Inside the Engine Room: Investigating Steam's Content Delivery Platform Infrastructure in the Era of 100GB Games

Christoff Visser[0009-0005-3511-6796] and Romain Fontugne[0000-0002-9816-5625]

IIJ Research Lab, Tokyo, Japan {christoff,romain}@iij.ad.jp

Abstract. As the size of video games continues to get bigger, new games and updates are becoming more visible in network operations. This research, coinciding with the 20th anniversary of the Steam store, provides an insightful exploration of a large-scale video game distribution platform. We place the operations of Steam under the lens and break down the details of its content delivery infrastructure. As part of this, we undertake a deep analysis of its data centres and cache locations. Recognising the trends in game development, this investigation acknowledges the dawn of the 100GB game era and the increasing pressure on distribution systems as a result.

Our research showcases the significant impact of major video game releases and provides an extensive investigation into the capacity of Steam cache servers, illuminating the strategies deployed when demand overshadows capacity. Players downloaded a monumental 44.7 exabytes from Steam in 2022 alone. With no signs of slowing down in 2023, Steam served an average of 15 Tbps of traffic between February and October, with peaks of up to 146 Tbps. This study lays bare the intricacies and operational challenges inherent to the digital game distribution landscape.

1 Introduction

For almost as long as modern computing has existed, so have video games. The Manchester Baby [16] ran its first program in 1948, giving birth to what we now consider the first modern computer. Shortly thereafter, Josef Kates built Bertie the Brain [34], one of the earliest examples of a modern video game.

Both computers and video games have come a long way since their early days. The launch of the Atari Video Computer System in 1977 exemplifies this monumental progress [50]. Its introduction of 2 kB cartridges significantly increased the capability and complexity of what developers can include in their games. However, cartridges still limited the storage available to developers, packing only up to 4 MB of space in the early '90s [3]. The introduction of CD-ROMs dramatically expanded the available capacity to 650 MB. The move to CD-based storage also introduced the possibility of distributing a game over multiple discs. Notably, in 1995, Sierra released *Phantasmagoria* across eight discs for the Sega Saturn in Japan [32]. Later, the increase in Internet availability and bandwidth gave developers the ability to move beyond the limitations of physical storage media and transition to digital releases.

Valve was one of the first adopters of digital game distribution when it released Steam, its video game distribution platform and storefront , in 2003 [7]. Since the introduction of the first digital game in the store in 2004 [24], Steam has grown exponentially over the past 20 years. In 2022 alone, players downloaded 44.7 exabytes of data from Steam [56]. Due to their increased file sizes and rising popularity, the release and updates of video games can have a significant impact on Internet traffic, causing record traffic peaks in broadband networks [15] and Content Delivery Networks (CDN) [11].

To better understand how Steam manages video game distribution, this paper investigates Steam's infrastructure and traffic data. First, we examine the data made available by Steam's public REST API [71], as well as the statistics they provide on their storefront [68]. Furthermore, we explore the insights that can be obtained from these data regarding Steam's infrastructure and its correlation with their traffic statistics.

This study offers a rare insight into the operation of a large-scale video game distribution platform. Using publicly available data from Steam, this study investigates how Steam manages the ever-growing traffic demand for the distribution of video games. We explore various aspects of the Steam infrastructure, from the data center level up to how they offload traffic to third-party CDNs. This gives us a better understanding of how Steam sustained a 36% growth in total content delivery between 2021 and 2022.

The main contributions of this study are:

- Characterising infrastructure requirements for a game distribution platform.
- Mapping Steam's presence at data centres and cache locations.
- Show the impact of big game releases and how Steam load balances traffic.
- Quantifying the effectiveness of Steam's infrastructure to serve their content.
- Estimating the capacity of Steam cache servers and quantifying Valve's usage of third-party CDNs to cope with peak traffic.
- Finally, we publish nine months of Steam aggregated traffic and cache load data [73].

2 Characterising content distribution

The Internet has enabled the digital distribution of software updates and media content. This section discusses the characteristics of digital distribution and the differences between popular content platforms (summarised in Table 1)

Urgency: A key benefit of digital distribution is the immediacy of content availability. Users no longer need to wait for a CD to update their system or a DVD to watch a movie. Instead, companies can now break down software and system updates into smaller parts and gradually roll them out. Microsoft, for example, famously has "Patch Tuesday" to roll out security updates [39]. For system administrators, this is a critical day to patch their systems. However,

	Steam	**Sports**	**Windows**	**Netflix**
Urgency	ASAP	1 h+ / streamed	'Patch Tuesday'	1 h+ / streamed
Time	World Wide Sync On-demand	World Wide Sync	Rolling Update	On-demand
Localization	Steam cache + CDN	Multi CDN	Multi CDN	Netflix cache
Catalogue	++++	-	-	++
Users	132+ Million	3 Billion	1.4 Billion	232.5 Million
Size	Up to 128GB	1 to 7GB/h	Avg 114MB - up to 5.4GB	1 to 7GB/h
Frequency	1 shot	1 shot	Low frequency	Daily

Table 1: Characteristics of different content delivery providers

most end users will patch at their convenience. Services like Netflix and sports broadcasting use streaming to allow users to start watching videos right away and conserve traffic. Consequently, the video download is transparent to the user, as long as their throughput can provide a smooth experience. *Contrarily, upon purchasing a new video game, or during the release of a significant update for a multiplayer game, users are eager to play right away. However, since the user needs to download the entire game, the traffic load cannot be spread over a longer period of time, as seen in video streaming. The only limitation on the download time is the bandwidth available to the user.*

Time: User requests exhibit different temporal characteristics depending on the content provider. Netflix users request videos on-demand, meaning that Netflix traffic occurs when both content and users are available. Microsoft moved towards a rolling updates model, spreading downloads out over a longer period of time. Similar to Steam game releases, sports streaming is usually independent of time zones, as audiences prefer to watch events in real time. *Users download Steam games on-demand, but Steam can automatically dispatch pre-ordered games and updates to users globally. This enables users to consume the latest content as soon as it is available.*

Localization: To improve user experience, content providers aim to bring the content as close as possible to their end users. Sports streaming and Windows Updates use multiple CDNs for widespread availability. Netflix relies mainly on its Open Connect Appliance (OCA) program [43] to deploy its caches within Internet Service Providers and Internet Exchange Points [20,8]. *Valve takes a different approach, using a combination of its own infrastructure (Steam caches) and CDNs. We investigate this further in Section 4.5.*

Catalogue: Content variety is a crucial factor in the allocation of cache resources. Content providers need to carefully balance storage space and the availability and regional preferences of items as the variety increases. Due to the live aspect of sports events, the variety of events per provider is quite limited. Likewise, the most frequently downloaded Windows updates are typically the latest ones. As Netflix and Steam focus more on content consumption, they offer a significantly larger library for their users. In March 2023, Slovakia had the largest Netflix library with more than 8,400 titles [64]. *In comparison, Steam has more than 150,000 products listed in its global store [70].*

Users: When comparing the total number of users for each platform, we found that the number of users is inversely proportional to the variety of content. According to Microsoft, there are currently 1.4 billion [41] monthly active devices running Windows 10 or 11. For sports, the number of unique viewers for the Tokyo Olympics 2020 reached 3.05 billion [31]. Unlike other content providers, sports broadcasting includes both TV and streaming. At the beginning of 2023, Netflix reported 232.5 million subscribers [65]. *In 2021, Steam reported an average of 132 million active monthly users [53].*

Size: The shift away from physical media has had varying effects for each of these providers. Streaming platforms are striving to optimise the size of the content they deliver, without compromising quality. Netflix, for example, typically uses a maximum of 7GB of traffic per hour [42]. Microsoft has been working over the past few years to optimise its cumulative update process for Windows. This includes reducing the size by up to 40%, with the average update size now around 114MB [19]. Although Microsoft occasionally provides larger updates, we estimate that these updates should be smaller than the Windows 11 installation media (5.4GB [40]). *Conversely, the size of video games has increased significantly in the past few years [47]. Games exceeding 100GB are becoming more common, requiring users to download very large files. For example, the download size for the base file of Baldur's Gate 3 on Windows is 105GB. However, this is only one of the files required for the game, and the game requires 150GB once installed [61].*

Frequency: Finally, the frequency of content acquisition by users can vary significantly between content providers. Even during repeated viewings, viewers must stream content from Netflix servers, although download options are available for mobile devices. Additionally, a viewer can typically only watch one title at a time. This encourages frequent streaming for users who want to explore the variety of content available. Due to the live aspect of sporting events, most people will watch the event as it occurs. Microsoft deploys its updates gradually, keeping the frequency of its updates relatively low. Although the size of video games is significantly larger than that of video content, the frequency of content acquisition differs significantly. *Users are only required to download a video game once, with the vast majority of games requiring only occasional updates.*

Summary: Considering all the characteristics described above, Steam has unique cache requirements. They must accommodate a wide variety of products while also providing sufficient global bandwidth for major game releases. We classify a major game release by download size and global popularity. With video games having a global release and update time, combined with the necessity that users need to download the entire game, distinctive traffic patterns emerge when compared to content from other content providers. We discuss these patterns in more detail in Section 4.

Parameter	Description
cell_id	Region ID
type	SteamCache or CDN
load	Load
weighted_load	Weighted load
host	hostname
preferred_server	Preferred or not
https_support	optional or Mandatory

(a) Cache Return

Parameter	Description
endpoint	hostname:port
type	websockets or netfilter
load	Load
wtd_load	Weighted load
dc	Data Centre code
realm	Realm identifier

(b) CM Return

Table 2: REST API return table

3 Peeking into the engine

In this section, we describe our methodology to collect the data used in this study. We also describe the data itself and how we processed it.

3.1 Web API

Steam provides a public API [71] that game developers [59], third-party sites [62,23,26], and the Steam client itself use. The parameter names are self-explanatory; however, documentation lacks details about the returned values. Of the 26 top-level interfaces that the API exposes, 2 are relevant to our study.

