игры до полного экрана, и игра изображена с полной разрешающей способностью дисплея 9 пользователя, как показано посредством изображения, на которое указывает стрелка 1. Сервер 6-8 приложения/игры, исполняющий приложение средства поиска игры, отправляет сообщения в другие серверы 6-8 приложения/игры, которые обеспечивали миниатюры, о том, что они больше не требуются, и сообщения в сервер 3 управления службы хостинга, о том, что другие игры больше не просматриваются. В это момент единственным изображением, которое он формирует, является наложение 2 другого графического изображения наверху экрана, которое предоставляет информацию и управляющие элементы меню пользователю. Отметим, что по мере продвижения этой игры, аудитория увеличилась до 2 503 зрителей. При таком большом количестве зрителей должно существовать много зрителей с дисплеями 13, разрешающие способности которых идентичны или близки .

Поскольку изображенная игра является игрой с несколькими участниками, пользователь может принять решение о присоединении к игре в некоторый момент времени. Служба 1 хостинга может обеспечить или может не обеспечить возможность пользователю присоединиться к игре по разным причинам. Например, пользователь, возможно, должен заплатить, чтобы вести игру, и он решил не платить, пользователь может не иметь достаточного рейтинга, чтобы присоединиться к этой конкретной игре (например, он не может конкурировать с другими игроками), или подключение к Internet пользователя

может не иметь достаточно малого времени ожидания для обеспечения возможности пользователю вести игру (например, не существует ограничения по времени ожидания для просмотра игры, соответственно, игру, которую проводят далеко (безусловно, на другом континенте), можно смотреть, не заботясь о времени ожидания, но для того, чтобы вести игру, время ожидания должно быть достаточно мало, чтобы пользователь (а) получал удовольствие от игры и (б) находился в равных условиях с другими игроками, которые могут иметь соединения с меньшим временем ожидания). Если пользователю обеспечена возможность вести игру, то сервер 6-8 приложения/игры, который обеспечивает пользователю интерфейс пользователя Game Finder (средства поиска игры), посылает запрос, чтобы сервер 3 управления службы хостинга инициировал (т.е. определил местоположение и запустил) сервер 6-8 приложения/игры, который соответственно сконфигурирован для ведения конкретной игры, для загрузки игры из дискового массива 3-4 типа RAID, и после этого сервер 3 управления службы хостинга выдает команду во входящую маршрутизацию 2 для передачи управляющих сигналов от пользователя в игровой сервер приложения/игры, который в настоящее время осуществляет хостинг игры, и он выдает команду в совместно используемое аппаратное сжатие 10 для переключения от сжатия видео/аудио из сервера приложения/игры, который осуществляет хостинг приложения Game Finder (средства поиска игры) к сжатию видео/аудио из сервера приложения/игры, который в настоящее время осуществляет хостинг игры. Кадровый синхронизирующий импульс службы приложения/игры Game Finder (средство поиска игры) и нового сервера приложения/игры, который осуществляет хостинг игры, не синхронизированы, и в результате, вероятно, существует разность времен между двумя синхронизирующими импульсами. Поскольку совместно используемое аппаратное обеспечение 10 сжатия видео начнет сжатие видео после того, как сервер 6-8 приложения/игры завершает видеокادر, то первый кадр из нового сервера может быть завершен раньше, чем (пройдет) полный период кадра старого сервера, что может произойти до того, как

завершится передача предыдущего сжатого кадра (например, рассмотрим период 23 передачи по рис. 2.18., если бы несжатый кадр 963 3 был закончен на пол периода кадра раньше, то он столкнулся бы с периодом 13 передачи). В такой ситуации совместно используемое аппаратное обеспечение 10 сжатия видео игнорирует первый кадр из нового сервера (например, подобно тому, как игнорируется 974 кадр 964 4) и клиент 7 удерживает последний кадр из старого сервера дополнительный период кадра, и совместно используемое аппаратное обеспечение 10 сжатия видео начинает сжатие видео следующего периода кадра из нового сервера приложения/игры, который осуществляет хостинг игры. Визуально для пользователя переход от одного сервера приложения/игры на другой будет бесперебойным. Сервер 3 управления службы хостинга после этого уведомляет игровой сервер 6-8 приложения/игры, который осуществлял хостинг Game Finder (средство поиска игры), для переключения в состояние ожидания до тех пор, пока он не потребуется снова.

После этого пользователь может вести игру. Исключительным является то, что создается впечатление, что игру можно вести немедленно (так как она загружается на игровой сервер 6-8 приложения/игры из дискового массива 3-4 типа RAID со скоростью гигабит/секунда), и игра загружается на сервер, полностью пригодной для игры, вместе с операционной системой, полностью сконфигурированной для игры с идеальными драйверами, конфигурацией системного реестра (в случае Windows), и при этом другие приложения, которые могут мешать функционированию игры, на сервере не исполняются.

Кроме того, по мере продвижения пользователя в игре каждый из сегментов игры загружается в сервер со скоростью гигабит/секунда (т.е. 1 гигабайт загружается за 8 секунд) из дискового массива 3-14 типа RAID, и из-за огромной емкости памяти дискового массива 3-4 типа RAID (так как он является разделяемым ресурсом среди многих пользователей, он может быть очень большим и тем не менее быть эффективным по стоимости), настройка геометрии или настройка другого сегмента игры может предварительно вычисляться и сохраняться на дисковом массиве 3-4 типа RAID и загружаться

чрезвычайно быстро. Кроме того, вследствие того что конфигурация аппаратного обеспечения и вычислительные мощности каждого сервера 6-8 приложения/игры известны, могут быть предварительно проведены расчеты (с использованием) шейдеров вершин и пикселей. Соответственно, игра запускается почти немедленно, она исполняется в идеальной среде, и последующие сегменты загружаются почти немедленно. Но кроме этих преимуществ пользователь может смотреть, как другие ведут игру (посредством Game Finder (средство поиска игры), описанного ранее, и других средств), а также принимать решение о том, является ли игра интересной, и если это так, то изучать рекомендации при наблюдении за другими. И пользователь может испытать демонстрационную версию игры немедленно, и при этом нет необходимости ожидать длительной загрузки и/или инсталляции, и пользователь может вести игру немедленно, возможно, для испытания за небольшую плату, или на долгосрочной основе. И пользователь может вести игру на PC Windows, Macintosh, по телевизору, дома, во время поездок, и даже по мобильному телефону посредством беспроводного 3G соединения с достаточно малым временем ожидания. И это все можно выполнять и при этом никогда не иметь физическую копию игры.

Как упоминалось ранее, пользователь может принять решение не обеспечивать другим (пользователям) возможность просмотра его геймплей, обеспечить возможность просмотра его игры после задержки, обеспечить возможность просмотра его игры выбранным пользователям или обеспечить возможность просмотра его игры всем пользователям. Независимо ни от чего видео/аудио будут сохраняться в одном варианте осуществления в течение 15 минут в буфере 5 задержки и пользователь сможет "перематывать" и просматривать свою предшествующую игру, и приостанавливать, воспроизводить ее медленно, быстро перематывать вперед и т.д. так, как он может делать при просмотре телевизора с цифровым видеомagneитофоном (DVR). Несмотря на то что в этом примере пользователь ведет игру, идентичное средство "DVR" доступно, если пользователь использует приложение. Это может

быть полезно при просмотре предшествующей работы и для других применений, как подробно описано ниже. Кроме того, если игра была разработана со средством перемотки на основе использования информации о состоянии игры так, что поле зрения камеры может быть изменено и т.д., то это средство "3D DVR" также будет поддерживаться, но потребуются, чтобы игра была разработана с его поддержкой. Средство "DVR" с использованием буфера 5 задержки работает с любыми игрой или приложением, конечно, с ограничениями для видео, которое сформировано во время использования игры или приложения, но в случае игр со средством DVR 3D пользователь может управлять "сквозным пролетом" (fly through) в 3D ранее проведенного сегмента и осуществлять запись получающегося в результате видео в буфер 5 задержки, и осуществлять запись состояния игры для сегмента игры. Соответственно, конкретное "сквозное быстрое перемещение" записывается как сжатое видео, но так как состояние игры также записывается, то другое сквозное быстрое перемещение для идентичного сегмента игры впоследствии возможно. Как описано ниже, каждый пользователь имеет на службе хостинга 1 страницу пользователя (User Page), куда он может помещать информацию о себе и другие данные. Помимо прочего пользователи могут помещать сегменты видео из геймплей (процесс игры), которые они сохранили. Например, если пользователь преодолел особенно трудное серьезное испытание в игре, то пользователь может "перемотать" до места непосредственно перед своим большим достижением в игре и затем выдать команду в службу 1 хостинга для сохранения сегмента видео некоторой продолжительности (например, 30 секунд) на странице пользователя (User Page) этого пользователя для просмотра другими пользователями. Для реализации этого, это - просто вопрос сервера 6-8 приложения/игры, который пользователь использует, считать видео, сохраненное в буфере 5 задержки, в дисковый массив 3-4-типа RAID и затем сделать ссылку на этот сегмент видео на странице пользователя (User Page) упомянутого пользователя.

