Internet Engineering Task Force (IETF)
Request for Comments: 7971
Category: Informational
ISSN: 2070-1721

M. Stiemerling
Hochschule Darmstadt
S. Kiesel
University of Stuttgart
M. Scharf
Nokia
H. Seidel
BENOCS
S. Previdi
Cisco

October 2016

Application-Layer Traffic Optimization (ALTO) Deployment Considerations

Abstract

Many Internet applications are used to access resources such as pieces of information or server processes that are available in several equivalent replicas on different hosts. This includes, but is not limited to, peer-to-peer file sharing applications. The goal of Application-Layer Traffic Optimization (ALTO) is to provide guidance to applications that have to select one or several hosts from a set of candidates capable of providing a desired resource. This memo discusses deployment-related issues of ALTO. It addresses different use cases of ALTO such as peer-to-peer file sharing and Content Delivery Networks (CDNs) and presents corresponding examples. The document also includes recommendations for network administrators and application designers planning to deploy ALTO, such as recommendations on how to generate ALTO map information.

Status of This Memo

This document is not an Internet Standards Track specification; it is published for informational purposes.

This document is a product of the Internet Engineering Task Force (IETF). It represents the consensus of the IETF community. It has received public review and has been approved for publication by the Internet Engineering Steering Group (IESG). Not all documents approved by the IESG are a candidate for any level of Internet Standard; see Section 2 of RFC 7841.

Information about the current status of this document, any errata, and how to provide feedback on it may be obtained at http://www.rfc-editor.org/info/rfc7971.

Table of Contents

1. Introduction

Many Internet applications are used to access resources such as pieces of information or server processes that are available in several equivalent replicas on different hosts. This includes, but is not limited to, peer-to-peer (P2P) file sharing applications and Content Delivery Networks (CDNs). The goal of Application-Layer Traffic Optimization (ALTO) is to provide guidance to applications that have to select one or several hosts from a set of candidates capable of providing a desired resource. The basic ideas and problem space of ALTO is described in [RFC5693] and the set of requirements is discussed in [RFC6708]. The ALTO protocol is specified in [RFC7285]. An ALTO server discovery procedure is defined in [RFC7286].

This document discusses use cases and operational issues that can be expected when ALTO gets deployed. This includes, but is not limited to, location of the ALTO server, imposed load to the ALTO server, and who initiates the queries. This document provides guidance on which ALTO services to use, and it summarizes known challenges as well as deployment experiences, including potential processes to generate ALTO network and cost maps. It thereby complements the management considerations in the protocol specification [RFC7285], which are independent of any specific use of ALTO.

2. General Considerations

2.1. ALTO Entities

2.1.1. Baseline Scenario

The ALTO protocol [RFC7285] is a client/server protocol, operating between a number of ALTO clients and an ALTO server, as sketched in Figure 1. Below, the baseline deployment scenario for ALTO entities is first reviewed independently of the actual use case. Specific examples are then discussed in the remainder of this document.

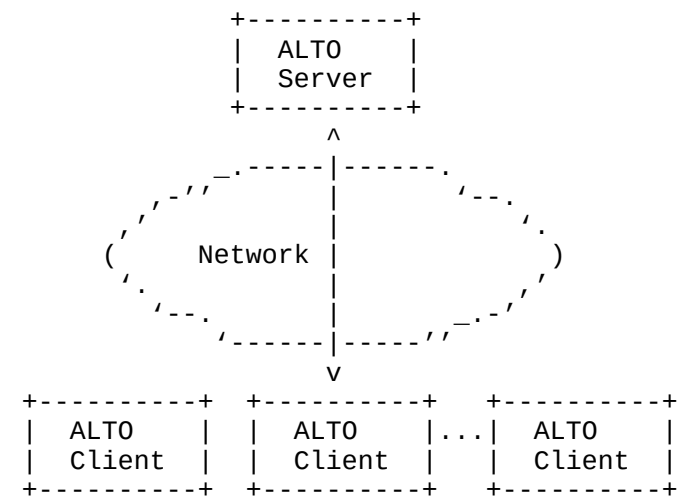

Figure 1: Baseline Deployment Scenario of the ALTO Protocol

This document uses the terminology introduced in [RFC5693]. In particular, the following terms are defined by [RFC5693]:

- ALTO Service: Several resource providers may be able to provide the same resource. The ALTO service gives guidance to a resource consumer and/or resource directory about which resource provider(s) to select in order to optimize the client's performance or quality of experience, while improving resource consumption in the underlying network infrastructure.

- ALTO Server: A logical entity that provides interfaces to satisfy the queries about a particular ALTO service.

- ALTO Client: The logical entity that sends ALTO queries. Depending on the architecture of the application, one may embed it in the resource consumer and/or in the resource directory.

- Resource Consumer: For P2P applications, a resource consumer is a specific peer that needs to access resources. For client-server or hybrid applications, a consumer is a client that needs to access resources.

- Resource Directory: An entity that is logically separate from the resource consumer and that assists the resource consumer to identify a set of resource providers. Some P2P applications refer to the resource directory as a P2P tracker.

We differentiate between an ALTO Client and a Resource Consumer as follows: the resource consumer is specific instance of a software ("process") running on a specific host. It is a client instance of a client/server application or a peer of a peer-to-peer application. It is the given (constant) endpoint of the data transmissions to be optimized using ALTO. The optimization is done by wisely choosing the other ends of these data flows (i.e., the server(s) in a client/server application or the peers in a peer-to-peer application), using guidance from ALTO and possibly other information. An ALTO client is a piece of software (e.g., a software library) that implements the client entity of the ALTO protocol as specified in [RFC7285]. It consists of data structures that are suitable for representing ALTO queries, replies, network and cost maps, etc. Furthermore, it has to implement an HTTP client and a JSON encoder/decoder, or it has to include other software libraries that provide these building blocks. In the simplest case, this ALTO client library can be linked (or otherwise incorporated) into the resource consumer, in order to retrieve information from an ALTO server and thereby satisfy the resource consumer's need for guidance. However, other configurations are possible as well, as discussed in Section 2.1.2 and other sections of this document.

According to these definitions, both an ALTO server and an ALTO client are logical entities. A particular ALTO service may be offered by more than one ALTO server. In ALTO deployments, the functionality of an ALTO server can therefore be realized by several server instances, e.g., by using load balancing between different physical servers. The term ALTO server should not be confused with use of a single physical server.

This document uses the term "Resource Directory" as defined in [RFC5693]. This term and its meaning is not to be confused with the "Information Resource Directory (IRD)" defined as a part of the ALTO protocol [RFC7285], i.e., a list of available information resources offered by a specific ALTO service and the URIs at which each can be accessed.

2.1.2. Placement of ALTO Entities

The ALTO server and ALTO clients may be situated at various places in a network topology. An important differentiation is whether the ALTO client is located on the host that is the endpoint of the data transmissions to be optimized with ALTO (see Figure 2) or whether the ALTO client is located on a resource directory, which assists peers or clients in finding other peers or servers, respectively, but does not directly take part in the data transmission (see Figure 3).

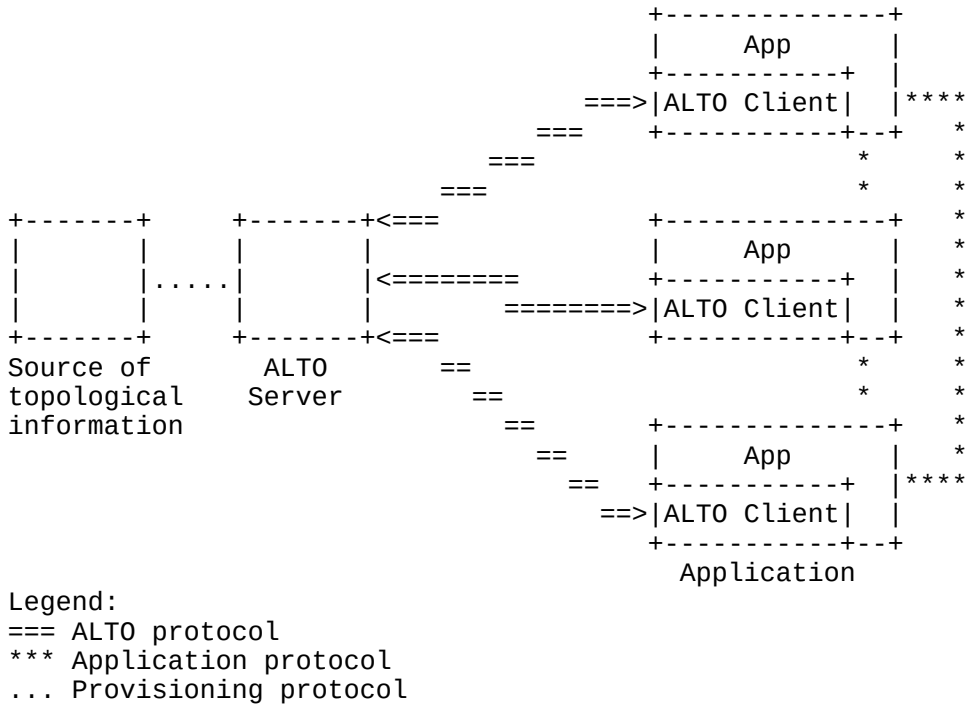

Figure 2: Overview of Protocol Interaction between ALTO Elements without a Resource Directory

Figure 2 shows the operational model for an ALTO client running at endpoints. An example would be a peer-to-peer file sharing application that does not use a tracker, such as edonkey. In addition, ALTO clients at peers could also be used in a similar way even if there is a tracker, as further discussed in Section 4.1.2.

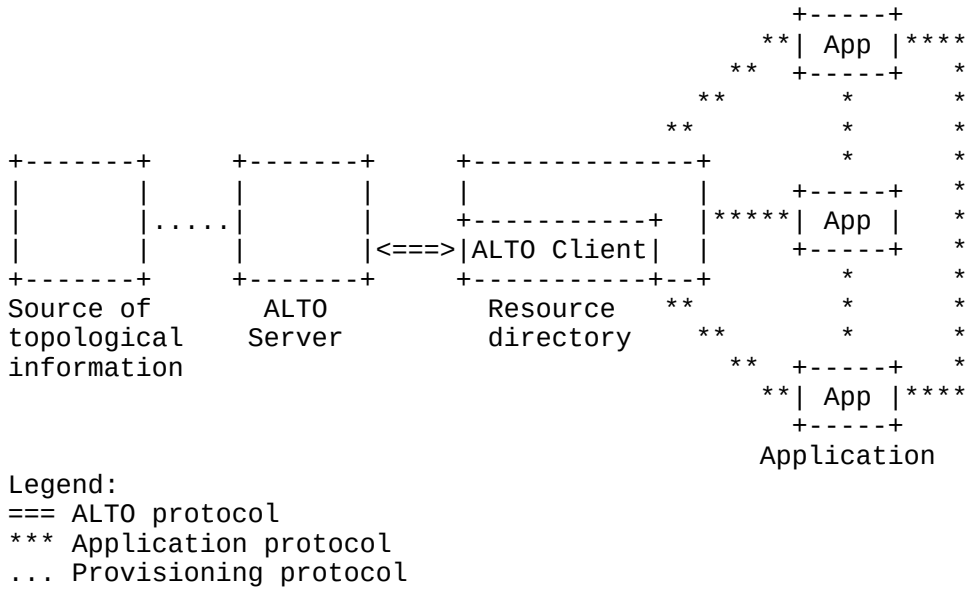

Figure 3: Overview of Protocol Interaction between ALTO Elements with a Resource Directory

In Figure 3, a use case with a resource directory is illustrated, e.g., a tracker in a peer-to-peer file-sharing application such as BitTorrent. Both deployment scenarios may differ in the number of ALTO clients that access an ALTO service. If an ALTO client is implemented in a resource directory, an ALTO server may be accessed by a limited and less dynamic set of clients, whereas in the general case any host could be an ALTO client. This use case is further detailed in Section 4.