The `ISteamDirectory` interface provides the `GetCMListForConnect` method, which provides a list of *connection managers* (CM). The Steam client application contacts these CM endpoints for user authentication and manages updates and content for the application itself. The configuration file of the Steam client (‘`config.vdf`’) confirms these findings. This file also includes a parameter `CMWebSocket` with a list of CMs and a `LastPingTimestamp` for each CM.

Furthermore, the `IContentServerDirectoryService` interface offers the method `GetServersForSteamPipe`, which returns a list of Steam cache servers and a few third-party CDN endpoints that are used to download games and updates. Both methods use an optional `cell_id` parameter, a unique identifier for the location of the user. Querying these methods without a `cell_id` parameter instead returns results based on the IP of the client. Specifying an id value overrides this behaviour and returns the servers for that location instead. For example, using a `cell_id` parameter value of 4 returns the same results as clients querying from London. To confirm how the Steam client uses this, we refer to an older Steam client that includes a CellMap.vdf file with a list of cells [63]. Starting from January 2022, Steam clients no longer include this mapping file.

Table 2a presents the cache format from the list returned by the `GetServersForSteamPipe` method. Each server includes a `cell_id` indicating the Steam cache location. Table 2b illustrates how the `GetCMListForConnect` method returns a DC value, but not the `cell_id`.

cell_id	CM	Cache	Interconnects in PeeringDB	City	Country
116	eze1		Cirion	Buenos Aires	Argentina
52	syd1	syd1	Equinix	Sydney	Australia
92	vie1	vie1	Digital Realty	Vienna	Austria
25	gru1	gru1	Equinix	São Paulo	Brazil
117	scl1		Cirion	Santiago	Chile
14	par1	par1	Digital Realty	Paris	France
5	fra1	fra1	Digital Realty, Equinix	Frankfurt	Germany
5	fra2	fra2	Digital Realty, Equinix	Frankfurt	Germany
33	hkg1	hkg1	Equinix, Equinix	Hong Kong	Hong Kong
32	tyo1	tyo1	Equinix, Equinix	Tokyo	Japan
32	tyo2	tyo2	Equinix, Equinix	Tokyo	Japan
15	ams1	ams1	Equinix	Amsterdam	Netherlands
118	lim1		Cirion	Lima	Peru
38	waw1	waw1	Equinix	Warsaw	Poland
35	sgp1	sgp1	Equinix, Equinix	Singapore	Singapore
26	jnb1	jnb1	Teraco	Johannesburg	South Africa
8	seo1			Seoul	South Korea
40	mad1	mad1	Digital Realty	Madrid	Spain
66	sto1	sto1	Digital Realty, Equinix	Stockholm	Sweden
66	sto2	sto2	Digital Realty, Equinix	Stockholm	Sweden
4	lhr1	lhr1	Telehouse	London	United Kingdom
50	atl1	atl1	Digital Realty, Digital Realty	Atlanta	United States
65	dfw1	dfw1	Equinix	Dallas	United States
63	iad1	iad1	Equinix	Ashburn	United States
64	lax1	lax1	CoreSite, Equinix	Los Angeles	United States
1	ord1	ord1	Equinix	Chicago	United States
31	sea1	sea1	Equinix	Seattle	United States
			Equinix	Dubai	United Arab Emirates

Table 3: Cache and Connection Manager locations

We map each DC to the `cell_id` by examining the hostnames of the caches and connection managers. For example, `cache1-sea1.steamcontent.com` shows the type of server, followed by the DC, where the DC is associated with a 3-letter IATA location code. Table 3 contains a detailed list of all caches and connection managers.

Server preference Both caches and CMs contain a load and a weighted load value. Steam's API always returns the results in order of the weighted load, with the lowest value on top. Based on our findings discussed in the next section, we can infer that Steam clients use this to prioritise which server to use.

The load value shows the server's current load, with the highest observed value being 95, while the weighted load peaked at 280. Although the weighted load calculation is unclear, we found that it appears to be related to server load and the distance between client and server, as it increases when queried from a `cell_id` that differs from the Steam cache `cell_id`.

Steam caches also include an additional field, `preferred_server`, with a boolean true or false value. We found that Steam considers servers to be preferred when their weighted load value is less than 130. Unlike Steam cache servers, third-party CDN endpoints always have a `prefered_server` value of false and a fixed weighted load value of 130. In Section 4.5 we discuss how Steam uses CDNs when no preferred caches are available. Although each of the Steam cache servers has a unique hostname, a CDN endpoint always has the same hostname endpoint regardless of the location queried. To map which `cell_ids` have caches, we query the ids found in `CellMap.vdf` using the `GetServersForSteampipe` method. We also query IDs of up to 100,000 to find undocumented cells; however, we were unable to find additional locations. In total, we found 20 locations that contained caches. The same process was repeated to find the CMs. In general, we found 27 DCs that contained CMS, compared to 23 that contained caches.

```
[08-24 07:57:48] AppID 740 state changed : Update Required,
[08-24 07:57:48] AppID 740 update changed : Running Update,Reconfiguring,
[08-24 07:57:48] Got 5 download sources and 0 caching proxies via
ContentServerDirectoryService::BYieldingGetServersForSteamPipe (CellID 32 / Launcher 3)
[08-24 07:57:48] Created download interface of type 'SteamCache' (7) to host cache4-tyo2.steamcontent.com
[08-24 07:57:48] Created download interface of type 'SteamCache' (7) to host cache1-tyo2.steamcontent.com
[08-24 07:57:48] Created download interface of type 'SteamCache' (7) to host cache3-tyo2.steamcontent.com
[08-24 07:58:01] Increasing target number of download connections to 4 (rate was 0.000, now 537.271)
[08-24 07:58:01] Created download interface of type 'SteamCache' (7) to host cache4-tyo1.steamcontent.com
[08-24 07:58:13] Increasing target number of download connections to 5 (rate was 537.271, now 807.634)
[08-24 08:00:19] AppID 740 state changed : Fully Installed,
[08-24 08:01:08] HTTP/2 (SteamCache,67) - cache3-tyo2.steamcontent.com: Closing connection
[08-24 08:01:14] HTTP/2 (SteamCache,94) - cache4-tyo1.steamcontent.com : Closing connection
[08-24 08:03:19] stats: (SteamCache, 67) cache3-tyo2.steamcontent.com: 2721675904 Bytes, 105 sec
(206.63 Mbps). 19295 Hits / 538 Misses (97 %, 96 % bytes)
[08-24 08:03:19] stats: (SteamCache, 68) cache4-tyo2.steamcontent.com: 2231609202 Bytes, 78 sec
(229.86 Mbps). 12637 Hits / 4 Misses (100 %, 100 % bytes)


[08-24 08:29:31] AppID 740 state changed : Update Required,
[08-24 08:29:31] AppID 740 update changed : Running Update,Reconfiguring,
[08-24 08:29:31] Got 2 download sources and 0 caching proxies via
ContentServerDirectoryService::BYieldingGetServersForSteamPipe (CellID 32 / Launcher 3)
[08-24 08:29:31] Created download interface of type 'CDN' (2) to host steampipe.akamaized.net
[08-24 08:29:31] Created download interface of type 'SteamCache' (7) to host cache6-hkg1.steamcontent.com
[08-24 08:32:11] AppID 740 state changed : Fully Installed,
[08-24 08:33:06] HTTP/2 (SteamCache,290) - cache6-hkg1.steamcontent.com: Closing connection
[08-24 08:33:11] HTTP/2 (CDN,11) - steampipe.akamaized.net : Closing connection
[08-24 08:35:02] stats: (Invalid, 0) : 14707081285 Bytes, 157 sec
(750.42 Mbps). 72373 Hits / 2129 Misses (97 %, 96 % bytes)
[08-24 08:35:02] stats: (SteamCache, 290) cache6-hkg1.steamcontent.com: 703437715 Bytes, 142 sec
(39.55 Mbps). 2377 Hits / 9 Misses (100 %, 100 % bytes)
[08-24 08:35:02] stats: (CDN, 11) steampipe.akamaized.net: 14003643570 Bytes, 156 sec
(718.19 Mbps). 0 Hits / 0 Misses (0 %, 0 % bytes)
```

Fig. 1: Condensed download logs for Steam Client in Tokyo as traffic and load increases. Although the first download is made exclusively from caches in Tokyo, the second download 30 minutes later comes from Hong Kong and a CDN hosted by Akamai.

Validation with Steam CLI To confirm our findings, we set up five virtual machines (VMs) in the following cities, each of which had a `cell_id`: Chicago, São Paolo, Frankfurt, Tokyo, and Sydney. Using Steam's command line interface [67] we downloaded an instance of *Counter Strike: Global Offensive (CS:GO)* on each VM. During the download, Steam provided verbose logging for the CMs it connects to, as well as the cache servers it uses. Although we were only able to retrieve a maximum of 20 download sources via the web API, we found that the Steam client can reach up to 30 download sources. However, when the available sources dropped below 20, we were able to match the number of sources with what we see in the Web API.

Figure 1 shows the Tokyo VM logs during a CS: GO download, highlighting increased traffic and load. The logs' `cell_id` correspond to Tokyo's data in Table 3. Initial download data shed light on Steam's cache load balancing. The download speed correlates with the intended number of connections, which increased to five, although only four Steam caches were used without initiating the fifth. Once done, a summary provided detailed statistics for each connection, including bytes downloaded, time taken, speeds, and cache hits.

Just 30 minutes after the initial download, Tokyo's Steam cache load had intensified. This time, the client connected to only two sources, a CDN endpoint from Akamai and a Steam Cache server in Hong Kong, compared to the previous five. The logs also showed an 'Invalid' entry, with no further explanation from Steam. Notably, unlike Steam Caches, it was not possible to ascertain cache hit or miss information from CDNs. Discussions on how Steam uses caches from various cities and their use of CDN are presented in Sections 4.2 and 4.5.

3.2 Steam Stats

The Steam store provides an extensive list of statistics [68] about its games and players. These include the number of players online, current players engaged in a game, and the maximum daily player count per game. Steam also conducts monthly hardware and software surveys [69]. Steam publishes the results of these surveys to help both users and developers gain a better understanding of the landscape surrounding PC gaming.