Если в игре существует средство DVR 3D, как описано выше, то информация состояния игры, требуемая для DVR 3D также может быть записана пользователем и сделана доступной на странице пользователя (User Page) упомянутого пользователя.

В случае когда игра разработана с возможностью наличия "посетителей" (т.е. пользователей, которые могут перемещаться через мир 3D и наблюдать за действием без участия в нем) наряду с активными игроками, то приложение Game Finder (средство поиска игры) обеспечивает возможность пользователям присоединяться к играм в качестве посетителей, а также игроков. С точки зрения реализации, для системы 1 хостинга не существует разницы, является ли пользователь посетителем или активным игроком. Игра загружается на сервер 6-8 приложения/игры, и пользователь управляет игрой (например, посредством управления виртуальной камерой, которая смотрит в мир). Единственным различием будет опыт использования игры пользователем.

Совместная работа множества пользователей.

Другим признаком службы 1 хостинга является возможность совместной работы множества пользователей, когда они просматривают видео в реальном масштабе времени, даже при использовании значительно различающихся устройств для просмотра. Это полезно и при ведении игры и при использовании приложений.

Многие PC и мобильные телефоны оснащены видеокамерами и в них существует средство для выполнения сжатия видео в реальном времени, в частности, когда изображение является маленьким. Кроме того, в продаже имеются маленькие камеры, которые могут быть подсоединены к телевизору, и не трудно реализовать сжатие в реальном времени или в программном обеспечении или с использованием одного из многих аппаратных устройств сжатия для сжатия видео. Кроме того, во многих PC и во всех мобильных телефонах существуют микрофоны, и в продаже имеются наушники с микрофонами.

Такие камеры и/или микрофоны, объединенные с локальным средством сжатия видео/аудио (в частности, с применением способов сжатия видео с малым временем ожидания, описанные в этом документе), обеспечивают пользователю возможность передачи видео и/или аудио с территории 2 пользователя в службу 2 хостинга вместе с данными управления устройства ввода. Когда применяются такие способы, то можно получить средство, иллюстрируемое на рис. 2.33 пользователь может выводить свое видео и аудио 2 на экран внутри приложения или игры другого пользователя. Этот пример является игрой с несколькими участниками, где товарищи по команде совместно участвуют в автомобильной гонке. Видео/аудио пользователя может быть выборочно просматриваемым/прослушиваемым только своими товарищами по команде. И так как фактически время ожидания отсутствует при использовании способов, описанных выше, то игроки могут разговаривать друг с другом или показывать жестом друг другу в реальном масштабе времени без ощутимой задержки. Эта интеграция видео/аудио выполняется посредством поступления сжатого видео и/или аудио из камеры/микрофона пользователя как входящий трафик 1 в Internet. После этого входящая маршрутизация 2 маршрутизирует видео и/или аудио в игровые серверы 6-8 приложения/игры, которым обеспечена возможность просмотра/прослушивания видео и/или аудио. Далее пользователи соответствующих игровых серверов 6-8 приложения/игры, которые решили использовать видео и/или аудио, восстанавливают его и интегрируют, как требуется для вывода внутри игры или приложения, например как изображено в позиции 2.

В примере по рис. 2.33 показано, как такая совместная работа используется в игре, но такая совместная работа может быть мощным инструментом для приложений.

РАЗДЕЛ 3. ЭКОНОМИЧЕСКАЯ ЧАСТЬ

Целью дипломной работы является разработка системы обработки данных, которая позволит уменьшить время ожидания при получении интерактивного потокового видео.

Данная техническая разработка относится к области систем обработки данных, которые улучшают способность пользователей манипулировать аудио и видеоносителями и подключаться к ним.

Технический результат заключается в уменьшении времени ожидания получения интерактивного потокового сжатого интерактивного видео в

клиентские устройства пользователей по сети Интернет, даже при ограниченной полосе пропускания.

Экономические затраты на проведение компьютерной игры с использованием интерактивного потокового видео включают в себя затраты на приобретение необходимого оборудования и его установку и настройку.

3.1 Экономические затраты на приобретение оборудования

Для проведения компьютерной игры с использованием интерактивного потокового видео состоящей из трех точек понадобится следующее оборудование:

- средства вывода видео- и аудиоинформации - плазменные или жидкокристаллические панели, мониторы, телевизоры, проекторы, усилители звука, аудиокolonки и т.п;

- серверное оборудование ;

- центральный процессор (CPU - central processing unit) - центральное обрабатывающее устройство ;

- графический процессор (англ. graphics processing unit, GPU) — отдельное устройство персонального компьютера или игровой приставки, выполняющее графический рендеринг;

- игровой контроллер - устройство ввода информации, которое используется в консольных и компьютерных играх.

- накопитель на оптических носителях информации — оптический диск (англ. optical disc) — носитель информации, выполненный в виде дисков, чтение с которых ведётся с помощью оптического (лазерного) излучения.

Специализированное программное обеспечение.

Запоминающее устройство - устройство для хранения информации.

Оперативная память (RAM - Random Access Memory).

Телекоммуникационная сеть (Интернет), посредством которой все эти элементы объединяются в единую систему.

Понадобятся услуги специалистов по настройке и установке оборудования.

3.2 Экономические затраты на установку и настройку оборудования

Экономические затраты на установку и настройку оборудования включают в себя затраты на привлечение соответствующего специалиста, который сможет установить и настроить необходимое оборудование.

Заработная плата такого специалиста в Анголе в месяц составляет 1268 \$ при пятидневном восьмичасовом рабочем дне.

Отсюда выведем заработную плату в час при условии 20 восьмичасовых рабочих дней в месяц:

$$20 \cdot 8 = 160 \text{ рабочих часов в месяц}$$

$$1268 : 160 = 7,9 \text{ \$ заработная плата в час}$$

Таким образом, заработная плата специалиста в сфере телекоммуникаций в Анголе составляет 7,9 \$ в час.

На установку необходимого оборудования понадобится 12 часов. Отсюда выведем сумму которую необходимо будет потратить на услуги специалиста по установке оборудования:

$$7,9 \cdot 12 = 94,8 \text{ \$}$$

Таким образом, на установку оборудования в одной точке понадобится 94,8 \$

На настройку необходимого оборудования понадобится 10 часов. Отсюда выведем сумму которую необходимо будет потратить на услуги специалиста по настройке оборудования:

$$7,9 \cdot 10 = 79 \text{ \$}$$

Таким образом, на настройку оборудования в одной точке понадобится 79 \$

Таблица 3.1 - Экономические затраты на покупку, установку и настройку необходимого оборудования проведения компьютерной игры с использованием интерактивного потокового видео (3 точки)

Экономические затраты	Количество	Цена за единицу	Сумма за 3 единицы
Монитор AOC e970Swn	1	\$ 83	\$ 249
Накопитель на оптических носителях информации(Hitachi-LG Blu-ray writer SATA INT BULK BLACK(BH 16 NSHO)	1	\$ 74	\$ 222
Жесткие диски (Toshiba 500 GB)	1	\$ 47	\$ 141
RAM(16GB)	1	\$ 75	\$ 225
Сервер Ustream	1	\$ 499	\$ 1497
CPU (Процессор intel windows 7 или 10)	1	\$ 365	\$ 1095
GPU 8GB (MSI Geforte gtx 60 drex armor 06	1	\$ 819	\$ 2457

Сетевое Соединение (Internet Unitel)	1	\$119	\$ 357
Игровой Контроллер(Game pad F710)	1	\$50	\$ 150
Установка оборудования	1	\$ 94.8	\$ 284.4
Настройка оборудования	1	\$ 79	\$ 237

Зная стоимость необходимого оборудования, его установки и настройки мы можем посчитать капитальные затраты:

$$З. \text{ кап.} = 6393 + 284.4 + 237 = 6914.4$$

Где: 6393 - стоимость оборудования в 3 точках;

284.4 - стоимость установки оборудования в 3 точках;

237 - стоимость настройки оборудования в 3 точках.

Итого: 6914.4 \$

Согласно проведенным экономическим затратам можно сделать вывод, что для проведения компьютерной игры с использованием интерактивного потокового видео состоящей из трех точек в Анголе понадобится 6914.4 \$

ЗАКЛЮЧЕНИЕ

Данная техническая разработка относится к области систем обработки данных, которые улучшают способность пользователей манипулировать аудио и видеоносителями и подключаться к ним.

Технический результат заключается в уменьшении времени ожидания получения интерактивного потокового сжатого интерактивного видео в клиентские устройства пользователей по сети Интернет, даже при ограниченной полосе пропускания.

Основные технические решения заключаются в следующем:

- Принимают в конкретном серверном центре службы хостинга передачу из клиентского устройства по пакетной сети, причем передача включает в себя запрос на воспроизведение выбранной твич видеоигры или приложения.

- Исполняют в конкретном серверном центре службы хостинга выбранную твич видеоигру или приложение на одном или нескольких серверах и, таким образом, формируют поток несжатого потокового интерактивного видео.

- Сжимают в конкретном серверном центре службы хостинга поток несжатого потокового интерактивного видео и, таким образом, создают поток сжатого с малым временем ожидания потокового интерактивного видео.

