Roland Austinat, Philipp Fechteler, Hartmut Gieselmann

Über den Wolken

Wie Cloud Gaming den Spielemarkt revolutioniert

Künftig sollen Spiele gar nicht mehr installiert werden. Sie laufen auf riesigen Server-Farmen im Internet, die nur noch Bild und Ton zum Spieler streamen. Eine ganze Wolkenfront von Cloud-Diensten kommt auf Europa zu und nimmt Herstellern von Grafikkarten und aufgemotzten Gaming-PCs die Sonne.

Sie gehören zu den Dinosauriern: Klobige Spiele-PCs mit sündhaft teuren SLI-Grafikkarten. Für derlei laute und energiehungrige Monster interessiert sich mittlerweile nur eine kleine, finanzkräftige Gruppe von Spiele-Fans. Die breite Masse greift zum Notebook oder spielt am Büro-PC mit Onboard-Grafik. Auf der anderen Seite sind in den vergangenen Jahren die Kosten der Spiele-Großproduktionen geradezu explodiert und erreichen zuweilen dreistellige Millionenbeträge. Je weiter die Schere zwischen technischem Aufwand der Produktion und bei den Kunden vorhandener Rechenleistung aufgeht, desto dringender benötigt die Spieleindustrie Lösungen, die diese Abhängigkeiten entkoppeln. Der Effekt wird sich weiter beschleunigen, je mehr schlanke Endgeräte wie Netbooks, iPad und andere Tablets in Mode kommen. Bei diesen zählen vor allem günstige Produktionskosten, Energie-Effizienz und Akkulaufzeit – dafür sind rechenintensive Spiele Gift.

Zudem hat die Bedrohung durch illegale Kopien zum Teil existenzielle Ausmaße angenommen. In den USA klagen Hersteller darüber hinaus über den Gebrauchtmarkt, aus dem Ladenketten wie Gamestop den Löwenanteil ihre Umsätze erzielen, ohne dass die Hersteller daran beteiligt werden.

Um die Vertriebskosten zu senken, hat sich neben dem Retail-Geschäft mit Discs inzwischen der Download-Vertrieb etabliert. Bleiben einem amerikanischen Hersteller etwa 45 Prozent des Verkaufspreises einer Retail-Disc, so ließe sich dieser Anteil nach Versprechen der Download-Anbieter auf bis zu 75 Prozent steigern.

Doch selbst der Download-Vertrieb verlangt noch immer, dass der Spieler zunächst die Dateien lädt und installiert, bevor er spielen kann. Es gibt Anbieter wie Instant Action und Gametap, die immerhin die Lade- und Installationszeit verkürzen, indem sie dem Spieler erlauben, bereits während des Downloads loszuspielen. Aber auch hier muss der Spieler nach wie vor leistungsfähige Hardware vorhalten.

Als nächster logischer Schritt soll deshalb die Installation komplett wegfallen. Beim Cloud Gaming empfängt der Client nur noch die Audio- und Video-Ausgabe eines Spiels, das auf einem Server im Netz läuft, an den die Befehle von Maus, Tastatur und Gamepad übertragen werden. Hierfür muss der Client lediglich den Video- und Audio-Stream dekodieren. Die Hauptbelastung wird den Servern aufgebürdet und dem Netz, das sehr große Datenströme mit möglichst geringer Laufzeit übertragen muss.

Wie groß die Verzögerung sein darf, um das Spielerlebnis nicht zu beeinträchtigen, hängt stark vom jeweiligen Genre ab. Am Fraunhofer Heinrich-Hertz-Institut in Berlin peilen die Entwickler für Cloud-basierte Umsetzungen bei Ego-Shootern und anderen First-Person-Spielen Verzögerungen unter 0,1 Sekunden

an. In der Third-Person-Perspektive von Sportspielen sei eine halbe Sekunde erträglich, während Strategie-Spiele und andere sogenannte omnipräsente Titel sogar noch ein bis zwei Sekunden Latenz verkraften würden. Für ambitionierte Starcraft-Spieler wären solche Verzögerungen allerdings inakzeptabel.

OnLive

Von den rund einem halben Dutzend Cloud-Gaming-Diensten, die derzeit in den neuen Markt drängen, hat der US-Anbieter OnLive am meisten für Furore gesorgt. In den USA ist er am 17. Juni offiziell gestartet. OnLive-Chef Steve Perlman verspricht, neueste PC-Spiele in HD-Auflösung mit allen Schikanen selbst auf schwachbrüstigen PCs und Settop-Boxen laufen zu lassen. Was sich zunächst nach Quacksalberei anhört, hat jedoch Hand und Fuß. Den Wert des Startup-Unternehmens schätzen Finanzexperten inzwischen auf rund eine Milliarde US-Dollar.

Über den eigenen PC steuert der Abonnent des Dienstes ein Spiel und schickt seine Eingabebefehle in eines von fünf OnLive-Datenzentren in den USA. Die Spielgrafik wird dort komprimiert und über das Internet zurück zum Nutzer gestreamt. Ein spezieller Videocodec und optimierte Server-Hardware wie auch neue Netzwerk-Routing-Algorithmen sollen geringe Laufzeiten von unter 80 ms garantieren, die selbst für schnelle Action-Spiele genügen. Ein zweiter, stärker komprimierter Grafik-Stream erlaubt anderen Kunden, Spiele-Sessions ihrer Freunde zu beobachten. Auch sogenannte „Brag Clips", serverseitig gespeicherte, zehn Sekunden lange Videoschnipsel mit besonders gelungenen Aktionen, nutzen diesen zweiten Stream.

