Network Working Group
Request for Comments: 5632
Category: Informational

C. Griffiths
J. Livingood
Comcast
L. Popkin
Pando
R. Woundy
Comcast
Y. Yang
Yale

September 2009

Comcast's ISP Experiences in a Proactive Network Provider Participation for P2P (P4P) Technical Trial

Abstract

This document describes the experiences of Comcast, a large cable broadband Internet Service Provider (ISP) in the U.S., in a Proactive Network Provider Participation for P2P (P4P) technical trial in July 2008. This trial used P4P iTracker technology, which is being considered by the IETF as part of the Application Layer Transport Optimization (ALTO) working group.

Status of This Memo

This memo provides information for the Internet community. It does not specify an Internet standard of any kind. Distribution of this memo is unlimited.

Copyright Notice

Table of Contents

1. Introduction

Comcast is a large broadband Internet Service Provider (ISP), based in the U.S., serving the majority of its customers via cable modem technology. A trial was conducted in July 2008 with Pando Networks, Yale, and several ISP members of the P4P working group, which is part of the Distributed Computing Industry Association (DCIA). Comcast is a member of the DCIA's P4P Working Group, whose mission is to work with Internet Service Providers (ISPs), peer-to-peer (P2P) companies, and technology researchers to develop "P4P" mechanisms, such as so-called "iTrackers" (hereafter P4P iTrackers), that accelerate distribution of content and optimize utilization of ISP network resources. P4P iTrackers theoretically allow P2P networks to optimize traffic within each ISP, reducing the volume of data traversing the ISP's infrastructure and creating a more manageable flow of data. P4P iTrackers can also accelerate P2P downloads for end users.

P4P's iTracker technology [SIGCOMM] was conceptually discussed with the IETF at the Peer-to-Peer Infrastructure (P2Pi) Workshop held on May 28, 2008, at the Massachusetts Institute of Technology (MIT), as documented in [RFC5594]. This work was discussed in greater detail at the 72nd meeting of the IETF, in Dublin, Ireland, in the ALTO BoF (Birds of a Feather meeting) on July 29, 2008. Due to interest from the community, Comcast shared P4P iTracker trial data at the 73rd meeting of the IETF, in Minneapolis, Minnesota, in the ALTO BoF on November 18, 2008. Since that time, discussion of P4P iTrackers and alternative technologies has continued among participants of the ALTO working group.

The P4P iTracker trial was conducted, in cooperation with Pando, Yale, and three other P4P member ISPs, from July 2 to July 17, 2008. This was the first P4P iTracker trial over a cable broadband network. The trial used a Pando P2P client, and Pando distributed a special 21-MB licensed video file in order to measure the effectiveness of P4P iTrackers. A primary objective of the trial was to measure the effects that increasing the localization of P2P swarms would have on P2P uploads, P2P downloads, and ISP networks, in comparison to normal P2P activity.

2. High-Level Details

As noted in Section 1 of [DynamicSwarmMgmt], a swarm is defined in the following way:

> The content and the set of peers distributing it [a file] is usually called a torrent. A peer that only uploads content is called a seed, while a peer that uploads and downloads at the same time is called a leecher. The connected set of peers participating in the piece exchanges of a torrent is referred to as a swarm.

There were five different swarms for the content used in the trial. The second, third, and fourth used different P4P iTrackers: Generic, Coarse Grained, and Fine Grained, all of which are described in Section 3. The fifth was a proprietary Pando mechanism. (The results of the fifth swarm, while satisfactory, are not included here since our focus is on open standards and a mechanism that may be leveraged for the benefit of the entire community of P2P clients.) Comcast deployed a P4P iTracker server in its production network to support this trial, and configured multiple iTracker files to provide varying levels of localization to clients.

In the trial itself, a P2P client begins a P2P session by querying a pTracker, which runs and manages the P2P network. The pTracker occasionally queries the P4P iTracker, which in this case was maintained by Comcast, the ISP. Other ISPs either managed their own P4P iTracker or used Pando or Yale to host their P4P iTracker files. The P4P iTracker returns network topology information to the pTracker, which then communicates with P2P clients, in order to enable P2P clients to make network-aware decisions regarding peers.

The Pando client was enabled to capture extended logging, when the version of the client included support for it. The extended logging included the source and destination IP address of all P2P transfers, the number of bytes transferred, and the start and end timestamps. This information gives a precise measurement of the data flow in the network, allowing computation of data transfer volumes as well as

data flow rates at each point in time. With standard logging, Pando captured the start and completion times of every download, as well as the average transfer rate observed by the client for the download.