- Передают поток сжатого с малым временем ожидания потокового интерактивного видео в клиентское устройство по пакетной сети, причем сжатое с малым временем ожидания потоковое интерактивное видео является твич видеоигры или приложения.

- Исполняют в конкретном серверном центре службы хостинга выбранную твич видеоигру или приложение на одном или нескольких серверах и, таким образом, формируют поток несжатого потокового интерактивного видео.

- Сжимают в конкретном серверном центре службы хостинга поток несжатого потокового интерактивного видео и, таким образом, создают поток сжатого с малым сжатым, в худшем случае, с временем ожидания передачи туда и обратно в 90 мс, отсчитываемым от приема пользовательского ввода на клиентском устройстве до появления отклика на связанном с клиентским устройством дисплее, при расстоянии передачи в 1800 километров или менее.

СПИСОК ИСПОЛЬЗОВАННОЙ ЛИТЕРАТУРЫ

1. <http://www3.ncsu.edu/dox/video/survey.html>
2. <http://www.video-conferencing.com/>
3. http://www.aps.anl.gov/helpdocs/mbone_conf/vconf-faq.html
4. http://www.kn.pacbeil.com/edfirst/dt_video.html
<http://www.visc.vt.edu/succeed/conferencing/index2.html>
5. <http://www.brookwood.com/desktop.htm>
6. <http://www.soi.city.ac.uk/homes/ef516/vidconf>
7. ММКР. Параметры кодирования сигналов цифрового телевидения для студий. - Рекомендации 601. – 450 с.
8. <http://drogo.cselst.stet.it/mpeg/>
9. <http://www.mpeg.org/index.html/>

10. <http://search.zdnet.com/pcmag/features/vidconf/>
11. <http://mice.ed.ac.uk/>
12. Золотов С.В. Протоколы INTERNET. – СПб.: BHV, 1998. – 304 с.
13. <http://www.cs.sfu.ca/CourseCentral/365/li/>
14. <http://www2.echo.lu/oii/en/oiistand.html>
15. <http://mice.ed.ac.uk/>
16. <http://www.ipmulticast.com>
17. <http://www.imtc.org>
18. Прэтт У. Цифровая обработка изображений. Т.1, Т.2. – М.:Мир, 1982. – 150 с.
19. Годьберг Л.М. , Матюшкин Б.Д. Цифровая обработка сигналов. Справочник. – М.:Радио и связь, 1995. – 235 с.
20. <ftp://havefun.stanford.edu/pub/>
21. <http://www.vtel.com/vcnews/>

ПРИЛОЖЕНИЕ Ж

Ж 1 Обширные общественные видеослужбы

Служба 1 хостинга обеспечивает беспрецедентную возможность для установки обширных общественных видеослужб в Internet. На рис. 2.34 изображена иллюстративная страница пользователя (User Page) для игрока в службе 1 хостинга. Как в случае с приложением Game Finder (средство поиска игры), User Page (страница пользователя) является приложением, которое исполняется на одном из серверов 6-8 приложения/игры. Во всех миниатюрах и видеоокнах на

этой странице постоянно показывается движущееся видео (если сегменты являются короткими, то они циклически повторяются). С использованием видеокамеры или посредством выгрузки видео пользователь (именем пользователя которого является "KILLHAZARD") может поместить видео самого себя 2000, которое могут просматривать другие пользователи. Видео хранится в дисковом массиве 3-4 типа RAID. Кроме того, когда другие пользователи переходят на страницу пользователя (User Page) KILLHAZARD, если KILLHAZARD использует службу 1 хостинга в это время, то будет показано видео 1 в реальном масштабе времени того, что он делает (при предположении, что он обеспечивает возможность пользователям просматривать его страницу пользователя (User Page) для наблюдения за ним). Это выполняется сервером 6-8 ия/игры, который осуществляет хостинг приложения User Page (страница пользователя), с запросом у системы 3 управления службой, является ли KILLHAZARD активным, и если это так, то сервер 6-8 приложения/игры, который он использует. После этого, с использованием идентичных способов, используемых Game Finder (средство поиска игры), сжатый видеопоток с соответствующими разрешающей способностью и форматом отправятся в сервер 6-8 приложения/игры, исполняющий приложение User Page (страница пользователя), и оно будет выведено на экран. Если пользователь выбирает окно геймплей в реальном масштабе времени KILLHAZARD и затем надлежащим образом щелкает по своему устройству ввода, то окно увеличивается (и опять же с использованием идентичных способов, как в приложениях Game Finder (средства поиска игры)) и видео в реальном масштабе времени заполняет экран с разрешающей способностью дисплея 16 наблюдающего пользователя, соответствующее характеристикам подключения к Internet наблюдающего пользователя.

Главным преимуществом этого подхода перед подходами предшествующего уровня техники является то, что пользователь, просматривающий User Page (страницу пользователя), может видеть игру, воспроизводящуюся в режиме реального времени, которая не принадлежит

пользователю, и он может не иметь локального компьютера или игровой консоли с возможностью воспроизведения игры. То, что пользователь видит пользователя в User Page (страница пользователя) ”в действии”, ведущего игры, предоставляет ему большие возможности, и это является возможностью узнать об игре, которую просматривающему пользователю, возможно, требуется испытать или улучшить свои результаты в ней. Записанные камерой или выгруженные видеоклипы от друзей 2 KILLHAZARD также показываются в User Page (на странице пользователя), и под каждым видеоклипом существует текст, который указывает, находится ли этот друг в режиме онлайн и ведет игру (например, six_shot ведет игру ”Eragon” и MrSnuggles99 находится в режиме офлайн и т.д.). При щелчке на пункте меню (не изображен) видеоклипы друга переключаются с показа записанных или выгруженных видео на видео в реальном масштабе времени того, что друзья, которые в настоящее время ведут игры в службе 1 хостинга, делают в этот момент в своих играх. Соответственно, это становится Game Finder (средством поиска игры), выполняющим группировку по друзьям. Если выбирается игра друга и пользователь щелкает по ней, то она увеличится до полного экрана и пользователь сможет смотреть игру, воспроизводимую во весь экран в режиме реального времени. И опять же, пользователю, просматривающему игру друга, не принадлежит копия этой (игры), а также не принадлежат локальные вычислительные ресурсы/ресурсы игровой консоли для ведения этой игры. Просмотр игры является фактически незамедлительным. Как ранее описано выше, когда пользователь ведет игру в службе 1 хостинга, пользователь может ”перематывать” игру и найти сегмент видео, который ему требуется сохранить, и затем сохраняет этот сегмент видео на своей странице пользователя (User Page). Их называют ”Brag Clip” (”клипы хвастовства”). Все сегменты 3 видео являются Brag Clip 3, сохраненными KILLHAZARD из предыдущих игр, которые он вел. В позиции 4 показывается, сколько раз Brag Clip был просмотрен, и когда Brag Clip просматривается, у пользователей есть возможность оценить его, и количество оранжевых пиктограмм 5 в виде замочной скважины указывает, насколько высоким является

рейтинг. Brag Clip 3 постоянно циклически повторяются, когда пользователь просматривает страницу пользователя (User Page), вместе с остальной частью видео на странице. Если пользователь выбирает и щелкает по одному из Brag Clip 3, он увеличивается для представления Brag Clip 3, причем элементы управления DVR обеспечивают возможность воспроизведения клипа, его приостановки, перемотки, быстрой перемотки вперед, прохождения по этапам и т.д.

Воспроизведение Brag Clip 3 реализуется сервером 6-8 приложения/игры, загружающим сегмент сжатого видео, сохраненный в дисковом массиве 3-4 типа RAID, когда пользователь записывал Brag Clip, и восстанавливающим его и воспроизводящим его. Brag Clip 3 могут также быть сегментами видео "DVR 3D" (т.е. последовательностью состояний игры из игры, которая может быть повторно воспроизведена и обеспечивает возможность пользователю изменять точку обзора камеры) из игр, которые поддерживают такое средство. В этом случае информация о состоянии игры сохраняется наряду с записью сжатого видео конкретного "сквозного пролета", которую сделал пользователь, когда записывался сегмент игры. Когда просматривается User Page (страница пользователя) и все миниатюры и видеоокна постоянно циклически повторяются, Brag Clip 3 DVR 3D постоянно циклически повторяет Brag Clip 3, который был записан как сжатое видео, когда пользователь записывал "сквозной пролет" сегмента игры. Но когда пользователь выбирает Brag Clip 3 DVR 3D и щелкает на нем, наряду с элементами управления DVR, обеспечивающими возможность воспроизведения сжатого видео Brag Clip, пользователь сможет щелкнуть на кнопке, которая предоставляет ему средство DVR 3D для сегмента игры. Он сможет управлять "сквозным пролетом" камеры в течение сегмента игры самостоятельно, и, если ему потребуется (и пользователь, которому принадлежит эта страница пользователя, обеспечивает эту возможность), то он сможет записать альтернативный "сквозной пролет" Brag Clip в форме сжатого видео, (которое) после этого будет доступно другим зрителям страницы

пользователя (или немедленно, или после того, как у владельца страницы пользователя будет возможность просмотреть этот Brag Clip).

