ECHOS: Edge Capacity Hosting Overlays of Nano Data Centers

Nikolaos Laoutaris
Telefonica Research
nikos@tid.es

Pablo Rodriguez
Telefonica Research
pablorr@tid.es

Laurent Massoulie
Thomson Research
laurent.massoulie@thomson.net

This article is an editorial note submitted to CCR. It has NOT been peer reviewed

Authors take full responsibility for this article's technical content. Comments can be posted through CCR Online

ABSTRACT

In this paper we propose a radical solution to data hosting and delivery for the Internet of the future. The current data delivery architecture is "network centric", with content stored in data centers connected directly to Internet backbones. This approach has multiple drawbacks among which complexity of deploying data centers, power consumption, and lack of scalability are the most critical ones. We propose a totally innovative and orthogonal approach to traditional data centers, through what we call "nano" data centers, which are essentially boxes deployed at the edge of the network (e.g., in home gateways, set-top-boxes, etc.) that cooperate in a peer-to-peer manner. Unlike traditional peer-to-peer clients, however, our nano data centers operate under a common management authority, e.g., the ISP who installs and maintains the set-top-boxes, and can thus cooperate more effectively and achieve a higher aggregate performance. Nano data centers are, therefore, better suited for providing guaranteed quality to new emerging applications such as online gaming, interactive IPTV and VoD, and user generated content.

Categories and Subject Descriptors

C.2.4 [**Distributed Systems**]: Distributed applications

General Terms

Design

Keywords

data center, peer-to-peer, set-top-box

1. INTRODUCTION

With the rise of peer-to-peer (P2P) applications at the beginning of the decade, multimedia content including movies and music has quickly become the dominating traffic contributor on the Internet.[1] For this reasons many have started seeing the Internet not only as a communication network, but also as a content dissemination network in which "Content is King" [12]. This seems to be a persisting trend as in the last couple of years several new multimedia applications have pushed the importance of content even further.

[1]http://www.cachelogic.com/

1.1 Content is Emperor

New types of multimedia content have gaining popularity next to traditional P2P content. Examples include:

Online multiplayer games: Online gaming has had a formidable success, with some games pulling in several million players, such as World of Warcraft (WoW) which has 8 million subscribers, up to half a million of them could be playing simultaneously. In 2006, WoW generated 750 million euros worth of revenue [1]. The vast scale of online gaming has brought new system design challenges. Today, online games are centralized and thus suffer from the well known problems of robustness and scalability. They also require a huge investment in terms of equipment and maintenance (WoW requires 1700 employees [1]).

Personalized TV & Video-on-Demand: Another increasingly important application is IP-TV and Video-on-Demand that supports real time interactive user behaviors. While there are a handful of commercial such systems,[2] the current client-server architectures are inherently not scalable and too expensive in bandwidth resources.

User Generated Content: User Generated Content (UGC) is re-shaping the way people watch video and TV, with millions of video producers and consumers. In particular, UGC sites are creating new viewing patterns and social interactions [3], empowering users to be more creative, and developing new business opportunities. YouTube and similar sites have had a huge success over the last few years and are now responsible for generating an identifiable percentage of Internet's traffic.

1.2 From client-server to P2P through CDN

Historically, content distribution in the Internet has relied on a client-server model. This model has shaped all Internet legacy applications such as the web, electronic mail messaging, and FTP. From a content creator's point of view, achieving scalability under the client-server model amounts to deploying and managing a data center. Due to the difficulties of deploying and managing an own data center (see next section) several companies have offered data center hosting as a service to content creators. In the most literal realization, this meant hosting and managing server and network equipment on the behalf of their clients[3]. Content Distribu-

[2]http://www.iptv-news.com/content/view/606/64/
[3]http://www.rackspace.com/

tion Networks (CDN)[4] took a different approach, in which the client does not own or lease any equipment but instead is charged for the amount of service offered on his behalf on the various points of presence of the CDN. CDN's are tightly tailored to the distribution of web content, and thus cannot be seen as a general purpose infrastructure for hosting data center applications like the ones previously mentioned.

