# ARCHITECTURES AND PROTOCOLS POWERING ILLEGAL CONTENT STREAMING OVER THE INTERNET

D. Leporini

Viaccess-Orca, France

## ABSTRACT

Over recent years, a major shift has occurred in piracy of paid-for content services toward illegal redistribution of live content in real-time over the Internet. This paper will provide insight into pirate content platforms, covering the various architectures and protocols used, from peer-to-peer protocols adapted for live streaming to more traditional Web streaming protocols. More specifically, it will focus on the methods generally employed to set up and scale ad-based illegal services using some of the above-mentioned protocols with streaming media platforms, while securing streaming servers, enabling these sites to remain hidden. A thorough analysis of the used architectures and protocols makes it possible to measure the actual audience viewing illegal streams, typically leveraging peer-to-peer networks data. This enables content service providers to assess the piracy threat level of any content, while illustrating the need for a business intelligence tool that provides relevant information on viewers' behavior.

## INTRODUCTION

The history of pay-TV [1, 2, 3], considering the business at stake, is unsurprisingly tightly coupled with the history of content services piracy, effectively proving the saying that "security is a process, not a product" in this industry. Content services piracy has evolved year after year, mainly adapting to both the solutions developed by content security vendors and technologies available to circumvent them while offering an alternative solution intended to generate parallel business. Piracy forms have ranged from video receivers and smartcard piracy, to the sharing of subscription rights through service access credentials (login and password) or conditional access smartcard sharing to content decryption keys (known as Control Words) redistribution both over the Internet and satellite feeds to cover wide distribution regions. Those Control Words feeds have even been delivered over the very satellites whose capacity was legitimately used by operators to broadcast their channel signals.

A major shift in paid-for content piracy has occurred in recent years. There has been, in particular, an increase toward illegal direct redistribution of live content in real-time over the Internet, most notably for content with very high "live value" such as sports events.

## THE PIRACY LANDSCAPE

Illegal live content redistribution over the Internet has been following two main distinct approaches: peer-to-peer (P2P) live streaming and Web streaming with and without the use of Content Delivery Networks (CDN).

### Peer-to-peer streaming

The origin of P2P as a technology dates back from the early days of Napster in 1999: people used their "own" bandwidth to share content (music back then) long before CDNs had become the standard for content delivery.

The origin of P2P streaming, although based on the same idea of limiting servers' bandwidth, is somewhat unclear. The Chinese P2PTV protocols (e.g., PPLive or PPStream) [4] began to deliver content in the second part of the 2000 decade. Most of these protocols were real-time-enabled BitTorrent derivatives. These protocols have evolved, with new improved implementations having emerged, paving the way for today's landscape.

Different P2P streaming platforms are used nowadays, with the two main ones being SopCast [5, 6] and AceStream (formerly known as TorrentStream). SopCast is a Chinese university project that became a widespread piece of software, loosely maintained. AceStream is certainly the latest and most innovative method of P2P streaming, apparently maintaining a legitimate side to its operations. Other P2P protocols have emerged and gained a certain level of exposure, most notably BitTorrent Live announced by BitTorrent Inc. at the Consumer Electronics Show (CES) in 2012, until a complete shutdown of the trials in February 2014. BitTorrent Live's new target seems to be mobile live broadcast through P2P.

P2P streaming principles are similar to P2P file exchange ones: P2P users sharing the same content form a loosely connected mesh network compared with a full mesh network where all peers are connected to all other peers. This structure gives the P2P network reliability and resilience. If one network node stops working, the remaining nodes can still work together, provided they are able to reconnect, if needed.

The benefits of P2P streaming, from a viewers' standpoint, are essentially twofold. Firstly, the size of a P2P network is virtually "limitless". The capacity to deliver content to a significant amount of viewers, widely acknowledged for P2P file sharing, therefore also applies to live streaming P2P protocols. Some of our recent measures over such P2P streaming networks indeed confirm audiences of over 30,000 viewers on each of selected streams on a regular basis. Secondly, the quality of streamed content is usually noticeably higher in terms of achievable bitrate. Whereas the majority of direct Web streaming bitrates are lower than 600Kbps, our knowledge base shows that 60 percent of the P2P streams have bitrates below 2Mbps, 30 percent between 2Mbps and 4Mbps, and the remaining streams with bitrates above 4Mbps.

The average ratio of P2P streams over all available streams for any given event still remains low as of today (on average less than 10 percent), and P2P streaming usage tracking does not turn out to be necessarily straightforward since P2P streaming protocols are generally closed-source. However, such a tracking does represent a tremendous opportunity for business intelligence purposes. P2P users in such streaming networks are not anonymous, and a lot of information can be collected and analyzed, including streaming usage data, audience measurement information in real-time, and user

geolocation for any monitored event. The rest of the paper focuses on Web streaming protocols and architectures that represent the vast majority of live content illegal redistribution at stake.

