# Peer the Peers: An Overlay ID Assignment Service at Internet Exchange Points[*]

Matthias Wählisch
Freie Universität Berlin, Inst. für Informatik
Takustr. 9
D–14195 Berlin, Germany
waehlisch@ieee.org

Thomas C. Schmidt
HAW Hamburg, Department Informatik
Berliner Tor 7
D–20099 Hamburg, Germany
t.schmidt@ieee.org

## ABSTRACT

P2P networks enable end users to establish services relying neither on a dedicated infrastructure nor on an ISP deployment of enhanced services at the network layer. Regrettably, overlay traffic is not optimal with respect to native connections and peering agreements, but may decrease network quality at end users at increased transit costs. This issue has been addressed by traffic localization approaches: The general objective is to keep overlay traffic local and to minimize provider crossing. Current efforts foster provider-assisted solutions. Overlays, which approximate network paths from the underlay, promise to significantly limit interdomain traffic. However, ISPs offering transits rejoice in additional traffic and may provide localization data to delocalize peers. In this paper, we argue that ISP interaction should be provided by neutral authorities, namely the Internet exchange points. We present an architecture which serves unstructured and structured overlay peers with a generic overlay ID that jointly reflects AS-paths *and* peering topologies, and is unbiased by unilateral ISP interests.

## Categories and Subject Descriptors

C.2.1 [**Computer-Communication Networks**]: Network Architecture and Design—*Network topology, Network communications*

## General Terms

Design, Economics, Management, Performance

## Keywords

P2P, ISP interaction, IXP, cooperation, traffic localization

## 1. INTRODUCTION

Localization of overlay traffic is a lively discussed topic, as it may reconcile P2P networks with ISPs. Triggered by several research projects (e.g., [1], [2]), also the IETF (ALTO) works towards a collaboration between ISPs and P2P users. ISP-specific knowledge about the Internet topology could improve overlay network performance, e.g., reduce delays. In return, ISPs are enabled to partially control P2P traffic in a natural way, e.g., without deep packet inspection. There are, however, two major problems: (a) P2P applications need to rely on the trustworthiness of the ISP, and (b) ISPs envision only their own, local topological neighborhood without knowledge of remote transit structures.

In this paper, we introduce a novel architecture that deploys the ISP/P2P interaction service at Internet Exchange Points (IXPs). This service provides overlay peers with an identifier prefix that simultaneously reflects the AS and the local peering topology, and thus reduces inter-provider traffic.

In the remainder, we briefly present the core components of our address assignment architecture (§2) and discuss pros and cons as well as future steps (§3).

## 2. ARCHITECTURE

It is common practice that Internet Exchange Points offer a route server (or reflector) to their customers to reduce BGP peering sessions. BGP peers may establish a (single) connection with this server, which on behalf of such nodes announces corresponding paths to further BGP speakers. We suggest to deploy an address assignment service at IXPs as they (a) have detailed regional knowledge about the AS paths *and* the peering topology and (b) represent a neutral authority between ISPs and in relation to their end users.

The overlay ID assignment procedure comprises of three building blocks:

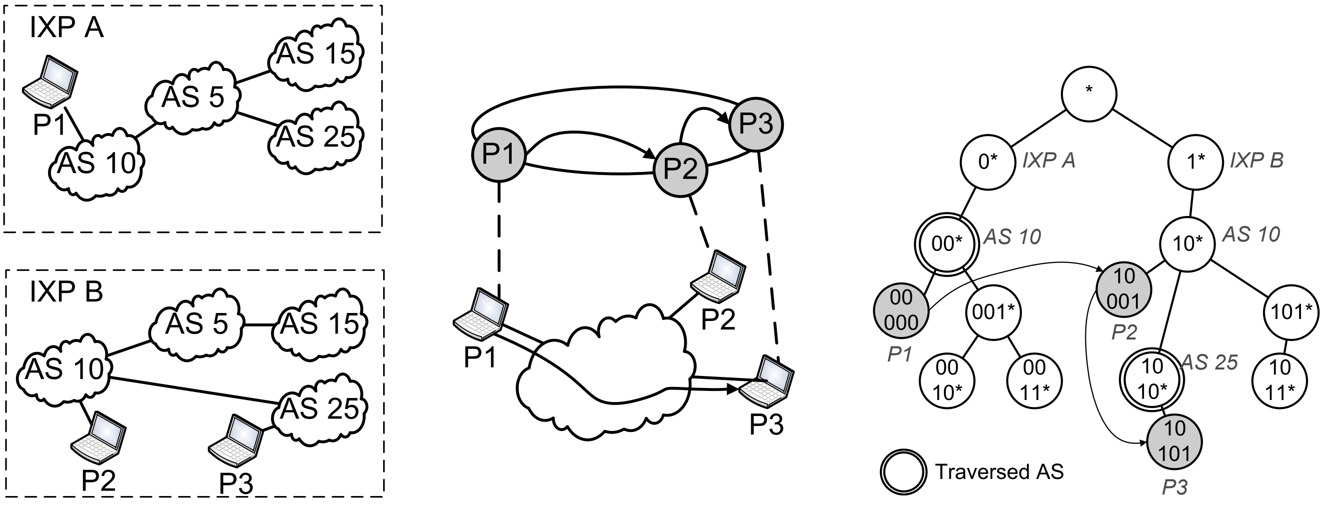
**Service discovery** Each overlay peer needs to contact the topologically nearest ID assignment service. This is realized by anycast routing implemented at the ISPs.

