NetEx: Efficient and Cost-effective Internet Bulk Content Delivery

Massimiliano Marcon, Nuno Santos, Krishna P. Gummadi
MPI-SWS
{mmarcon,nsantos,gummadi}@mpi-sws.org

Nikolaos Laoutaris, Pablo Rodriguez
Telefonica Research
{nikos,pablorr}@tid.es

Amin Vahdat
UCSD
vahdat@cs.ucsd.edu

ABSTRACT

The Internet is witnessing explosive growth in traffic due to bulk content transfers, such as multimedia and software downloads, and online sharing of personal, commercial, and scientific data. Yet bulk data transfers remain very expensive and inefficient. As a result, huge amounts of digital data continue to be delivered outside of the Internet using hard drives, optical media or tapes. Meanwhile, large reserves of spare bandwidth lie unutilized in today's networks, where links are overprovisioned for peak load. We designed NetEx, a bulk transfer system that opportunistically exploits the excess capacities of network links to deliver bulk content cheaply and efficiently. Our results based on data from both a commercial tier-1 ISP and the Abilene network suggest that NetEx can considerably increase the capacity of the network, and at the same time it can provide good average performance to bulk transfers.

Categories and Subject Descriptors

C.2.3 [**Computer-Communication Networks**]: Network Operations

General Terms

Design, Performance, Management

Keywords

Bulk Data Transfers, Traffic Differentiation, Traffic Engineering

1. INTRODUCTION

The Internet is witnessing explosive growth in demand for bulk content. Examples of bulk content transfers include downloads of music and movie files, distribution of large software and games, online backups of personal and commercial data, and sharing of huge scientific data repositories. Recent studies of Internet traffic in commercial and research backbones [3] and residential [2] access networks show that such bulk transfers account for a large and rapidly growing fraction of bytes transferred across the Internet.

ANCS'10, October 25-26, 2010, La Jolla, CA, USA.

Despite rising demand, the cost of transit bandwidth in the Internet remains high, and consequently wide-area bulk data transfers remain expensive. For example, the monthly cost of raw transit bandwidth varies between $30K and $90K per Gbps depending on the region of the world [5]. Data transfers to and from Amazon's S3 online storage service cost 10 to 17 cents per GByte without any bandwidth guarantees [1].

While the high cost and delay of large data transfers discourages bulk content delivery over the Internet, tremendous amounts of digital data are being delivered outside of the Internet. Data delivery through hard drives or optical media is currently cheaper and faster–though usually not more convenient or secure–than using the Internet. On an average day, Netflix, ships 2 million movie DVDs [7], or 8 petabytes of data. This is a substantial fraction (25%) of the estimated traffic exchanged between ISPs in the U.S. [6].

The transit ISPs that run national or intercontinental backbones generally overprovision their links to avoid congestion and thus satisfy the performance guarantees they offer to their customers in the form of service level agreements (SLAs). Furthermore, network traffic varies considerably over time exhibiting periodic diurnal and weekly patterns, which creates additional spare capacity, especially during non-peak hours. As a result, even though transit bandwidth remains expensive, large reserves of spare bandwidth continue to lie unutilized in regional, continental, and inter-continental backbones today.

We propose NetEx, a bulk transfer system that exploits spare network resources to deliver bulk content cheaply and efficiently. NetEx is based on the observation that bulk data transfers are not sensitive to the latency of individual packets, i.e., their main performance metric is completion time. Therefore, bulk data transfers can exploit spare resources *opportunistically*, i.e., data is sent only when spare bandwidth is available. To exploit spare bandwidth opportunistically, NetEx differentiates traffic into normal and bulk traffic classes, and forwards bulk traffic with strictly lower priority. By tapping previously unused capacity, NetEx increases link utilization and delivers bulk content at lower cost.

2. NETEX DESIGN

NetEx has two primary components: *traffic differentiation* and *bandwidth-aware traffic engineering*. The first is necessary to exploit spare bandwidth without interfering with existing traffic, while the second is required to achieve efficient use of spare resources.

1. Traffic differentiation: Traffic differentiation is necessary to allow bulk transfers routed through NetEx to use left-over bandwidth without affecting normal Internet traffic. NetEx differentiates traffic into two classes: normal traffic, which is treated as delay sen-

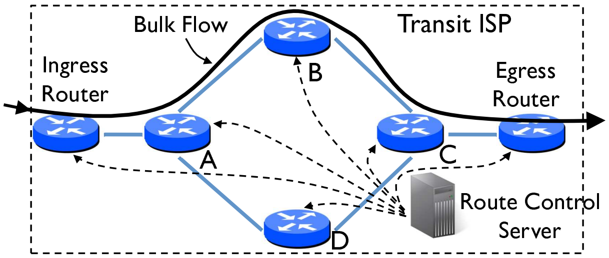

Figure 1: Deployment of NetEx in a transit ISP: Bulk transfers are routed through the ISP along the routes computed by the Route Control Server (RCS).

sitive, and bulk traffic, which is treated as delay tolerant. When forwarding network packets, routers in the ISPs send bulk traffic packets only if there are no packets from normal traffic waiting in the router queues. Thus, normal traffic is always forwarded with strictly higher priority than bulk traffic. As a result, normal traffic is completely isolated from bulk traffic and is never delayed even when bulk traffic completely saturates the capacity on a link.