Это средство Brag Clip 3 DVR 3D обеспечивается посредством активации игры, которая готова начать повторное воспроизведение записанной информации о состоянии игры на другом сервере 6-8 приложения/игры. Так как игра может быть активирована почти мгновенно (как описано ранее), ее нетрудно активировать, причем воспроизведение ее ограничивается до состояния игры, записанного сегментом Brag Clip, и после этого обеспечить возможность пользователю выполнять "сквозной пролет", причем камера записывает сжатое видео в буфер 5 задержки. После завершения пользователем выполнения "сквозного пролета" игра отключается.

С точки зрения пользователя активация "сквозного пролета" посредством Brag Clip 3 DVR 3D не требует большего усилия, чем управление элементами управления DVR линейного Brag Clip 3. Он может ничего не знать об игре или даже то, как вести игру. Он является только оператором виртуальной камеры, смотрящей в мир 3D в течение сегмента игры, записанного другим. Пользователи также могут накладывать их собственный аудио на Brag Clip, которое является или записанным с микрофонов, или выгруженным. Следовательно, Brag Clip могут использоваться для создания пользовательской анимации с использованием персонажей и действий из игр. Этот способ анимации обычно известен как "машинима". По мере продвижения пользователей в играх они достигают отличающихся уровней мастерства. Воспроизведенные игры сообщают об этих достижениях в систему управления 3 службы, и эти уровни мастерства будут показаны на User Page (страницах пользователя).

Ж 2 Интерактивные анимированные рекламные объявления

Осуществлялся переход онлайн-рекламных объявлений от текста к неподвижным изображениям, к видео и в настоящее время к интерактивным сегментам, как правило, реализуемым с использованием тонких клиентов

анимации, подобных Adobe Flash. Причиной, по которой используются тонкие клиенты анимации, является то, что пользователи испытывают нетерпение при задержке (осуществления преимущественного права получения) продукта или услуги, передаваемых им. Кроме того, тонкие клиенты исполняются на оченьнизкопроизводительных РС и, по существу, лицо, дающее объявление, может быть совершенно уверено, что интерактивное рекламное объявление будет работать должным образом. К сожалению, тонкие клиенты анимации, например Adobe Flash, ограничены по степени интерактивности и продолжительности опыта использования (для уменьшения времени загрузки).

На рис. 2.35 изображена интерактивная реклама, в которой пользователь должен выбрать внешние и внутренние цвета автомобиля во время вращения автомобиля в демонстрационном зале, когда посредством трассировки лучей в реальном времени демонстрируется, как выглядит автомобиль. Далее пользователь выбирает "аватара" для управления автомобилем, и после этого пользователь может взять автомобиль для поездки или по трассе для гонок, или по экзотической местности, например, Монако. Пользователь может выбрать двигатель с большим рабочим объемом или лучшие шины и после этого может проверить, как измененная конфигурация влияет на возможность автомобиля ускоряться или держать дорогу. Несомненно, упомянутая реклама фактически является усложненной видеоигрой 3D. Но для воспроизведения такой рекламы на РС или видеоигровой консоли может потребоваться загрузка 100 мегабайт, и в случае РС может потребоваться установка специальных драйверов, и может вообще оказаться невозможным ее исполнение, если в РС отсутствуют отвечающая требованиям вычислительная возможность GPU или CPU. Соответственно, такие рекламные объявления практически невыполнимы в конфигурациях предшествующего уровня техники.

В службе 1 хостинга такие рекламные объявления запускаются почти мгновенно и исполняются превосходно независимо от возможностей клиента 7пользователя.

Соответственно, они запускаются более быстро, чем интерактивные рекламные объявления тонкого клиента, со значительно более обширным опытом использования и с высокой степенью надежности.

Ж 3 Потокковая геометрия во время анимации в реальном времени

Дисковый массив 3-4 типа RAID и входящая маршрутизация 2 могут обеспечивать скорости передачи данных, которые являются такими большими и с величинами времени ожидания такими малыми, что можно проектировать видеоигры и приложения, которые основаны на дисковых массивах 3-4 типа RAID и входящей маршрутизации 2 для надежной доставки геометрии на лету в середине геймплея или в приложении во время анимации в реальном времени (например сквозной пролет” со сложной базой данных).

С системами предшествующего уровня техники, например, игровая видеосистема, изображенная на рис. 1.1, доступные устройства массовой памяти, в частности, в применяемых на практике домашних устройствах, являются слишком медленными для передачи потоком геометрии во время геймплей, за исключением ситуаций, когда требуемая геометрия является до некоторой степени предсказуемой. Например, в игре с вождением, где существует заданное шоссе, геометрия для зданий, которые появляются в поле зрения, может быть приемлемо предсказуемой, и устройства массовой памяти могут вести поиск заранее до места, где находится приближающаяся геометрия.

Но в сложной сцене с непредсказуемыми изменениями (например, в сцене сражения со сложными персонажами повсюду), если RAM в PC или игровой видеосистеме полностью заполнена геометрией для объектов, в настоящее время находящихся в поле зрения, и далее пользователь внезапно поворачивает своего персонажа, чтобы видеть то, что находится сзади него, если геометрия не была предварительно загружена в RAM, то может произойти задержка перед выводом ее на экран. В службе 1 хостинга дисковые массивы 3-4 типа RAID могут потоком передавать данные со скоростью, превышающей скорость по технологии Gigabit Ethernet, и в случае с сетью SAN можно достичь скорости 10 гигабит/секунда по технологии 10 Gigabit Ethernet или по другим сетевым

технологиям. При 10 гигабит/секунда гигабайт данных загружается менее чем за секунду. В период кадра 60 кадр/с (16,67 мс) могут быть загружены приблизительно 170 мегабитов (21 мегабайт) данных. Вращающиеся носители информации, несомненно, даже в конфигурации RAID по-прежнему вызывают величины времени ожидания, превышающие период кадра, но запоминающее устройство RAID на основе флэш-памяти в итоге будет такого размера, как дисковые массивы типа RAID вращающихся носителей информации, и не будет вызывать такого большого времени ожидания. В одном варианте осуществления используется кэширование со сквозной записью в массивную RAM для обеспечения доступа с очень малым временем ожидания. Соответственно, с достаточно высокой скоростью сети и достаточной массовой памятью с достаточно малым временем ожидания геометрию можно передавать потоком в игровые серверы 6-8 приложения/игры с такой скоростью, с которой CPU и/или GPU могут обрабатывать данные 3D. Соответственно, в примере, приведенном ранее, где пользователь внезапно поворачивает своего персонажа и смотрит назад, геометрия для всех персонажей, находящихся сзади, может быть загружена до того, как персонаж завершает поворот, и, соответственно, пользователю будет казаться, как будто он или она находятся в фотореалистичном мире, который настолько реальный, как игра актеров. Как обсуждалось ранее, одной из последних границ в фотореалистичной компьютерной анимации является человеческое лицо, и из-за чувствительности человеческого глаза к несовершенствам малейшая ошибка в фотореальном лице может в результате привести к отрицательной реакции у зрителя. На рис. 2.36 изображено то, как живое исполнение, захваченное с использованием технологии Contour™ Reality Capture, в результате приводит к очень гладкой захваченной поверхности, затем к сетчатой поверхности, насчитывающей большое количество многоугольников (т.е. движение многоугольника точно следует за движением лица). Наконец, когда видео живого исполнения отображают на сетчатую поверхность для вывода текстурированной поверхности, выводится фотореальный результат.

Несмотря на то что посредством современной технологии GPU можно визуализировать количество многоугольников в сетчатой поверхности и текстурировать и осветить поверхность в реальном времени, если многоугольники и текстуры изменяются каждый период кадра (что приводит к наиболее фотореальным результатам), то вся доступная RAM современного PC или видеоигровой консоли быстро израсходуется.

С использованием способов потоковой геометрии, описанных выше, становится возможным применение на практике непрерывной подачи геометрии в игровые серверы 6-8 приложения/игры, чтобы они непрерывно могли анимировать фотореальные лица с обеспечением возможности создания видеоигр с лицами, которые почти нельзя отличить от лиц при игре актеров.

Интеграция линейного содержимого с интерактивными признаками.

Кинофильмы, телевизионные программы и аудиоматериал (обобщенно, "линейное содержимое") являются широко доступными для домашних и офисных пользователей во многих видах. Линейное содержимое может быть приобретено на физических носителях информации, подобных носителям информации CD, DVD, HD-DVD и Blu-ray. Оно также может быть записано посредством DVR из вещательной передачи кабельного телевидения и спутниковой вещательной передачи. И оно доступно как содержимое с платой за просмотр (PPV) через спутниковое и кабельное телевидение и как видео по запросу (VOD) по кабельному телевидению. Все больше линейное содержимое доступно через Internet и как загружаемое и как потоковое содержимое. В настоящее время на самом деле не существует ни одного места, где можно испытать все признаки, связанные с линейным мультимедиа. Онлайн-музыкальные сайты охватывают художественную и песенную информацию, как правило, не доступную на CD, но не все CD доступны в режиме онлайн. И на Web-сайтах, связанных с телевизионными программами, часто существуют дополнительные признаки, блоги и иногда комментарии актеров или творческого персонала.