The study concentrates on Steam traffic data, the data displayed on their web statistics Web page [68] is also accessible via a Cloudflare endpoint. This displays regional bandwidth, country-specific total bytes, and top Autonomous System Numbers (ASNs) per country. Each endpoint has a date parameter; however, only data for the last 48 hours are available. This restricts the historical analysis to the data collected during this study. The bandwidth API provides granular results, showing bandwidth usage per region every 10 minutes. Steam's website uses this API to plot the download bandwidth used by Steam users [68]. The total number of bytes per country and the top ASNs per country contain only the daily volume of traffic.

Table 4 summarises the countries that consume the most bandwidth and the availability of the cache and connection manager servers per country. In particular, despite Chan's ranking as the top consumer of Steam traffic, its caches and connection managers are absent from the dataset. Section 5 discusses in more detail the unique way that Steam manages their infrastructure, explaining their absence. Similarly, no API endpoints were found within Russia.

Rank	Country	Cache	CM
1	CHN	False	False
2	USA	True	True
3	RUS	False	False
4	DEU	True	True
5	BRA	True	True
6	GBR	True	True
7	CAN	False	False
8	FRA	True	True
10	KOR	False	True
11	JPN	True	True
12	POL	True	True

Rank	Country	Cache	CM
13	AUS	True	True
16	ESP	True	True
17	ARG	False	True
19	NLD	True	True
20	SWE	True	True
24	HKG	True	True
25	CHL	False	True
34	AUT	True	True
38	PER	False	True
44	SGP	True	True
45	ZAF	True	True

Table 4: Examining the countries with the highest download rates on Steam and the presence of cache and connection managers.

3.3 PeeringDB

Steam's API offers insight into the distribution of Valve's facilities, particularly its content delivery network. PeeringDB [49], another tool for researching Internet infrastructures. Table 3 cross-references the data centre locations found in the Web API with the facilities listed for Valve (AS32590) [48]. This cross-referencing reveals a near perfect match, with the exception of Seoul and Dubai.

Although Seoul does contain CM servers, they do not have any Steam cache servers. Dubai is a more interesting case study. Not only could we not find any CM or Steam Cache servers related to Dubai, but it is also the only presence that we see in the Middle East region.

As of December 2023 there were 1296 ASes in PeeringDB with a 'Content' network type. When ranking ASes based on the number of IXPs the AS is a member of, Valve comes in 21st with 64 IXPs. It ranks just below Apple, which is a member of 67 IXPs, and above Sony Interactive Entertainment (PlayStation) and Twitter, which are members of 38 IXPs each.

Valve's notes section in PeeringDB confirms the findings of Section 3.1, noting that "Clients automatically find the closest location".

3.4 Data Collection

We queried each of the API endpoints every 30 minutes from February 17, 2023, to October 31, 2023. Each API endpoint was queried to obtain the weighted load values experienced within all locations, we queried all APIs multiple times using all valid values of `cell_id` found in Section 3.1. We stored the collected data in a MongoDB database to ease data cleaning and processing. As part of the study, both the code used for the collection and the dataset itself are made publicly available [73].

Figure 2 shows the overall trend of Steam traffic. Steam sees an overall increase in traffic throughout the year. Of the five distinct traffic peaks depicted

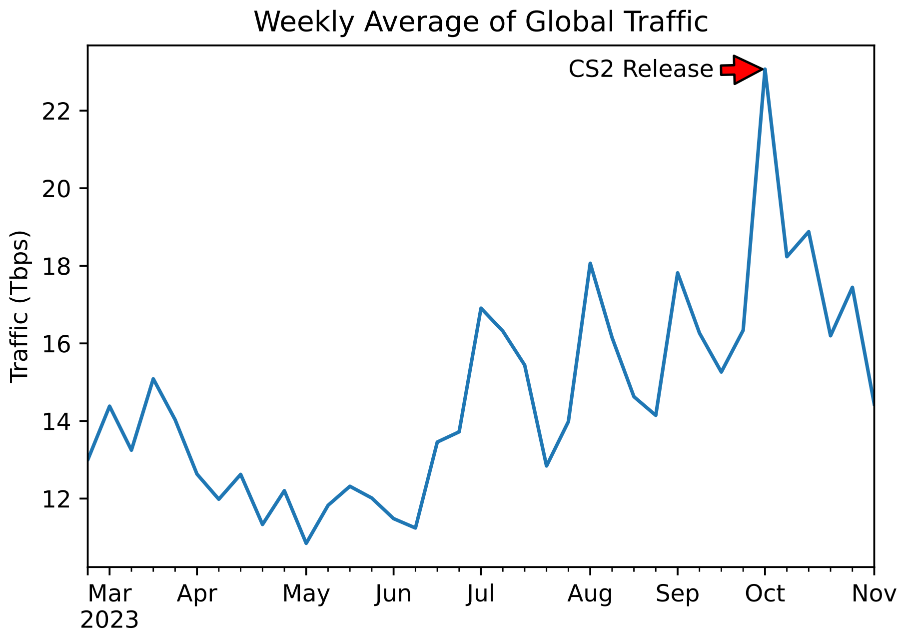


Fig. 2: Weekly average of traffic sent from Steam to end users.

in Figure 2, four are attributed to a week that contains one or two popular video game releases. The only exception is the peak seen in July, which is attributed to the annual Steam Summer Sale [54]. The large spike at the end of September/start of October is due to the release and update of multiple games [58,57,21], including Counter Strike 2.

Table 5 contains the time intervals used throughout this paper and the names of the relevant datasets. The release of *Counter Strike 2 (CS2)* is a great case study to illustrate how Steam handles sudden increases in traffic, reaching a peak of 85 Tbps. For the baseline, we used two time frames from the dataset. The first is `Prerelease`, based on the same period of three days but two weeks before the release of CS2. The week prior to the CS2 release has multiple popular game releases that made it ineligible to use as a baseline. The other baseline, `Postrelease`, is set a week after the release of CS2. In mid-September we added the metadata of `cell_id` from which the query originated into our dataset. This means that unlike the `Prerelease` dataset, the `Postrelease` dataset includes the origin `cell_id` used to poll the Web API. The origin `cell_id` in `Origin id` is particularly useful to determine how Steam caches are shared between `cell_id`s. Finally, the `Latest` dataset is the most recent 2 months of data representing the most up-to-date information on Steam's infrastructure.

Dataset Name	Start	End	Notes
`Whole dataset`	2023/02/18	2023/10/31	Complete dataset
`Prerelease`	2023/09/13	2023/09/15	Two weeks prior to CS2 release
`CS2 release`	2023/09/27	2023/09/29	Release period of Counter Strike 2
`Postrelease`	2023/10/04	2023/10/06	One week after the CS2 release
`Origin id`	2023/09/25	2023/10/31	Origin `cell_id` present in metadata
`Latest`	2023/09/01	2023/10/31	Most recent 2 months

Table 5: Data set properties and notable dates studied.

4 Steam under pressure

4.1 Load balancing within a city

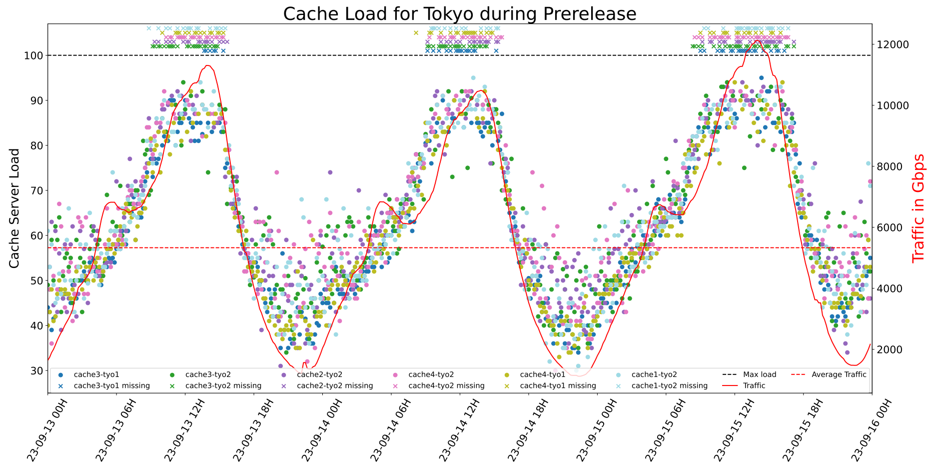
Methodology Our analysis begins by looking at the traffic and cache load patterns seen during popular game releases. As stated in Section 3.2, the Steam API returns a varying number of caches for a DC. To illustrate this behaviour, the load values of each cache server are plotted over time. When a cache is not visible, we assign an artificial load value greater than 100 and mark these events as 'x' in Figure 3. The study also examines traffic data observed during this period. While the Steam Stats page updates daily for each country, a more detailed analysis utilises regional traffic statistics, available at 10-minute intervals. However, these statistics offer a broader geographic perspective, lacking the finer granularity of country-specific data.

Discussion Figure 3 depicts an example of traffic in the Asia region compared to the load on Steam cache servers in Tokyo, Japan. Figure 3a shows a baseline using the `Prerelease` times. The load values follow the typical daily peak and trough pattern expected with daily Internet use. Overlaying the traffic data for the Asia region shows a matching pattern emerges, which indicates that the loads on the servers are related to how much traffic Steam serves.

