RWW: World Wide Web Network Services Infrastructure

Decentralised Cloud Storage, Content Distribution, and Knowledge Economics

Overview

Rich Media Distribution accounts for a significant portion of Internet. These systems transport rich media in a variety of formats, encapsulated as media files or other formats, as to enable distribution across the web. In the SandVine report for the first half of 2014[1], The Peak Period Traffic Composition of Internet use in North America shows significant percentages of Real-Time Entertainment, File-sharing and other activities both upstream and downstream accounting for Peak-Period internet Use in the Region. These statistics are then backed-up by the same report, outlining the share of web-traffic accounted for by Bit-Torrent[2] and an array of popular Internet Media Services.

These statistics show the relatively high demand or use or Internet for Media.

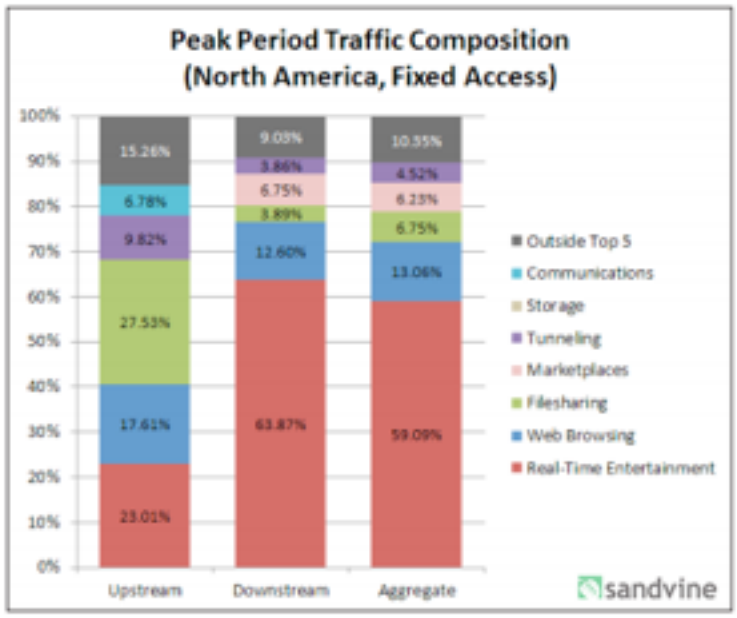

Figure 1 - Peak Period Aggregate Traffic Composition - North America, Fixed Access

Web 2.0[3] furthered internet website technology to provide an array of enhancements made possible by scripting[4] and RDBMS[5] Connectivity. These systems were first used for mainframe, workstation and local area networking. As Web 2.0 systems developed, these technologies were then applied to cre-

1 https://www.sandvine.com/downloads/general/global-internet-phenomena/2014/1h-2014-global-internet-phenomena-report.pdf

2 http://en.wikipedia.org/wiki/BitTorrent

3 http://en.wikipedia.org/wiki/Web_2.0

4 http://en.wikipedia.org/wiki/Scripting_language

5 http://en.wikipedia.org/wiki/Relational_database_management_system

	Upstream		Downstream		Aggregate	
Rank	Application	Share	Application	Share	Application	Share
1	BitTorrent	24.53%	Netflix	34.21%	Netflix	31.09%
2	HTTP	14.27%	YouTube	13.19%	YouTube	12.28%
3	SSL	6.54%	HTTP	11.65%	HTTP	11.84%
4	Netflix	6.44%	iTunes	3.64%	BitTorrent	5.96%
5	YouTube	5.52%	SSL	3.42%	SSL	3.80%
6	Skype	2.23%	BitTorrent	3.40%	iTunes	3.33%
7	Facebook	2.17%	MPEG	2.85%	MPEG	2.62%
8	FaceTime	1.50%	Facebook	1.99%	Facebook	1.83%
9	Dropbox	1.20%	Amazon Video	1.90%	Amazon Video	1.82%
10	iTunes	1.15%	Hulu	1.74%	Hulu	1.58%
		64.40%		76.24%		74.58%