Using ALTO in CDNs may be similar to a resource directory [CDN-USE]. The ALTO server can also be queried by CDN entities to get guidance about where a particular client accessing data in the CDN is located in the Internet Service Provider's network, as discussed in Section 5.

2.2. Classification of Deployment Scenarios

2.2.1. Roles in ALTO Deployments

ALTO is a general-purpose protocol and it is intended to be used by a wide range of applications. In different use cases, applications, resource directories, etc., can correspond to different functionality. The use cases listed in this document are not meant to be comprehensive. This also implies that there are different

possibilities where the ALTO entities are actually located, i.e., if the ALTO clients and the ALTO server are in the same Internet Service Provider (ISP) domain, or if the clients and the ALTO server are managed/owned/located in different domains.

An ALTO deployment involves four kinds of entities:

1. Source of topological information

2. ALTO server

3. ALTO client

4. Resource consumer

Each of these entities corresponds to a certain role, which results in requirements and constraints on the interaction between the entities.

A key design objective of the ALTO service is that each of these four roles can be separated, i.e., they can be realized by different organizations or disjoint system components. ALTO is inherently designed for use in multi-domain environments. Most importantly, ALTO is designed to enable deployments in which the ALTO server and the ALTO client are not located within the same administrative domain.

As explained in [RFC5693], from this follows that at least three different kinds of entities can operate an ALTO server:

1. Network operators. Network Service Providers (NSPs) such as ISPs may have detailed knowledge of their network topology and policies. In this case, the source of the topology information and the provider of the ALTO server may be part of the same organization.

2. Third parties. Topology information could also be collected by companies or organizations that are distinct from the network operators, yet have arranged certain legal agreements with one or more network operators, regarding access to their topology information and/or doing measurements in their networks. Examples of such entities could be CDN operators or companies specialized in offering ALTO services on behalf of ISPs.

3. User communities. User communities could run distributed measurements for estimating the topology of the Internet. In this case, the topology information may not originate from ISP data.

Regarding the interaction between ALTO server and client, ALTO deployments can be differentiated according to the following aspects:

1. Applicable trust model: The deployment of ALTO can differ depending on whether or not the ALTO client and ALTO server are operated within the same organization and/or network. This affects a number of constraints because the trust model is very different. For instance, as discussed later in this memo, the level of detail of maps can depend on who the involved parties actually are.

2. Composition of the user group: The main use case of ALTO is to provide guidance to any Internet application. However, an operator of an ALTO server could also decide to offer guidance only to a set of well-known ALTO clients, e.g., after authentication and authorization. In the peer-to-peer application use case, this could imply that only selected trackers are allowed to access the ALTO server. The security implications of using ALTO in closed groups differ from the public Internet.

3. Covered destinations: In general, an ALTO server has to be able to provide guidance for all potential destinations. Yet, in practice, a given ALTO client may only be interested in a subset of destinations, e.g., only in the network cost between a limited set of resource providers. For instance, CDN optimization may not need the full ALTO cost maps because traffic between individual residential users is not in scope. This may imply that an ALTO server only has to provide the costs that matter for a given user, e.g., by customized maps.

The following sections enumerate different classes of use cases for ALTO, and they discuss deployment implications of each of them. In principle, an ALTO server can be operated by any organization, and there is no requirement that an ALTO server be deployed and operated by an ISP. Yet, since the ALTO solution is designed for ISPs, most examples in this document assume that the operator of an ALTO server is a network operator (e.g., an ISP or the network department in a large enterprise) that offers ALTO guidance in particular to users of this network.

It must be emphasized that any application using ALTO must also work if no ALTO servers can be found or if no responses to ALTO queries are received, e.g., due to connectivity problems or overload situations (see also [RFC6708]).

2.2.2. Information Exposure

There are basically two different approaches to how an ALTO server can provide network information and guidance:

1. The ALTO server provides maps that contain provider-defined cost values between network location groupings (e.g., sets of IP prefixes). These maps can be retrieved by clients via the ALTO protocol, and the actual processing of the map data is done inside the client. Since the maps contain (aggregated) cost information for all endpoints, the client does not have to reveal any internal operational data, such as the IP addresses of candidate resource providers. The ALTO protocol supports this mode of operation by the Network and Cost Map Service.

2. The ALTO server provides a query interface that returns costs or rankings for explicitly specified endpoints. This means that the query of the ALTO client has to include additional information (e.g., a list of IP addresses). The server then calculates and returns costs or rankings for the endpoints specified in the request (e.g., a sorted list of the IP addresses). In ALTO, this approach can be realized by the Endpoint Cost Service (ECS) and other related services.

Both approaches have different privacy implications for the server and client:

For the client, approach 1 has the advantage that all operational information stays within the client and is not revealed to the provider of the server. However, this service implies that a network operator providing an ALTO server has to expose a certain amount of information about its network structure (e.g., IP prefixes or topology information in general).

For the operator of a server, approach 2 has the advantage that the query responses reveal less topology information to ALTO clients. However, it should be noted that collaborating ALTO clients could gather more information than expected by aggregating and correlating responses to multiple queries sent to the ALTO server (see Section 5.2.1, item (3) of [RFC6708]). Furthermore, this method requires that clients send internal operational information to the server, such as the IP addresses of hosts also running the application. For clients, such data can be sensitive.

As a result, both approaches have their pros and cons, as further detailed in Section 3.3.

2.2.3. More-Advanced Deployments

From an ALTO client's perspective, there are different ways to use ALTO:

1. Single-service instance with single-metric guidance: An ALTO client only obtains guidance regarding a single metric (e.g., "routingcost") from a single ALTO service, e.g., an ALTO server that is offered by the network service provider of the corresponding access network. Corresponding ALTO server instances can be discovered, e.g., by ALTO server discovery [RFC7286] [XDOM-DISC]. Since the ALTO protocol is an HTTP-based, REST-ful (Representational State Transfer) protocol, the operator of an ALTO may use well-known techniques for serving large web sites, such as load balancers, in order to serve a large number of ALTO queries. The ALTO protocol also supports the use of different URIs for different ALTO features and thereby the distribution of the service onto several servers.

2. Single service instance with multiple metric guidance: An ALTO client could also query an ALTO service for different kinds of information, e.g., cost maps with different metrics. The ALTO protocol is extensible and permits such operation. However, ALTO does not define how a client shall deal with different forms of guidance, and it is up to the client to interpret the received information accordingly.

3. Multiple service instances: An ALTO client can also decide to access multiple ALTO servers providing guidance, possibly from different operators or organizations. Each of these services may only offer partial guidance, e.g., for a certain network partition. In that case, it may be difficult for an ALTO client to compare the guidance from different services. Different organization may use different methods to determine maps, and they may also have different (possibly even contradicting or competing) guidance objectives. How to discover multiple ALTO servers and how to deal with conflicting guidance is an open issue.

There are also different options regarding the synchronization of guidance offered by an ALTO service:

1. Authoritative servers: An ALTO server instance can provide guidance for all destinations for all kinds of ALTO clients.

2. Cascaded servers: An ALTO server may itself include an ALTO client and query other ALTO servers, e.g., for certain destinations. This results is a cascaded deployment of ALTO servers, as further explained below.

3. Inter-server synchronization: Different ALTO servers may communicate by other means. This approach is not further discussed in this document.

An assumption of the ALTO design is that ISPs operate ALTO servers independently, irrespective of other ISPs. This may be true for most envisioned deployments of ALTO, but there may be certain deployments that may have different settings. Figure 4 shows such a setting with a university network that is connected to two upstream providers. NREN is a National Research and Education Network, which provides cheap high-speed connectivity to specific destinations, e.g., other universities. ISP is a commercial upstream provider from which the university buys connectivity to all destinations that cannot be reached via the NREN. The university, as well as ISP, are operating their own ALTO server. The ALTO clients, located on the peers in the university network will contact the ALTO server located at the university.

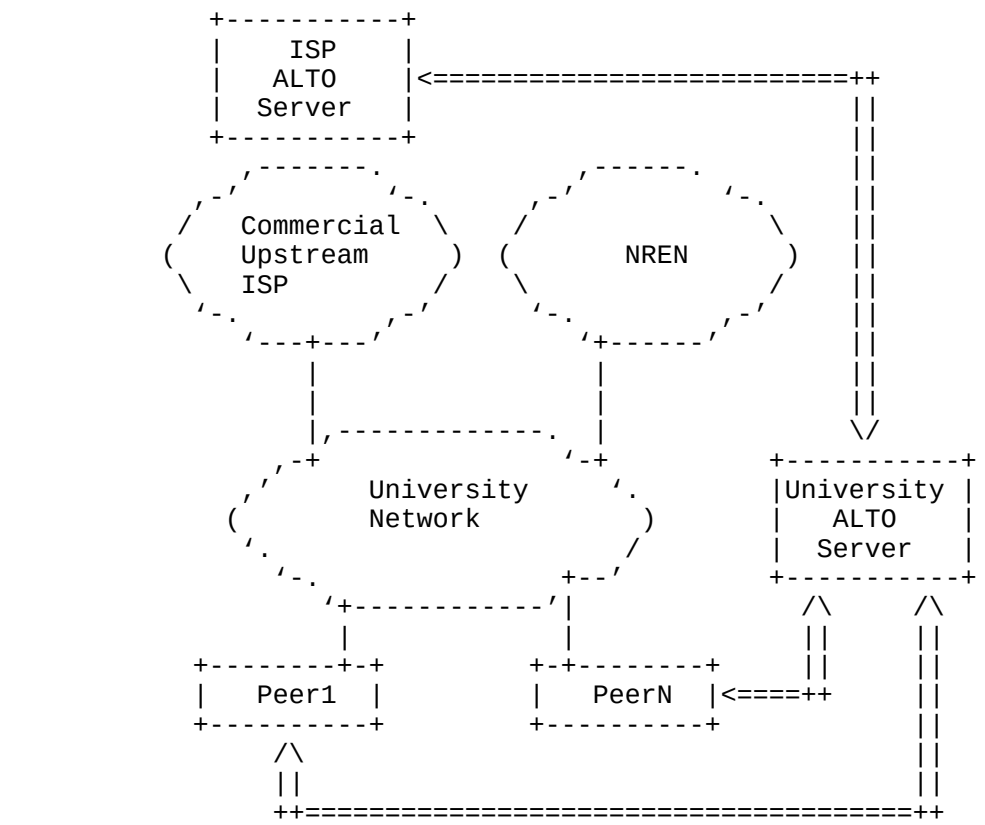

Figure 4: Example of a Cascaded ALTO Server

In this setting, all destinations that can be reached via the NREN are preferred in the rating of the university's ALTO server. In contrast, all traffic that is not routed via the NREN will be handled by the commercial upstream ISP and is in general less preferred due to the associated costs. Yet, there may be significant differences between various destinations reached via the ISP. Therefore, the ALTO server at the university may also include the guidance given by the ISP ALTO server in its replies to the ALTO clients. This is an example for cascaded ALTO servers.