In recent years, content distribution has gradually shifted towards a peer-to-peer (P2P) paradigm, which is based on the utilization of the resources of the end user. Despite its prevalence as the current number one contributor to Internet's traffic, P2P has several shortcoming including, legal, economic, and technical ones, that inhibit its use as a full substitute to data centers (details on Sect. 3).

1.3 The need for Edge Capacity Hosting Overlays (ECHOS)

In this paper we make the case for Edge Capacity Hosting OverlayS (ECHOS) as the next step in the content distribution paradigm. By enabling a distributed hosting edge infrastructure, ECHOS can enable the next generation of interactive services and applications to flourish, complementing existing data centers and reaching a massive number of users in a much more efficient manner. The key idea behind ECHOS is to create a fully distributed service platform based on managed boxes located at the edge of the network, where both the boxes and the access bandwidth to those boxes are controlled and managed by a given entity (e.g., by Telco, virtual operator or service provider). Such "boxes" exist everywhere on the Internet where broadband access is available, e.g., in TriplePlay gateways, DSL/cable modems, or set-top-boxes. ECHOS combines the advantages of both the data center and the P2P paradigms, and does away with their respective shortcomings. In the next two section we elaborate on the limitations of data centers and P2P and then move on to presenting how ECHOS is able to bypass them by combining only their positive features.

2. THE DATA CENTER APPROACH

A data center contains primarily electronic equipment used for data processing, storage, and communications. Data centers also usually contain specialized power conversion and backup equipment to maintain reliable, high-quality power, as well as environmental control equipment to maintain the proper temperature and humidity for the IT equipment.

2.1 Buy-at-bulk economics, workload trends, and peak rate dimensioning

Initially, the data center paradigm came into existence due to "economy-of-scale" considerations at times when high processor and storage costs encouraged "buy-at-bulk" investments. Legacy data centers were tailored towards servicing a few large corporate clients that outsourced to the data center their bulk storage and/or processing requirements. This model has changed a lot at recent times for multiple reasons. First, the constant decline in cost of processing and storage equipment has shifted the major cost associated with data centers to real estate, power, cooling, manning, etc. Second, the characteristics of demand have also changed drastically. Whereas demand has skyrocketed in terms of volume, it no longer flows in from a few, anticipated directions. Rather, it is the product of a mind boggling composition of myriad micro-flows coming from all around the wired and wireless network ecosystem. Such a diversified workload creates several challenges.

The first has to do with location. Since data centers are few in number, they have to be located near the core of the network, so as to be equidistant to most users. Therefore, even if they have the capacity in terms of bandwidth and processing rate to serve the requested demand, they have no means to reduce beyond a certain point the latency to the end costumers sitting at the edge of the network. While latency is not important for applications dealing with bulk data transfer, it is the critical parameter for new interactive applications such as online gaming and interactive IPTV. In the first one, the actions of any individual player have to be conveyed immediately to all others players so as to sustain an interactive gaming experience. In the second, the system has to match the channel switching speed of broadcast TV.

The second challenge has to do with dimensioning. Data centers need to be dimensioned for peak capacity, and this causes scalability and economic efficiency problems in view of flash-crowd events. The release of a new movie, software security patch or upgrade, cause sudden peaks on the demand put on data centers. To ensure that servers can cope with peak demand, most servers in a typical data center are only used at about 30% capacity [17]. Still, they need to be cooled down, powered-up and maintained almost as if they were being fully utilized. Therefore, dimensioning a data center around peak demand is quite expensive in terms of both purchase and maintenance.