Кроме того, в отношении многих кинофильмов или спортивных событий часто существуют видеоигры, которые выпускают (в случае кинофильмов) часто вместе с линейным мультимедиа или (в случае спортивных состязаний) которые могут быть близко связаны с реальными событиями (например, торговля игроками). Служба 1 хостинга хорошо подходит для доставки линейного содержимого в компоновке разрозненных форм взаимосвязанного содержимого. Конечно, доставка кинофильмов требует не больше усилий, чем доставка в высокой степени интерактивных видеоигр, и служба 1 хостинга может доставлять линейное содержимое в широкий диапазон устройств, находящихся в домах или офисах, или в мобильные устройства. На рис. 2.36 изображена иллюстративная страница интерфейса пользователя для службы 1 хостинга, на которой изображена выборка линейного содержимого.

Но в отличие от большинства систем доставки линейного содержимого служба 1 хостинга также может доставлять взаимосвязанные интерактивные компоненты (например, меню и признаки на DVD, интерактивные наложения одного графического изображения на другое на HD-DVD и анимацию Adobe Flash, как объясняется ниже) на Web-сайтах. Соответственно, ограничения клиентского устройства больше не вводят ограничений относительно того, какие признаки являются доступными.

Кроме того, система 1 хостинга может динамически и в реальном времени компоновать линейное содержимое с содержимым видеоигры. Например, если пользователь смотрит матч Quidditch в фильме Гарри Поттер и решает, что она бы хотела попробовать сыграть в Quidditch, то она может просто щелкнуть на кнопке и фильм будет приостановлен, и она немедленно переместится в сегмент Quidditch видеоигры Гарри Поттер. После проведения матча Quidditch - еще один щелчок кнопкой и фильм немедленно будет продолжен. С фотореальной графикой и технологией производства, при которой фотографически захваченное видео нельзя отличить от персонажей с игрой актеров, когда пользователь осуществляет переход от игры Quidditch в фильме с игрой актеров к игре Quidditch в видеоигре на службе хостинга, как описано в

этом документе, эти две сцены фактически нельзя отличить. Это обеспечивает полностью новыми творческими возможными вариантами для режиссеров и линейного содержимого и интерактивного содержимого (например, видеоигры), поскольку границы между этими двумя мирами становятся неразличимыми. С использованием архитектуры службы хостинга, изображенной на рис. 2.28, зрителю может быть предоставлено управление виртуальной камерой в фильме 3D. Например, в сцене, которая имеет место внутри вагона поезда, можно обеспечить возможность зрителю управлять виртуальной камерой и осматривать вагон по ходу рассказа.

Тогда различные кадры могут быть преобразованы так, что изображения не будут искажены и после этого сохранены в дисковом массиве 3-4 типа RAID вместе с фильмом и воспроизведены, когда пользователь решит отправиться в виртуальное путешествие. В отношении спортивных событий, спортивное событие, происходящее в реальном масштабе времени, например, игра в баскетбол, может передаваться через службу 1 хостинга, чтобы пользователи смотрели ее так, как они могут смотреть ее по обычному телевизору. После того, как пользователи посмотрели конкретную игру, видеоигра этой игры (в итоге, с баскетболистами, которые выглядят также фотореально, как реальные игроки) может появиться на экране, причем игроки начинают с идентичной позиции, и пользователи (возможно, каждый при этом получает управление одним игроком) могут восстановить игру и посмотреть, могут ли они добиться большего успеха, чем эти игроки.

Служба 1 хостинга, описанная в этом документе, очень хорошо подходит для поддержки этого мира будущего, потому что она может привлекать вычислительные возможности и ресурсы массовой памяти, которые практически нецелесообразно устанавливать в домашней обстановке или в большинстве офисных обстановок, и также ее вычислительные ресурсы всегда являются современными, с последним вычислительным доступным аппаратным обеспечением, тогда как в домашней обстановке всегда будут существовать дома с видеоиграми и РС предыдущих поколений. И в службе 1 хостинга вся эта

вычислительная сложность скрыта от пользователя, поэтому несмотря на то что они могут использовать очень усложненные системы, с точки зрения пользователя, это является простым, как переключение каналов на телевизоре. Кроме того, пользователи могут получать доступ ко всей вычислительной мощности и опыт использования, который может предоставить эта вычислительная мощность из любого клиента 7.

Ж 4 Игры с несколькими участниками.

В тех случаях, когда игра является игрой с несколькими участниками, тогда она может осуществлять связь и с игровыми серверами 31-4 приложения/игры через сеть входящей маршрутизации 2 и посредством сетевого моста с Internet (не изображен) с серверами или игровыми машинами, которые не функционируют в службе 1 хостинга. При ведении игры с несколькими участниками посредством компьютеров в общем Internet игровые серверы 6-8 приложения/игры будут иметь преимущество с чрезвычайно быстрым доступом к Internet (по сравнению с тем, если игра исполняется на сервере дома), но они будут ограничены возможностями других компьютеров, которые ведут игру на более медленных соединениях, и также, возможно, ограничены тем, что игровые серверы в Internet предназначены для обеспечения наименьшего общего знаменателя, который может быть домашними компьютерами на относительно медленных потребительских подключениях к Internet. Но когда игра с несколькими участниками полностью ведется внутри серверного центра службы 1 хостинга, тогда можно достичь большого различия. Каждый игровой сервер 6-8 приложения/игры, который осуществляет хостинг игры для пользователя, будет соединен с другими игровыми серверами 6-8 приложения/игры, а также с любыми серверами, которые осуществляют хостинг центрального управления для игры с несколькими участниками с чрезвычайно высокой скоростью, с соединением с чрезвычайно малым временем ожидания и с огромными, очень быстрыми массивами запоминающих устройств. Например, если используется

технология Gigabit Ethernet для сети входящей маршрутизации 2, то игровые серверы 6-8 приложения/игры осуществляют связь друг с другом и осуществляют связь с любыми серверами, которые осуществляют хостинг центрального управления для игры с несколькими участниками на скорости гигабит/секунда с, возможно, временем ожидания только 1 мс или меньше. Кроме того, дисковые массивы 3-4 типа RAID смогут отвечать очень быстро и затем передать данные со скоростями гигабит/секунда. В качестве примера, если пользователь настраивает самостоятельно персонажа с точки зрения внешнего вида и снаряжения так, что этот персонаж имеет большой объем геометрии и линий поведения, которые являются уникальными для персонажа, то с системами предшествующего уровня техники, которые ограничены клиентом игры, исполняющимся дома на PC или игровой консоли, если этот персонаж должен появиться в поле зрения другого пользователя, то этот пользователь должен ждать, пока не завершится длинная, медленная загрузка, чтобы все данные линии поведения и геометрии загрузились в его компьютер. В пределах службы 1 хостинга идентичная загрузка может осуществляться по Gigabit Ethernet с обслуживанием из дискового массива 3-4 типа RAID со скоростью гигабит/секунда. Даже если домашний пользователь имеет подключение к Internet 8 Мбит/с (которое является чрезвычайно быстрым по современным стандартам), то Gigabit Ethernet является в 100 раз быстрее. Соответственно, то, что занимает одну минуту по быстрому подключению к Internet, займет меньше чем одна секунда по Gigabit Ethernet.

Ж 5 Группы лучших игроков и соревнования

Служба 1 хостинга очень хорошо подходит для соревнований. Поскольку игра не исполняется в локальном клиенте, для пользователей не существует возможности обманывать. Кроме того, из-за того, что выходная маршрутизация 12 может осуществлять групповую передачу потоков UDP, Служба 1 хостинга может транслировать основные соревнования тысячам людей в аудитории одновременно. Фактически, когда существуют определенные видеопотоки, которые настолько популярны, что тысячи пользователей принимают

идентичный поток (например, показ видов основного соревнования), может оказаться более эффективным отправить этот видеопоток в сеть доставки контента (CDN), например Akamai или Limelight, для массового распределения во многие клиентские устройства 7.

Аналогичный уровень эффективности может быть получен, когда CDN используют для показа страниц Game Finder (средство поиска игры) групп лучших игроков. Для основных соревнований можно использовать реального спортивного комментатора знаменитостей для обеспечения комментариев во время определенных матчей. Хотя большое количество пользователей будут смотреть основное соревнование и относительно маленькое количество пользователей будут играть на соревновании. Аудио от спортивного комментатора знаменитостей может маршрутизироваться в игровые серверы 68 приложения/игры, осуществляющие хостинг пользователей, играющих на соревновании, и хостинг любых копий режимов посетителя игры во время соревнования, и это аудио может быть наложено поверх аудиоигры. Видео спортивного комментатора знаменитостей может также быть наложено на игры, возможно, только на отображения на экране посетителя.

Ж 6 Ускорение загрузки web-страницы

Исходный транспортный протокол всемирной паутины, протокол передачи гипертекстовых файлов (HTTP), был задуман и определен в период, когда только в компаниях были высокоскоростные подключения к Internet, и потребители, которые были в режиме онлайн, использовали модемы коммутируемой линии передачи или ISDN. В это время "золотым стандартом" для высокоскоростного соединения была линия T1, которая обеспечивала скорость передачи данных 1,5 Мбит/с симметрично (т.е. с равной скоростью передачи данных в обоих направлениях). В настоящее время ситуация является совершенно другой. Средняя скорость домашнего соединения через соединения кабельной модемной связи или DSL в большей части развитых стран имеет гораздо более высокую скорость передачи данных нисходящего потока данных, чем линия T1. Фактически в некоторых частях мира посредством технологии ввода в здания

оптического кабеля (fiber-to-the-curb) в дом подаются скорости передачи данных от 50 до 100 Мбит/с.