Über die genauen Hardware-Spezifikationen der Server schweigt sich OnLive hartnäckig aus. Bekannt ist lediglich, dass in mehrjähriger Zusammenarbeit mit dem US-Hersteller Dell entwickelte Systeme zum Einsatz kommen. Diese verfügen im Unterschied zu normalen Servern über Grafikchips der Marktführer AMD und Nvidia. Außerdem sind sie mit einem von OnLive entwickelten Chip bestückt, der die Kompression der Grafik übernimmt. Grafisch aufwendige Titel wie „Mafia 2" nutzen nach Angaben von Steve Perlman eine GPU und damit einen Server pro Spieler, Indie-Hits wie „World of Goo" laufen virtualisiert in mehreren Ausführungen pro System. OnLive plant, einen Teil der Server im Halbjahresrhythmus aufzurüsten, um für stetig steigende Hardware-Anforderungen gewappnet zu sein. Die heute im Single-GPU-Modus laufenden Rechner könnten dann in ein paar Jahren primär als Virtualisierungsmaschinen für zukünftige Spiele mit eher einfacher Grafik genutzt werden.

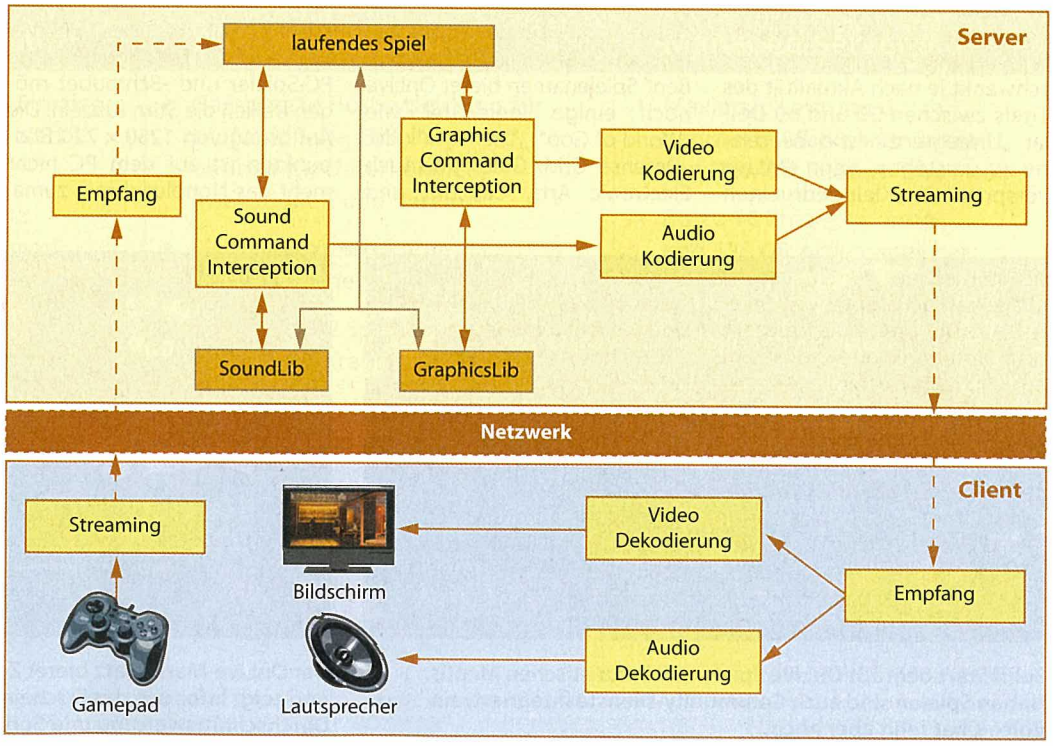
Tschüß 3D-Grafikkarte

Die Hardware-Anforderungen auf der Client-Seite sind moderat. OnLive empfiehlt derzeit lediglich eine Dual-Core-CPU und eins der Betriebssysteme Windows 7 oder Vista in den 32- oder 64-Bit-Fassungen, Windows XP (32-Bit-Version) und Mac OS X 10.6 (Snow Leopard). Für das derzeit unterstützte HD-Format 720p muss das heimische System eine Auflösung von 1280 × 720 Bildpunkten darstellen können und über eine 5-MBit-Leitung mit dem Internet verbunden sein. Nachdem experimentierfreudige OnLive-Kunden von erfolgreichen Versuchen mit Asus Netbooks (Atom-CPU) und lediglich 1 GByte RAM berichtet hatten, verkündete Steve Perlman Anfang September zur Penny Arcade Expo, dass der Dienst jetzt auch offiziell mit Netbooks, einer 3 MBit/s schnellen Verbindung sowie mit Mac OS X 10.5.8 (Leopard) funktioniere. Für größere Monitore gelte jedoch die ursprüngliche Empfehlung.