2.2 Energy consumption and real estate cost

Data centers have received a lot of scrutiny because of the increasing amounts of energy they consume [10]. From 2000 to 2006, the energy used by U.S. servers and data centers and the power and cooling infrastructure that supports them has doubled, raising alarming concern. Among others, it was realized that data centers can be more than 40 times as energy intensive as conventional office buildings [7]. For instance, in 2006, the electricity consumed by servers at U.S. data centers represented about 1.5 percent of the total electricity used nationwide. In an attempt to address the mounting concern, the US Congress has instituted Public Law 109-431 for controlling energy wastage.

Although a great deal of R&D effort is being devoted towards improving the efficiency of existing data centers [4] there are limits to what can be achieved. Service availability has become a critical parameter, and so data centers must always be built with considerable redundancy, in particular for network access, storage capacity, and emergency power systems. This keeps energy and management costs high. Providing such redundancy to the ever increasing demand requires having huge facilities for housing all the equipment[5]. Granted the prohibiting real estate costs of urban/sub-urban areas, many data center have been ostracized to inexpensive and less crowded areas[6], hence increasing the network delay to the end user and the complexity associated with providing support and maintenance.

[4]http://www.akamai.com/

[5]http://www.eweek.com/article2/0,1759,1996919,00.asp

[6]http://weblog.philringnalda.com/2005/03/17/just-how-much-power-does-google-need

3. THE P2P APPROACH

In the last decade, the low cost of computing capacity and the commoditization of broadband access (whether cable or DSL), have given rise to edge-initiated content distribution in the form of the peer-to-peer (P2P) paradigm.

3.1 P2P was supposed to be a temporary patch

The P2P paradigm, from the original Napster to the currently prevailing BitTorrent [6], represents a revolution on the part of the end-user community. Users thirsty for new content and service got tired of waiting for ISPs to provide capabilities like IP multicast and QoS. Thus they took matters in their own hands and came up with innovative solutions for realizing such capabilities using only their own resources at the edge of the network. However, due to its very origin, the P2P paradigm has several inherent limitations:

- Lack of service guarantees due to uncontrolled interference between different applications. This is less of a problem for elastic application like file-swapping, but becomes critical for interactive ones. Although ongoing research is trying to improve what is possible with only statistical guarantees [11], it is clear that much better service can be offered when the resources of nodes are controlled deterministically.
- Inefficient use of network's and other peer's resources and consequently suboptimal performance. Granted that the state of the underlying network and much of the state of other peers are unknown, it is difficult for a P2P application to make optimized decisions. Consequently, P2P overlay topologies are either completely random [14], or at best optimized locally [5]. This has been shown to be quite far from what a careful, cross-peer coordinated overlay optimization can achieve [8]. Similarly, there's quite a lot of room for optimizing content replication [9] and scheduling decisions.
- Even if sufficient state information is in place, still P2P is inherently unable to fully utilize it as it designed around selfish user behavior and free-riding prevention mechanism, rather than well thought out resource scheduling for maximizing the performance of the overall system. Current P2P seems to be trapped in an ever going campaign against selfishness. The implicit cooperation achieved through bilateral tit-for-tat schemes used e.g., in BitTorrent, seems to be the limit of what one could hope for under the selfish user assumption.
- Absence of security and control make it impossible to guarantee the integrity and security of content.

In short, although P2P started as a patch, it is here to stay and will keep growing as a low cost alternative to client-server delivery. However, its value is currently limited by the fact that it cannot guarantee a certain quality of service, it is unsecured, and more fundamentally "uncontrolled".

3.2 A pirate's reputation and the new conflict with ISPs

Content owners such as TV broadcasters, movie studios, game designers, although intrigued by the potential of P2P as a distribution mechanism, view most P2P systems as tainted by their piracy history, and are also sceptical about their ability to provide guaranteed service to end users. To overcome some of these shortcomings, there have recently been a number of legal P2P proposals that tackle a particular application (e.g. Microsoft's Avalanche[7] for distributing Visual Studio files, Skype for Internet telephony, Kontiki for delivering BBC's videos, or Joost for providing TV and VoD services). These systems employ closed-clients which are difficult to hack, and are used for the dissemination of legal content.