3. Deployment Considerations by ISPs

3.1. Objectives for the Guidance to Applications

3.1.1. General Objectives for Traffic Optimization

The Internet consists of many networks. The networks are owned and managed by different network operators, such as commercial ISPs, enterprise IT departments, universities, and other organizations. These network operators provide network connectivity, e.g., by access networks, such as cable networks, xDSL networks, 3G/4G mobile networks, etc. Network operators need to manage, control, and audit the traffic. Therefore, it is important to understand how to deploy an ALTO service and what its expected impact might be.

The general objective of ALTO is to give guidance to applications on what endpoints (e.g., IP addresses or IP prefixes) are to be preferred according to the operator of the ALTO server. The ALTO protocol gives means to let the ALTO server operator express its preference, whatever this preference is.

ALTO enables network operators to support application-level traffic engineering by influencing application resource provider selection. This traffic engineering can have different objectives:

1. Inter-network traffic localization: ALTO can help to reduce inter-domain traffic. The networks of different network operators are interconnected through peering points. From a business view, the inter-network settlement is needed for exchanging traffic between these networks. These peering agreements can be costly. To reduce these costs, a simple objective is to decrease the traffic exchange across the peering points and thus keep the traffic in the own network or Autonomous System (AS) as far as possible.

2. Intra-network traffic localization: In case of large network operators, the network may be grouped into several networks, domains, or ASes. The core network includes one or several backbone networks, which are connected to multiple aggregation, metro, and access networks. If traffic can be limited to certain areas such as access networks, this decreases the usage of backbone and thus helps to save resources and costs.

3. Network offloading: Compared to fixed networks, mobile networks have some special characteristics, including lower link bandwidth, high cost, limited radio frequency resource, and limited terminal battery. In mobile networks, wireless links should be used efficiently. For example, in the case of a P2P

service, it is likely that hosts should prefer retrieving data from hosts in fixed networks, and avoid retrieving data from mobile hosts.

4. Application tuning: ALTO is also a tool to optimize the performance of applications that depend on the network and perform resource provider selection decisions among network endpoints; an example is the network-aware selection of CDN caches.

In the following, these objectives are explained in more detail with examples.

3.1.2. Inter-Network Traffic Localization

ALTO guidance can be used to keep traffic local in a network, for instance, in order to reduce peering costs. An ALTO server can let applications prefer other hosts within the same network operator's network instead of randomly connecting to other hosts that are located in another operator's network. Here, a network operator would always express its preference for hosts in its own network, while hosts located outside its own network are to be avoided (i.e., they are undesired to be considered by the applications). Figure 5 shows such a scenario where hosts prefer hosts in the same network (e.g., Host 1 and Host 2 in ISP1 and Host 3 and Host 4 in ISP2).

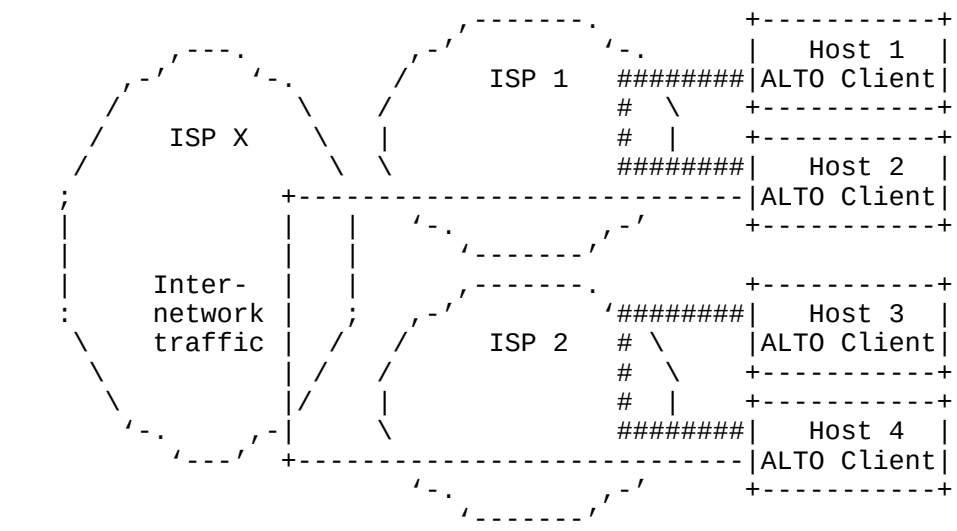

Figure 5: Inter-Network Traffic Localization

Examples for corresponding ALTO maps can be found in Section 3.5. Depending on the application characteristics, it may not be possible or even desirable to completely localize all traffic.

3.1.3. Intra-Network Traffic Localization

The previous section describes the results of the ALTO guidance on an inter-network level. In the same way, ALTO can also be used for intra-network localization. In this case, ALTO provides guidance on which internal hosts are to be preferred inside a single network (e.g., one AS). This application-level traffic engineering can reduce the capacity requirements in the core network of an ISP. Figure 6 shows such a scenario where Host 1 and Host 2 are located in an access net 1 of ISP 1 and connect via a low capacity link to the core of the same ISP 1. If Host 1 and Host 2 exchange their data with remote hosts, they would probably congest the bottleneck link.

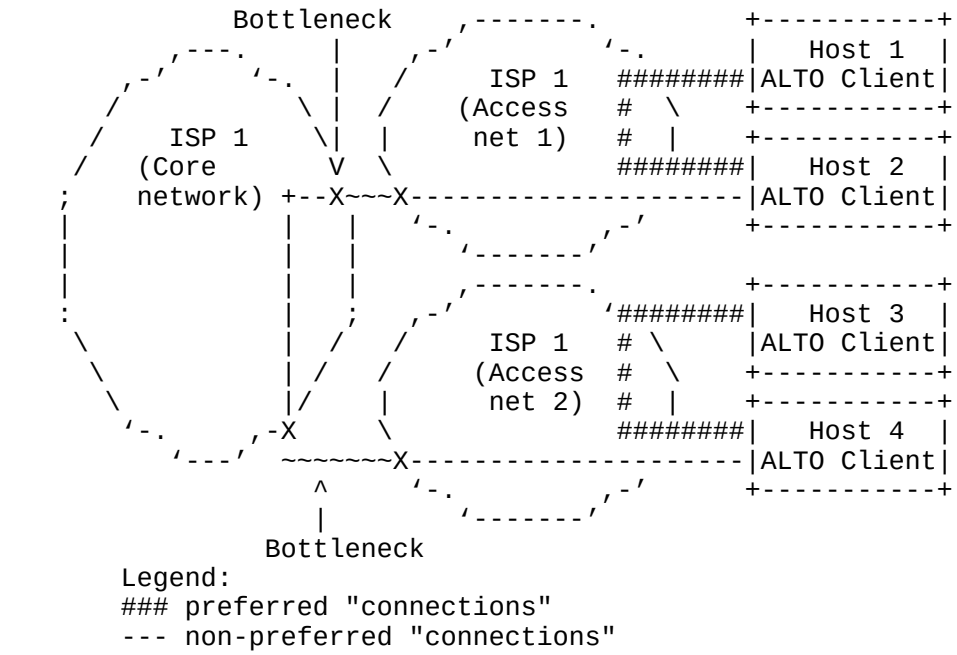

Figure 6: Intra-Network Traffic Localization

In such a situation, the operator can guide the hosts to try local hosts in the same network islands first, avoiding or at least lowering the effect on the bottleneck link, as shown in Figure 6.

The objective is to avoid bottlenecks by optimized endpoint selection at the application level. That said, it must be understood that ALTO is not a general-purpose method to deal with the congestion at the bottleneck.

3.1.4. Network Offloading

Another scenario is offloading traffic from networks. This use of ALTO can be beneficial in particular in mobile networks. A network operator may have the desire to guide hosts in its mobile network to use hosts outside this mobile network. One reason could be that the wireless network or the mobile hosts were not designed for direct peer-to-peer communications between mobile hosts, and therefore, it makes sense for peers to fetch content from remote peers in other parts of the Internet.

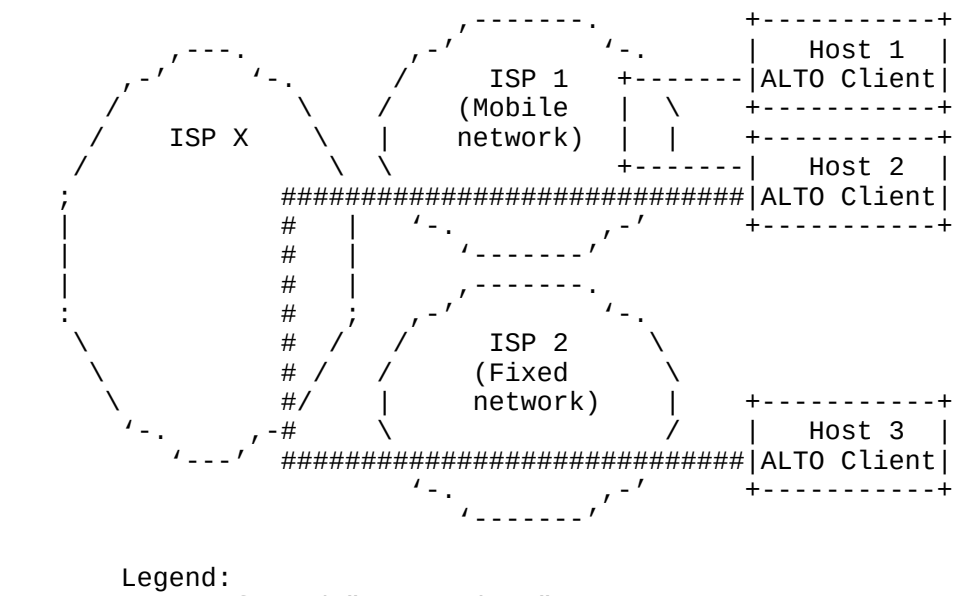

Figure 7: ALTO Traffic Network De-localization

Figure 7 shows the result of such a guidance process where Host 2 prefers a connection with Host 3 instead of Host 1, as shown in Figure 5.

A realization of this scenario may have certain limitations and may not be possible in all cases. For instance, it may require the ALTO server to distinguish mobile and non-mobile hosts based on their IP address. This may depend on mobility solutions and may not be possible or accurate. In general, ALTO is not intended as a fine-grained traffic engineering solution for individual hosts. Instead, it typically works on aggregates (e.g., if it is known that certain IP prefixes are often assigned to mobile users).

3.1.5. Application Tuning

ALTO can also provide guidance to optimize the application-level topology of networked applications, e.g., by exposing network performance information. Applications can often run their own measurements to determine network performance, e.g., by active delay measurements or bandwidth probing, but such measurements result in overhead and complexity. Accessing an ALTO server can be a simpler

alternative. In addition, an ALTO server may also expose network information that applications cannot easily measure or reverse-engineer.

3.2. Provisioning of ALTO Topology Data

3.2.1. High-Level Process and Requirements

A process to generate ALTO topology information typically comprises several steps. The first step is to gather information, which is described in the following section. The subsequent sections describe how the gathered data can be processed and which methods can be applied to generate the information exposed by ALTO, such as network and cost maps.

Providing ALTO guidance can result in a win-win situation for network providers and users of the ALTO information. Applications possibly get a better performance, while the network provider has means to optimize the traffic engineering and thus its costs. Yet, there can be security concerns with exposing topology data. Corresponding limitations are discussed in Section 7.2.