**AS graph calculation** The route server maintains BGP path information. Based on this data, a directed AS-level graph will be constructed using common best path selection mechanisms and BGP tie-breaking rules. This graph is updated infrequently, as core properties of the AS structure do not change very often.

**Overlay ID creation** The AS topology will be mapped on an overlay identifier space. An overlay ID is composed of the following parts: `<IXP ID>:<AS path>:<subnet>:<endhost ID>`. To identify peering locations, we use the AS number of an IXP. The first three blocks represent the overlay prefix, which will be announced to a requesting peer. Each node creates a complete overlay address by the concatenation of this prefix with a unique ID, e.g., the hash value of its public key, see Figure 1. It is worth noting that the ID space predefined by an IXP can be transformed to an arbitrary alphabet at the end hosts.

[*]This work is supported by the German Bundesministerium für Bildung und Forschung within the project H∀Mcast (http://hamcast.realmv6.org).

*CoNEXT Student Workshop'09,* December 1, 2009, Rome, Italy.
ACM 978-1-60558-751-6/09/12.



(a) Peers at different IXPs (b) P1 sends data to P3 via overlay (c) Overlay prefix structure

**Figure 1: An exemplary AS topology at different IXPs and the corresponding overlay construction**

## 3. DISCUSSION

**Overlay structure** The introduced ID creation scheme groups end users according to an equal overlay prefix with respect to their provider *and* current IXP location. Prefix changes correspond to inter provider transitions according to BGP data. In contrast to existing approaches, our scheme introduces an IXP level to the overlay structure, which also uncovers the next-hop IXP and localizes traffic regionally. The evaluation of the IXP-based address assignment requires access to complete BGP data sets available at IXP route servers. We cooperate with the Berlin Internet Exchange Point (BCIX) to gather this information, which then will be used to analyze our scheme in detail. It is worth noting that these data cannot be acquired by open route information services such as RIPE or Route Views. These projects peer with BGP speakers that simply announce all their BGP routes. The resulting AS graph neither reflects locally valid routing policies, nor does it provide a notion of peering topologies.

As network sizes vary between different ASs, AS-dependent schemes may cause unbalanced ID distributions. In general, this problem can be solved by regrouping strategies (e.g. [3]). Furthermore, our IXP-related aggregation of overlay peers most likely helps to distribute identifiers more evenly: Peers that belong to the same, large ISP may be located under different IXP-prefixes. The performance evaluation will be part of future work.

**Stability** Previous studies have shown that BGP paths are instable. This may have negative impact on the overlay structure. However, most of the BGP updates are triggered by unpopular sites and concern only few prefixes [4], or result from regular maintenance intervals. The main inter-AS connections, though, do not change during the lifetime of an overlay peer in typical P2P scenarios.

**Assignment complexity** For a given BGP path set, the ID assignment service required at the IXP consists in a static lookup of AS paths. The underlying path tree can be pre-cached, no individual calculations are required. Efforts needed at the IXP to provide an ID assignment service are thus limited and should scale with IXP size.

**Inter-domain traffic** Overlay paths may be chosen from IDs according to the AS topology. Routing along overlay prefixes can traverse autonomous systems without introducing additional provider crossing. However, there are prefixes that aggregate several ISPs (e.g. IXP change), and usually the overlay routing determines one peer sharing this prefix as next hop. An overlay peer then should select a node that is located in its own AS or in its vicinity, guided by a common proximity neighbor selection scheme.

**Deployment** There are political and economic incentives for IXP and ISP to deploy the proposed architecture. ISPs and IXPs are already in a contractual relationship. While an IXP has an economic interest to extend its services, non-transit ISPs may want to rely on the neutral role of an IXP to prevent P2P users from misleading information by transit providers.

The presented approach allows for an incremental deployment per IXP. Additionally, our scheme leaves the complexity of path acquisition and calculation at the infrastructure level, which makes it also suitable for P2P networks deployed on lightweight mobile devices.

## 4. REFERENCES

[1] V. Aggarwal, A. Feldmann, and C. Scheideler, “Can ISPs and P2P users cooperate for improved performance?” *Comput. Commun. Rev.*, vol. 37, no. 3, pp. 29–40, 2007.

[2] H. Xie, Y. R. Yang, A. Krishnamurthy, Y. G. Liu, and A. Silberschatz, “P4P: Provider Portal for Applications,” in *Proc. of SIGCOMM 2008*. New York, NY, USA: ACM, 2008, pp. 351–362.

[3] L. Cheng, M. R. Ito, and N. Hutchinson, “Internet Topology Based Identifier Assignment for Tree-based DHTs,” in *Proc. of IFIP NTMS’07*, Dordrecht: Springer, 2007, pp. 607–616.

[4] J. Rexford, J. Wang, Z. Xiao, and Y. Zhang, “BGP Routing Stability of Popular Destinations,” in *Proc. of IMW ’02*. New York: ACM, 2002, pp. 197–202.