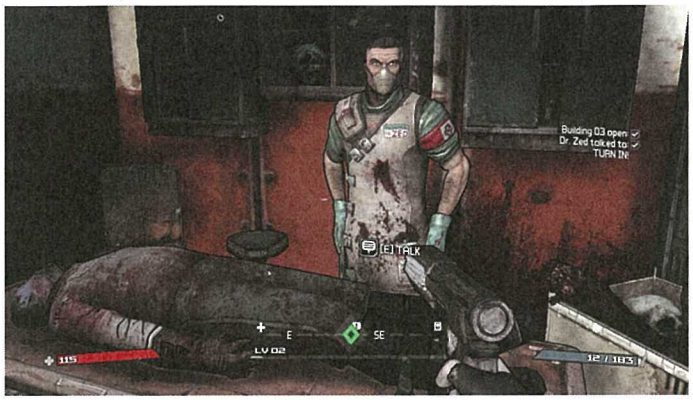
Wir testeten den Dienst mit einem knapp zwei Jahre alten MacBook White (OS X 10.6.4, 4 GByte RAM, Core 2 Duo T8100) sowie einem PC mit Conroe-Kern (E6600), Windows XP und 3 GByte RAM. Beide Rechner waren über einen 12-MBit-Zugang des US-Anbieters Comcast mit dem Internet verbunden. Während der Testphase unterstützte der Dienst noch kein WLAN. Weil der Client prüft, ob eine LAN-Verbindung tatsächlich vorhanden ist, lässt sich das schnurlose Spielvergnügen nur über eine Windows-Netzwerkbrücke erzwingen. Für einen solchen Umweg gibt OnLive aber keine Garantie, dass die Spiele flüssig laufen – es kann funktionieren, muss es aber nicht. Mit Erscheinen dieses Heftes soll eine offizielle WiFi-Beta-Phase bereits begonnen haben. Etwaige 3D-Grafikkarten sowie die Größe des Arbeitsspeichers spielen nach wie vor keine Rolle.

Einloggen, losspielen

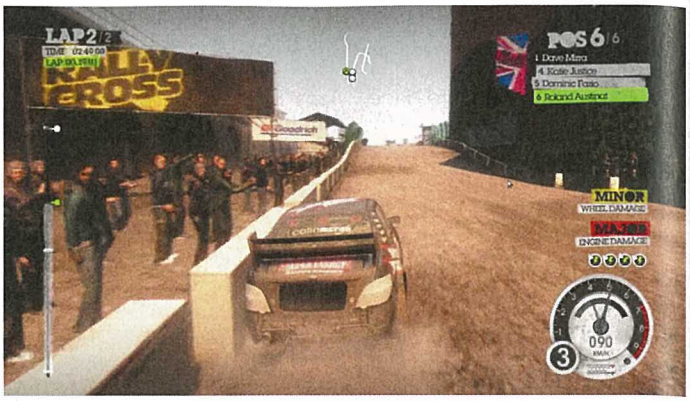
Die Kommunikation mit dem am schnellsten erreichbaren OnLive-Server übernimmt ein rund 500 KByte großes Programm, das Abonnenten auf der OnLive-Webseite herunterladen und installieren. Nach dem Aufruf beginnt es damit, den per Real-Time Transport Protocol (RTP) übertragenen Datenstrom zu dekodieren und zeigt nach kurzem Intro das futuristische Hauptmenü, umrahmt von Fenstern, in denen ohne Unterlass Spielsequenzen zu sehen sind. Im Menüpunkt „Arena" darf man anderen Mitgliedern über die Schulter schauen, so sie diese die Funktion freigeschaltet haben. In den „Brag Clips" versammeln sich die bereits zuvor erwähnten Angeber-Videos. Der Menüpunkt „Friends" erlaubt es, den eigenen Freunden beim Spielen zuzuschauen. Dazu gibt es die Möglichkeit, 140 Zeichen lange Nachrichten an Freunde zu schreiben sowie einige Grafik- und Ton-Op-



Das Cloud-Konzept am Beispiel von Games@Large: Der Client leitet die Eingabebefehle des Gamepad über das Netzwerk an den Server weiter, wo das eigentliche Spiel läuft und die Audio- und Video-Ausgabe komprimiert über das Netz wieder an den Client zurückgeschickt wird.



In ruhigeren Spielmomenten wie hier im Action-Rollenspiel „Borderlands" unterscheidet sich die Grafik der OnLive-Spiele nicht wesentlich von lokal installierten Titeln.



In schnellen Rennspielen treten bei OnLive zuweilen Kompressionsartefakte auf, wie in der 30-Minuten-Demo von „DiRT 2". Das Spiel läuft aber flüssig weiter.

tionen festzulegen. In den Spielen werden sämtliche Setup-Einstellungen vom Server vorgegeben, die örtlichen Änderungsmöglichkeiten beschränken sich auf den Helligkeitswert des Bildes sowie das Ein- und Ausschalten des Tons. Ein integrierter Sprach-Chat ist geplant.

Was kostet?

Die derzeit 28 verfügbaren Titel lassen sich bis auf Ubisofts neuesten „Prince of Persia" beliebig oft je eine halbe Stunde lang testen. Dabei werden allerdings keine Fortschritte gespeichert. Kaufwillige Abonnenten wählen dann bei Interesse ähnlich wie in einer Videothek zwischen einem Drei- oder Fünftagesticket für fünf bis neun Dollar oder erwerben ein unbegrenztes Nutzungsrecht. Dessen Preis schwankt je nach Aktualität des Titels zwischen 20 und 60 Dollar. „Unbegrenzt" ist dabei relativ zu verstehen, denn OnLive verspricht im Kleingedruckten derzeit nur den Zugang für „mindestens drei Jahre nach dem Erstverkaufstag" des jeweiligen Spiels – sicher ausreichend, um bis zum Abspann zu gelangen, ein Archiv von Spieleklassikern lässt sich damit aber nicht aufbauen.