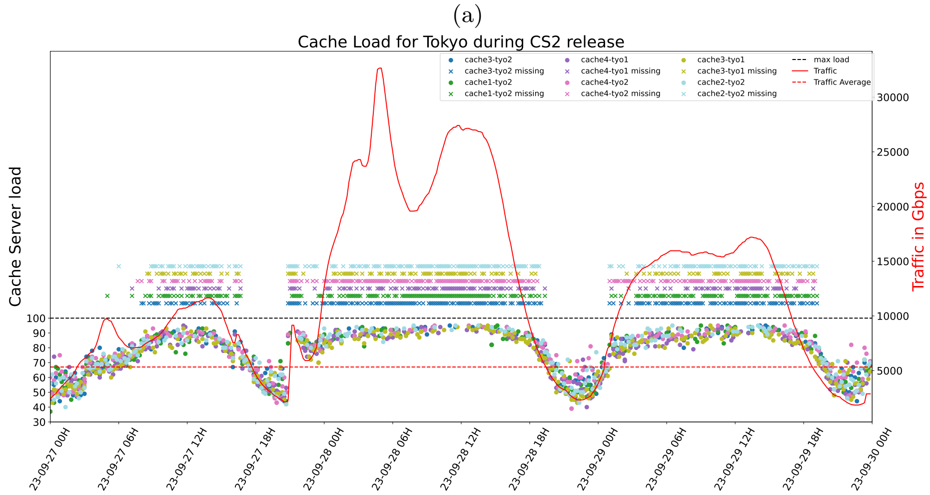
The loads on all the cache servers remain uniform to each other throughout the day. This demonstrates that Steam shares traffic fairly between each of their cache servers. The other observation we make is that less Steam cache servers are visible in the Steam API during peak times, which suggests that the API is not returning overloaded cache servers. However, while there are fewer Steam cache servers advertised, we do not observe traffic volume flattening, or even decreasing. This is another indication that Valve effectively offloads traffic when there are no Steam cache servers available within a city.

Figure 3b illustrates the traffic and cache load during the Counter Strike 2 (CS2) release. This release sees Steam serving nearly six times more traffic compared to the average traffic they serve. Although traffic volume follows the trend of cache load during off-peak times, the two patterns are different during large releases.

Steam usually sees a significant increase in traffic for new games or updates, but the release of CS2 had a combination of factors that contributed to the



(a)



(b)

Fig. 3: Load on Steam's cache servers for Tokyo against traffic for Asia. Figure 3a represents a typical week (Sept 13 to Sept 15) while Figure 3b shows the behaviour for the same period two weeks later during the release of CS2.

large increase in traffic shown in Figure 3b. First, CS2 completely replaces its predecessor Counter Strike: Global Offensive. This means that it behaves more like an update and that Steam clients will download the update as soon as it is available. Whereas new game releases require user interaction to start the download when available. Second, since CS2 does not use any of the already downloaded data from its predecessor, the client needs to download the entire game. The storage requirements for CS2 also increased from 15GB to 85GB, resulting in a significantly larger download compared to a normal update. Finally, Counter Strike is consistently the most played game on Steam, peaking daily at just over a million concurrent players. During the week of CS2's release, the peak concurrent player count grew to over 1.4 million.

Figure 3b shows a sharp increase in traffic and load at 21:00 UTC on September 27 for the release of CS2. However, peak traffic is not reached until 06:00 on September 28. This can be attributed to time zone differences, with it releasing at 06:00 local time in Japan. The traffic volume continues to grow as the rest of Asia wakes up and updates the game. Although the traffic decreases after the peak, it does not drop below the typical peak that normally allows the Steam Cache servers to recover. Traffic returns to its normal pattern in the evening, but at double the traffic levels compared to the previous evening.

The popularity of the game extends to the next day. However, instead of a sharp spike, the traffic is a lot flatter but still sits well above the peak traffic of the 27th (i.e. the day before the game release).

4.2 Load balancing between cities

Methodology This section investigates how and when Steam redirects users to caches located in other cities and regions. However, the geographic isolation of Tokyo makes it difficult to study how Steam does load balancing between cities. Thus, the focus is on the two regions that have the most Steam cache servers, namely Europe and North America.

During our experiments, we found that Steam often returns servers from other `cell_id`s than the one specified in the request query. Starting in mid-September, the origin `cell_id` was added to the metadata to keep track of servers not based in the origin city.

We create a mapping between the city we searched for and the city of the caches returned by the API. In this study, we focus only on queries that originate from Europe and North America. These regions have the highest number of caches and are therefore more likely to serve other parts of the world. In order to prevent any favoritism towards cities that have a high number of Steam cache servers, such as Frankfurt which has 30 servers, only the top five servers returned from Steam are considered. Due to the large number of servers in Frankfurt, it frequently acts as a secondary cache destination for various European cities. By limiting the selection to the top five entries, it also enhances the assurance that the Steam servers are utilized by the Steam client to download content.

Because the query id metadata was not added until mid-September, we are unable to use the same time frames as the load data in Section 4.1 and instead focus on the datasets `CS2 Release` and `Postrelease`. However, to increase readability and maintain consistency with our other figures, Figure 4.3 shows the baseline week first followed by the release week.

Discussion Figure 4 shows a heat map of the cities that request Steam caches versus the cities that advertise available caches. Each row represents the preferred cache servers used to offload traffic by certain cities. As expected, Steam caches mostly serve clients querying from the same `cell_id`. However, when the

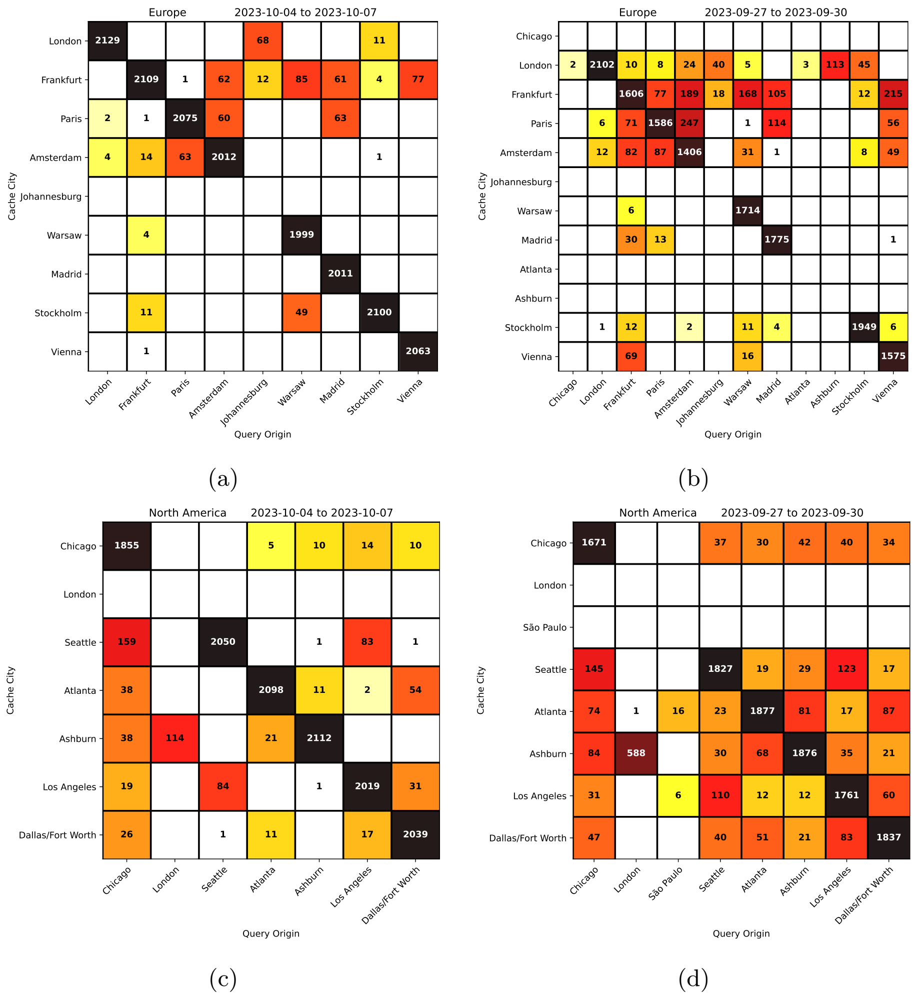


Fig. 4: Load balancing between cities within the same region. Figures 4b and 4d show the state during the release of CS2. Figures 4a and 4c one week after release.

cache servers of the origin city become overloaded, the Steam API also returns cache servers from other cities. We found that Steam mainly advertises cache servers which are geographically close to the origin `cell_id`. For example, during the `Postrelease` baseline dataset, both Europe and North America contain each only one city, Johannesburg and London, querying from outside their region (Figures 4a and 4c).

First, we look at Johannesburg, which is the only city in the African region that contains Steam caches. The geographic isolation of Johannesburg in the Steam infrastructure means that there are no other cache servers within the

Africa region to offload traffic when overloaded. Although Europe is the closest region to Africa, there is still a large geographical distance between Johannesburg and London or Frankfurt. Despite this, London and Frankfurt caches frequently serve clients in Johannesburg.

London itself, however, does not follow the same distance logic when overloaded. Unlike Johannesburg, London is not a geographically isolated city. The expectation is that traffic would be offloaded to Paris or Amsterdam because of the close proximity of these cities. However, Figures 4c and 4d, featuring the North American region, indicate that London is significantly more likely to offload traffic to Ashburn instead.

Figure 4 shows a clear geographic preference for cache entries that appear outside of the origin city. The clearest examples of this are Frankfurt, Paris, and Amsterdam. All three cities are geographically close to each other, and cache entries for each show up in each other's queries. Amsterdam and Paris stay exclusively within this cluster, while Frankfurt also offloads traffic to Warsaw, Vienna, and Stockholm. Even with the limitation of the top five preferred servers, Frankfurt is still the most seen backup cache location in Europe. Madrid also shows a geographic preference, as it often offloads traffic to Paris and Frankfurt, but given its limited number of cache servers, Madrid does not serve other cities.

For North America, there is a geographical preference amongst the East and West Coast servers. Figure 4c shows this best with Seattle on the West Coast during the usual load, where the queries do not return any servers from Chicago, Ashburn, or Atlanta.

The plots on the right of Figure 4 show the state during the release of CS2. As the cache servers become overloaded, there is an increase in the number of servers retrieved outside the query city. Figure 4b shows that in Europe, cities such as Vienna, Warsaw, and Amsterdam go from two backup cities to five. The North American plots in Figure 4d show a more uniform distribution of the servers seen from the city from which the query originates. Despite the remaining geographical preferences, when under heavy load, Steam increasingly serves caches to clients on the opposite side of the region. In addition, both regions see new cities from which queries originate. Europe only sees additional requests from the North American region, receiving servers from London. In North America, there is only one new origin city, São Paulo, which received servers from Los Angeles and Atlanta.