**Direct Web streaming**

Contrary to P2P streaming, direct Web streaming is point-to-point in nature, meaning all users are provided with a separate stream by the video streaming architecture (either via a single streaming server or a CDN). The server bandwidth consumption, along with the related costs, increase linearly with the number of users. It is worth noting that most of the time CDNs are incorrectly and misleadingly referenced as regular streaming servers in the context of illegal content streaming. Indeed, the use of real CDNs, although occasionally occurring, is certainly not the standard case today as far as pirated content redistribution architectures are concerned. Using a legitimate CDN could turn out to be a vulnerability for pirates as mainstream CDNs like the ones provided by Akamai and Amazon would enforce Digital Millennium Copyright Act (DMCA) requests and thus shutdown the whole pirate infrastructure swiftly, should a complaint by right owners be issued.

Consequently, the model for illegal direct Web streaming of live content over the Internet is typically multi-tiered. Content discovery platforms first promote pirated content. Such promotion happens through event-exposing websites known as “link farms” acting as an online Electronic Program Guide (EPG), or through link-exposing social networks such as Twitter or Facebook. Actual content redistribution then occurs through streaming platforms delivering video content over the Internet to a variety of devices. Hundreds of such illegal content platforms are currently active, with legitimate websites such as youtube.com, justin.tv, or ustream.tv remaining exceptions, despite the possibility for premium accounts to lift some limitations to the streaming regarding the number of simultaneous viewers and/or the bitrate of the stream. In other words, most of the time, the streaming infrastructure is pirate.

Such discovery and content platforms for illegal streaming of live sport events are very popular. They happen to be extremely easy to find through popular search engines and social networks. For example, a typical search on Google with simple keywords like “live streaming football” will return millions of pages. Most importantly, the top 20 results from Google all correspond to relevant results.

Illegal content platforms are, in turn, popular and numerous because they are easy to setup and can generate high advertising revenues. Associated risk for the platform is relatively low and distributed among the various layers of the distribution model, which include the “link farm” website advertising links to the video streams, the streaming platform, and the streaming server. Pirates are generally moving away from using mainstream streaming platforms for several reasons:

- Mainstream platforms typically enforce DMCA rights and will allow copyright holders to remove content and close channels very quickly.
- Pirates will not benefit as much from advertising as they would if they ran their own ad-supporting streaming platform.

**Audience measurement and risk assessment**

The audience of illegal streams, possibly coupled with the geolocation of viewers in areas of interest, is one of the key indicators for content service providers as far as business risk

assessment is concerned. Some relevant proxy metrics include the number of illegal streams available along with the number of “link farms” and content platforms for a given content. Video stream bitrates and actual quality of experience typically complement such risk assessment metrics.

As mentioned above, the audience on P2P streaming networks can actually be measured when proper technologies and tools are in place. An overall estimation of audience for a given event typically combines an accurate measurement for P2P streams relying on SopCast and/or AceStream P2P streaming platforms with actual measurements or estimations from streams distributed from content platforms. Measurements rely on dedicated counters available on certain websites and audience per stream provided by some service portals. Estimations make use of a combination model taking into account various parameters, including website ranking (e.g., Alexa and information from ISPs and CDNs), network caching data, popularity statistics (e.g., type of sports, championship, game, etc.), and social networks statistics (e.g., popularity and geolocation based on Twitter keywords and hashtags used during an event).

Figure 1 shows an example of audience estimation for the 2015 ICC Cricket World Cup final. The cumulated audience for illegal streams of the top content platform exceeded 10 million viewers during this event.


Figure 1 – 2015 ICC Cricket World Cup audience measurement.

## ANALYSIS AND INSIGHTS ON ILLEGAL CONTENT REDISTRIBUTION

In the previous section, various architectures and protocols for illegal content redistribution over the Internet have been introduced, from P2P protocols adapted for live streaming to

more traditional Web streaming protocols. This section presents a more detailed analysis of a typical pirate streaming platform.

### Content platform perspective

A significant number of content streaming platforms are available today for anyone willing to illegally stream live content (e.g., mips.tv, leton.tv, jjcast.com, hdcast.org, etc.). Streaming content to any of these platforms simply requires a video capture device along with software that is compatible with protocols like Real-Time Streaming Protocol (RTSP) and Real-Time Messaging Protocol (RTMP) used by Flash players.

Software like the one provided by the Open Broadcaster Software Project [7] is compatible with all DirectShow capture devices and allows to stream captured content to any RTMP-compatible server (e.g., YouTube, Dailymotion and any Wowza Media streaming servers [8]). Alternatively, the popular VLC player and Adobe Flash Media Live Encoder can achieve the same results.

Content streaming itself then becomes as easy as entering a channel name on the chosen streaming platform, and then pointing the video capture software to the indicated streaming URL. Upon creation of a channel, the provided Web link or HTML code snippet for the Flash video player can then be embedded by the streamer in a website in order to advertise the corresponding channel. People behind the streaming platform and websites containing the corresponding links to the streams (“link farms”) can obviously be the same.