К сожалению, HTTP не был спроектирован (он также не был реализован) для фактического использования этих существенных увеличений скорости. Web-сайт является совокупностью файлов на удаленном сервере. В самых простых словах, HTTP запрашивает первый файл, ожидает его загрузки и после этого запрашивает второй файл, ожидает его загрузки и т.д. Фактически HTTP обеспечивает возможность нескольких ”открытых соединений”, т.е. запрос нескольких файлов за один раз, но из-за установленных стандартов (и стремления предотвратить перегрузку web-серверов) обеспечивается возможность только очень малого количества открытых соединений.

Кроме того, из-за способа создания Web-страниц браузерам часто не известно, что множество страниц могут быть одновременно доступны для немедленной загрузки (т. е. только после анализа (парсинга) страницы становится очевидно, что должен быть загружен новый файл, например изображение). Соответственно, файлы на web-сайте по существу загружаются один за другим.

В отношении соединений с относительно низкой скоростью это не представляет большую проблему, потому что время загрузки самих файлов доминирует над временем ожидания для веб-страниц. Но, поскольку скорости соединения растут, особенно в отношении сложных веб-страниц, начинают возникать проблемы.

В примере, изображенном на рис. 2.38, представлен типичный коммерческий web- сайт (этот конкретный web-сайт являлся сайтом от лидирующей торговой марки спортивной обуви). На этом web-сайте существует 54 файла. Файлы включают в себя HTML, CSS, JPEG, PHP, JavaScript и Flash файлы и включают в себя содержимое видео. В общей сложности должно быть загружено 1,5 мегабайта до того, как страница станет активной (т.е. пользователь может щелкнуть по ней и начать использовать ее). Существует несколько причин для большого количества файлов. Однако она является очень типичной коммерческой веб- страницей.

На рис. 2.38 представлено количество времени, которое протекает до того, как веб-страница становится активной, по мере того как скорость соединения возрастает. При скорости 2401 соединения 1,5 Мбит/с с использованием обычного web-сервера посредством обычного web-браузера требуется 13,5 секунд, чтобы веб-страница стала активной. При скорости 2402 соединения 12 Мбит/с время загрузки уменьшается до 6,5 секунд или почти в два раза быстрее. Но при скорости 2403 соединения 96 Мбит/с, время загрузки уменьшается только приблизительно до 5,5 секунд. Причина, по которой это происходит, заключается в том, что при такой высокой скорости загрузки время загрузки самих файлов является минимальным, но время ожидания для каждого файла, примерно 100 мс, по-прежнему, остается, что в результате приводит к $54 \text{ файла} \times 100 \text{ мс} = 5,4 \text{ секунд}$ времени ожидания. Соответственно, независимо от скорости соединения до дома время для того, чтобы этот вебсайт стал активным, всегда будет равно, по меньшей мере, 5,4 секунды. Другим фактором является организация очереди со стороны сервера. Каждый запрос HTTP добавляется в конец очереди, соответственно, на загруженном сервере это оказывает существенное влияние, потому что, чтобы получить из web-сервера каждый маленький элемент, запросы HTTP должны ждать своей очереди. Одним способом решения этой проблемы является отказ от HTTP или его переопределение. Или, возможно, лучше, чтобы владелец web-сайта объединил свои файлы в единый файл (например, в формате Adobe Flash). Но на практике эта компания, а также многие другие (компании) много инвестировали в архитектуру своего веб-сайта. Кроме того, в то время как в некоторых домах существуют соединения 12-100 Мбит/с, в большинстве домов все еще существуют более медленные скорости, а HTTP хорошо работает при медленной скорости.

Одним альтернативным вариантом является хостинг web-браузеров на серверах 6-8 приложения/игры и хостинг файлов для web-серверов на дисковых массивах 3-4-типа RAID (или, возможно, в RAM или на локальном запоминающем устройстве на серверах 6-8 приложения/игры, осуществляющих

хостинг web-браузеров). Вследствие очень быстрой взаимосвязи через входящую маршрутизацию 2 (или с локальным запоминающим устройством) вместо 100 мс времени ожидания при использовании НТТР для каждого файла, будет существовать незначительное время ожидания при использовании НТТР для каждого файла. Тогда вместо того, чтобы пользователь в своем доме получал доступ к веб-странице через НТТР, пользователь может получать доступ к веб-странице через клиента 415. Тогда даже с соединением 1,5 Мбит/с (потому что эта веб-страница не требует большой полосы пропускания для своего видео) web-страница будет активной меньше чем через 1 секунду для каждой линии 2400 связи. По существу, не будет времени ожидания до вывода на экран активной страницы web-браузером, исполняющимся на сервере 6-8 приложения/игры, и не будет заметного времени ожидания до вывода на экран клиентом 7 видеовыхода из web-браузера. По мере того как пользователь водит мышью по веб-странице и/или набирает текст на ней, информация ввода пользователя отправляется в web-браузер, исполняющийся на сервере 6-8 приложения/игры, и web-браузер отвечает соответственно.

Одним недостатком этого подхода является то, что, если устройство сжатия постоянно передает видеоданные, то полоса пропускания используется, даже если веб- страница становится неподвижной. Это можно исправить посредством выполнения устройства сжатия с возможностью передавать данные только тогда, когда (и если) веб-страница изменяется, и тогда передавать данные только в части страницы, которые изменились. Несмотря на то что существуют некоторые веб-страницы с флэш- баннерами и т.д., которые постоянно изменяются, такие веб-страницы, как правило, раздражают, и обычно веб-страницы являются неподвижными, если не существует причины для какого-либо движения (например, видеоклип). Для таких веб-страниц это, вероятно, тот случай, когда с использованием службы 1 хостинга будет передаваться меньше данных, чем для обычного web-сервера, потому что будут передаваться только фактически выводимые на экран изображения, никакого исполняемого

кода тонкого клиента и никаких больших объектов, которые могут никогда не просматриваться, например переключающиеся картинки.

Соответственно, с использованием службы 1 хостинга для осуществления хостинга унаследованных веб-страниц время загрузки веб-страниц может быть уменьшено настолько, что открытие веб-страницы будет похоже на переключение телевизионных каналов, веб-страница становится активной фактически немедленно.

Обеспечение отладки игр и приложений

Как упоминалось ранее, видеоигры и приложения с графикой в реальном времени являются очень сложными приложениями, и, как правило, когда их выпускают в реальные условия эксплуатации, они содержат ошибки. Хотя разработчики программного обеспечения получают с обратной связью от пользователей (сообщения) об ошибках и у них могут быть некоторые средства для обратной передачи машинного состояния после аварийных отказов, очень трудно точно идентифицировать то, что вызвало аварийный отказ или ненадлежащее выполнение игры или приложения реального времени.

Когда игра или приложение исполняются в службе 1 хостинга, видео/аудиовыход игры или приложения постоянно записывается в буфер 5 задержки. Кроме того, сторожевой процесс выполняется каждым сервером 6-8 приложения/игры, который регулярно сообщает в систему 3 управления службы хостинга о том, что сервер 6-8 приложения/игры выполняется гладко. Если сторожевой процесс не передает сообщение, то система 3 управления сервером пытается связаться с сервером 6-8 приложения/игры и в случае успеха собирает все доступное машинное состояние. Вся доступная информация вместе с видео/аудио, записанным буфером 5 задержки, будет отправлена разработчику программного обеспечения.

Соответственно, когда разработчик прикладного программного обеспечения или игры получает уведомление об аварийном отказе из службы 1 хостинга, (он) получает покадровую запись того, что привело к этому аварийному отказу.

Эта информация может быть очень ценной в обнаружении ошибок и исправлении их.

Также отметим, что, когда происходит аварийный отказ сервера 6-8 приложения/игры, сервер повторно запускается с самой последней точки перезапуска и пользователю передается сообщение с извинениями за техническую проблему.

Совместное использование ресурсов и снижение расходов

Система, изображенная на рис. 2.4 и рис. 2.5, обеспечивает множество преимуществ как для конечных пользователей, так и для разработчиков приложений и игр. Например, как правило, домашние и офисные клиентские системы (например, PC или игровые консоли) используются только в течение небольшого количества времени в неделю. Согласно пресс-релизу, опубликованному 5 октября 2006 г., Nielsen Entertainment "Active Gamer Benchmark Study" активные игроки тратят в среднем 14 часов в неделю на игру на видеоигровых консолях и приблизительно 17 часов в неделю на карманных компьютерах. В этом сообщении также утверждается, что вся игровая активность активных игроков в среднем составляет 13 часов в неделю. С учетом более высокий показатель времени ведения видеоигр на консоли, $24 \times 7 = 168$ часов в неделю, которая подразумевает, что в доме активного игрока видеоигровая консоль используется только $17/168 = 10\%$ времени в неделю. Или 90% времени видеоигровая консоль не используется. С учетом высокой стоимости видеоигровых консолей и того, что производители финансируют такие устройства, это является очень неэффективным использованием дорогостоящего ресурса. PC, находящиеся в компаниях, также, как правило, используются только часть времени в неделю, особенно стационарные настольные PC, часто требуемые для приложений с широкими функциональными возможностями, например, Maya Autodesk. Несмотря на то что некоторые компании работают все время и в праздники и некоторые PC, используются все время и в праздники, большая часть коммерческой

деятельности, как правило, сосредоточена примерно от 9 часов до 17 часов в часовом поясе данной компании с понедельника по пятницу, кроме выходных и перерывов, и так как большинство РС используются, когда пользователь активно работает с РС, из этого следует, что настольные РС, как правило, используются в это время работы. Если предположить, что РС постоянно используются от 9 часов до 17 часов 5 дней в неделю, то это означает, что РС используются $40/168=24\%$ времени в неделю. Высокопроизводительные настольные РС являются очень дорогостоящими инвестициями для компаний, и это отражается в очень низком уровне использования.