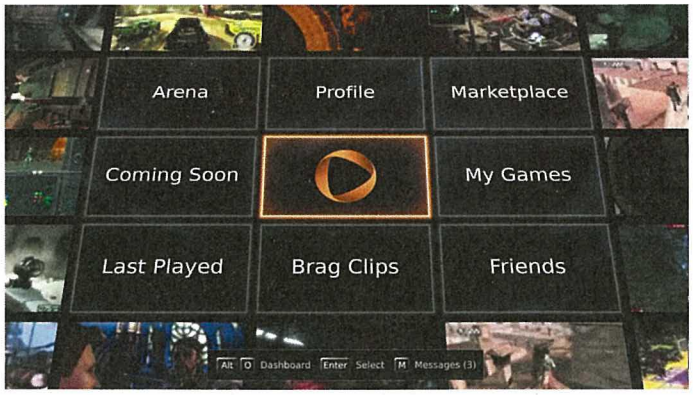
Immerhin: Die Liste der Partner-Publisher liest sich wie das „Who is Who" der Branche. Zudem vertrauen die Anbieter dem Dienst meist große Namen an, zum Beispiel „Assassin's Creed 2", „Splinter Cell: Conviction" (beide Ubisoft), „Just Cause 2" (Square Enix), „Lego Harry Potter", „Batman: Arkham Asylum" (beide Warner Brothers Interactive), „Red Faction: Guerilla" (THQ), „Mafia 2", „Borderlands" (beide Take Two Interactive), „Unreal Tournament 3" (Epic Games) und „Puzzle Chronicles" (Konami). Zusätzlich zu den „großen" Spielenamen bietet OnLive noch einige Indie-Titel wie „World of Goo", „The Maw" oder „Defense Grid: Gold" an. Auch Electronic Arts, Capcom und Atari wollen Spiele via OnLive veröffentlichen. Einzig Activision Blizzard hält sich derzeit noch vornehm zurück.

Zu den Leihgebühren der einzelnen Spiele kommt noch eine monatliche Grundgebühr für den Dienst, die ab dem zweiten Jahr 15 US-Dollar betragen soll. Die Kosten der ersten zwölf Monate sowie zwei Drittel der künftigen Gebühren übernimmt derzeit das Partnerunternehmen AT&T. Trotz des offiziellen Starts lassen die Entwickler noch nicht jeden Interessenten sofort auf die Server los, sondern bestimmen, wer ab wann auf den Dienst zugreifen kann.

Flink, doch pixelig

Auch wenn Skeptiker große Zweifel haben: Das OnLive-Spielgefühl ist gut. Eingefleischte PC-Spieler und -Schrauber mögen freilich die Stirn runzeln: Die Auflösung von 1280 × 720 Bildpunkten ist auf dem PC nicht mehr das Nonplusultra – zumal je nach Spielgenre und -geschwindigkeit teils heftige Kompressionsartefakte sichtbar werden. Das fällt besonders stark bei Rennspielen („DiRT 2") und 3D-Shootern („Unreal Tournament 3", „FEAR 2") auf. Allerdings ist der subjektive Grafikeindruck während des Spielgeschehens besser als das, was die teils arg verpixelten Screenshots abbilden – der ausreichend hohen Bildfrequenz von durchschnittlich 35 bis 40 FPS der OnLive-Server sei's gedankt. Action-Adventures („Splinter Cell: Conviction") beziehungsweise -Rollenspiele („Borderlands") machen einen besseren Eindruck, wobei auch hier auffällt, dass die komprimierte Grafik etwas farbärmer als die eines lokal installierten Spiels erscheint. Dies deutet auf einen reduzierten Farbraum bei der Videokodierung hin (Farb-Subsampling). Die Full-HD-Auflösung 1080p ist noch im Planungsstadium.

Latenzprobleme gab es während unseres Tests so gut wie



Beim Start begrüßt OnLive Spieler mit futuristischen Menüs. Neben Spielen sind auch Community-Dienste integriert, ein Voice-Chat fehlt aber noch.

Der OnLive-Marktplatz bietet Zugriff auf alle verfügbaren Spiele und zeigt Infos wie das Erscheinungsdatum, die Metacritic-Durchschnittswertung und Sonderangebote.