ISPs may have important privacy requirements when deploying ALTO, which have to be taken into account when processing ALTO topology data. In particular, an ISP may not be willing to expose sensitive operational details of its network. The topology abstraction of ALTO enables an ISP to expose the network topology at a desired granularity only, determined by security policies.

With the ECS, the ALTO client does not have to implement any specific algorithm or mechanism in order to retrieve, maintain and process network topology information (of any kind). The complexity of the network topology (computation, maintenance and distribution) is kept in the ALTO server and ECS is delivered on demand. This allows the ALTO server to enhance and modify the way the topology information sources are used and combined. This simplifies the enforcement of privacy policies of the ISP.

The ALTO Network and Cost Map Service expose an abstract view on the ISP network topology. Therefore, care is needed when constructing those maps in order to take privacy policies into account, as further discussed in Section 3.2.3. The ALTO protocol also supports further features such as endpoint properties, which could also be used to expose topology guidance. The privacy considerations for ALTO maps also apply to such ALTO extensions.

3.2.2. Data Collection from Data Sources

The first step in the process of generating ALTO information is to gather the required information from the network. An ALTO server can collect topological information from a variety of sources in the network and provides a cohesive, abstract view of the network topology to applications using an ALTO client. Topology data sources may include routing protocols, network policies, state and performance information, geolocation, etc. An ALTO server requires at least some topology and/or routing information, i.e., information about existing endpoints and their interconnection. With this information, it is in principle possible to compute paths between all known endpoints. Based on such basic data, the ALTO server builds an ALTO-specific network topology that represents the network as it should be understood and utilized by applications (resource consumers) at endpoints using ALTO services (e.g., Network and Cost Map Service or ECS). A basic dataset can be extended by many other information obtainable from the network.

The ALTO protocol does not assume a specific network technology or topology. In principle, ALTO can be used with various types of addresses (Endpoint Addresses). [RFC7285] defines the use of IPv4/IPv6 addresses or prefixes in ALTO, but further address types could be added by extensions. In this document, only the use of IPv4/IPv6 addresses is considered.

The exposure of network topology information is controlled and managed by the ALTO server. ALTO abstract network topologies can be automatically generated from the physical or logical topology of the network, e.g., using "live" network data. The generation would typically be based on policies and rules set by the network operator. The maps and the guidance can significantly differ depending on the use case, the network architecture, and the trust relationship between ALTO server and ALTO client, etc. Besides the security requirements that consist of not delivering any confidential or critical information about the infrastructure, there are efficiency requirements in terms of what aspects of the network are visible and required by the given use case and/or application.

The ALTO server operator has to ensure that the ALTO topology does not reveal any details that would endanger the network integrity and security. For instance, ALTO is not intended to leak raw Interior Gateway Protocol (IGP) or Border Gateway Protocol (BGP) databases to ALTO clients.

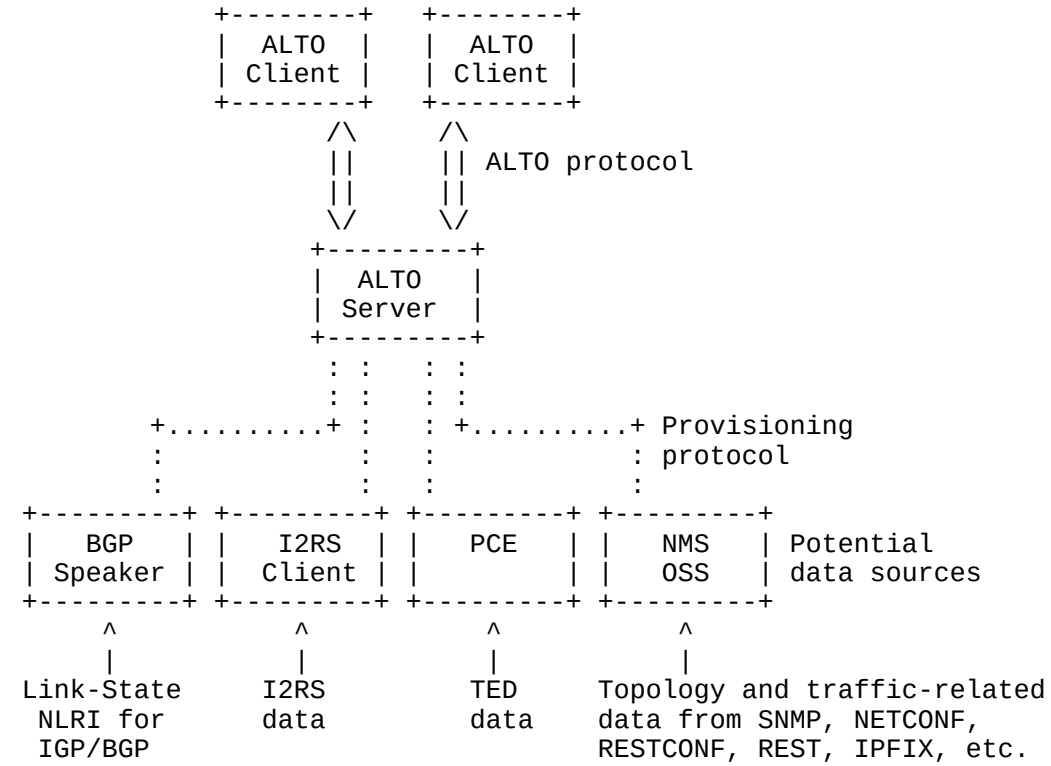

Figure 8: Potential Data Sources for ALTO

As illustrated in Figure 8, the topology data used by an ALTO server can originate from different data sources:

- Relevant information sources are IGPs or BGP. An ALTO server could get network routing information by listening to IGPs and/or peering with BGP speakers. For data collection, link-state protocols are more suitable since every router propagates its information throughout the whole network. Hence, it is possible to obtain information about all routers and their neighbors from one single router in the network. In contrast, distance-vector protocols are less suitable since routing information is only shared among neighbors. To obtain the whole topology with distance-vector routing protocols it is necessary to retrieve routing information from every router in the network.

- [RFC7752] describes a mechanism by which link-state and Traffic Engineering (TE) information can be collected from networks and shared with external components using the BGP routing protocol. This is achieved using a new BGP Network Layer Reachability

Information (NLRI) encoding format. The mechanism is applicable to physical and virtual IGP links and can also include TE data. For instance, prefix data can be carried and originated in BGP, while TE data is originated and carried in an IGP. The mechanism described is subject to policy control.

- The Interface to the Routing System (I2RS) is a solution for state transfer in and out of the Internet's routing system [RFC7921]. An ALTO server could use an I2RS client to observe routing-related information. With the rise of Software-Defined Networking (SDN) and a decoupling of network data and control plane, topology information could also be fetched from an SDN controller. If I2RS is used, [RFC7922] provides traceability for these interactions. This scenario is not further discussed in the remainder of this document.

- Another potential source of topology information could be a Path Computation Element (PCE) [RFC4655]. Topology and traffic-related information can be retrieved from the Traffic Engineering Database (TED) and Label Switched Path Database (LSP-DB). This scenario is not further discussed in the remainder of this document.

- An ALTO server can also leverage a Network Management System (NMS) or an Operations Support System (OSS) as data sources. NMS or OSS solutions are used to control, operate, and manage a network, e.g., using the Simple Network Management Protocol (SNMP) or Network Configuration Protocol (NETCONF). As explained for instance in [RFC7491], the NMS and OSS can be consumers of network events reported and can act on these reports as well as displaying them to users and raising alarms. In addition, NMS and OSS systems may have access to routing information and network inventory data (e.g., links, nodes, or link properties not visible to routing protocols, such as Shared Risk Link Groups). Furthermore, Operations, Administration, and Maintenance (OAM) information can be leveraged, including traffic utilization obtained from IP Flow Information Export (IPFIX), event notifications (e.g., via syslog), liveness detection (e.g., bidirectional forwarding detection, BFD). NMS or OSS systems also may have functions to correlate and orchestrate information originating from other data sources. For instance, it could be required to correlate IP prefixes with routers (Provider, Provider Edge, Customer Edge, etc.), IGP areas, VLAN IDs, or policies.

In the context of the provisioning protocol, topology information could be modeled in a YANG data model [NETWORK-TOPO].

The data sources mentioned so far are only a subset of potential topology sources and protocols. Depending on the network type, (e.g., mobile, satellite network) different hardware and protocols are in operation to form and maintain the network.

In general, it is challenging to gather detailed information about the whole Internet, since the network consists of multiple domains and in many cases it is not possible to collect information across network borders. Hence, potential information sources may be limited to a certain domain.

3.2.3. Partitioning and Grouping of IP Address Ranges

ALTO introduces provider-defined network location identifiers called Provider-defined Identifiers (PIDs) to aggregate network endpoints in the Map Services. Endpoints within one PID may be treated as single entity, assuming proximity based on network topology or other similarity. A key use case of PIDs is to specify network preferences (costs) between PIDs instead of individual endpoints. It is up to the operator of the ALTO server how to group endpoints and how to assign PIDs. For example, a PID may denote a subnet, a set of subnets, a metropolitan area, a POP, an autonomous system, or a set of autonomous systems.

This document only considers deployment scenarios in which PIDs expand to a set of IP address ranges (CIDR). A PID is characterized by a string identifier and its associated set of endpoint addresses [RFC7285]. If an ALTO server offers the Map Service, corresponding identifiers have to be configured.

An automated ALTO implementation may use dynamic algorithms to aggregate network topology. However, it is often desirable to have a mechanism through which the network operator can control the level and details of network aggregation based on a set of requirements and constraints. This will typically be governed by policies that enforce a certain level of abstraction and prevent leakage of sensitive operational data.

For instance, an ALTO server may leverage BGP information that is available in a network’s service provider network layer and compute the group of prefix. An example being BGP communities, which are used in MPLS/IP networks as a common mechanism to aggregate and group prefixes. A BGP community is an attribute used to tag a prefix to group prefixes based on mostly any criteria (as an example, most ISP networks originate BGP prefixes with communities identifying the Point of Presence (PoP) where the prefix has been originated). These BGP communities could be used to map IP address ranges to PIDs. By an additional policy, the ALTO server operator may decide an

arbitrary cost defined between groups. Alternatively, there are algorithms that allow the dynamic computation of costs between groups. The ALTO protocol itself is independent of such algorithms and policies.

3.2.4. Rating Criteria and/or Cost Calculation

An ALTO server indicates preferences amongst network locations in the form of abstract costs. These costs are generic costs and can be internally computed by the operator of the ALTO server according to its own policy. For a given ALTO network map, an ALTO cost map defines directional costs pairwise amongst the set of source and destination network locations defined by the PIDs.

The ALTO protocol permits the use of different cost types. An ALTO cost type is defined by the combination of a cost metric and a cost mode. The cost metric identifies what the costs represent. The cost mode identifies how the costs should be interpreted, i.e., whether returned costs should be interpreted as numerical values or ordinal rankings. The ALTO protocol also allows the definition of additional constraints defining which elements of a cost map shall be returned.

The ALTO protocol specification [RFC7285] defines the "routingcost" cost metric as the basic set of rating criteria, which has to be supported by all implementations. This cost metric conveys a generic measure for the cost of routing traffic from a source to a destination. A lower value indicates a higher preference for traffic to be sent from a source to a destination. How that metric is calculated is up to the ALTO server.