4.3 Impact of video game releases for a region

Methodology Previous sections focus on the impact of a game release on a city and inter-city level. Where Figure 5 instead looks at the impact on a regional level. The average load for a region is calculated by combining the load values for Steam caches per region and calculating the average value per timestamp. As described in Section 4.1 during peak times, caches no longer advertise being available. In these circumstances, we designate a load value of 100% to that cache.

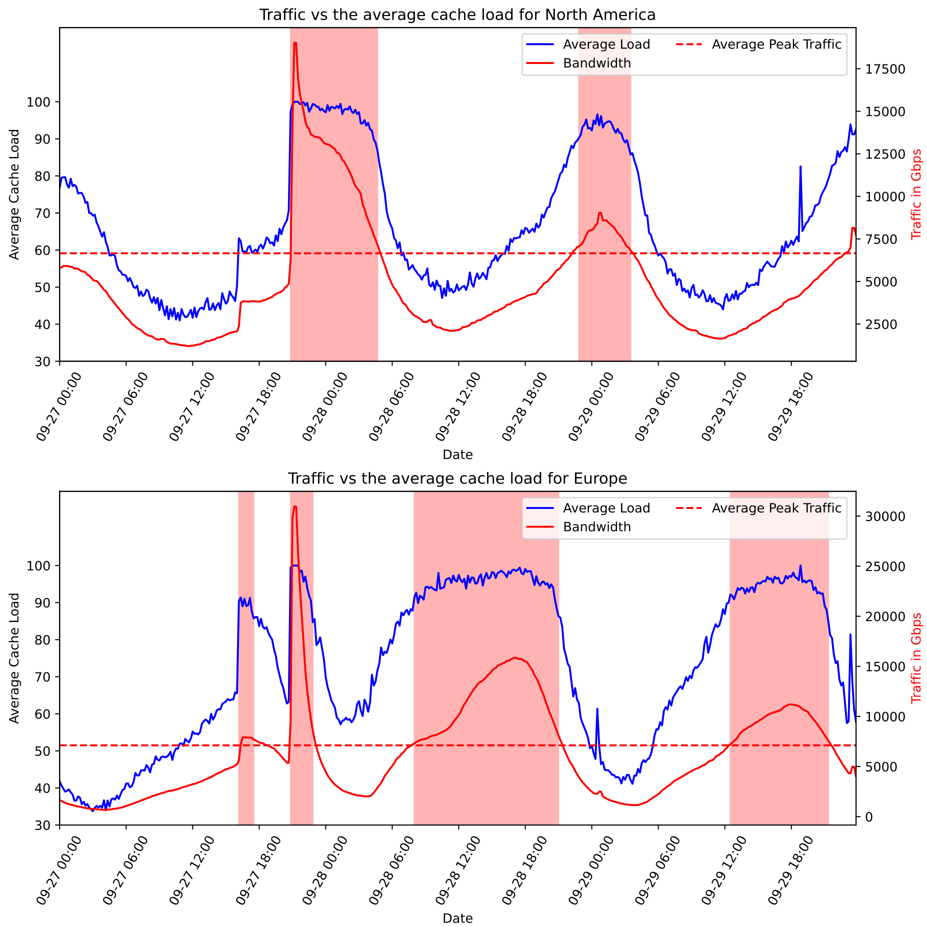


Fig. 5: Traffic and cache load for Europe and North America during the release of the game *Counter Strike 2* on March 28.

Discussion Unlike in Figure 3, Figure 5 shows that both Europe and North America have a much more dramatic increases in traffic. In Europe, there is an almost immediate drop in traffic after the release that can be attributed to the game releasing at 11 PM Central European time. Compared to the North American region, where the game is releasing in the early evening.

We found that when the average cache load for a region reaches around 90, we start to encounter cache servers as no longer being available. The areas highlighted in red show when cache servers start to get overloaded. At this point, Steam no longer advertises these as being available. The cache load starts to plateau during these periods of time, although the traffic still increases. Thus, Steam offloads traffic to third-party CDNs.

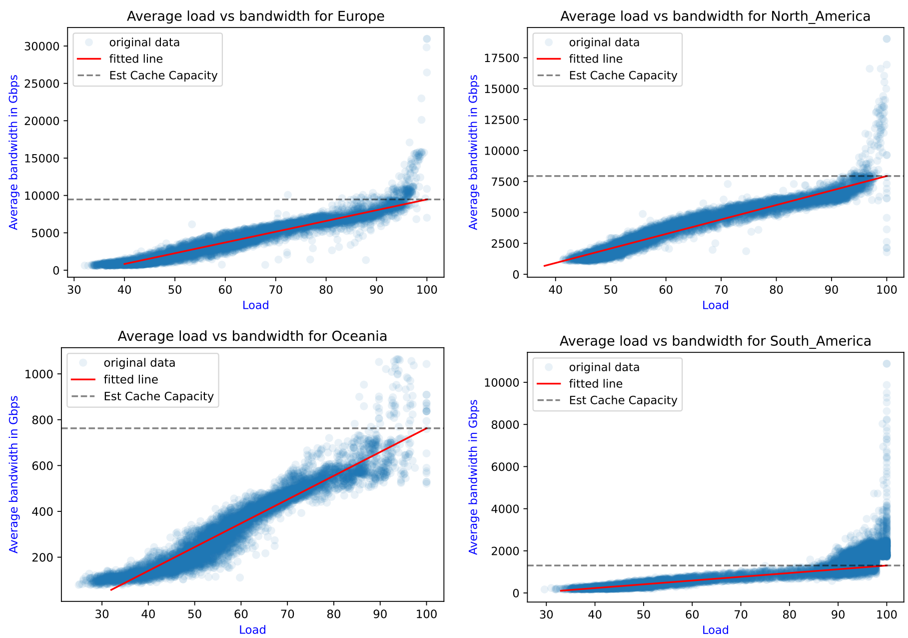


Fig. 6: Traffic against cache load per region. The linear regression is computed for the range $x = [0, 90]$. The cache capacity is derived from linear regression at $x = 100$.

4.4 Finding the limits of Steam Cache servers

Methodology Figure 5 shows the synchronised trends for the cache server load and the traffic data per region. The relationship between these two metrics reflects the capacity of the cache servers in terms of traffic. We compute the linear regression between the load and traffic values for each region and exclude the periods when Steam sends traffic to third-party CDNs.

Similarly to Section 4.3, when a cache is no longer available, a load value of 100 is assigned. Although the highest load returned was 95, the inflated load value of 100 helps to increase the average load for the region when caches are no longer present. To find the maximum capacity that Steam caches can provide, we calculate the linear regression against a maximum load value of 80 to 100. Using a maximum load value of 90 gives the highest average r^2 value across all regions of 0.931. We plot the average load and traffic values for each region using data collected from September 1 to October 30.

Discussion Figure 6 shows the linear regression results along with the data collected from server load and traffic. Each region has a strong positive linear relationship until the cache loads reach 90 and then only the traffic data continue to increase. This is more evidence that Steam sends traffic to third-party CDNs

Region	Cache served %	Cache traffic %	Est Cache Cap	Peak traffic
Africa	90.95	82.88	106	281
Asia	96.09	89.50	13556	32657
Europe	97.75	93.42	9470	30954
North America	98.60	96.24	7940	19025
Oceania	99.17	97.66	763	1064
South America	59.22	33.79	1303	10883

Table 6: Estimated effectiveness and capacity (Gbps) of Steam caches per region.

when their cache servers are overloaded. For South America, where load values are routinely high, we assume a limited cache capacity in that region. In the next section, we quantify the Steam cache capacity for each region and the amount of traffic offloaded to CDNs.

4.5 Pushing past the limits

Methodology To estimate the maximum capacity of a region's cache servers, we use the linear regression results from the previous section and estimate the traffic value for a load value of 100. Figure 6 shows the estimated cache capacity with a dashed line, which corresponds to the change in traffic behaviour shown in Figure 3. The volume of traffic served increases significantly after crossing the cache capacity line. This indicates that Steam uses CDNs when its cache servers are unable to cope with the increase in traffic.

Based on the estimated cache capacity, we infer the effectiveness of Steam cache servers in distribution over time and in terms of traffic volume. The cache served percentage is identified as the percentage of time when traffic remains within the cache's handling capacity. The cache traffic percentage represents the ratio of the total estimated traffic managed by the caches to the aggregate traffic observed in that region.

Discussion Table 6 shows the computed cache served and traffic percentages, as well as the estimated cache capacity and the peak traffic seen for each region. These results highlight the effectiveness of Steam caches in Oceania, Europe, and Africa. This is especially apparent in Oceania, where we see limited use of third-party CDNs. CDNs offer a solution to alleviate cache load, allowing increased capacity during periods of high traffic, such as game releases or significant updates. This is particularly beneficial when there is a sudden surge in traffic in a short period of time. The clearest example is South America, where the observed peak traffic is 8.3 times the estimated cache capacity, giving a much flatter representation in Figure 6. All other regions see peaks between 2 (North America) and 3 (Europe) times their cache capacity.

The South American region would benefit the most from adding more Steam cache servers. Only 59% of the traffic values fall below the cache capacity line. More importantly, caches can only serve 34% of the total traffic in that region.