Einfluss der Netzwerkstruktur

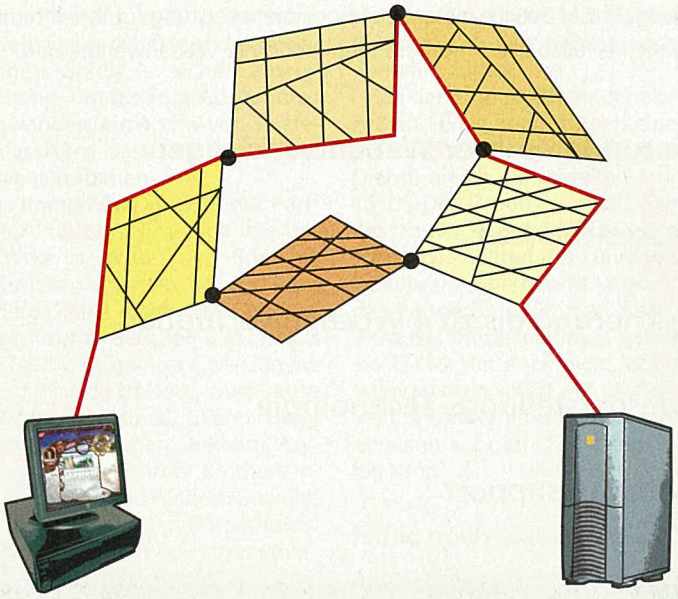
Eine Netzwerkverbindungen kann auf zweierlei Weise schnell oder langsam sein: Die Bandbreite gibt an, wie viele Bytes pro Sekunde durch die Leitung gehen. Die Latenz sagt, wie lange die einzelnen Pakete unterwegs sind. Klassische Online-Spiele sind auf möglichst geringes Datenvolumen optimiert. Die Game-Engines der Teilnehmer tauschen nur kleine Päckchen mit Daten wie Koordinaten, Schussrichtung, Laufgeschwindigkeit und derlei aus. Daher spielt die Bandbreite keine wesentliche Rolle; nur die Latenz ist wichtig.

Zur Latenz tragen alle Leitungen, Funkstrecken und die an deren Enden stehenden Switches, Router, Access Points und Modems bei. In Kupfer- und Glasfaserleitungen rasen die Signale mit rund zwei Drittel der Vakuumlichtgeschwindigkeit, sodass allein deshalb 200 Kilometer Kabel rund 1 ms zusätzliche Latenz bedeuten. Etwas größer ist der Einfluss der Bandbreite auf die Laufzeit. So braucht beispielsweise ein volles Ethernet-Paket von 1500 Bytes rechnerisch mindestens 3 ms, um sich durch eine 4 MBit/s langsame DSL-Leitung zu quälen. Doch die Latenzen auf echten Leitungen liegen oft viel höher. Das liegt an den zusätzlichen Verwaltungsdaten, die mit übertragen werden, sowie an der verwendeten Kodierung und der Bearbeitungszeit in den Geräten auf beiden Seiten der Leitung. So tragen übliche DSL-Leitungen meist zwischen 20 und 40 ms zur Latenz bei.

Oft sind der Client und der Spiele-Server nicht über denselben Provider ans Internet angebunden. Damit die Datenpakete aus dem einen Provider-Netz ins andere gelangen, müssen sie entweder durch einen „Peering Point", in dem die beiden direkt verbunden sind, oder durch weitere Netze anderer Carrier. Wie und wo sie ihre Netze zusammenschalten, verhandeln die Carrier untereinander. Welchen Weg ein Paket nimmt, entscheiden die Buchhalter der Carrier mit. Denn manche verlangen für den durchgeleiteten Verkehr Geld, sodass die Router eventuell darauf eingestellt werden, statt der Route mit der geringsten Latenz die billigste zu wählen. Ideal ist es, wenn Client und Server im Netzwerk desselben Providers stehen. Denn dann ist nicht nur die Zahl der durchquerten Netzwerk-Abschnitte (Hops) gering, sondern auch das ganze Netz unter der Kontrolle eines Anbieters.

Auf diesem Gedanken beruhen Content Distribution Networks (CDN). Große Inhalte-Anbieter verteilen ihre Server in die Netze der Provider und bezahlen sie für deren Betrieb. Deshalb kommen beispielsweise die Windows-Updates von ganz unterschiedlichen Servern, je nachdem, bei welchem Provider man Kunde ist. Dies ist auch ein Kerngedanke des Cloud-Computing. Für den Zugriff muss der Nutzer gar nicht wissen, wo der Server steht. Der Anbieter kümmert sich darum, dass es jeweils der günstigst gelegene ist.

Aber nicht nur die Netzstrukturen, sondern auch die Knotenpunkte beeinflussen die Laufzeit. Die Verweilzeit eines Pakets in einem Router ist nicht konstant, sondern hängt vom sonstigen Datenverkehr ab. Die meisten Heimrouter schicken die Pakete in der Reihenfolge weiter, in der sie sie bekommen. In den Routern innerhalb der Provider-Netze kommen aber fast immer viele Pakete quasi gleichzeitig an. Die Geräte entscheiden dann anhand der Paketeigenschaften, welches Paket sie zuerst behandeln und welche sie noch etwas aufhalten. Bei dieser Quality-of-Service-Einteilung werden zum Beispiel Voice-over-IP-Daten bevorzugt, damit es nicht zu Aussetzern in Gesprächen kommt. Auch die ICMP-Pakete werden oft schneller befördert, da sie auch zur Signalisierung von Fehlern im Netzwerk dienen. Das Programm Ping benutzt solche ICMP-Pakete.



IP-Pakete durchlaufen normalerweise mehrere Netze mit unterschiedlichen Eigenschaften und Priorisierungsregeln. Daher ist der räumliche Abstand zwischen Client und Server für die Latenz weniger aussagekräftig als die Netzwerkstruktur.

Allerdings steht es jedem Carrier frei, welche Daten er mit welcher Priorität transportiert. Da der Datenstrom eines Spiels normalerweise durch mehrere Carrier-Netze läuft, kann der Betreiber des Spiele-Servers nicht allein dafür sorgen, dass seine Datenströme überall Vorfahrt bekommen.