It is possible to calculate the "routingcost" cost metric based on actual routing protocol information. Typically, IGPs provide details about endpoints and links within a given network, while the BGP is used to provide details about links to endpoints in other networks. Besides topology and routing information, networks have a multitude of other attributes about their state, condition, and operation that comprises but is not limited to attributes like link utilization, bandwidth and delay, ingress/egress points of data flows from/towards endpoints outside of the network up to the location of nodes and endpoints.

In order to enable use of extended information, there is a protocol extension procedure to add new ALTO cost types. The following list gives an overview on further rating criteria that have been proposed or that are in use by ALTO-related prototype implementations. This list is not intended as normative text. Instead, its only purpose is to document and discuss rating criteria that have been proposed so far. Whether such rating criteria are useful and whether the

corresponding information would actually be made available by ISPs can also depend on the use case of ALTO. A list of rating criteria for which normative specifications exist and which have successfully passed the IETF review process can be found at IANA's "ALTO Cost Metric Registry", available from [ALTO-REG].

Distance-related rating criteria:

- Relative topological distance: The term relative means that a larger numerical value means greater distance, but it is up to the ALTO service how to compute the values, and the ALTO client will not be informed about the nature of the computation. One way to determine relative topological distance may be counting AS hops, but when querying this parameter, the ALTO client must not assume that the numbers actually are AS hops. In addition to the AS path, a relative cost value could also be calculated taking into account other routing protocol parameters, such as BGP local preference or Multi-Exit Discriminator (MED) attributes.

- Absolute topological distance, expressed in the number of traversed autonomous systems.

- Absolute topological distance, expressed in the number of router hops (i.e., how much the TTL value of an IP packet will be decreased during transit).

- Absolute physical distance, based on knowledge of the approximate geolocation (e.g., continent, country) of an IP address.

Performance-related rating criteria:

- The minimum achievable throughput between the resource consumer and the candidate resource provider, which is considered useful by the application (only in ALTO queries).

- An arbitrary upper bound for the throughput from/to the candidate resource provider (only in ALTO responses). This may be, but is not necessarily, the provisioned access bandwidth of the candidate resource provider.

- The maximum Round-Trip Time (RTT) between resource consumer and the candidate resource provider, which is acceptable for the application for useful communication with the candidate resource provider (only in ALTO queries).

- An arbitrary lower bound for the RTT between resource consumer and the candidate resource provider (only in ALTO responses). This may be, for example, based on measurements of the propagation delay in a completely unloaded network.

Charging-related rating criteria:

- Metrics representing an abstract cost, e.g., determined by policies that distinguish "cheap" from "expensive" IP subnet ranges without detailing the cost function. According to [RFC7285], the abstract metric "routingcost" is an example for a metric for which the cost function does not have to be disclosed.

- Traffic volume caps, in case the Internet access of the resource consumer is not charged with a "flat rate". For each candidate resource location, the ALTO service could indicate the amount of data or the bitrate that may be transferred from/to this resource location until a given point in time, and how much of this amount has already been consumed. Furthermore, an ALTO server may have to indicate how excess traffic would be handled (e.g., blocked, throttled, or charged separately at an indicated price), e.g., by a new endpoint property. This is outside the scope of this document. Also, it is left for further study how several applications would interact if only some of them use this criterion. Also left for further study is the use of such a criterion in resource directories that issue ALTO queries on behalf of other endpoints.

All the above-listed rating criteria are subject to the remarks below:

The ALTO client must be aware that with high probability the actual performance values will differ from whatever an ALTO server exposes. In particular, an ALTO client must not consider a throughput parameter as a permission to send data at the indicated rate without using congestion control mechanisms.

The discrepancies are due to various reasons, including, but not limited to the following facts:

- The ALTO service is not an admission control system.

- The ALTO service may not know the instantaneous congestion status of the network.

- The ALTO service may not know all link bandwidths, i.e., where the bottleneck really is, and there may be shared bottlenecks.

- The ALTO service may not have all information about the actual routing.
- The ALTO service may not know whether the candidate endpoint itself is overloaded.
- The ALTO service may not know whether the candidate endpoint throttles the bandwidth it devotes for the considered application.
- The ALTO service may not know whether the candidate endpoint will throttle the data it sends to the client (e.g., because of some fairness algorithm, such as tit for tat).

Because of these inaccuracies and the lack of complete, instantaneous state information, which are inherent to the ALTO service, the application must use other mechanisms (such as passive measurements on actual data transmissions) to assess the currently achievable throughput, and it must use appropriate congestion control mechanisms in order to avoid a congestion collapse. Nevertheless, the rating criteria may provide a useful shortcut for quickly excluding candidate resource providers from such probing, if it is known in advance that connectivity is in any case worse than what is considered the minimum useful value by the respective application.

Rating criteria that should not be defined for and used by the ALTO service include:

- Performance metrics that are closely related to the instantaneous congestion status. The definition of alternate approaches for congestion control is explicitly out of the scope of ALTO. Instead, other appropriate means, such as using TCP-based transport, have to be used to avoid congestion. In other words, ALTO is a service to provide network and policy information, with update intervals that are possibly several orders of magnitude slower than congestion-control loops (e.g., in TCP) can react on changes in network congestion state. This clear separation of responsibilities avoids traffic oscillations and can help for network stability and cost optimization.
- Performance metrics that raise privacy concerns. For instance, it has been questioned whether an ALTO service should publicly expose the provisioned access bandwidth of cable/DSL customers, as this could enable identification of "premium customers" of an ISP.

3.3. ALTO Focus and Scope

The purpose of this section is ensure that administrators and users of ALTO services are aware of the objectives of the ALTO protocol design. Using ALTO beyond this scope may limit its efficiency. Likewise, Map-based and Endpoint-based ALTO Services may face certain issues during deployment. This section explains these limitations and also outlines potential solutions.

3.3.1. Limitations of Using ALTO beyond Design Assumptions

ALTO is designed as a protocol between clients integrated in applications and servers that provide network information and guidance (e.g., basic network location structure and preferences of network paths). The objective is to modify network resource consumption patterns at application level while maintaining or improving application performance. This design focus results in a number of characteristics of ALTO:

- Endpoint focus: In typical ALTO use cases, neither the consumer of the topology information (i.e., the ALTO client) nor the considered resources (e.g., files at endpoints) are part of the network. The ALTO server presents an abstract network topology containing only information relevant to an application overlay for better-than-random resource provider selection among its endpoints. The ALTO protocol specification [RFC7285] is not designed to expose network internals such as routing tables or configuration data that are not relevant for application-level resource provider selection decisions in network endpoints.

- Abstraction: The ALTO services such as the Network and Cost Map Service or the ECS provide an abstract view of the network only. The operator of the ALTO server has full control over the granularity (e.g., by defining policies how to aggregate subnets into PIDs) and the level of detail of the abstract network representation (e.g., by deciding what cost types to support).

- Multiple administrative domains: The ALTO protocol is designed for use cases where the ALTO server and client can be located in different organizations or trust domains. ALTO assumes a loose coupling between server and client. In addition, ALTO does not assume that an ALTO client has any a priori knowledge about the ALTO server and its supported features. An ALTO server can be discovered automatically.

- Read-only: ALTO is a query/response protocol to retrieve guidance information. Neither network/cost map queries nor queries to the ECS are designed to affect state in the network.

If ALTO shall be deployed for use cases beyond the scope defined by these assumptions, the protocol design may result in limitations.

For instance, in an Application-Based Network Operations (ABNO) environment, the application could issue an explicit service request to the network [RFC7491]. In this case, the application would require detailed knowledge about the internal network topology and the actual state. A network configuration would also require a corresponding security solution for authentication and authorization. ALTO is not designed for operations to control, operate, and manage a network.

Such deployments could be addressed by network management solutions, e.g., based on SNMP [RFC3411] or NETCONF [RFC6241] and YANG [RFC6020], that are typically designed to manipulate configuration state. [RFC7491] contains a more detailed discussion of interfaces between components such as Element Management System (EMS), Network Management System (NMS), Operational Support System (OSS), Traffic Engineering Database (TED), Label Switched Path Database (LSP-DB), Path Computation Element (PCE), and other Operations, Administration, and Maintenance (OAM) components.

3.3.2. Limitations of Map-Based Services and Potential Solutions

The specification of the Map Service in the ALTO protocol [RFC7285] is based on the concept of network maps. A network map partitions the network into PIDs that group one or more endpoints (e.g., subnetworks) to a single aggregate. The "costs" between the various PIDs are stored in a cost map. Map-based approaches such as the ALTO Network and Cost Map Service lower the signaling load on the server as maps have to be retrieved only if they change.

One main assumption for map-based approaches is that the information provided in these maps is static for a long period of time. This assumption is fine as long as the network operator does not change any parameter, e.g., routing within the network and to the upstream peers, and IP address assignment stays stable (and thus the mapping to the partitions). However, there are several cases where this assumption is not valid:

1. ISPs reallocate IP subnets from time to time.

2. ISPs reallocate IP subnets on short notice.

3. IP prefix blocks may be assigned to a router that serves a variety of access networks.

4. Network costs between IP prefixes may change depending on the ISP's routing and traffic engineering.

These effects can be explained as follows:

Case 1: ISPs may reallocate IP subnets within their infrastructure from time to time, partly to ensure the efficient usage of IPv4 addresses (a scarce resource), and partly to enable efficient route tables within their network routers. The frequency of these "renumbering events" depends on the growth in number of subscribers and the availability of address space within the ISP. As a result, a subscriber's household device could retain an IP address for as short as a few minutes or for months at a time or even longer.

It has been suggested that ISPs providing ALTO services could subdivide their subscribers' devices into different IP subnets (or certain IP address ranges) based on the purchased service tier, as well as based on the location in the network topology. The problem is that this sub-allocation of IP subnets tends to decrease the efficiency of IP address allocation, in particular for IPv4. A growing ISP that needs to maintain high efficiency of IP address utilization may be reluctant to jeopardize their future acquisition of IP address space.

However, this is not an issue for map-based approaches if changes are applied in the order of days.

Case 2: ISPs can use techniques that allow the reallocation of IP prefixes on very short notice, i.e., within minutes. An IP prefix that has no IP address assignment to a host anymore can be reallocated to areas where there is currently a high demand for IP addresses.

Case 3: In residential access networks (e.g., DSL, cable), IP prefixes are assigned to broadband gateways, which are the first IP-hop in the access-network between the Customer Premises Equipment (CPE) and the Internet. The access-network between CPE and broadband gateway (called aggregation network) can have varying characteristics (and thus associated costs), but still using the same IP prefix. For instance, one IP address IP1 out of a given CIDR prefix can be assigned to a VDSL access line (e.g., 2 Mbit/s uplink) while another IP address IP2 within the same given CIDR prefix is assigned to a slow ADSL line (e.g., 128 kbit/s uplink). These IP addresses may be assigned on a first come first served basis, i.e., a single IP address out of the same CIDR prefix can change its associated costs quite fast. This may not be an issue with respect to the used upstream provider (thus the cross ISP traffic), but, depending on the capacity of the aggregation network, this may raise to an issue.

Case 4: The routing and traffic engineering inside an ISP network, as well as the peering with other autonomous systems, can change dynamically and affect the information exposed by an ALTO server. As a result, cost maps and possibly also network maps can change.

One solution to deal with map changes is to use incremental ALTO updates [UPDATE-SSE].