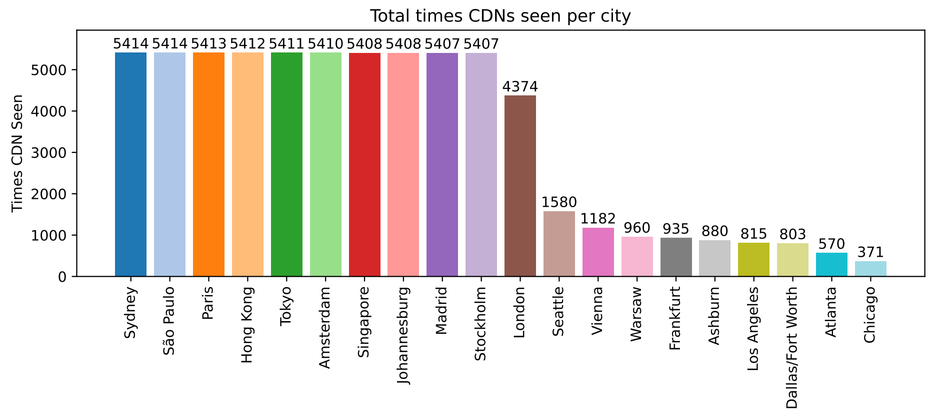


Fig. 7: Frequency of CDNs seen per city

4.6 Third-party CDNs per city

Methodology The `Origin id` dataset allows us to gain insight into how Steam uses third-party CDNs around the world. By querying the Steam API with different `cell_id` we can discover where Steam advertises different third-party CDNs to clients and the discrepancies between different cities.

Discussion Figure 7 depicts the number of third-party CDNs returned per city. Although we performed the same number of queries for each city, the total number of third-party CDNs seen in our queries varies drastically from one city to another. The Steam API appears to be configured to return no more than four CDNs in each query, accounting for the peak values observed near 5400, and does not consistently advertise CDNs for certain cities, with 63% of requests including at least one CDN. The findings of Section 4.5 suggest that regions like South America experience frequent traffic offloading, in contrast to Oceania, where this is a rarity. This would lead to the expectation that a city like São Paulo would receive the highest number of CDNs in queries, while Sydney would see the fewest. However, both see the same number of CDNs. Manual inspection of the data reveals that Steam's API only returns third-party CDNs either (1) during peak hours when all the caches in a city are overloaded or (2) for cities that are far from other Steam caches; hence, CDNs are preferred backups. São Paulo and Sydney are both geographically isolated, which explains their high values in Figure 7. In contrast, cities where Steam rarely advertises CDNs (e.g., Frankfurt and Ashburn) commonly host a larger number of Steam caches and have a close geographical proximity to other host caches.

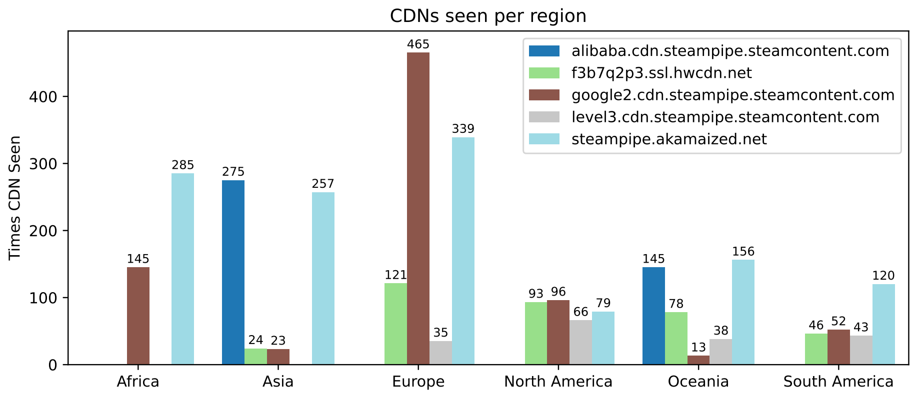


Fig. 8: Number of times CDNs are encountered in each region over a three day period during the release of Counter Strike 2. (27-29 September 2023)

4.7 Global CDN trends

Methodology Our dataset reveals that Valve relies on multiple third-party CDNs and that the advertised CDN endpoints may vary depending on the queried `cell_id`. To understand the mechanisms behind this multi-CDN setup, we now focus on which CDNs Steam advertise for different regions and during different loads. In this section, we leverage the `Origin id` dataset to correlate the queried `cell_id` with the domain name of the CDN endpoints returned by the Steam API and look for discrepancies between regions and time periods. Similarly to Section 4.3, we limit our results to the top five entries to increase the confidence that the Steam client will download content from the CDNs.

Discussion Figure 8 shows the third-party CDN providers seen per region during the release of CS2 and the frequency with which each provider was seen. Steam uses five different third-party CDN providers: Google, Akamai, Alibaba, Highwinds, and Lumen (Level 3). However, Google and Akamai are the only providers that Steam returns for all regions. Only Oceania has access to all five CDN providers. In contrast, Africa only receives endpoints for Google and Akamai. Alibaba is the least prevalent CDN in all regions, being present only in Asia and Oceania. Notably, Alibaba is the only provider that lacks HTTPS support.

The preference for CDN providers per region is apparent in Figure 8. Although Google and Akamai are available in all regions, Steam shows a preference for Google in Europe and North America. Akamai consistently ranks as the first or second most preferred CDN provider. Although Alibaba is only present in two regions, it is also highly preferred as a CDN provider for both of these regions. North America stands out with a more distributed preference among CDN providers. Whereas South America has a strong preference towards Akamai and a balanced preference between the other providers.

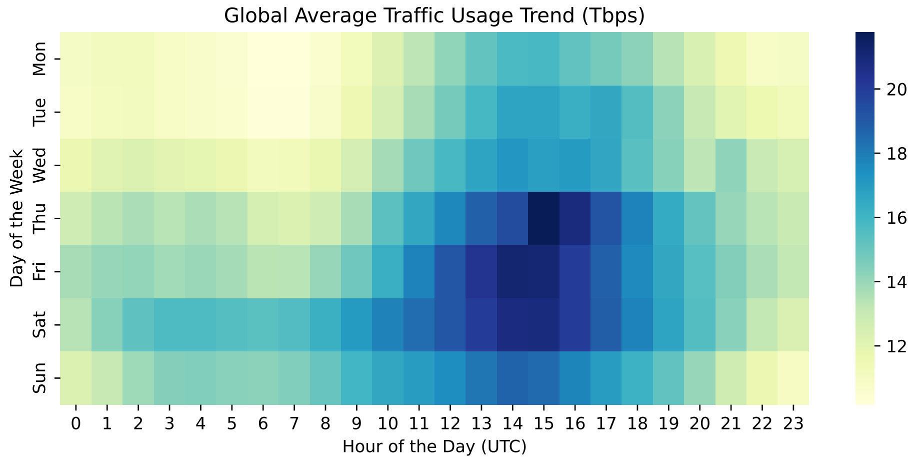


Fig. 9: Traffic trends for Steam based on the time of day and the day of the week. Users are most likely to download from Steam at 15:00 on a Wednesday (UTC), and least likely at 06:00 on a Monday.

4.8 Hourly/Weekly Traffic Insights

Methodology In previous sections, we compared the traffic within a region during a game release with a baseline time period. Traffic patterns per region tend to follow the normal day-night cycle, except during game releases and updates.

Instead, this section focuses on the trend of traffic usage on a global scale. Mapping the average traffic per hour of the day and per day of the week highlights the most popular times worldwide to download content from Steam.

Discussion Figure 9 illustrates average traffic usage patterns by time of day and day of the week. The data indicates that Thursdays at 15:00 (UTC) experience the highest average traffic volume, reaching 21.8 Tbps. A significant increase compared to the average of 10.2 Tbps observed on Mondays at 06:00. Analysing the Steam charts for new releases in 2023 offers insight into this trend. In particular, of the twelve top new games based on gross revenue in 2023, five were released on Thursdays and another four on Fridays. Interestingly, Counter Strike 2, despite being the most played game, is absent from the list of new releases in 2023. This is attributed to its classification as an update rather than a new release, underlining the importance of considering game updates as a factor in traffic spikes. Although updates contribute to increased traffic, tracking them poses challenges. Unfortunately, Steam does not provide specific data on update sizes or release timings. However, through SteamDB, we can monitor when new builds are submitted [60], although these do not always correlate with the updates distributed to users.

5 Limitations

Steam data:

The main constraints of this research are related to the reliance on the data provided by Steam's API. Although our findings regarding Steam's promotion of their Steam cache are supported by the Steam client, we lack information about the inner workings of the Steam caches themselves. For instance, the significance of the 'weighted load' remains unclear, and we were unable to get further clarification from Valve regarding this. However, based on our experiments, we could determine that this value fluctuates depending on the load and proximity to the cache and that it serves as a priority indicator. Unfortunately, the exact method used by Steam to calculate this value remains unknown on the basis of the data accessible through the APIs.

There are also insufficient data on cache infrastructure in the Middle East, Central America, and Russia regions. By examining the Cisco Umbrella [12] top 1 million domains and subdomains, we identified two companies that host caches using the same top-level domain as Steam's caches. These companies include RETN ltd [52], which specialises in connecting networks between Europe and Asia, and CTM [17], a Macau-based telecommunications company.

Another constraint is the expansive interpretation of the term 'region'. Consequently, it becomes challenging to determine the specific geographical location where user traffic is attributed to one region rather than another. This issue is particularly significant in regions with limited caches, such as Africa, where only one datacenter is available in South Africa.

Splinternet: In our study, fully incorporating or disregarding China is a challenging task. The launch of Steam China in early 2021 makes it a relatively new platform. However, unlike Steam Global, Steam China lacks access to community features, such as discussions or the Steam marketplace. Furthermore, the public statistics we use in this study do not include data from Steam China. As shown in Table 2b, the only 'realms' present are 'steamglobal' and 'steamchina'. This limitation is the result of the strict regulations imposed on game releases in China [45].

According to the public statistics provided by Steam [68], China is often ranked as the top country for global Steam traffic. However, when using the Web API, no Chinese cache servers or connection managers are returned. By changing the API domain name from *steampowered.com* to *steamchina.com*, we can access an API similar to that described in Section 3.2. However, the naming convention of the returned values is significantly different, making it difficult to include them in our study. We also attempted to query the `cell_id`s we found in Section 3.1 against the Steam China API. However, the API only returned CDNs, similar to cases where there are no Steam caches for a `cell_id`. Given the increasing awareness of the 'Splinternet', we believe that investigating the differences between Steam China and Steam Global could serve as an interesting case study.