Für die gesamte Latenz ist also die räumliche Entfernung zwischen Client und Server weniger wichtig als die Struktur des Netzwerkes dazwischen. Generell ist statt der räumlichen Nähe wichtiger, dass möglichst wenige Hops auf dem Weg liegen. Ob ein Server in diesem Netzwerksinn „in der Nähe" liegt, verrät das Programm traceroute (unter Windows tracert), das die Hops anzeigt.

Bandbreite

Klassische Online-Spiele sind auf möglichst kleine Datenmengen optimiert, unter anderem, weil kleinere Pakete oft mit geringerer Latenz befördert werden. Doch beim Cloud Gaming fließen prinzipiell umfangreiche Video-Daten vom Server zum Client. Außer geringer Latenz ist also auch eine höhere und gleichmäßige Bandbreite gefordert. Die OnLive-Server senden rund 700 KByte/s, was mit einer üblichen 6-MBit/s-Leitung leicht abzufeiern ist. Auch aktuelle WLANs stellen genug Bandbreite zur Verfügung. Allerdings schwankt sie stark, weil sich alle Stationen das Medium Funk teilen und ein Datenpaket eines anderen PCs den nächsten Frame des Cloud-Spiels ausbremsen könnten. Dagegen hilft auf WLAN-Ebene ein Wi-Fi Multimedia (WMM) genanntes Verfahren für die nötige Quality-of-Service im Funknetz.

Außerdem hilft das Übertragungsverfahren Real Time Transport Protocol (RTP). Denn anders als bei einer E-Mail spielt es beim Game-Stream eine kleinere Rolle, dass jedes abgeschickte Paket auch wirklich ankommt. Wichtiger ist, dass es rechtzeitig ankommt. Ein verzögerter Frame, der vom nächsten Bild des Videos im Netz überholt wurde, wird ohnehin nicht mehr angezeigt. Daher setzt RTP auf das effiziente Transportprotokoll UDP auf, bei dem Quittungen und erneute Übertragungen entfallen. Wenn das Netzwerk stockt oder Pakete verloren gehen, läuft das Video trotzdem weiter, notfalls mit geringerer Qualität. (je)

GoD hat es eilig

Anders als beim herkömmlichen Live-Videostreaming kommt es beim Gaming on Demand (GoD) darauf an, den jeweiligen Bildinhalt in Echtzeit und aufgrund der Internet-Bandbreitenbeschränkungen möglichst effizient zu kodieren. Da der Spieler den Bildinhalt durch seine Eingaben steuert, muss all dies auch mit einer möglichst geringen zeitlichen Verzögerung (Latenz) geschehen, ähnlich wie bei Videokonferenzen. Naheliegend ist die Verwendung des Video-Industriestandard MPEG-4 AVC (H.264) zur Videokompression. Er zeichnet sich durch eine hervorragende Kodiereffizienz aus. Viele der Optimierungen lassen sich jedoch nicht für den Echtzeitbetrieb nutzen, weil das En- und Decoding zu viel Zeit in Anspruch nimmt. Wie der Hauptentwickler des freien H.264-Encoders x264, Jason Garrett-Glaser, ausführt, setzt sich die Gesamtlatenz beim H.264-Encoding (inklusive Decoder-Pufferung) aus folgenden Komponenten zusammen: Zahl der B-Frames (jeweils ein Bild Latenz), Zahl der verwendeten Encoder-Threads (pro Thread ein Bild Verzögerung), Ratecontrol- und Sync-Lookahead (jeweils in Frames), VBV-Puffergröße (in Sekunden) sowie die eigentliche Kodierzeit. Noch im Oktober 2009 war x264 für Echtzeitbetrieb, wie ihn GoD erfordert, nicht brauchbar. In knapp drei Monaten optimierte Garrett-Glaser den Encoder, der nun auf einem Core-i7-Prozessor eine Gesamtlatenz von unter zehn Millisekunden bei einer Auflösung von 800 × 600 erreichen kann.

Um dies zu ermöglichen, beschränkt man sich beispielsweise auf Prädiktionen aus dem vorangegangenen Frame; zeitraubende Multi-Frame-Referenzen und bidirektional-prädizierte Frames (B-Frames) bleiben außen vor. Auch auf Key-Frames, die als Anker- und Einsprungpunkte dienen, verzichtet man. Hierbei stehen weniger deren (ohnehin nicht allzu große) Komplexität, sondern vielmehr die beträchtlichen resultierenden Datenraten im Vordergrund. Durch geschicktes Intra-Kodieren von Frame-Teilen kann man die periodisch anfallenden Spitzen in den Bitraten komplett vermeiden (Periodic Intra Refresh); mehr Details hierzu finden sich unter http://x264dev.multimedia.cx/?p=249.