Соответственно, в общем, РС и видеоигровые консоли используются только небольшую часть времени в неделю. То есть так как многие люди работают в компаниях или в школе в дневное время с понедельника по пятницу, кроме выходных, эти люди, как правило, не играют в видеоигры в это время, и поэтому, на самом деле, они играют в видеоигры, как правило, в другое время, например по вечерам, в выходные дни и в праздники.

С учетом конфигурации службы хостинга, изображенной на рис. 2.4, схемы использования, описанные в двух абзацах выше, в результате приводят к очень эффективному использованию ресурсов. Очевидно, что существует ограничение на количество пользователей, которые могут обслуживаться службой 1 хостинга в данное время, особенно, если пользователи требуют быстрого реагирования в реальном времени для сложных приложений, аналогичных усложненным видеоиграм 3D. Но в отличие от видеоигровой консоли в доме или РС, используемой компанией, которые, как правило, не используются большую часть времени, серверы 4 могут быть повторно использованы разными пользователями в разное время. Например, высокопроизводительный сервер 2 с высокопроизводительными сдвоенным CPU и сдвоенными GPU и большим количеством RAM может использоваться компаниями и школами с 9 часов до 5 часов по рабочим дням, но использоваться игроками, ведущими усложненные видеоигры по вечерам, в выходные дни и в праздники. Аналогично, приложения с низкими характеристиками могут использоваться компаниями и школами на

низкопроизводительном сервере 2 с CPU Celeron, без GPU (или с очень низкопроизводительным GPU) и с ограниченной RAM в течение рабочего времени, и игра с низкими характеристиками может использовать низкопроизводительный сервер 2 в течение нерабочего времени.

В общем, только небольшой процент от общего количества пользователей онлайн-услуг используют службу в данный момент. Если учитывать статистику Nielsen использования видеоигр, перечисленную ранее, то легко понять почему. Если активные игроки играют в консольные игры только 17 часов в неделю и если предположить, что пиковое время использования игры обычно приходится на нерабочее время, нерабочее время по вечерам (с 17 часов до 24 часов, 7×5 дней = 35 часов/неделя) и по выходным дням (с 8 часов до 24 часов, $16 \times 2 = 32$ часов/неделя), то есть $35 + 32 = 65$ пиковых часов в неделю для 17 часов геймплея. Точную пиковую пользовательскую нагрузку в системе трудно оценить по многим причинам, некоторые пользователи играют в непиковые периоды времени, могут существовать периоды дня, когда существуют пики разбиения на группы пользователей, причем на эти пиковые периоды может влиять тип ведущейся игры (например, в детские игры будут, вероятно, играть вечером в более раннее время), и т.д. Но с учетом того, что среднее количество часов, которые игрок проводит за игрой, является гораздо меньшим, чем количество часов дня, когда игрок, вероятно, играет в игру, только часть от количества пользователей службы 1 хостинга будут использовать ее в данный момент. Для этого анализа предположим, что пиковая нагрузка составляет 12,5%. Соответственно, только 12,5% ресурсов полосы пропускания, сжатия и вычисления используются в данный момент, что в результате приводит только к 12,5% от стоимости аппаратного обеспечения для поддержки данного пользователя для ведения данного уровня выполнения игры благодаря повторному использованию ресурсов.

Кроме того, с учетом того, что некоторые игры и приложения требуют большей вычислительной мощности, чем другие, ресурсы могут выделяться динамически на основе ведущейся игры или приложений, исполняемых

пользователями. Соответственно, пользователю, выбирающему приложение или игру с низкими характеристиками, будет выделен низкопроизводительный (менее дорогой) сервер 4, а пользователю, выбирающему приложения или игру с высокими характеристиками, будет выделен высокопроизводительный (более дорогой) сервер 4. Безусловно, данная игра или приложение могут иметь секции с более высокими характеристиками и более низкими характеристиками игры или приложений, и пользователя могут переключать с одного сервера 4 на другой сервер 4 между секциями игры или приложения для поддержки работы пользователя на сервере 4 с самой низкой стоимостью, который удовлетворяет потребностям приложения или игры. Отметим, что дисковые массивы 7 типа RAID, которые намного быстрее, чем один диск, доступны даже низкопроизводительным серверам 4, которые будут иметь преимущество более быстрых дисковых скоростей передачи. Соответственно, средняя стоимость каждого сервера 4 по всем ведущимся играм или используемым приложениям намного меньше, чем стоимость наиболее дорогого сервера 4, на котором исполняются приложения или игра с самыми высокими характеристиками, однако даже низкопроизводительные серверы 4 извлекают преимущества производительности диска из дисковых массивов 7 типа RAID. Кроме того, сервер 40в службе 1 хостинга может быть не чем иным как системной платой PC без интерфейсов периферийных или дисковых запоминающих устройств за исключением сетевого интерфейса и со временем может быть интегрирован в однокристалльную интегральную схему только с быстрым сетевым интерфейсом с SAN 5. Кроме того, дисковые массивы 7 типа RAID, вероятно, будут совместно использоваться намного большим количеством пользователей, чем существует дисков, поэтому стоимость дискового запоминающего устройства для каждого активного пользователя будет намного меньше, чем один накопитель на магнитных дисках. Все это оборудование, вероятно, будет находиться в стойке в обстановке помещения для установки серверов с регулируемыми характеристиками окружающей среды. Если произойдет отказ сервера 4, то он может быть быстро исправлен или заменен в службе 1

хостинга. Напротив, РС или игровая консоль в доме или офисе должны быть жестким автономным аппаратом, который должен выдерживать приемлемый износ из-за ударов или падений, требует корпуса, имеет, по меньшей мере, один накопитель на магнитных дисках, должен выдерживать неблагоприятные окружающие условия (например, будучи втиснутым в перегретый аудио- и видеокорпус вместе с другим оборудованием), требует гарантийного обслуживания, должен быть упакован и отправлен и продается розничным торговцем, который, вероятно, получает розничную наценку. Кроме того, РС или игровая консоль должны быть выполнены в соответствии с пиковыми характеристиками ожидаемых приложения или игры, требующих наиболее большого объема вычислений, которые будут использоваться в некоторый момент в будущем, даже несмотря на то что приложение или игры с более низкими характеристиками (или секции игр или приложений) могут исполняться большую часть времени. И если произойдет отказ РС или консоли, то их починка является дорогостоящим процессом, отнимающим много времени (с неблагоприятным влиянием на производителя, пользователя и разработчика программного обеспечения).

Соответственно, с учетом того, что система, изображенная на рис. 2.4, обеспечивает пользователю опыт использования, сопоставимый с опытом использования локального вычислительного ресурса, для пользователя в доме, офисе или школе, чтобы испытывать данный уровень вычислительных возможностей, намного дешевле обеспечивать эти вычислительные возможности посредством архитектуры, изображенной на рис. 2.4.

Ж 7 Устранение необходимости модернизации

Кроме того, пользователи больше не должны заботиться о модернизации РС и/или консолей для игры в новые игры или управления новыми приложениями с более высокими характеристиками. Любая игра или приложения в службе 1 хостинга независимо от того, какой тип сервера 4 требуется для этой игры или приложений, доступны пользователю, и все игры и приложения исполняются почти немедленно и должным образом с последними обновлениями и

исправлениями ошибок. Безусловно, после того как пользователь начинает использовать службу 1 хостинга, пользователь, вероятно, обнаружит, что игры и приложения продолжают обеспечивать лучший опыт использования (например, посредством обновлений и/или исправлений ошибок, и может оказаться так, что пользователь обнаружит год спустя, что новая игра или приложение стали доступны в службе 1, которая использует технологию вычислений (например, более высокопроизводительное GPU), которая даже не существовала за год до этого, соответственно, за год до этого было невозможно, чтобы пользователь купил упомянутую технологию, которая могла бы исполнять игру или приложения год спустя. Так как вычислительный ресурс, на котором ведется игра или который исполняет приложение, является невидимым для пользователя, аппаратное обеспечение пользователя будет ”модернизировано” без пользователя, даже если ему известно о модернизации.