Pando served the data from an origin server external to Comcast's network. This server served about 10 copies of the file, after which all transfers (about 1 million downloads across all ISPs) were performed purely via P2P.

The P2P clients in the trial start with tracker-provided peers, then use peer exchange to discover additional peers. Thus, the initial peers were provided according to P4P iTracker guidance (90% guidance based on P4P iTracker topology and 10% random guidance), then later peers discover the entire swarm via either additional announces or peer exchange.

3. Differences between the P4P iTrackers Used

Given the size of the Comcast network, it was felt that in order to truly evaluate the P4P iTracker application we would need to test various network topologies that reflected its network and would help gauge the level of effort and design requirements necessary to get correct statistical data out of the trial. In all cases, P4P iTrackers were configured with automation in mind, so that any successful P4P iTracker configuration would be automatically updating, rather than manually configured on an ongoing basis. All P4P iTrackers were hosted on the same small server, and it appeared to be relatively easy and inexpensive to scale up a P4P iTracker infrastructure should P4P iTracker-like mechanisms become standardized and widely adopted.

3.1. P4P Fine Grain

The Fine Grain topology was the first and most complex P4P iTracker that we built for this trial. It was a detailed mapping of Comcast backbone-connected network Autonomous System Numbers (ASNs) to IP Aggregates, which were weighted based on priority and distance from each other. Included in this design was a prioritization of all Peer and Internet transit connected ASNs to the Comcast backbone to ensure that P4P traffic would prefer settlement-free and lower-cost networks first, and then more expensive transit links. This attempted to optimize and lower transit costs associated with this traffic. We then took the additional step of detailing each ASN and IP Aggregate into IP subnets down to Optical Transport Nodes (OTNs) where all Cable Modem Termination Systems (CMTS, as briefly defined in Section 2.6 of [RFC3083]) reside . This design gave a highly localized and detailed description of the Comcast network for the iTracker to disseminate. This design defined 1,182 P4P iTracker node

identifiers, and resulted in a 107,357-line configuration file.

This P4P iTracker was obviously the most time-consuming to create and the most complex to maintain. Trial results indicated that this level of localization was too high, and was less effective compared to lower levels of localization.

3.2. P4P Coarse Grain

Given the level of detail in the Fine Grain design, it was important that we also enable a high-level design, which still used priority and weighting mechanisms for the Comcast backbone and transit links. The Coarse Grain design was a limited or summarized version of the Fine Grain design, which used the ASN to IP Aggregate and weighted data for transit links, but removed all additional localization data. This ensured we would get similar data sets from the Fine Grain design, but without the more detailed localization of each of the networks attached to the Comcast backbone. This design defined 22 P4P iTracker node identifiers, and resulted in a 998-line configuration file.

From an overall cost, complexity, risk, and effectiveness standpoint, this was judged to be the optimal P4P iTracker for Comcast. Importantly, this did not require revealing the complex, internal network topology that the Fine Grain did. Updates to this iTracker were also far simpler to automate, which will better ensure that it is accurate over time, and keeps administrative overhead relatively low. However, the differences, costs, and benefits of Coarse Grain and Generic Weighted (see below) likely merit further study.

3.3. P4P Generic Weighted

The Generic Weighted design was a copy of the Coarse Grained design, but instead of using ISP-designated priority and weights, all weights were defaulted to pre-determined parameters that the Yale team had designed. All other data was replicated from the Coarse Grain design. Gathering and providing the information necessary to support the Generic Weighted iTracker was roughly the same level of effort as for Coarse Grain.

4. High-Level Trial Results

Trial data was collected by Pando Networks and Yale University, and raw trial results were shared with Comcast and all of the other ISPs involved in the trial. Analysis of the raw results was performed by Pando and Yale, and these organizations delivered an analysis of the P4P iTracker trial. Using the raw data, Comcast also analyzed the trial results. Furthermore, the raw trial results for Comcast were

shared with Net Forecast, Inc., which performed an independent analysis of the trial for Comcast.

4.1. Swarm Size

During the trial, downloads peaked at 24,728 per day, per swarm, or nearly 124,000 per day for all five swarms. The swarm size peaked at 11,703 peers per swarm, or nearly 57,000 peers for all five swarms. We observed a comparable number of downloads in each of the five swarms.

For each swarm, Table 1 below gives the number of downloads per swarm from Comcast that finished downloading, and the number of downloads from Comcast that canceled downloading before finishing.