sandvine

Table 2 - Top 10 Peak Period Applications - North America, Fixed Access

ate 'web applications'[6] used on the World Wide Web[7]. In the Late 1990's, Tim Berners-Lee[8] and others, started reinventing[9] these systems. This course produced the 'Semantic Web'[10] concept. Incorporating a new type of database methodology, embedded within the concept of Linked-Data[11]. Considerations surrounding the application of these technologies have been examined for over a decade for an array of applications, including cloud-storage for the social-web[12]. A recent PHD Thesis by Andrei Sambra[13], under the direction of Tim Berners-Lee, examines data Ownership and Interoperability for a Decentralised Social Semantic Web. This paper outlines

[6] http://en.wikipedia.org/wiki/Web_application

[7] http://en.wikipedia.org/wiki/World_Wide_Web

[8] http://www.w3.org/People/Berners-Lee/

[9] http://www.w3.org/DesignIssues/

[10] http://www.w3.org/DesignIssues/Semantic.html

[11] http://www.w3.org/DesignIssues/LinkedData.html

[12] http://www.w3.org/DesignIssues/CloudStorage.html

[13] http://myprofile-project.org/thesis/manuscript_en.pdf

considerations relating to Privacy, Access Control and decentralised storage as is exampled by the works produced in relation to the thesis rww.io[14].

Rww.io is an example of a linked-data platform. The system is designed to authenticate a user (person / machine) using an X.509v3 Certificate (WebID-TLS[15]) which in-turn authorises access to materials stored on that server.

The Server itself is an open-source implementation, which can support a multitude of users and applied business systems. The server implementation incorporates the use of SPARQL[16], JSON-LD, RDF and an array of read/write web functions (DAV, AJAX, JSONP, CORS).

Existing Research and Development has very much focused on user-interaction with these systems as a means to support decentralisation of user-data, from traditional Web 2.0 application, user-experiences. These new-forms of 'cloud storage' models support the development of applications whereby the application provider does not store user-data, but rather references the user-data-storage location.

The Era of Web 3.0 Technology has been contemplated by Jeff Sayre in his paper "Web 3.0: Powering Startups to Become Smartups"[17]. As noted by Jeff in this paper, the term "smart-ups" was originally coined by Kingsley Idehen.

Web 3.0 denotes the use of Semantic Web Technology; where structured data is used within web-documents to make declarations of meaning for information entered into web-documents, in such a way that the web itself becomes a database.

---

[14] http://rww.io/

[15] http://www.w3.org/2005/Incubator/webid/spec/tls/

[16] http://www.w3.org/TR/sparql11-query/

[17] http://jeffsayre.com/2010/09/13/web-3-0-powering-startups-to-become-smartups/

The culmination of these concepts have traditionally been considered in terms of the 'social web'[18]. Indeed, Tim Berners-Lee's Paper "Socially Aware Cloud Storage"[19] documents the concept of "social network silos" which when combined with Read-Write Linked-Data[20] created an established way of thinking surrounding the use of Read-Write Web Cloud Storage Platforms, extending the standards documentation to the creation of "Linked Data Platform"[21] from an initial standpoint of considerations surrounding the Social Web.

Content Distribution Services

Content Distribution Services or "CDN's" have traditionally involved a syndicated system of servers and software, created to distribute content in large volumes across the web to a plurality of users. The Basic Concept of a CDN is that it transports web-resources from a point of origin, through a distribution framework, that then enables end-users to access those resources from a location that is local to that user. This has the Network Effect of distributing requests across a multitude of server-locations whilst facilitating access to particular resources by particular users from nodes within this distribution service, that are most effective for that user to access.