Beim GoD-Projekt Games@Large nutzen die Forscher des Fraunhofer HHI unter anderem die bei vielen 3D-Spielen verwendete Skybox-Technik zur optimierten Videokodierung – hier im Falle des OpenGL-Spiels „Extreme Tux Racer": Original-Ausgabe (links), Skybox-Bereich und z-Buffer (Mitte), Differenzbild (rechts)

Im Rahmen des Projekts Games @Large haben die Forscher des Heinrich-Hertz-Instituts (HHI) sich weitere Optimierungen überlegt [1]. So nutzt ein Ansatz aus, dass bei 3D-Spielen meist die Skybox-Rendering-Technik (auch Skydome, Skysphere genannt) verwendet wird. Hierbei rendert das Spiel für jedes Szenen-Update zunächst den Hintergrund mit Himmel, Bergen oder Panorama-Ansichten. Da dies bei deaktiviertem Z-Puffer geschieht, lassen sich die der Skybox zugeordneten Pixel eindeutig erkennen. Bei der H.264-Kodierung kann man die komplett in der Skybox liegenden Makroblock-Partitionen einem gemeinsamen Bewegungsvektor zuordnen und so den Rechenaufwand minimieren. Des Weiteren vereinfacht die HHI-Lösung die zeitraubende Bewegungssuche (Motion Estimation), indem sie Transformationsmatrizen der Zeichen-Befehle abfängt und für die Bewegungsvektorbestimmung zweckentfremdet. Nur wenn diese Methode unzureichende Ergebnisse liefert, wird auf die H.264-Bewegungssuche zurückgegriffen, um zu vermeiden, dass die Kodiereffizienz abnimmt. Zur weiteren Beschleunigung lassen sich all diese Berechnungen als Shader-Programm auf der GPU ausführen.

Grafik-Streaming

Statt einen Videostrom zu übertragen, könnte man für Endgeräte mit entsprechender Grafikleistung die Grafikbefehle des Spiels abfangen, zum Client übertragen und ausgeben [2]. Der Vorteil dieser Technik ist, dass die resultierende Bitrate nicht von der Größe des anzeigenden Displays abhängt. Allerdings muss in dem Client eine vollwertige Grafikkarte stecken, die alle gestreamten Grafikbefehle beherrscht und schnell genug abarbeitet.

Das für das Projekt entwickelte Streaming-Protokoll wurde im Wesentlichen von GLX (OpenGL Extension to the X Window System) abgeleitet und überträgt die Grafik-Kommandos in einem von der verwendeten Grafikbibliothek unabhängigen Format. So können beispielsweise auch DirectX-Spiele auf einem OpenGL-fähigen Client gerendert werden. Um die Reaktionszeiten in tolerierbaren Schranken zu halten, wurden verschiedene Optimierungen entwickelt, die berücksichtigen, dass Computerspiele häufig einfach vorhersagbare Zustände oder Matrizen des Grafik-Kontexts abfragen. Die zu übertragende Datenmenge wird darüber hinaus über ein intelligentes Caching sowie Entropiekodierung drastisch verringert. Im Schnitt ließen sich so 80 Prozent der Datenrate einsparen.

Literatur

[1] P. Fechteler, P. Eisert: Accelerated Video Encoding Using Render Context Information

[2] P. Eisert, P. Fechteler: Low Delay Streaming of Computer Graphics

keine. Das Datenzentrum im südlich von San Francisco gelegenen Mountain View antwortete innerhalb von 19 Millisekunden auf einen Ping. Dazu kommt im Spielbetrieb freilich noch die serverinterne Kompressionszeit. Allerdings ist OnLive ein eifersüchtiger Zeitgenosse: Lief auf unserem Testrechner im Hintergrund beispielsweise ein Download, warnte der Dienst öfter vor einer unzureichenden Verbindung oder brach diese gleich komplett ab. Kein Wunder, denn durch die Leitungen fließen beachtliche Datenmengen. Unsere Messungen ergaben einen Downstream von rund 700 KByte/s. In die andere Richtung ist der OnLive-Client sparsam und sendet im Durchschnitt nur knapp 10 KByte/s. Zum Vergleich: Ein Online-Rollenspiel wie „World of WarCraft", das komplett auf dem Client-Rechner läuft, die Grafik lokal berechnet und nur die Aktions-Daten der anderen Spieler lädt, kommt mit einem Downstream von 35 KByte/s aus. Während in Deutschland echte Flatrates Usus sind, bieten viele Provider in den USA nur Tarife mit Volumenbegrenzung an. US-Anbieter Comcast erlaubt derzeit nur 250 GByte pro Monat. Bei OnLive fallen pro Stunde 2,5 GByte Download-Daten an, somit reicht der Comcast-Anschluss immerhin für drei bis vier Stunden Spielen pro Tag.

Wolkenfront

Auch wenn OnLive mit noch nicht ganz artefaktfreier Grafik und eher niedriger 720p-Auflösung derzeit nicht mit lokalen Spiele-Installationen auf perfor-

Mitspieler können über OnLive Angeber-Videos aus Spielszenen erstellen und sich in fremde Partien als Zuschauer einklinken.

Die Grafikdetails aller OnLive-Spiele werden vom Server vorgegeben – lokal darf man lediglich die Helligkeit anpassen.

manten PCs konkurrieren kann: Das ferngesteuerte Spielen funktioniert so, wie es seine Macher versprochen haben.