Characteristics of P4P iTracker Swarms:

Swarm	Completed Downloads	Cancellations	Total Attempts	Cancellation Rate
Random (Control)	2,719	89	2,808	3.17%
P4P Fine Grained	2,846	64	2,910	2.20%
P4P Generic Weight	2,775	63	2,838	2.22%
P4P Coarse Grained	2,886	52	2,938	1.77%

Table 1: Per-Swarm Size and Cancellation Rates

4.2. Impact on Download Speed

The results of the trial indicated that P4P iTrackers can improve the speed of downloads to P2P clients. In addition, P4P iTrackers were effective in localizing P2P traffic within the Comcast network.

Impact of P4P iTrackers on Downloads:

Swarm	Global Avg bps	Change	Comcast Avg bps	Change
Random (Control)	144,045 bps	n/a	254,671 bps	n/a
P4P Fine Grained	162,344 bps	+13%	402,043 bps	+57%
P4P Generic Weight	163,205 bps	+13%	463,782 bps	+82%
P4P Coarse Grained	166,273 bps	+15%	471,218 bps	+85%

Table 2: Per-Swarm Global and Comcast Download Speeds

4.3. General Impacts on Upstream and Downstream Traffic and Other Interesting Data

An analysis of the effects of P4P iTracker use on upstream utilization and Internet transit was also interesting. It did not appear that P4P iTrackers significantly increased upstream utilization in the Comcast access network; in essence, uploading was already occurring no matter what and a P4P iTracker in and of itself did not appear to materially increase uploading for this specific, licensed content. (A P4P iTracker is not intended as a solution for the potential of network congestion to occur.) Random was 143,236 MB and P4P Generic Weight was 143,143 MB, while P4P Coarse Grained was 139,669 MB. We also observed that using a P4P iTracker reduced outgoing Internet traffic by an average of 34% at peering points. Random was 134,219 MB and P4P Generic Weight was 91,979 MB, while P4P Coarse Grained was 86,652 MB.

In terms of downstream utilization, we observed that the use of a P4P iTracker reduced incoming Internet traffic by an average of 80% at peering points. Random was 47,013 MB, P4P Generic Weight was 8,610 MB, and P4P Coarse Grained was 7,764 MB. However, we did notice that

download activity in the Comcast access network increased somewhat, from 56,030 MB for Random, to 59,765 MB for P4P Generic Weight, and 60,781 MB for P4P Coarse Grained. Note that for each swarm, the number of downloaded bytes according to logging reports is very close to the number of downloads multiplied by file size. But they do not exactly match due to log report errors and duplicated chunks. One factor contributing to the differences in access network download activity is that different swarms have different numbers of downloaders, due to random variations during uniform random assignment of downloaders to swarms (see Table 1). One interesting observation is that Random has higher cancellation rate (3.17%) than that of the guided swarms (1.77%-2.22%). Whether guided swarms achieve lower cancellation rate is an interesting issue for future research.

5. Important Notes on Data Collected

Raw data is presented in this document. We did not normalize traffic volume data (e.g., upload and download) by the number of downloads in order to preserve this underlying raw data.

We also recommend that readers not focus too much on the absolute numbers, such as bytes downloaded from internal sources and bytes downloaded from external sources. Instead, we recommend readers focus on ratios such as the percentage of bytes downloaded that came from internal sources in each swarm. As a result, the small random variation between number of downloads of each swarm does not distract readers from important metrics like shifting traffic from external to internal sources, among other things.

We also wish to note that the data was collected from a sample of the total swarm. Specifically, there were some peers running older versions of the Pando client that did not implement the extended transfer logging. For those nodes, which participated in the swarms but did not report their data transfers, we have download counts. The result of this is that, for example, the download counts generated from the standard logging are a bit higher than the download counts generated by the extended logging. That being said, over 90% of downloads were by peers running the newer software, which we believe shows that the transfer records are highly representative of the total data flow.

In terms of which analysis was performed from the standard logging compared to extended logging, all of the data flow analysis was performed using the extended logging. Pando’s download counts and performance numbers were generated via standard logging (i.e., all peers report download complete/cancel, data volumes, and measured download speed on the client). Yale’s download counts and

performance numbers were derived via extended logging (e.g., by summing the transfer records, counting IP addresses reported, etc.).

One benefit of having two data sources is that we can compare the two. In this case, the two approaches both reported comparable impacts.