A Simple example would be that if an Australian User utilised a resource from an American Website, that resource may be forwarded by the CDN to an Australian location, so that when the

[18] http://en.wikipedia.org/wiki/Social_web

[19] http://www.w3.org/DesignIssues/CloudStorage.html

[20] http://www.w3.org/DesignIssues/ReadWriteLinkedData.html

[21] http://www.w3.org/TR/ldp/

Australian User sought to use that resource - the data-transfer between the Australian User and the resource would be facilitated via national networks.

RWW / LDP - Innovation Summary

The Principle concept is to apply RWW Cloud Server Technology for the purpose of CDN Deployment. This in-effect has the capacity to create a virtualised 'web-storage' platform whereby authenticated users can store data on the cloud, in a decentralised manner using Access Control whilst optimising the Web for Content Services.

Background

After many years working on Video On Demand Systems, a current project required investigation into the delivery of a Hybrid TV Platform, incorporating IPTV capabilities. These types of systems have very high throughput and data requirements. To facilitate these technical requirements infrastructure is deployed throughout the Internet Network, at scale, often including to Local Nodes such as DSLAM Locations as to facilitate high-throughput for connected subscribers incorporating QOS capabilities in relation to the delivery of a Infrastructure Grade Service. These systems have in-turn required operators to build and manage this sophisticated Infrastructure.

Proposal

Internet Service Providers ("ISP" / "ISP's") have engaged these network service providers in an array of ways, including through the provision of Rack-Space, Power and Connectivity to CDN providers, who in-turn manage the content services platform. This business

model often results in difficulties due to the ISP Infrastructure Managers (ISP's) providing the Infrastructure service to the CDN service provider, who in-turn manages the maintenance requirements of the CDN Equipment.

I then considered that perhaps, rather than acquiring and deploying this infrastructure solution for deployment into ISP Networking Infrastructure, Perhaps it is possible to refactor the use-case around RWW Cloud Storage Platforms as to provide this software platform to ISP's who can then, in-turn, internally manage the CDN Infrastructure. Further considering the opportunity, i then realised that whilst the initial business case for exploring this technical design / business opportunity was to lower the effective cost for Media Service Providers ("MSP's") in deploying infrastructure grade service - the methodology is not exclusively tied to the service provision of any one particular 'vendor', but rather, a cloud-services platform that is capable of virtualising the web from its existing state of 'silos' to 'smart storage'.

In one embodiment; the software embodiment is installed from an open-source package. This software allows a service-operator to prioritise the content stored on the network subject to the network load relating to the usage of assets over internet connected networks. In this way, "Web 3.0" software (including media files, as an example) can be cached on the network to the local node based on usage, or other prioritisation factors managed by the network provider.

FileSystem storage may include;

- Personal Files: for example, a user who works on the web from home, may utilises a specific exchange the majority of the time. Therein, the ability to store data at that exchange (rather than providing transit to an alternative peer) may increase performance whilst lowering cost for that users cloud storage Requirements.

- IPTV; An IPTV network is defined as an Internet Video Service that incorporates Quality of Service (QOS) Requirements, as to ensure the network performance is appropriately managed; in a way, for example, that ensures availability and managed buffering for usage of a qualified service. In this way, the services-platform may be provided by the ISP as to support enhanced services in a manner that was otherwise either unobtainable or more costly.

Economically, ISP's often manage their own infrastructure. It becomes expensive to engage external providers to manage what is arguably a practical requirement of ISP infrastructure. Expense drivers include network design, available co-location or rack-space (inclusive of power, etc.) in any specific environment (ie: inclusive of telecommunications exchanges) which may or may not have resources available to best cater for all network opportunities, as driven by a customer base.

By offering an open-source system that is capable of providing Access Control upon the Content Services Layer, rather than the equipment layer, it may be possible to create an abstraction between tele housing data for the purpose of network engineering and use-case fulfilment and the ownership of the content tele housed within said environment.