3.3.3. Limitations of Non-Map-Based Services and Potential Solutions

The specification of the ALTO protocol [RFC7285] also includes the ECS mechanism. ALTO clients can ask the ALTO server for guidance for specific IP addresses, thereby avoiding the need of processing maps. This can mitigate some of the problems mentioned in the previous section.

However, frequent requests, particularly with long lists of IP addresses, may overload the ALTO server. The server has to rank each received IP address, which causes load at the server. This may be amplified when a large number of ALTO clients are asking for guidance. The results of the ECS are also more difficult to cache than ALTO maps. Therefore, the ALTO client may have to await the server response before starting a communication, which results in an additional delay.

Caching of IP addresses at the ALTO client or the use of the H12 approach [ALTO-H12] in conjunction with caching may lower the query load on the ALTO server.

When an ALTO server receives an ECS request, it may not have the most appropriate topology information in order to accurately determine the ranking. [RFC7285] generally assumes that a server can always offer some guidance. In such a case, the ALTO server could adopt one of the following strategies:

- Reply with available information (best effort).
- Query another ALTO server presumed to have better topology information and return that response (cascaded servers).
- Redirect the request to another ALTO server presumed to have better topology information (redirection).

The protocol mechanisms and decision processes that would be used to determine if redirection is necessary and which mode to use is out of the scope of this document, since protocol extensions could be required.

3.4. Monitoring ALTO

3.4.1. Impact and Observation on Network Operation

ALTO presents a new opportunity for managing network traffic by providing additional information to clients. In particular, the deployment of an ALTO server may shift network traffic patterns, and the potential impact to network operation can be large. An ISP providing ALTO may want to assess the benefits of ALTO as part of the management and operations (cf. [RFC7285]). For instance, the ISP might be interested in understanding whether the provided ALTO maps are effective in order to decide whether an adjustment of the ALTO configuration would be useful. Such insight can be obtained from a monitoring infrastructure. An ISP offering ALTO could consider the impact on (or integration with) traffic engineering and the deployment of a monitoring service to observe the effects of ALTO operations. The measurement of impacts can be challenging because ALTO-enabled applications may not provide related information back to the ALTO service provider.

To construct an effective monitoring infrastructure, the ALTO service provider should decide how to monitor the performance of ALTO and identify and deploy data sources to collect data to compute the performance metrics. In certain trusted deployment environments, it may be possible to collect information directly from ALTO clients. It may also be possible to vary or selectively disable ALTO guidance for a portion of ALTO clients either by time, geographical region, or some other criteria to compare the network traffic characteristics with and without ALTO. Monitoring an ALTO service could also be realized by third parties. In this case, insight into ALTO data may require a trust relationship between the monitoring system operator and the network service provider offering an ALTO service.

The required monitoring depends on the network infrastructure and the use of ALTO, and an exhaustive description is outside the scope of this document.

3.4.2. Measurement of the Impact

ALTO realizes an interface between the network and applications. This implies that an effective monitoring infrastructure may have to deal with both network and application performance metrics. This document does not comprehensively list all performance metrics that could be relevant, nor does it formally specify metrics.

The impact of ALTO can be classified regarding a number of different criteria:

- Total amount and distribution of traffic: ALTO enables ISPs to influence and localize traffic of applications that use the ALTO service. Therefore, an ISP may be interested in analyzing the impact on the traffic, i.e., whether network traffic patterns are shifted. For instance, if ALTO shall be used to reduce the inter-domain P2P traffic, it makes sense to evaluate the total amount of inter-domain traffic of an ISP. Then, one possibility is to study how the introduction of ALTO reduces the total inter-domain traffic (inbound and/our outbound). If the ISP's intention is to localize the traffic inside his network, the network-internal traffic distribution will be of interest. Effectiveness of localization can be quantified in different ways, e.g., by the load on core routers and backbone links or by considering more-advanced effects, such as the average number of hops that traffic traverses inside a domain.

- Application performance: The objective of ALTO is to improve application performance. ALTO can be used by very different types of applications, with different communication characteristics and requirements. For instance, if ALTO guidance achieves traffic localization, one would expect that applications achieve a higher throughput and/or smaller delays to retrieve data. If application-specific performance characteristics (e.g., video or audio quality) can be monitored, such metrics related to user experience could also help to analyze the benefit of an ALTO deployment. If available, selected statistics from the TCP/IP stack in hosts could be leveraged, too.

Of potential interest can also be the share of applications or customers that actually use an offered ALTO service, i.e., the adoption of the service.

Monitoring statistics can be aggregated, averaged, and normalized in different ways. This document does not mandate specific ways how to calculate metrics.

3.4.3. System and Service Performance

A number of interesting parameters can be measured at the ALTO server. [RFC7285] suggests certain ALTO-specific metrics to be monitored:

- Requests and responses for each service listed in an Information Directory (total counts and size in bytes).

- CPU and memory utilization
- ALTO map updates
- Number of PIDs
- ALTO map sizes (in-memory size, encoded size, number of entries)

This data characterizes the workload, the system performance as well as the map data. Obviously, such data will depend on the implementation and the actual deployment of the ALTO service. Logging is also recommended in [RFC7285].

3.4.4. Monitoring Infrastructures

Understanding the impact of ALTO may require interaction between different systems operating at different layers. Some information discussed in the preceding sections is only visible to an ISP, while application-level performance can hardly be measured inside the network. It is possible that not all information of potential interest can directly be measured, either because no corresponding monitoring infrastructure or measurement method exists or because it is not easily accessible.

One way to quantify the benefit of deploying ALTO is to measure before and after enabling the ALTO service. In addition to passive monitoring, some data could also be obtained by active measurements, but due to the resulting overhead, the latter should be used with care. Yet, in all monitoring activities, an ALTO service provider has to take into account that ALTO clients are not bound to ALTO server guidance as ALTO is only one source of information, and any measurement result may thus be biased.

Potential sources for monitoring the use of ALTO include:

- Network monitoring and performance management systems: Many ISPs deploy systems to monitor the network traffic, which may have insight into traffic volumes, network topology, bandwidth information inside the management area. Data can be obtained by SNMP, NETCONF, IP Flow Information Export (IPFIX), syslog, etc. On-demand OAM tests (such as Ping or BDF) could also be used.
- Applications/clients: Relevant data could be obtained by instrumentation of applications.
- ALTO server: If available, log files or other statistics data could be analyzed.

- Other application entities: In several use cases, there are other application entities that could provide data as well. For instance, there may be centralized log servers that collect data.

In many ALTO use cases, some data sources are located within an ISP network while some other data is gathered at the application level. Correlation of data could require a collaboration agreement between the ISP and an application owner, including agreements of data interchange formats, methods of delivery, etc. In practice, such a collaboration may not be possible in all use cases of ALTO, because the monitoring data can be sensitive and because the interacting entities may have different priorities. Details of how to build an overarching monitoring system for evaluating the benefits of ALTO are outside the scope of this memo.

3.5. Abstract Map Examples for Different Types of ISPs

3.5.1. Small ISP with Single Internet Uplink

The ALTO protocol does not mandate how to determine costs between endpoints and/or determine map data. In complex usage scenarios, this can be a non-trivial problem. In order to show the basic principle, this and the following sections explain for different deployment scenarios how ALTO maps could be structured.

For a small ISP, the inter-domain traffic optimizing problem is how to decrease the traffic exchanged with other ISPs, because of high settlement costs. By using the ALTO service to optimize traffic, a small ISP can define two "optimization areas": one is its own network and the other one consists of all other network destinations. The cost map can be defined as follows: the cost of a link between clients of the inner ISP’s network is lower than between clients of the outer ISP’s network and clients of inner ISP’s network. As a result, a host with an ALTO client inside the network of this ISP will prefer retrieving data from hosts connected to the same ISP.

An example is given in Figure 9. It is assumed that ISP A is a small ISP only having one access network. As operator of the ALTO service, ISP A can define its network to be one optimization area, named as PID1, and define other networks to be the other optimization area, named as PID2. C1 is denoted as the cost inside the network of ISP A. C2 is denoted as the cost from PID2 to PID1, and C3 from PID1 to PID2. In the following, C2=C3 is assumed for the sake of simplicity. In order to keep traffic local inside ISP A, it makes sense to define C1<C2.

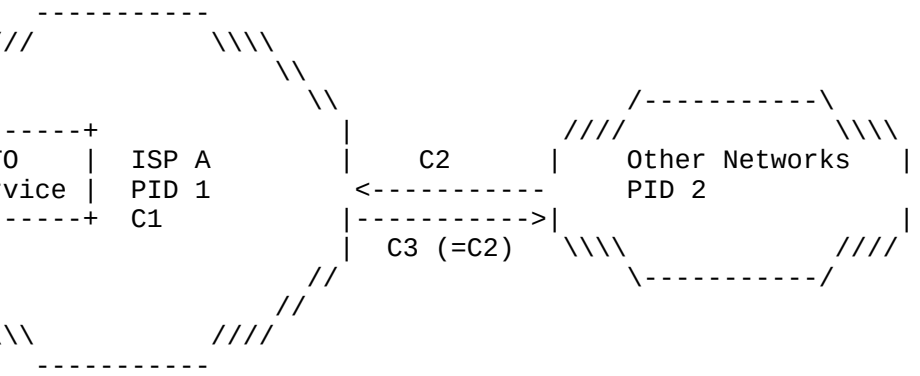

Figure 9: Example ALTO Deployment for a Small ISP

A simplified extract of the corresponding ALTO network and cost maps is listed in Figures 10 and 11, assuming that the network of ISP A has the IPv4 address ranges 192.0.2.0/24 and 198.51.100.0/25, as well as the IPv6 address range 2001:db8:100::/48. In this example, the cost values C1 and C2 can be set to any number C1<C2.

```
HTTP/1.1 200 OK
...
Content-Type: application/alto-networkmap+json

{
 ...
  "network-map" : {
    "PID1" : {
      "ipv4" : [
        "192.0.2.0/24",
        "198.51.100.0/25"
      ],
      "ipv6" : [
        "2001:db8:100::/48"
      ]
    },
    "PID2" : {
      "ipv4" : [
        "0.0.0.0/0"
      ],
      "ipv6" : [
        "::/0"
      ]
    }
  }
}
```

Figure 10: Example ALTO Network Map

```
HTTP/1.1 200 OK
...
Content-Type: application/alto-costmap+json

{
    ...
    "cost-type" : {"cost-mode"  : "numerical",
                   "cost-metric": "routingcost"
    }
  },
  "cost-map" : {
    "PID1": { "PID1": C1,  "PID2": C2 },
    "PID2": { "PID1": C2,  "PID2": 0 },
  }
}
```

Figure 11: Example ALTO Cost Map

3.5.2. ISP with Several Fixed-Access Networks

This example discusses a P2P application traffic optimization use case for a larger ISP with a fixed network comprising several access networks and a core network. The traffic optimizing objectives include (1) using the backbone network efficiently, (2) adjusting the traffic balance in different access networks according to traffic conditions and management policies, and (3) achieving a reduction of settlement costs with other ISPs.

Such a large ISP deploying an ALTO service may want to optimize its traffic according to the network topology of its access networks. For example, each access network could be defined to be one optimization area, i.e., traffic should be kept local withing that area if possible. This can be achieved by mapping each area to a PID. Then, the costs between those access networks can be defined according to a corresponding traffic optimizing requirement by this ISP. One example setup is further described below and also shown in Figure 12.