6 Related Work

6.1 Gaming

Past networking studies on video games have focused on the characterisation of game traffic [10,22,27]. More recent studies have focused on the impact of cloud gaming on Internet traffic [9,18,76,74,5,14]. Others focus on game security issues, such as DoS attacks [33], account theft [51], and cheating detection [35].

To the best of our knowledge, there is only a limited number of research papers on the Steam platform. Using Steam's API, studies focus primarily on the social impact of Steam users [46], and Steam communities [6]. Lin et al. [36] analysed the growing trend of popular games that require urgent updates. While they are using Steam as an example, this work's primary focus is on the patch update cycle of games.

6.2 CDNs

The research community has dedicated substantial efforts to uncovering the infrastructure of large content distribution networks. Gigis et al. [25] studied hypergiant off-net cache deployments and demonstrated the massive deployment of cache servers in access networks to bring content closer to users and reduce inter-domain traffic.

The hybrid-CDN approach used by Steam is not uncommon for large content providers to use, with multiple companies advertising such services [4,72,38]. One of the key benefits of using this approach is to increase availability in geographically distributed regions. The other benefit, which is one of the primary focuses of our study, is to use third-party CDNs as an overflow as caches reach their maximum capacity.

Similar multi-CDN solutions are commercially available [13,2,37], with the same goal of making content available in more regions. Hohlfeld et al. [29] studied one of these multi-CDN solutions, providing insights into the infrastructure employed, its usage, and performance (mainly in terms of RTTs). We found that Valve closely follows the strategies described in the survey among EBU members surrounding the use of CDNs [66]. In particular the following:

- Use own CDN until max capacity and CDNs as overflow.
- Use a content management system to provide the CDN location to the client.
- Use Multi-CDNs.

Online video streaming services, and in particular Netflix, have also received a lot of attention in recent research. A measurement study [20] from residential networks has shown the extensive deployment of Netflix caches and their effectiveness. Other studies have emphasised how Netflix leverages IXPs [8] and third-party CDNs [1] to cope with the ever increasing amount of video traffic.

Another approach to hybrid CDNs is to incorporate peer-to-peer into the design architecture, creating a Hybrid CDN-P2P. Research into this design aims to lower cost, increase scalability, and reduce latency for content delivery, especially in video content streaming [44,28,30,75]. In March 2023 Steam introduced a feature that allows users to copy existing game installations and update files from one PC to another over a local network [55]. Although this is on a much smaller scale compared to other P2P networks, this reduces Internet traffic in networks with multiple Steam clients. We can foresee a possibility of Steam expanding the capabilities of this to work outside the local network, further reducing the load on their infrastructure.

7 Conclusion

In this paper, we characterise the properties of video game distribution on the Internet and investigate the case of the Steam platform. Our study documents the physical infrastructure of Steam, the impact of video game releases, the capacity of Steam cache servers, and its usage of third-party CDNs. We found that Steam infrastructure handles most of the traffic load but offloads to CDNs the substantial traffic peaks due to popular game releases and updates. Overall, our study shows that Valve does a good job with its Steam distribution platform, especially during game releases. However, regions such as Africa and especially South America could benefit from more caches within the region.

Infrastructure design: The notably large spikes observed for video game distribution are mainly due to the size of the video game files, the synchronised worldwide releases, the urgency to download all the files to play a game and the increasing popularity of video games. Handling these relatively short but intense traffic spikes with only Steam infrastructure would require them to more than double their infrastructure capacity. This is an interesting example that illustrates a cost trade-off to develop an infrastructure that is effective about 95% of the time and leverages the flexibility of CDNs to deal with overloads. Valve serves as an exemplary case study, illustrating why EBU members advocate for their specific strategies in the management of CDNs [66].

Public dataset: The ephemeral nature of our data availability during collection makes our dataset a valuable resource for insights into Steam's inner workings. We actively maintain and update this dataset and have made it publicly available[73]. We believe that the combination of granular traffic statistics, infrastructure overview, and Internet performance per country and per Internet Service Provider gives a better understanding of global Internet usage trends.

Future work The main objective of this exploratory study is to collect information about Steam's infrastructure from data that Steam provides. However, this is not the end, as Steam is continually growing. Since the start of our study, Steam has increased the total number of caches they operate from 152 up to 193. Over time, we plan to document and monitor further changes to their infrastructure and the resulting impact.

Other future work also includes monitoring long-term trends of performance by Internet Service Providers and download statistics per country. Although these were not granular enough to include for this study, they could prove valuable for following long-term trends.

References

1. Adhikari, V.K., Guo, Y., Hao, F., Varvello, M., Hilt, V., Steiner, M., Zhang, Z.L.: Unreeling Netflix: Understanding and improving multi-CDN movie delivery. In: 2012 Proceedings IEEE INFOCOM. pp. 1620–1628 (2012). https://doi.org/10.1109/INFCOM.2012.6195531
2. Altomare, F.: Multi Content Delivery Network Explained. https://www.globaldots.com/resources/blog/multi-content-delivery-network-explained (Apr 2021), accessed: 2023-10-31
3. Arsenault, D.: Super power, spoony bards, and silverware: The Super Nintendo Entertainment System. MIT Press (2017)
4. Barcelona, A.: Why and When Do You Need a Hybrid or Multi-CDN Strategy?, https://info.varnish-software.com/blog/why-and-when-do-you-need-a-hybrid-or-multi-cdn-strategy, accessed: 2023-10-31
5. Barman, N., Zadtootaghaj, S., Schmidt, S., Martini, M.G., Möller, S.: GamingVideoSET: A Dataset for Gaming Video Streaming Applications. In: 2018 16th Annual Workshop on Network and Systems Support for Games (NetGames). pp. 1–6 (2018). https://doi.org/10.1109/NetGames.2018.8463362
6. Becker, R., Chernihov, Y., Shavitt, Y., Zilberman, N.: An analysis of the Steam community network evolution. In: 2012 IEEE 27th Convention of Electrical and Electronics Engineers in Israel. pp. 1–5 (2012). https://doi.org/10.1109/EEEI.2012.6377133
7. Bode, K.: Steam Powered - Broadband distribution system to go live | DSLReports, ISP Information. http://www.dslreports.com/shownews/32475 (Sep 2003), accessed: 2023-05-25
8. Böttger, T., Cuadrado, F., Tyson, G., Castro, I., Uhlig, S.: Open connect everywhere: A glimpse at the internet ecosystem through the lens of the Netflix CDN. ACM SIGCOMM Computer Communication Review **48**(1), 28–34 (2018)
9. Carrascosa, M., Bellalta, B.: Cloud-gaming: Analysis of Google Stadia traffic. Computer Communications **188**, 99–116 (2022)
10. Chen, K.T., Huang, P., Huang, C.Y., Lei, C.L.: Game traffic analysis: An MMORPG perspective. In: Proceedings of the international workshop on Network and operating systems support for digital audio and video. pp. 19–24 (2005)
11. Cimpanu, C.: Last week's Fortnite update helped Akamai set a new CDN traffic record . https://www.zdnet.com/article/last-weeks-fortnite-update-helped-akamai-set-a-new-cdn-traffic-record (2019), accessed: 2023-05-25
12. Cisco Umbrella: Cisco Popularity List. https://s3-us-west-1.amazonaws.com/umbrella-static/index.html, accessed: 2023-10-30
13. Citrix: Best practices for evaluating and implementing a multi-CDN strategy. https://www.citrix.com/content/dam/citrix/en_us/documents/white-paper/best-practices-for-evaluating-and-implementing-a-multi-cdn-strategy.pdf (2023), accessed: 2023-10-31
14. Claypool, M., Finkel, D., Grant, A., Solano, M.: Thin to win? network performance analysis of the OnLive thin client game system. In: 2012 11th Annual Workshop on Network and Systems Support for Games (NetGames). pp. 1–6 (2012). https://doi.org/10.1109/NetGames.2012.6404013
15. Collins, B.: What's Caused The Biggest UK Broadband Traffic Spike In The Past Week? 'Call Of Duty,' Not Coronavirus (2020), https://www.forbes.com/sites/barrycollins/2020/03/19/whats-caused-the-biggest-broadband-traffic-spike-in-the-past-week-call-of-duty-not-coronavirus, accessed: 2023-05-25