Neben OnLive sucht derzeit Gaikai nach einem alternativen Vermarktungskonzept für Cloud-Spiele. Gaikai-Chef David Perry plant keinen Abo-Service, sondern will die Cloud-Spiele per Flash-Plug-in in Webseiten einbauen, die mit nur einem Mausklick starten. Geplant sind zunächst Demos neuer PC-Spiele, später auch Vollversionen, Casual Games und Indie-Spiele. Zahlen muss nicht der Spieler, sondern der Spieleanbieter – 1 Cent pro gestreamter Minute. Während OnLive auf wenige Datenzentren setzt, will Gaikai seinen Dienst über das ganze Land verteilen. Gegen Ende des Jahres soll der Service mit 14 Server-Zentren in den USA starten. Um die Verbindungswege und -zeiten zum Anwender möglichst kurz zu halten, sind bis zu 50 Standorte im Land der unbegrenzten Möglichkeiten geplant. Gaikai wird wahrscheinlich nicht so aufwendige Titel wie OnLive hosten. Laut Perry würden auf jedem Server zehn HD- oder 14 SD-Spiele in virtuellen Maschinen laufen. Die Bildqualität kann sich dynamisch an die Qualität der Netzwerkverbindung anpassen. So soll die Framerate bei Störungen stabil bleiben, der Anwender sieht stattdessen mehr Klötzchenartefakte. „Wir wollen nicht mit der PS3 konkurrieren, sondern das System möglichst einfach halten", sagte Perry.

Etwas still geworden ist es mittlerweile um Otoy, die ebenso wie OnLive Spiele auf alle nur erdenklichen Endgeräte streamen wollen. Dazu nutzen sie einen proprietären Videocodec namens ORBX, dessen Kodierung komplett per OpenCL auf den GPUs der Server läuft. Das Startup-Unternehmen ist eine Kooperation mit Supermicro und AMD eingegangen, auf dessen Fusion Render Cloud der Otoy-Dienst eines Tages starten soll.

Der in Israel entwickelte Playcast-Dienst (www.playcast-media.com) will Spiele nicht über das Internet, sondern die Netzwerke der Kabelfernseh-Betreiber auf Settop-Boxen streamen. In Japan ist seit einigen Jahren G-Cluster aktiv, die bislang aber nur ein Fitness-Spiel (http://minnafit.jp) hostet.

In Deutschland beteiligt sich das Fraunhofer Heinrich-Hertz-Institut an dem nichtkommerziellen EU-Projekt Games@Large, das derzeit in Italien, Israel und Brasilien getestet wird. Der Schwerpunkt dieses Game-on-Demand-Systems liegt auf der Übertragung in WLAN-Netzwerken mit geringeren Bandbreiten. Um das Datenvolumen klein zu halten, kann das System neben dem komprimierten Video-Stream auch lediglich Grafikkarten-Befehle an den PC-Client senden (siehe Kasten auf S. 82).

Das entlastet zwar Server und die Netzwerke, dafür benötigt der Client-Rechner noch immer eine leistungsfähige GPU.

Spiele als Dienst

Der Umstieg vom Retail-Verkauf auf Cloud-Dienste bedeutet einen Paradigmenwechsel. War man bisher als Käufer auch der Besitzer der Spiele-Disc, so erwirbt man nunmehr bei OnLive nur ein nicht übertragbares Nutzungsrecht für bis zu drei Jahre. Spiele sind keine Ware mehr, sondern eine Unterhaltungs-Dienstleistung. Ähnlich wie man sich bei einem Fernseher nicht um die Leistung sorgen muss, ob er denn auch die neuesten Filme ruckelfrei wiedergibt, könnte auch bei Spielen ein Umdenken einsetzen. Sie sind auf Knopfdruck einfach da und lassen sich genauso konsumieren wie Pay-TV. Dass sie flüssig laufen und alle Patches eingespielt sind, ist ein Problem des Dienstleisters und nicht mehr des Kunden.

Während in der Anfangsphase bei OnLive gewöhnliche PC-Spiele 1:1 gestreamt werden, wären zukünftig auch speziell für den Dienst entwickelte Spiele mit Supergrafik vorstellbar, für die Endanwender zuhause sonst ein ganzes Server-Rack mit Grafikkarten-Batterie im Wohnzimmer haben müssten. Die Trennung von Spieleleistung und Anwender-Hardware hat natürlich auch direkte Auswirkungen auf den Grafikkartenmarkt. Nvidia und die GPU-Sparte von AMD haben schon heute Probleme, genügend Hochleitungs-Grafikkarten an private Gamer zu verkaufen, weshalb sie mit CUDA und Stream händeringend nach neuen Anwendungsfeldern suchen. Je mehr sich Cloud-Gaming-Dienste etablieren, desto weiter schrumpft die Nachfrage nach potenten Grafikchips – das kann auch der Ausbau der Server-Zentren nicht wett machen.

Bis jedoch die Netzwerk-Technik und Server-Infrastruktur soweit sind, dass Spiele genauso flüssig und hochauflösend gestreamt werden können wie das Fernsehprogramm, werden noch einige Jahre ins Land gehen. Sony Computer Entertainment peilt beispielsweise einen Zeithorizont von zehn Jahren an, bevor eine Konsole ohne Disc-Einschub auf den Markt kommt, die Spiele nur noch aus dem Netz lädt. Während in den USA Provider landesweit tätig sind und dort die politischen und wirtschaftlichen Unterschiede zwischen den einzelnen Bundesstaaten nicht ins Gewicht fallen, gleicht Europa einem Flickenteppich, was die Verbreitung hierzulande zusätzlich bremst. Hier wird die Spieleindustrie mit ähnlichen Problemen zu kämpfen haben wie die Filmindustrie, die Jahre benötigt hat, bis Video-On-Demand-Angebote flächendeckend verfügbar waren. (hag) c't

Ende 2010 soll in den USA eine Mikrokonsole von OnLive erscheinen, die einen HD-Fernseher direkt mit dem Cloud-Gaming-Dienst verbindet.