Their main problem, however, is that they interfere with the traffic engineering of most ISPs. As a result, for instance, soon after the release of the BBC iPlayer, British Telecom threatened to rate limit, or even completely block it, as it flooded the network with content that did not generate any financial compensation for the ISP [16] that devoted significant resources to handling it. Not only there is no alignment of incentives, but P2P systems, owing to their lack of knowledge about the internal network structure of ISPs, employ mechanisms that are particularly harsh on the ISP. For example, by selecting neighbors randomly, a P2P client is allowed to exchange traffic through an expensive transit link, when there are alternative local clients that could offer the content faster and at no additional cost for the ISP [2]. For this reason ISPs have started dropping BitTorrent and other P2P traffic at their peering points with other networks [13].

4. THE ECHOS APPROACH

The key idea behind ECHOS is to create a fully distributed service platform based on managed boxes located at the edge of the network, where both the boxes and the access bandwidth to those boxes are controlled and managed by a given entity (e.g., by Telco, virtual operator or service provider) like in [15]. Such "boxes" exist everywhere on the Internet where broadband access is available, e.g., in TriplePlay gateways, DSL/cable modems, or Set-top-Boxes.

4.1 Learning from both David and Goliath

The proposed ECHOS paradigm draws the best lessons out of both the data center and the P2P paradigms. It is trusted and controlled like a data center but scalable, inexpensive, and close to the end users like a P2P application. Table 1 provides an overview of the comparison between the three paradigms. The main advantages of ECHOS boil down to the following list.

- Requires almost no capital expenditure as it is composed of components which are already deployed, monitored, cooled and powered. Furthermore, these components are deployed near the end users.
- By operating under a single authority ECHOS solves almost all of P2P's shortcomings. It doesn't have to worry about piracy or free-riding. It can employ near optimal overlay construction, replication and routing to offer guaranteed performance. Finally it can be integrated to the broader business strategy of the ISP and make careful use of its resources.

Such a solution is better suited towards servicing next-generation high quality multimedia services.

[7]http://research.microsoft.com/camsys/avalanche/

	Traditional Data Center	P2P	ECHOS
Data Plane	Centralized	Distributed	Distributed
Control Plane	Centralized	Distributed Uncoordinated	Distributed Coordinated
QoS	Guaranteed	Best Effort	Guaranteed
Capital intensive	Yes	No	No
Distance from end-user	Large	Small	Small
ISP friendly	Yes	No	Yes
Security	Strong	Weak	Strong
Applications	All (static and interactive)	Mostly static	All
Design around incentives	Not required	Required	Desirable

Table 1: Overview and comparison between different content distribution paradigms.

4.2 How to implement a vast data center at home in your free time

All the above can be accomplished without explicitly requiring end-users to contribute their resources (e.g., storage or bandwidth), which often complicates the system design of traditional P2P systems. Instead, under ECHOS, box and network resources will be controlled by a single service provider who can provision and isolate different parts of the box, as well as the link connecting the box to the Internet. By utilizing virtualization technology the ECHOS application running on a box can be completely transparent to the end user. The provider needs only to incentivise the user to leave the box on (which is what most users do anyway), e.g., by providing discounts or bonus content.

4.3 Integrating ECHOS with existing data centers and CDNs

ECHOS can be deployed alone or in conjunction to existing data centers. In the later case its goal would be to absorb the biggest part of the load and leave to the data centers only rare demand spikes that cannot be serviced by ECHOS within the pre-agreed service level agreement. This might seem as a retreat to the problems of traditional data centers but this is not the case. The huge capacity of ECHOS with thousand of boxes deployed throughout the edge filters the peak load and thus requires maintaining only a small backup centralized facility for handling special cases.