6. Next Steps

One objective of this document is to share with the IETF community the results of one P4P iTracker trial in a large broadband network, given skepticism regarding the benefits to P2P users as well as to ISPs. From the perspective of P2P users, P4P iTrackers potentially deliver faster P2P downloads. At the same time, ISPs can increase the localization of swarms, enabling them to reduce bytes flowing over transit points, while also delivering an optimized P2P experience to customers. However, an internal analysis of varying levels of P4P iTracker adoption by ISPs leads us to believe that, while P4P iTracker-type mechanisms are valuable on a single ISP basis, the value of P4P iTrackers increases dramatically as many ISPs choose to deploy it.

We believe these results can inform the technical discussion in the IETF over how to use P4P iTracker mechanisms. Should such a mechanism be standardized, the use of ISP-provided P4P iTrackers should probably be an opt-in feature for P2P users, or at least a feature of which they are explicitly aware of and which has been enabled by default in a particular P2P client. In this way, P2P users could choose to opt-in either explicitly or by their choice of P2P client in order to choose to use the P4P iTracker to improve performance, which benefits both the user and the ISP at the same time. Importantly in terms of privacy, the P4P iTracker makes available only network topology information, and would not in its current form enable an ISP, via the P4P iTracker, to determine which P2P clients were downloading any specific content, whether to determine, for example, if content was a song or a movie or even the title.

It is also possible that a P4P iTracker type of mechanism, in combination with a P2P cache, could further improve P2P download performance, which merits further study. In addition, this was a limited trial that, while very promising, indicates a need for additional technical investigation and trial work. Such a follow-up study should explore the effects of P4P iTrackers when more P2P client software variants are involved, with larger swarms, and with additional and more technically diverse content (file size, file type, duration of content, etc.).

7. Security Considerations

This document does not propose any kind of protocol, practice or standard.

The experiment did show that an ISP can improve performance without exposing fine-grained details about network structure, which might otherwise be a security concern (see Section 3.1 (P4P Fine Grain) and Section 3.2 (P4P Coarse Grain). Section 6 (Next Steps) mentions that the opt-in architecture allows P2P users to maintain privacy.

Other security aspects were not considered in the experiment, which focused on performance measurements.

8. Acknowledgements

The authors wish to acknowledge the hard work of all of the P4P working group members, and specifically the focused efforts of the teams at both Pando and Yale for the trial itself. Finally, the authors recognize and appreciate Peter Sevcik and John Bartlett of NetForecast, Inc., for their valued independent analysis of the trial results.

9. Informative References

[DynamicSwarmMgmt] Carlsson, N. and G. Dan, "Dynamic Swarm Management for Improved BitTorrent Performance", USENIX 8th International Workshop on Peer-to-Peer Systems, March 2009, <http://www.usenix.org/events/iptps09/tech/full_papers/dan/dan_html/>.

[RFC3083] Woundy, R., "Baseline Privacy Interface Management Information Base for DOCSIS Compliant Cable Modems and Cable Modem Termination Systems", RFC 3083, March 2001.

[RFC5594] Peterson, J. and A. Cooper, "Report from the IETF Workshop on Peer-to-Peer (P2P) Infrastructure, May 28, 2008", RFC 5594, July 2009.

[SIGCOMM] Xie, H., Yang, Y., Krishnamurthy, A., Liu, Y., and A. Silberschatz, "ACM SIGCOMM 2008 - P4P: Provider Portal for Applications", Association for Computing Machinery SIGCOMM 2008 Proceedings, August 2008, <http://ccr.sigcomm.org/online/files/p351-xieA.pdf>.

Authors' Addresses

Chris Griffiths
Comcast Cable Communications
One Comcast Center
1701 John F. Kennedy Boulevard
Philadelphia, PA 19103
US

EMail: chris_griffiths@cable.comcast.com
URI: http://www.comcast.com

Jason Livingood
Comcast Cable Communications
One Comcast Center
1701 John F. Kennedy Boulevard
Philadelphia, PA 19103
US

EMail: jason_livingood@cable.comcast.com
URI: http://www.comcast.com

Laird Popkin
Pando Networks
520 Broadway Street
10th Floor
New York, NY 10012
US

EMail: laird@pando.com
URI: http://www.pando.com

Richard Woundy
Comcast Cable Communications
27 Industrial Avenue
Chelmsford, MA 01824
US

EMail: richard_woundy@cable.comcast.com
URI: http://www.comcast.com

Richard Yang
Yale University
51 Prospect Street
New Haven, CT 06520
US

EMail: yry@cs.yale.edu
URI: http://www.cs.yale.edu