The Web is a Peer-to-Peer based system. Initial transit models over Internet Protocol facilitated the 'peer-to-peer' networking capability via mapping addresses, in-turn creating a carriage service, without inter-mediatory storage. As the Internet exponentially grew, utility problems emerged through the lack of bandwidth between points (ie: international carriage) as required for every request from one source-point, to another; related, was the issue of centralised resourcing and the ability for a system to facilitate access simultaneously to more individual requests than any single computing device was capable of servicing alone.

In the initial stages of developing solutions for this problem, systems called 'cache servers'[22] were implemented, which created rule-sets, allowing content to be stored in a location closer to any requesting entity; which in-turn reduced the requirement for accessing content from an original source location.

As Web 2.0 solutions emerged, furthering online use of rich-media and database driven web-systems, more advanced, yet privately managed systems were further developed; these systems were called "Content Distribution Networks" or 'CDN's'[23].

CDN's changed the way websites were able to manage the distribution of their content, with the interest of controlling use of that content within their own systems or 'web portals' as a medium of intellectual property management, offering enhancements to quality of service without the need to deploy and manage hardware infrastructure throughout the region of use. The other effect surrounding the business model of a CDN was that rather than the ISP managing their own cache systems (as required prior to the advent of CDN's) which cost the ISP funds, in exchange for network optimisation benefits; the CDN Businesses PAID the ISP's to manage new, more sophisticated means of managing this content networking issue on the behalf of websites, rather than ISPs.

CDN's and related infrastructure are still fundamentally based upon the use of centralised relational database systems. Inbuilt into the design, are factors relating to the architecture of these relational database technology which are different to the traits of graph database technology; across many levels, including that of security.

[22] http://en.wikipedia.org/wiki/Web_cache

[23] http://en.wikipedia.org/wiki/Content_delivery_network

Within the sphere of RDF, RWW, TURTLE, WebID-TLS, and JSON-LD based decentralised web-technologies, new paradigms have been developed where the principle focus technically, has been formatted to consider economics and the social-web. Problematically, these infrastructure solutions, that have much greater capacities to offer communities or individual entities a new means to manage data-rights[24], have suffered from a lack of deployment solutions, regardless of the benefits made available to developers should these systems become more easily usable by a sufficiently resource critical mass of users who in-turn are made capable of better controlling their own data; even where their influence of control, is to provide constituents or embodiments of their information uniformly or freely.

These technologies can make distinctions between a reference to a persons data, and a contribution of a person or entity; or embodiment through accumulative configuration and presentation, of a plurality of contributors concepts; into a unique format, that may be its own embodiment, as distinct from each and every constituent, without prejudice.

By adapting the before mentioned technologies, into a networking platform made available to infrastructure operators (such as ISP's and Data-centre locations), the capacity to further virtualise 'cloud storage' has distinct opportunities to provide remedy and/or opportunity, to a variety of problems (on a variety of levels); including but not exclusive to, initial platform frameworks required for the emerging 'knowledge-economy', establishing capacities for such systems to develop, with discourse through socio-economic frameworks, as a predicate to further application innovation, utility and related economic growth / resourcing measures.

[24] http://lists.w3.org/Archives/Public/public-webpayments/2014Jul/0043.html

Economic factors are an important manifest function of human sociology, or perhaps more succinctly - how the world works. Web-Payments[25] and Credentials[26] W3 Community Groups[27] [28] are actively working on systems capable of servicing economic solutions surrounding the application of these types of 'cloud storage' platforms, in the manner described by this paper.

Conclusion

There is a meaningful opportunity to refactor the RWW Social-Web Cloud-Storage Platforms into an open-source services platform that is capable of applying the same techniques, for network services, in a manner that does not simply service the needs of network providers but also - end-users. To do this, it would be best if a decentralised, cooperative framework of collaborators could work together, with transparency.

[25] https://web-payments.org/

[26] http://opencreds.org/

[27] http://lists.w3.org/Archives/Public/public-webpayments/

[28] http://lists.w3.org/Archives/Public/public-credentials/