16. Copeland, B.: The Manchester Computer: A Revised History Part 1: The Memory. IEEE Annals of the History of Computing **33**(1), 4–21 (2011). https://doi.org/10.1109/MAHC.2010.1
17. CTM: About CTM. https://www.ctm.net/en-US/company/about.html, accessed: 2023-10-30
18. Di Domenico, A., Perna, G., Trevisan, M., Vassio, L., Giordano, D.: A Network Analysis on Cloud Gaming: Stadia, GeForce Now and PSNow. Network **1**(3), 247–260 (2021). https://doi.org/10.3390/network1030015, https://www.mdpi.com/2673-8732/1/3/15
19. DiAcetis, S.: Windows 11 cumulative update improvements: an overview. https://techcommunity.microsoft.com/t5/windows-it-pro-blog/windows-11-cumulative-update-improvements-an-overview/ba-p/2842961 (Oct 2021), accessed: 2023-05-16
20. Doan, T.V., Bajpai, V., Crawford, S.: A longitudinal view of Netflix: Content Delivery over IPv6 and Content Cache Deployments. In: IEEE INFOCOM 2020 - IEEE Conference on Computer Communications. pp. 1073–1082 (2020). https://doi.org/10.1109/INFOCOM41043.2020.9155367
21. EA Sports: EA SPORTS FC™ 24, https://store.steampowered.com/app/2195250/EA_SPORTS_FC_24, accessed: 2024-01-11
22. Feng, W.c., Chang, F., Feng, W.c., Walpole, J.: A traffic characterization of popular on-line games. IEEE/ACM Transactions On Networking **13**(3), 488–500 (2005)
23. Galyonkin, S.: Games sales. https://steamspy.com, accessed: 2023-05-27
24. Gamespot: Half-Life 2 now preloading via Steam. https://www.gamespot.com/articles/half-life-2-now-preloading-via-steam/1100-6105848 (2004), accessed: 2023-05-25
25. Gigis, P., Calder, M., Manassakis, L., Nomikos, G., Kotronis, V., Dimitropoulos, X., Katz-Bassett, E., Smaragdakis, G.: Seven years in the life of Hypergiants' off-nets. In: Proceedings of the 2021 ACM SIGCOMM 2021 Conference. pp. 516–533 (2021)
26. Gray, J.: Steam Charts - Tracking What's Played. https://steamcharts.com, accessed: 2023-05-27
27. Gu, L., Jia, A.L.: Player Activity and Popularity in Online Social Games and their Implications for Player Retention. In: 2018 16th Annual Workshop on Network and Systems Support for Games (NetGames). pp. 1–6 (2018). https://doi.org/10.1109/NetGames.2018.8463415
28. Ha, D.H., Silverton, T., Fourmaux, O.: A novel Hybrid CDN-P2P mechanism For effective real-time media streaming. Semantic scholar pp. 1–8 (2008)
29. Hohlfeld, O., Rüth, J., Wolsing, K., Zimmermann, T.: Characterizing a meta-CDN. In: Passive and Active Measurement: 19th International Conference, PAM 2018, Berlin, Germany, March 26–27, 2018, Proceedings 19. pp. 114–128. Springer (2018)
30. Huang, C., Wang, A., Li, J., Ross, K.W.: Understanding Hybrid CDN-P2P: Why Limelight Needs its Own Red Swoosh. In: Proceedings of the 18th International Workshop on Network and Operating Systems Support for Digital Audio and Video. pp. 75–80 (2008)
31. International Olympic Committee: Olympic Games Tokyo 2020 watched by more than 3 billion people - Olympic News. https://olympics.com/ioc/news/olympic-games-tokyo-2020-watched-by-more-than-3-billion-people (Sep 2022), accessed: 2023-05-24
32. Kalata, K.: Hardcore Gaming 101: Phantasmagoria. https://web.archive.org/web/20170109174423/http://www.hardcoregaming101.net/phantasmagoria/phantasmagoria.htm, accessed: 2023-05-26

33. Karami, M., McCoy, D.: Rent to pwn: Analyzing commodity booter DDoS services. Usenix login **38**(6), 20–23 (2013)
34. Kates, J.: Bertie the Brain. Toronto: Canadian National Exhibition (1950)
35. Kim, H., Yang, S., Kim, H.K.: Crime scene re-investigation: a postmortem analysis of game account stealers' behaviors. In: 2017 15th Annual Workshop on Network and Systems Support for Games (NetGames). pp. 1–6 (Jun 2017). https://doi.org/10.1109/NetGames.2017.7991540, iSSN: 2156-8146
36. Lin, D., Bezemer, C.P., Hassan, A.E.: Studying the urgent updates of popular games on the Steam platform. Empirical Software Engineering **22**, 2095–2126 (2017)
37. Lumen: How does a multi-CDN strategy work for my business? `https://assets.lumen.com/is/content/Lumen/how-does-a-multi-cdn-strategy-work-for-my-business`, accessed: 2023-10-31
38. Medianova: Hybrid CDN: A combination of Private and Multi-CDN!, `https://www.medianova.com/hybrid-cdn-a-combination-of-private-and-multi-cdn`, accessed: 2023-10-31
39. Microsoft: Security Update Guide FAQs. `https://www.microsoft.com/en-us/msrc/faqs-security-update-guide`, accessed: 2023-05-24
40. Microsoft: Download Windows 11. `https://www.microsoft.com/software-download/windows11` (2023), accessed: 2023-05-24
41. Microsoft: Microsoft by the Numbers. `https://news.microsoft.com/bythenumbers/bythenumbers/en/windowsdevices` (2023), accessed: 2023-05-24
42. Netflix: How to control how much data Netflix uses. `https://help.netflix.com/en/node/87`, accessed: 2023-05-24
43. Netflix: Requirements for deploying embedded appliances. `https://openconnect.zendesk.com/hc/en-us/articles/360034538352-Requirements-for-deploying-embedded-appliances` (2023), accessed: 2023-05-24
44. Ni, J., Tsang, D.H., Yeung, I.S., Hei, X.: Hierarchical content routing in large-scale multimedia content delivery network. In: IEEE International Conference on Communications, 2003. ICC'03. vol. 2, pp. 854–859. IEEE (2003)
45. Niko: Game regulations in China: everything you need to know. `https://nikopartners.com/game-regulations-in-china-everything-you-need-to-know` (2020), accessed: 2023-10-31
46. O'Neill, M., Vaziripour, E., Wu, J., Zappala, D.: Condensing Steam: Distilling the diversity of gamer behavior. In: Proceedings of the 2016 internet measurement conference. pp. 81–95 (2016)
47. Orland, K.: Despite 100GB video games, average download times are decreasing. `https://arstechnica.com/gaming/2020/06/ars-analysis-were-spending-less-time-downloading-games-on-average` (2020), accessed: 2023-05-24
48. PeeringDB: AS32590 - Valve Corporation. `https://www.peeringdb.com/net/4782`, accessed: 2023-09-18
49. PeeringDB: PeeringDB. `https://www.peeringdb.com`, accessed: 2023-09-18
50. Perry, T.S., Wallich, P.: Microprocessors: Design case history: The Atari Video Computer System: By omitting lots of hardware, designers added flexibility and gave video-game programmers room to be creative. IEEE Spectrum **20**(3), 45–51 (1983). https://doi.org/10.1109/MSPEC.1983.6369841
51. Prather, J., Nix, R., Jessup, R.: Trust management for cheating detection in distributed massively multiplayer online games. In: 2017 15th Annual Workshop on Network and Systems Support for Games (NetGames). pp. 1–3 (Jun 2017). https://doi.org/10.1109/NetGames.2017.7991547, iSSN: 2156-8146

52. RETN: About RETN. https://retn.net/about, accessed: 2023-10-30
53. Steam: Steam - 2021 Year in Review. https://store.steampowered.com/news/group/4145017/view/3133946090937137590 (Mar 2022), accessed: 2023-05-24
54. Steam: 2023 Official Steam Sales and Fests Schedule (2023), https://store.steampowered.com/news/group/4145017/view/3645136358931000680, accessed: 2024-01-11
55. Steam: News - Steam Client Update Released. https://store.steampowered.com/oldnews/191068 (2023), accessed: 2024-03-11
56. Steam: Steam Year In Review 2022. https://store.steampowered.com/news/group/4145017/view/3677786186779762807 (Feb 2023), accessed: 2023-05-25
57. Steam Community: Cyberpunk 2077: Phantom Liberty (2023), https://steamcommunity.com/games/1091500/announcements/detail/3725096176290982673, accessed: 2024-01-11
58. Steam Community: Cyberpunk 2077: Update 2.0 (2023), https://steamcommunity.com/games/1091500/announcements/detail/3725096176290982673, accessed: 2024-11-01
59. Steam Support: Documentation Home Page. https://partner.steamgames.com/doc/home (2023), accessed: 2023-05-27
60. SteamDB: Patchnotes, https://steamdb.info/patchnotes, accessed: 2024-01-11
61. SteamDB: Baldur's Gate 3 - SteamDB. https://steamdb.info/app/1086940/depots (2023), accessed: 2023-09-31
62. SteamDB: SteamDB. https://steamdb.info (2023), accessed: 2023-05-27
63. SteamDB: Steamtracking - CellMap.vdf (2024), https://github.com/SteamDatabase/SteamTracking/blob/0aac5ea77a74acd76194b7b8eaca43ecd6c997d6/ClientExtracted/steam/cached/CellMap.vdf, accessed: 2024-01-11
64. Stoll, J.: Netflix library by country for March 2023. https://www.statista.com/statistics/1013571/netflix-library-size-worldwide (2023), accessed: 2023-05-16
65. Stoll, J.: Netflix: number of subscribers worldwide 2023. https://www.statista.com/statistics/250934/quarterly-number-of-netflix-streaming-subscribers-worldwide (Apr 2023)
66. Tullemans, B.: CDN Architectures Demystified. https://tech.ebu.ch/publications/tr068 (2022), accessed: 2023-10-31
67. Valve: SteamCMD: - Valve Developer Community. https://developer.valvesoftware.com/wiki/SteamCMD#Linux, accessed: 2023-09-20
68. Valve: Game and Player Statistics (2023), https://store.steampowered.com/stats/content, accessed: 2023-05-16
69. Valve: Steam Hardware & Software Survey (2023), https://store.steampowered.com/hwsurvey, accessed: 2023-05-16
70. Valve: Steam Search. https://store.steampowered.com/search?ndl=1 (2023), accessed: 2023-05-24
71. Valve: Steam Web API - Valve Developer Community. https://developer.valvesoftware.com/wiki/Steam_Web_API#GetNewsForApp_.28v0001.29 (2023), accessed: 2023-05-25
72. Vecima: Hybrid CDN — Vecima Networks, Inc. https://vecima.com/hybrid-cdn, accessed: 2023-10-31
73. Visser, C., Fontugne, R.: Steam Content Delivery Network Study (2024), https://belthazaar.github.io/ExploringSteamPage
74. Wang, P.C., Ellis, A.I., Hart, J.C., Hsu, C.H.: Optimizing next-generation cloud gaming platforms with planar map streaming and distributed rendering. In: 2017

15th Annual Workshop on Network and Systems Support for Games (NetGames). pp. 1–6 (Jun 2017). https://doi.org/10.1109/NetGames.2017.7991544, iSSN: 2156-8146
75. Xu, D., Kulkarni, S.S., Rosenberg, C., Chai, H.K.: Analysis of a CDN-P2P hybrid architecture for cost-effective streaming media distribution. Multimedia Systems **11**, 383–399 (2006)
76. Zhao, S., Abou-zeid, H., Atawia, R., Manjunath, Y.S.K., Sediq, A.B., Zhang, X.P.: Virtual reality gaming on the cloud: A reality check. In: 2021 IEEE Global Communications Conference (GLOBECOM). pp. 1–6. IEEE (2021)