### Viewer perspective

Watching a live stream redistributed illegally over the Internet in general does not require a direct knowledge of the back-end streaming platforms. Video players are directly embedded on websites that act as EPGs of illegal streams. As previously noted, such “link farms” can easily be found through Google or advertised on Facebook or Twitter.

### “Behind the scenes” perspective

An in-depth analysis of several illegal streaming platforms reveals the following generic patterns and insights.

Streaming platforms possess a public-facing Internet storefront with a well-identified domain name. The corresponding domain names are often protected by “whois privacy” for obvious privacy reasons and when not, the information contained in the “whois” records are generally fake.

Many streaming platforms are using proxy and Denial of Service (DoS) protection services such as CloudFlare [9], thereby making the real IP address of the platform difficult to obtain. Such services slow down legal processes and prevent copyright holders from immediately identifying the hosting provider used. When not using CloudFlare to hide the hosting servers, offshore hosting companies (called “bulletproof” hosting companies) are used by pirates due to their claim to ignore DMCA takedown notices or abuse reports. However, the cost for such hosting services is usually two to five times higher than traditional hosting. Some of these hosting companies accept bitcoin as a payment means.

Administration tasks are made through a simple Web interface. The whole service can be run by a single person, typically a skilled Linux administrator.

Back-end servers used to stream content to end-users generally rely on streaming servers such as Wowza Media Systems servers. In that context, content is first published on one of a few Wowza Media publishing servers and then duplicated on multiple Wowza streaming servers. Such streaming servers are typically hosted over multiple hosting providers and quickly deployed using bash scripts. This setup allows to simultaneously serve thousands of streams and can be scaled quickly and easily by increasing the number of back-end servers. Despite being relatively amateur, the infrastructure works properly enough to achieve the objectives it intends to meet.

The business model of such content streaming platforms is to gain revenue through advertising. Ads are displayed as an overlay on the Web player. After a certain period of time (typically between 10 to 20 seconds), ads can be hidden to watch the stream. It should be noted that the majority of ads displayed are used to distribute malware and adware (e.g., using fake messages like “You need to install Flash Player HD”).

For example, let’s consider the case of the “leton.tv” content platform [10]. Leton.tv uses several other linked domain names and websites with different layouts. All such linked domains however share the same infrastructure as leton.tv. The different websites may make stream takedowns more difficult. Moreover, having multiple domains also results in better ranking on “link farms” and increases the number of viewers. Such viewers are typically balanced over tens of active streaming servers. Matomy Media Group, using ad120m.com and rev2pub.com, is one the main ad providers used by leton.tv. It is worth mentioning that all ads served on ad120m.com are linked to viruses and malwares when searched on Google.

## CONCLUSION

This paper has presented the various architectures and protocols powering today’s content service piracy shift toward illegal redistribution of live content over the Internet. Most notably, it has provided insights on the typical simple Web streaming infrastructures that enable both the easy setup and the scaling of the streaming servers architecture to serve in real-time thousands of illegal video streams to hundreds of thousands of viewers. The rapid evolution of open technologies available and paradigms for content discovery and consumption therefore calls for efficient and reinforced ongoing monitoring and investigation approaches on the value chains used for illegal access to and delivery of content services. From content acquisition, preparation and distribution, to Web hosting and ad-based or subscription-based business model support, providing awareness to all players involved in such a value chain on the role they actually play becomes of paramount importance.

## REFERENCES

[1] Digital Video Broadcasting (DVB), A Guideline for the Use of DVB Specifications and Standards, DVB Document A020 Rev. 1, May 2000.

[2] Digital Video Broadcasting (DVB), Support for use of scrambling and Conditional Access (CA) within digital broadcasting systems, ETSI Technical Report ETR 289, European Telecommunications Standards Institute ETSI, October 1996.

[3] Tong Ho, “Digital Video Broadcasting Conditional Access Architecture”, A Report Prepared for CS265-Section 2, 2002 (http://www.cs.sjsu.edu/~stamp/CS265/projects/papers/dvbca.pdf).

[4] Akos Horvath, Miklos Telek, Dario Rossi, Paolo Veglia, Delia Ciullo, Maria Antonieta Garcia, Emilio Leonardi, Marco Mellia, “Dissecting PPLive, SopCast, TVAnts”, 2008 (http://www.napa-wine.eu/twiki/pub/Public/DocumentsOld/napa_techrep1.pdf).
[5] Benny Fallica, Yue Lu, Fernando Kuipers, Rob Kooij, Piet Van Mieghem, “On the Quality of Experience of SopCast”, 2007 (https://www.nas.ewi.tudelft.nl/people/Fernando/papers/sopcast.pdf).
[6] Alex Borges Vieira, Ana Paula Couto da Silva, Francisco Henrique,Glauber Goncalves, Pedro de Carvalho Gomes, “SopCast P2P Live Streaming: Live Session Traces and Analysis”, MMSys ’13, February 26-March 1, 2013 (http://netlab.ice.ufjf.br/publications/2013/mmsys_12.pdf).
[7] http://obsproject.com/
[8] http://www.wowza.com/
[9] http://www.cloudflare.com
[10] http://www.leton.tv