5. CONCLUSIONS

In this paper we have attempted to make the case for Edge Capacity Hosting OverlayS (ECHOS) that combine the positive feature of traditional data center solutions and P2P applications and thus are better than these earlier technologies in handling new types of multimedia content. Apart from the aforementioned business opportunities, we believe that ECHOS also offers a wealth of fresh challenges to the research community. Among these we would underline the question of how to fully utilize the P2P paradigm without the handicap of selfishness, and how to implement control and optimization at huge scales but with more detailed network and overlay information.

6. REFERENCES

[1] Le Monde 17.1.2007. Vivendi renaît grâce au phénoménal succés de World of Warcraft.

[2] R. Bindal, P. Cao, W. Chan, J. Medved, G. Suwala, T. Bates, and A. Zhang. Improving traffic locality in BitTorrent via biased neighbor selection. In *Proc. of IEEE ICDCS '06*, Lisbon, Portugal, 2006.

[3] Meeyoung Cha, Haewoon Kwak, Pablo Rodriguez, Yong-Yeol Ahn, and Sue Moon. I tube, you tube, everybody tubes: analyzing the world's largest user generated content video system. In *Proc. of ACM IMC '07*, pages 1–14, 2007.

[4] Jeffrey S. Chase, Darrell C. Anderson, Prachi N. Thakar, Amin M. Vahdat, and Ronald P. Doyle. Managing energy and server resources in hosting centers. In *Proc. of ACM SOSP '01*, 2001.

[5] Yatin Chawathe, Sylvia Ratnasamy, Lee Breslau, Nick Lanham, and Scott Shenker. Making Gnutella-like P2P systems scalable. In *Proc. of ACM SIGCOMM '03*, pages 407–418, Karlsruhe, Germany, 2003.

[6] Bram Cohen. Incentives build robustness in BitTorrent. In *Proc. of First Workshop on Economics of Peer-to-Peer Systems*, Berkeley, CA, Jun 2003.

[7] Steve Greenberg, Evan Mills, William Tschudi, Peter Rumsey, and Bruce Myatt. Best practices for data centers: Results from benchmarking 22 data centers. In *Proc. of ACEEE Summer Study on Energy Efficiency in Buildings*, Washington, DC, Aug 2006.

[8] Nikolaos Laoutaris, Georgios Smaragdakis, Azer Bestavros, and John Byers. Implications of selfish neighbor selection in overlay networks. In *Proc. of IEEE INFOCOM '07*, Anchorage, AK, May 2007.

[9] Nikolaos Laoutaris, Orestis Telelis, Vassilios Zissimopoulos, and Ioannis Stavrakakis. Distributed selfish replication. *IEEE Transactions on Parallel and Distributed Systems*, 17(12):1401–1413, Dec 2006.

[10] Joe Loper and Sara Parr. Alliance to save energy, Energy efficiency in data centers: A new policy frontier, 2007.

[11] Laurent Massoulie, Andy Twigg, Christos Gkantsidis, and Pablo Rodriguez. Randomized decentralized broadcasting algorithms. In *Proc. of IEEE INFOCOM '07*, Anchorage, AK, USA, 2007.

[12] Pablo Rodriguez, See-Mong Tan, and Christos Gkantsidis. On the feasibility of commercial, legal P2P content distribution. *SIGCOMM Comput. Commun. Rev.*, 36(1):75–78, 2006.

[13] Slashdot. Comcast Hinders BitTorrent Traffic, 2007.

[14] Daniel Stutzbach and Reza Rejaie. Characterizing the two-tier Gnutella topology. In *Proc. of ACM SIGMETRICS '05*, Banff, Alberta, Canada, 2005.

[15] K. Suh, M. Varvello, C. Neumann, L. Massoulie, J. Kurose, D. Towsley, and C. Diot. Push-to-peer video-on-demand system: design and evaluation. *IEEE Journal on Selected Areas in Communications*, 25(10), 2007.

[16] August 2007 The Independent, UK. Internet groups warn BBC over iPlayer plans.

[17] Cisco white paper. Data center virtualization and orchestration: Business and financial justification, 2007.