In this example, ISP A has one backbone network and three access networks, named as AN A, AN B, and AN C. A P2P application is used in this example. For a reasonable application-level traffic optimization, the first requirement could be a decrease of the P2P traffic on the backbone network inside the AS of ISP A and the second requirement could be a decrease of the P2P traffic to other ISPs, i.e., other ASes. The second requirement can be assumed to have priority over the first one. Also, we assume that the settlement rate with ISP B is lower than with other ISPs. ISP A can deploy an ALTO service to meet these traffic distribution requirements. In the following, we will give an example of an ALTO setting and configuration according to these requirements.

In the network of ISP A, the operator of the ALTO server can define each access network to be one optimization area, and assign one PID to each access network, such as PID 1, PID 2, and PID 3. Because of different peerings with different outer ISPs, one can define ISP B to be one additional optimization area and assign PID 4 to it. All other networks can be added to a PID to be one further optimization area (PID 5).

In the setup, costs (C1, C2, C3, C4, C5, C6, C7, C8) can be assigned as shown in Figure 12. Cost C1 is denoted as the link cost in inner AN A (PID 1), and C2 and C3 are defined accordingly. C4 is denoted as the link cost from PID 1 to PID 2, and C5 is the corresponding cost from PID 3, which is assumed to have a similar value. C6 is the cost between PID 1 and PID 3. For simplicity, this scenario assumes

symmetrical costs between the AN this example. C7 is denoted as the link cost from the ISP B to ISP A. C8 is the link cost from other networks to ISP A.

According to previous discussion of the first requirement and the second requirement, the relationship of these costs will be defined as: (C1, C2, C3) < (C4, C5, C6) < (C7) < (C8)

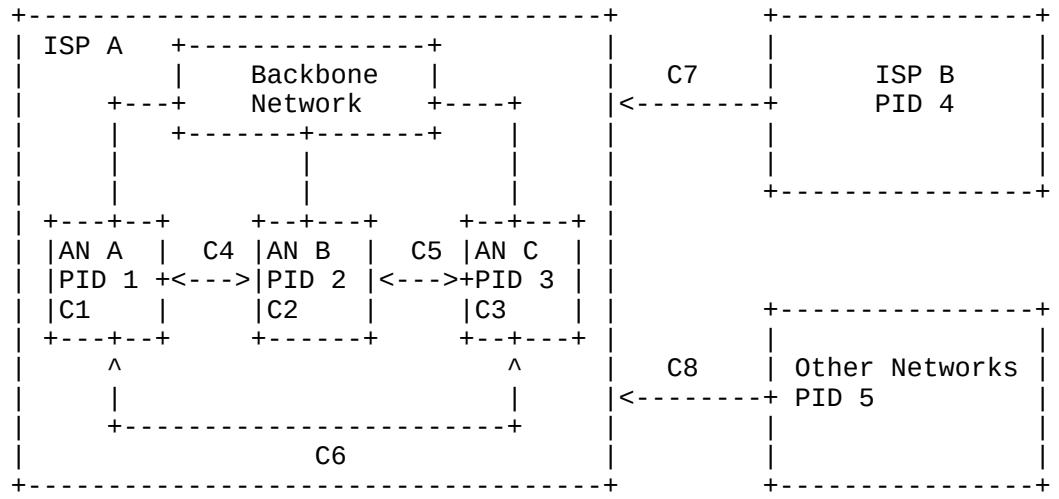

Figure 12: ALTO Deployment in Large ISPs with Layered Fixed-Network Structures

3.5.3. ISP with Fixed and Mobile Network

An ISP with both mobile network and fixed network may focus on optimizing the mobile traffic by keeping traffic in the fixed network as much as possible, because wireless bandwidth is a scarce resource and traffic is costly in mobile network. In such a case, the main requirement of traffic optimization could be decreasing the usage of radio resources in the mobile network. An ALTO service can be deployed to meet these needs.

Figure 13 shows an example: ISP A operates one mobile network, which is connected to a backbone network. The ISP also runs two fixed-access networks AN A and AN B, which are also connected to the backbone network. In this network structure, the mobile network can be defined as one optimization area, and PID 1 can be assigned to it. Access networks AN A and B can also be defined as optimization areas, and PID 2 and PID 3 can be assigned, respectively. The cost values are then defined as shown in Figure 13.

To decrease the usage of wireless link, the relationship of these costs can be defined as follows:

From view of mobile network: C4 < C1 and C4 = C8. This means that clients in mobile network requiring data resources from other clients will prefer clients in AN A or B to clients in the mobile network. This policy can decrease the usage of wireless link and power consumption in terminals.

From view of AN A: C2 < C6, C5 = maximum cost. This means that clients in other optimization area will avoid retrieving data from the mobile network.

From view of AN B: Analog to the view of AN A, C3 < C8 and C9 = maximum cost.

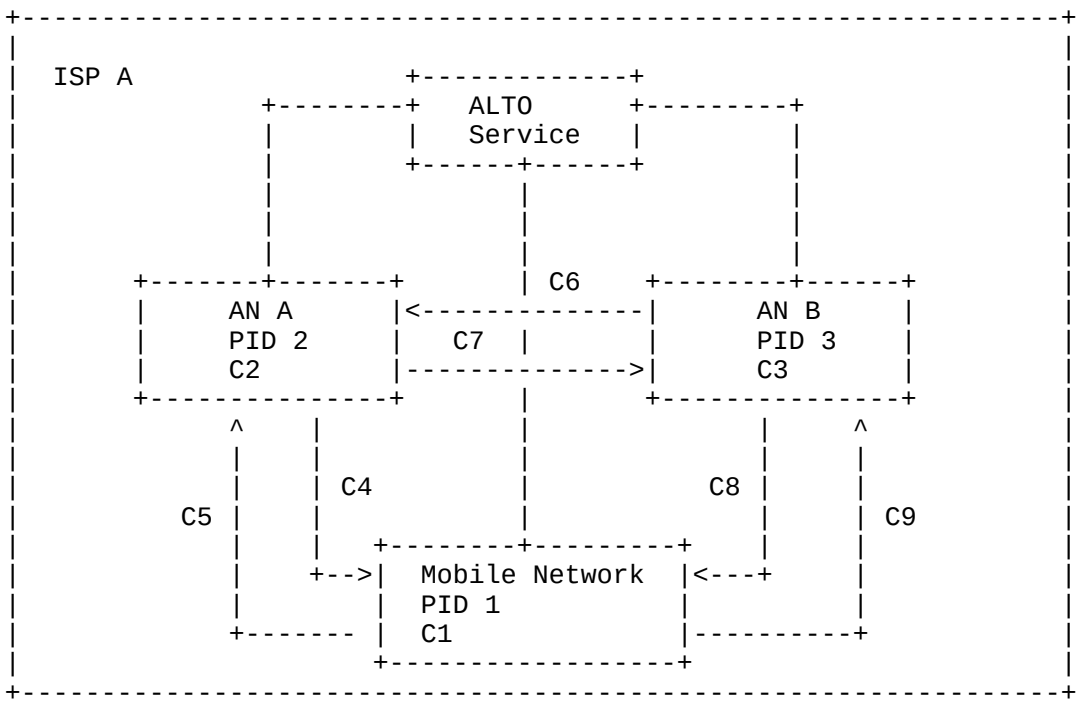

Figure 13: ALTO Deployment in ISPs with Mobile Network

These examples show that for ALTO in particular the relationships between different costs matter; the operator of the server has several degrees of freedom how to set the absolute values.

3.6. Comprehensive Example for Map Calculation

In addition to the previous, abstract examples, this section presents a more detailed scenario with a realistic IGP and BGP routing protocol configuration. This example was first described in [MAP-CALC].

3.6.1. Example Network

Figure 14 depicts a network that is used to explain the steps carried out in the course of this example. The network consists of nine routers (R1 to R9). Two of them are border routers (R1 + R8) connected to neighbored networks (AS 2 to AS 4). Furthermore, AS 4 is not directly connected to the local network, but has AS 3 as transit network. The links between the routers are point-to-point connections. These connections also form the core network with the 2001:db8:1:0::/56 prefix. This prefix is large enough to provide addresses for all router interconnections. In addition to the core network, the local network also has five client networks attached to five different routers (R2, R5, R6, R7 and R9). Each client network has a /56 prefix with 2001:db8:1:x00:: (x = [1..5]) as network address.

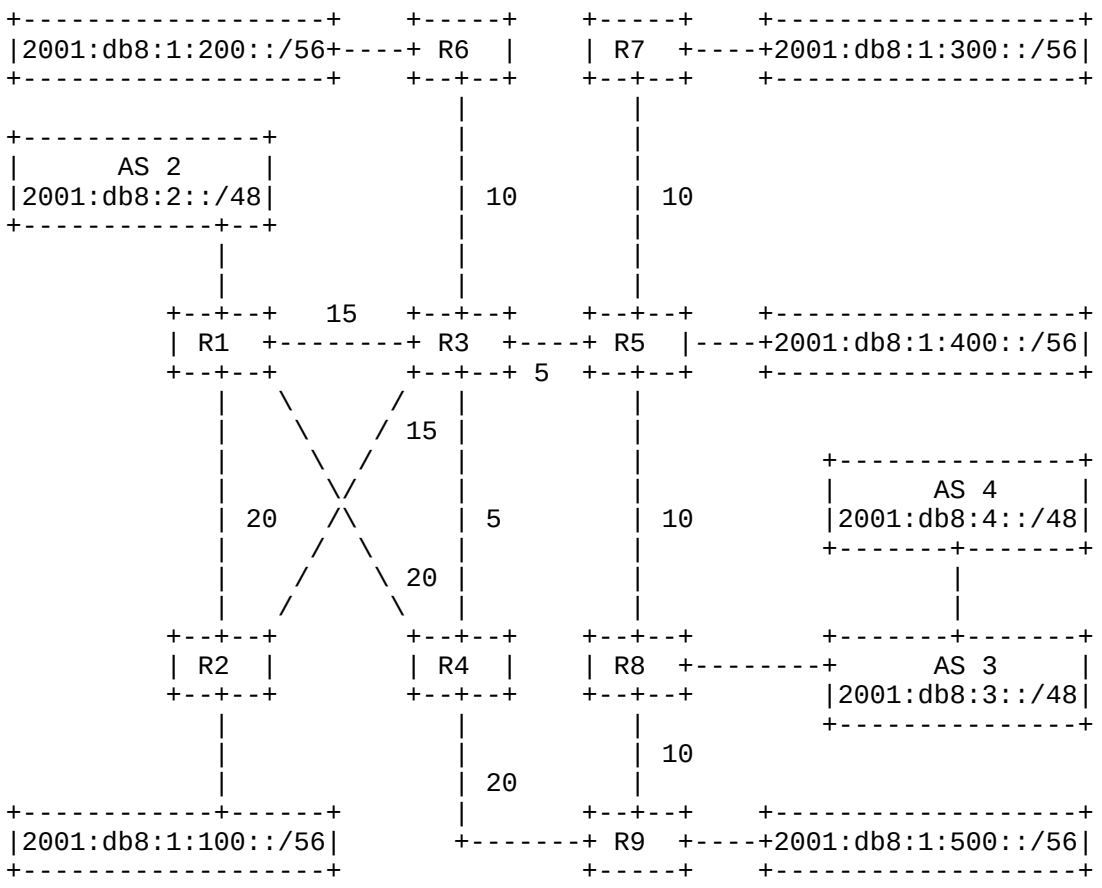

Figure 14: Example Network

The example network utilizes two different routing protocols, one for IGP and another for EGP routing. The used IGP is a link-state protocol such as IS-IS. The applied link weights are annotated in the graph and additionally shown in Figure 15. All links are bidirectional and their weights are symmetric. To obtain the topology and routing information from the network, the topology data source must be connected directly to one of the routers (R1...R9). Furthermore, the topology data source must be enabled to communicate with the router and vice versa.