2. Bandwidth-aware traffic engineering: To achieve efficient use of spare resources, transit ISPs would have to modify the default routing within their networks. This is because intra-domain routing today is not optimized for maximum bandwidth usage: ISPs do not necessarily pick the paths with the most available bandwidth, do not use all possible paths between a pair of nodes, and do not dynamically adapt their paths as the available bandwidths on the links vary. Therefore, NetEx employs a bandwidth-aware routing algorithm that maximizes usage of available spare capacity. The algorithm computes its routes by solving a standard maximum concurrent flow problem [8] by means of a linear solver. The linear solver takes as input short-term estimates of future spare bandwidth and traffic load, which are directly computed from the network traffic observed in the recent past [4, 5]. The resulting routes maximize usage of spare bandwidth and comprise potentially multiple paths between a source and a destination. NetEx periodically executes the routing algorithm to account for changes in network traffic. Our evaluation suggests that, in order to achieve maximum efficiency, routes need to be recomputed only every 30 minutes.

2.1 Deployment

Figure 1 illustrates how a transit ISP could deploy NetEx within its backbone. At a high-level, NetEx identifies and bulk traffic as soon as it enters the ISP's network through an ingress router. The bulk traffic thus identified is then routed according to routing tables computed and disseminated by a central Route Control Server (RCS). RCS computes the routes using NetEx's bandwidth-aware traffic engineering algorithm. When forwarding packets along the routes, NetEx routers forward bulk traffic at a strictly lower priority than normal Internet traffic.

3. EVALUATION

We used real traffic matrices and link loads from a commercial Tier-1 ISP and from the Abilene network to estimate how much additional bulk traffic NetEx could transfer when deployed in an ISP backbone. Our evaluation shows that NetEx can send a considerable amount of additional bulk traffic through the network, with increases between 60% and 170% relative to the current traffic amount, depending on the network topology. We found that the amount of additional bulk traffic sent is close to optimal, meaning that it is not possible to noticeably increase the level of bulk traffic without exceeding the capacity of some link. To understand how much NetEx's bandwidth-aware traffic engineering contributes to NetEx's performance, we ran NetEx with only traffic differentiation but no bandwidth-aware traffic engineering. This means that NetEx sends its bulk traffic along the native minimum-weight routes used by the ISP. In this case, we observed that the amount of additional bulk traffic is roughly half of what is sent with bandwidth-aware routing, thus justifying the need for bandwidth-aware traffic engineering in NetEx.

Even though NetEx uses only spare resources, we found that bulk flows sent through NetEx achieve good average performance, with throughputs that are good enough for most bulk applications. For example, our evaluation using data from the Tier-1 ISP shows that the median completion time for 4GB transfers (comparable to the size of a DVD) is low as 17 minutes, while the median completion time for 100GB transfers (comparable to the remote backup of a hard disk) is less than 1.5 hours. [1]

While NetEx flows achieve good average throughput, their instantaneous throughput can fluctuate greatly. In fact, we found that NetEx flows occasionally experience periods of very poor performance that would be unacceptable for many latency-sensitive, interactive applications. For example, some flows occasionally experience throughput as low as 20Kbps on a whole 5-minute interval, which would be very bad for applications like, say, Web browsing or VoIP. Thus, our results confirm the intuition that NetEx is only well-suited for relatively large bulk flows whose primary performance metric is average throughput.

References

[1] Amazon Simple Storage Service. `http://aws.amazon.com/s3/`.

[2] K. Cho, K. Fukuda, H. Esaki, and A. Kato. The Impact and Implications of the Growth in Residential User-to-user Traffic. In *Proc. of SIGCOMM'06*, 2006.

[3] T. Karagiannis, A. Broido, N. Brownlee, K. Claffy, and M. Faloutsos. Is P2P Dying or just Hiding? In *Proc. of IEEE Globecom*, 2004.

[4] A. Lakhina, K. Papagiannaki, M. Crovella, C. Diot, E. D. Kolaczyk, and N. Taft. Structural analysis of network traffic flows. 2004. Proc. of SIGMETRICS.

[5] N. Laoutaris, G. Smaragdakis, P. Rodriguez, and R. Sundaram. Delay tolerant bulk data transfers on the Internet. To appear in Proc. of SIGMETRICS'09.

[6] Minnesota Internet Traffic Studies. `http://www.dtc.umn.edu/mints/home.php`.

[7] Netflix Online Movie Rental. `http://www.netflix.com/MediaCenter?id=5379#facts`.

[8] F. Shahrokhi and D. W. Matula. The maximum concurrent flow problem. *Journal of the ACM*, 1990.

[1] Since we are interested in estimating the potential performance of NetEx in backbones, our evaluation assumes that bulk flows are not bottlenecked at access links to the backbones.