Ж 8 Устранение необходимости резервных копий

Другой важной проблемой для пользователей в компаниях, школах и домах являются резервные копии. Информация, хранящаяся в локальном РС или видеоигровой консоли, может быть утрачена в случае сбоя диска или случайного стирания. Существует много доступных приложений, которые обеспечивают ручное или автоматическое резервное копирование для РС, и состояние игровой консоли может выгружаться на онлайн-сервер для резервирования, но локальные резервные копии обычно копируются на другой локальный диск, который должен храниться в каком-нибудь безопасном месте и быть организован, и резервирование в онлайн-службы часто ограничено из-за низкой скорости восходящего потока данных, доступной через типичные низкочастотные подключения к Internet. Посредством службы 210 хостинга по рис. 2.4 данные, которые хранятся в дисковых массивах 7 типа RAID, могут быть сконфигурированы с использованием таких способов конфигурации RAID предшествующего уровня техники, известных специалистам в данной области техники, образом, что, если произойдет сбой диска, то данные не будут утрачены и техник в серверном центре, в котором находится поврежденный диск, будет

уведомлен и тогда (он) заменит диск, который после этого будет автоматически обновлен так, чтобы дисковый массив типа RAID снова был устойчивым к сбоям. Кроме того, так как все накопители на магнитных дисках находятся рядом друг с другом с (обеспечением) быстрых локальных сетей между ними посредством SAN 403, то в серверном центре нетрудно организовать, чтобы регулярно создавались резервные копии всех дисковых систем на вспомогательном запоминающем устройстве, которое может храниться в серверном

центре или перемещено за его пределы. С точки зрения пользователей службы 210 хостинга, их данные просто все время защищены и они никогда не должны думать о резервных копиях.

Ж 9 Доступ к демонстрационным версиям

Пользователям часто требуется испытать игры или приложения перед их покупкой. Как описано ранее, существуют средства предшествующего уровня техники для испытания демонстрационной версии анимацией службы 1 хостинга пользователям легко и удобно испытать демонстрационные версии. На самом деле, все, что пользователь делает, - это выбирает демонстрационную версию посредством интерфейса пользователя (например, интерфейса, описанного ниже) и испытывает ее.

Демонстрационная версия загружается на сервер 4 подходящий для демонстрационной версии, почти немедленно, и она исполняется подобно любой другой игре или приложению. Независимо от того, требуется ли для демонстрационной версии очень высокопроизводительный сервер 402 или низкопроизводительный сервер 4, и независимо от того, какого типа домашний или офисный клиент 7 использует пользователь, с точки зрения пользователя, демонстрационная версия просто работает. Издатель программного обеспечения демонстрационной версии или игры, или приложения может регулировать то, какую именно демонстрационную версию предоставить пользователю для испытания и на какое время, и, конечно, эта

демонстрационная версия может включать в себя элементы интерфейса пользователя, которые предоставляют пользователю возможность получить доступ к полной версии демонстрируемой игры или приложения.

Так как демонстрационные версии, вероятно, будут предложены за меньшую цену или бесплатно, то некоторые пользователи могут попытаться использовать демонстрационные версии повторно (особенно демонстрационные версии игр, повторная игра в которые может являться развлечением). Служба 1 хостинга может применять различные способы для ограничения использования демонстрационной версии для данного пользователя. Самым прямым подходом является установка ID пользователя для каждого пользователя и ограничение количества испытаний демонстрационной версии, обеспечиваемого для данного ID пользователя. Пользователь, однако, может установить множество ID пользователя, особенно если они являются бесплатными. Одним способом, направленным на решение этой проблемы является ограничение количества испытаний демонстрационной версии, обеспечиваемого для данного клиента 7. Если клиентом является автономное устройство, то у этого устройства существует серийный номер, и служба 1 хостинга может ограничить количество раз, которое клиент может получать доступ к демонстрационной версии с этим серийным номером. Если клиент 7 исполняется как программное обеспечение на PC или другом устройстве, то серийный номер может быть назначен службой 1 хостинга и сохранен на PC и использоваться для ограничения использования демонстрационной версии, но с учетом того, что PC могут быть перепрограммированы пользователями, и серийный номер может быть стерт или изменен, другим возможным вариантом для службы 1 хостинга является ведение учета адреса протокола управления доступом к среде (MAC) сетевого адаптера PC (и/или другие машинно-зависимые идентификаторы, например серийные номера накопителей на жестких дисках и т.д.) и ограничение использования демонстрационной версии для него. С учетом того что MAC-адреса сетевых адаптеров могут быть изменены, однако это не является надежным способом. Еще одним подходом является ограничение

количества испытаний демонстрационной версии для данного IP-адреса. Несмотря на то что IP-адреса могут периодически переназначаться провайдерами DSL и кабельной модемной связи, на практике это не происходит очень часто, и если можно определить (например, связавшись с ISP) то, что данный IP находится в блоке IP-адресов для доступов к DSL или кабельной модемной связи, связанных с жилыми домами, то для данного дома, как правило, может быть установлено небольшое количество использований демонстрационной версии. Кроме того, может быть множество устройств в доме за маршрутизатором NAT, совместно использующих идентичный IP-адрес, но, как правило, в жилом окружении существует ограниченное количество таких устройств. Если IP-адрес находится в блоке, обслуживаемом компанией, то для компании может быть установлено большее количество демонстрационных версий. Но, в конечном счете, комбинация всех ранее упомянутых подходов является лучшим способом ограничения количества демонстрационных версий для РС. Несмотря на то что не может существовать надежного способа, посредством которого полного решимости и технически искусного пользователя можно ограничить в количестве повторных использований демонстрационных версий, с созданием большого количества преград можно создать такое достаточное сдерживающее средство, что нарушение режима эксплуатации системы демонстрационной версии станет нецелесообразным для большинства пользователей РС и, скорее, они будут использовать демонстрационные версии так, как подразумевалось испытание новых игр и приложений.

Ж 10 Преимущества для школ, компаний и других учреждений

Значительные преимущества, в частности, получают компании, школы и другие учреждения, которые используют систему, изображенную на рис. 2.4. В компаниях и школах существуют существенные затраты, связанные с установкой, поддержанием в рабочем состоянии и модернизацией РС, особенно когда дело доходит до РС для исполнения приложений с высокими характеристиками, например, Maya. Как утверждалось ранее, РС, как правило, используются только часть времени в неделю, и как в доме, стоимость РС с

данным уровнем рабочих характеристик является намного выше в окружающей обстановке школы или офиса, чем в окружающей обстановке серверного центра.

В случае больших компаний или школ (например, большие университеты), может оказаться целесообразным, чтобы отделы ИТ таких организаций устанавливали серверные центры и поддерживали компьютеры, которые получают доступ удаленно через соединения уровня LAN. Существует несколько решений для удаленного доступа компьютеров по LAN или через частное широкополосное соединение между офисами. Например, посредством терминального сервера Microsoft Windows Terminal Server или через приложения удаленного доступа к рабочему столу

компьютера (virtual network computing), подобные VNC, от RealVNC, Ltd., или через средство тонкого клиента от Sun Microsystems пользователи могут получать удаленный доступ к PC или серверам с диапазоном качества по времени отклика графических устройств и практическому опыту пользователя. Кроме того, такие самоуправляемые серверные центры обычно закрепляются за одной компанией или школой и, по существу, не могут использовать преимущества совмещения использования, которое является возможным, когда различные приложения (например, развлекательные приложения и бизнес-приложения) используют идентичные вычислительные ресурсы в разное время недели. Соответственно, у многих компаний и школ отсутствует масштаб, ресурсы или специальные знания для установки серверного центра самостоятельно, в котором существует сетевое соединение со скоростью LAN с каждым пользователем. На самом деле, большой процент школ и компаний имеет идентичные подключения к Internet (например, DSL, кабельные модемы), как и в домах.

Однако в таких организациях может, тем не менее, существовать потребность очень высокопроизводительных вычислениях или постоянно, или периодически. Система, изображенная на рис. 2.4, очень хорошо подходит для таких организаций. Упомянутым организациям требуется не что иное, как сетевое соединение вида,

идентичного тем, которые предлагаются для домов (например, DSL, кабельные модемы) и, как правило, являются очень экономичными. Они могут или использовать экономичные PC, как клиент 14, или совершенно обходиться без PC и использовать экономичные специализированные устройства, которые просто реализуют логику 13 управляющего сигнала и восстановление 12 сжатого видео с малым временем ожидания. Эти признаки, в частности, привлекательны для школ, у которых могут быть проблемы с воровством PC или повреждением компонентов, требующих осторожного обращения, внутри PC.

Такая компоновка решает несколько проблем для таких организаций (и многие из этих преимуществ также совместно используются домашними пользователями, выполняющими универсальные вычисления. Например, эксплуатационные расходы (которые, в конечном счете, должны возвращаться в некоторой форме к пользователям, чтобы иметь жизнеспособный бизнес) могут быть намного меньше, потому что (а) вычислительные ресурсы совместно используются с другими приложениями, которые имеют разные пиковые периоды использования в течение недели, (b) организации могут получать доступ к (и нести расходы по) высокопроизводительным вычислительным ресурсам только тогда, когда требуется, организации не должны обеспечивать ресурсы для создания резервных копий или каким-либо иным способом поддерживать высокопроизводительные вычислительные ресурсы.