The BGP is used in this scenario to route between autonomous systems. External BGP is running on the two border routers R1 and R8. Furthermore, internal BGP is used to propagate external as well as internal prefixes within the network boundaries; it is running on every router with an attached client network (R2, R5, R6, R7 and R9).

Since no route reflector is present it is necessary to fetch routes from each BGP router separately.

	R1	R2	R3	R4	R5	R6	R7	R8	R9
R1	0	15	15	20	-	-	-	-	-
R2	15	0	20	-	-	-	-	-	-
R3	15	20	0	5	5	10	-	-	-
R4	20	-	5	0	5	-	-	-	20
R5	-	-	5	5	0	-	10	10	-
R6	-	-	10	-	-	0	-	-	-
R7	-	-	-	-	10	-	0	-	-
R8	-	-	-	-	10	-	-	0	10
R9	-	-	-	20	-	-	-	10	0

Figure 15: Example Network Link Weights

For monitoring purposes, it is possible to enable, e.g., SNMP or NETCONF on the routers within the network. This way an ALTO server may obtain several additional information about the state of the network. For example, utilization, latency, and bandwidth information could be retrieved periodically from the network components to get and keep an up-to-date view on the network situation.

In the following, it is assumed that the listed attributes are collected from the network:

- IS-IS: topology, link weights
- BGP: prefixes, AS numbers, AS distances, or other BGP metrics
- SNMP: latency, utilization, bandwidth

3.6.2. Potential Input Data Processing and Storage

Due to the variety of data sources available in a network, it may be necessary to aggregate the information and define a suitable data model that can hold the information efficiently and easily accessible. One potential model is an annotated directed graph that represents the topology. The attributes can be annotated at the corresponding positions in the graph. The following shows how such a topology graph could describe the example topology.

In the topology graph, a node represents a router in the network, while the edges stand for the links that connect the routers. Both routers and links have a set of attributes that store information gathered from the network.

Each router could be associated with a basic set of information, such as:

- ID: Unique ID within the network to identify the router.
- Neighbor IDs: List of directly connected routers.
- Endpoints: List of connected endpoints. The endpoints may also have further attributes themselves depending on the network and address type. Such potential attributes are costs for reaching the endpoint from the router, AS numbers, or AS distances.

In addition to the basic set, many more attributes may be assigned to router nodes. This mainly depends on the utilized data sources. Examples for such additional attributes are geographic location, host name and/or interface types, just to name a few.

The example network shown in Figure 14 represents such an internal network graph where the routers R1 to R9 represent the nodes and the connections between them are the links. For instance, R2 has one directly attached IPv6 endpoint that belongs to its own AS, as shown in Figure 16.

```
ID:  2

Neighbor IDs:  1,3 (R1, R3)

Endpoints:

   Endpoint:  2001:db8:1:100::/56

   Weight:  10 (e.g., the default IGP metric value)

   ASNumber:  1 (our own AS)

   ASDistance:  0

Host Name:  R2
```

Figure 16: Example Router R2

Router R8 has two attached IPv6 endpoints, as explained in Figure 17. The first one belongs to a directly neighbored AS with AS number 3. The AS distance from our network to AS3 is 1. The second endpoint belongs to an AS (AS4) that is no direct neighbor but directly connected to AS3. To reach endpoints in AS4, it is necessary to cross AS3, which increases the AS distance by one.

```
ID:  8

Neighbor IDs:  5,9 (R5, R9)

Endpoints:

   Endpoint:  2001:db8:3::/48

   Weight:  100

   ASNumber:  3

   ASDistance:  1

   Endpoint:  2001:db8:4::/48

   Weight:  200

   ASNumber:  4

   ASDistance:  2

Host Name:  R8
```

Figure 17: Example Router R8

A potential set of attributes for a link is described in the following list:

- Source ID: ID of the source router of the link.
- Destination ID: ID of the destination router of the link.
- Weight: The cost to cross the link, e.g., defined by the used IGP.

Additional attributes that provide technical details and state information can be assigned to links as well. The availability of such additional attributes depends on the utilized data sources. Such attributes can be characteristics like maximum bandwidth, utilization, or latency on the link as well as the link type.

In the example, the link attributes are equal for all links and only their values differ. It is assumed that the attributes utilization, bandwidth, and latency are collected, e.g., via SNMP or NETCONF. In the topology of Figure 14, the links between R1 and R2 would then have the following link attributes explained in Figure 18:

R1->R2:

Source ID: 1

Destination ID: 2

Weight: 15

Bandwidth: 10 Gbit/s

Utilization: 0.1

Latency: 2 ms

R2->R1:

Source ID: 2

Destination ID: 1

Weight: 15

Bandwidth: 10 Gbit/s

Utilization: 0.55

Latency: 5 ms

Figure 18: Link Attributes

It has to be emphasized that values for utilization and latency can be very volatile.

3.6.3. Calculation of Network Map from the Input Data

The goal of the ALTO map calculation process is to get from the graph representation of the network to a coarser-grained and abstract matrix representation. The first step is to generate the network map. Only after the network map has been generated is it possible to compute the cost map since it relies on the network map.

To generate an ALTO network map, a grouping function is required. A grouping function processes information from the network graph to group endpoints into PIDs. The way of grouping is manifold and algorithms can utilize any information provided by the network graph to perform the grouping. The functions may omit certain endpoints in

order to simplify the map or in order to hide details about the network that are not intended to be published in the resulting ALTO network map.

For IP endpoints, which are either an IP (version 4 or version 6) address or prefix, [RFC7285] requires the use of a longest-prefix matching algorithm to map IPs to PIDs. This requirement results in the constraints that every IP must be mapped to a PID and the same prefix or address not be mapped to more than one PID. To meet the first constraint, every calculated map must provide a default PID that contains the prefixes 0.0.0.0/0 for IPv4 and ::/0 for IPv6. Both prefixes cover their entire address space, and if no other PID matches an IP endpoint, the default PID will. The second constraint must be met by the grouping function that assigns endpoints to PIDs. In case of collision, the grouping function must decide to which PID an endpoint is assigned. These or other constraints may apply to other endpoint types depending on the used matching algorithm.

A simple example for such grouping is to compose PIDs by host names. For instance, each router's host name is selected as the name for a PID and the attached endpoints are the member endpoints of the corresponding PID. Additionally, backbone prefixes should not appear in the map so they are filtered out. The following table in Figure 19 shows the resulting ALTO network map, using the network in Figure 14 as example:

PID	Endpoints
R1	2001:db8:2::/48
R2	2001:db8:1:100::/56
R5	2001:db8:1:400::/56
R6	2001:db8:1:200::/56
R7	2001:db8:1:300::/56
R8	2001:db8:3::/48, 2001:db8:4::/48
R9	2001:db8:1:500::/56
default	0.0.0.0/0, ::/0

Figure 19: Example ALTO Network Map

Since router R3 and R4 have no endpoints assigned, they are not represented in the network map. Furthermore, as previously mentioned, the "default" PID was added to represent all endpoints that are not part of the example network.

3.6.4. Calculation of Cost Map

After successfully creating the network map, the typical next step is to calculate the costs between the generated PIDs, which form the cost map. Those costs are calculated by cost functions. A cost function may calculate unidirectional values, which means it is necessary to compute the costs from every PID to every PID. In general, it is possible to use all available information in the network graph to compute the costs. In case a PID contains more than one IP address or prefix, the cost function may first calculate a set of cost values for each source/destination IP pair. In that case, a tiebreaker function is required to decide the resulting cost value, as [RFC7285] allows one cost value only between two PIDs. Such a tiebreaker can be a simple function such as minimum, maximum, or average value.

No matter what metric the cost function uses, the path from source to destination is usually defined by the path with minimum weight. When the link weight is represented by an additive metric, the path weight is the sum of link weights of all traversed links. The path may be determined, for instance, with the Bellman-Ford or Dijkstra algorithms. The latter progressively builds the shortest path in terms of cumulated link lengths. In our example, the link lengths are link weights with values illustrated in Figure 15. Hence, the cost function generally extracts the optimal path with respect to a chosen metric, such as the IGP link weight. It is also possible that more than one path with the same minimum weight exists, which means it is not entirely clear which path is going to be selected by the network. Hence, a tiebreaker similar to the one used to resolve costs for PIDs with multiple endpoints is necessary.

An important note is that [RFC7285] does not require cost maps to provide costs for every PID pair, so if no path cost can be calculated for a certain pair, the corresponding field in the cost map is left out. Administrators may also not want to provide cost values for some PID pairs due to various reasons. Such pairs may be defined before the cost calculation is performed.

Based on the network map example shown in the previous section, it is possible to calculate the cost maps. Figure 20 provides an example where the selected metric for the cost map is the minimum number of hops necessary to get from the endpoints in the source PID to endpoints in the destination PID. Our chosen tiebreaker selects the minimum hop count when more than one value is returned by the cost function.

PID	default	R1	R2	R5	R6	R7	R8	R9
default	x	x	x	x	x	x	x	x
R1	x	0	2	3	3	4	4	3
R2	x	2	0	3	3	4	4	4
R5	x	3	3	0	3	2	2	3
R6	x	3	3	3	0	4	4	4
R7	x	4	4	2	4	0	3	4
R8	x	4	4	2	4	3	0	2
R9	x	3	4	3	4	4	2	0

Figure 20: Example ALTO Hop Count Cost Map

It should be mentioned that R1->R9 has several paths with equal path weights. The paths R1->R3->R5->R8->R9, R1->R3->R4->R9, and R1->R4->R9 all have a path weight of 40. Due to the minimum hop count value tiebreaker, 3 hops is chosen as value for the path R1->R4->R9. Furthermore, since the "default" PID is, in a sense, a virtual PID with no endpoints that are part of the example network, no cost values are calculated for other PIDs from or towards it.

3.7. Deployment Experiences

There are multiple interoperable implementations of the ALTO protocol. Some experiences in implementing and using ALTO for large-scale networks have been documented in [MAP-CALC] and are here summarized:

- Data collection: Retrieving topology information typically requires implementing several protocols other than ALTO for data collection. For such other protocols, ALTO deployments faced protocol behaviors that were different from what would be expected from the specification of the corresponding protocol. This includes behavior caused by older versions of the protocol specification, a lax interpretation on the remote side or simply incompatibility with the corresponding standard. This sort of problems in collecting data can make an ALTO deployment more complicated, even if it is unrelated to ALTO protocol itself.

- Data processing: Processing network information can be very complex and quite resource demanding. Gathering information from an autonomous system connected to the Internet may imply that a server must store and process hundreds of thousands of prefixes, several hundreds of megabytes of IPFIX/Netflow information per minute, and information from hundreds of routers and attributes of thousands of links. A lot of disk memory, RAM, and CPU cycles as well as efficient algorithms are required to process the

information. Operators of an ALTO server have to be aware that significant compute resources are not only required for the ALTO server, but also for the corresponding data collection.

- Network map calculation: Large IP-based networks consist of hundreds of thousands of prefixes that have to be mapped to PIDs in the process of network map calculation. As a result, network maps get very large (up to tens of megabytes). However, depending on the design of the network and the chosen grouping function the calculated network maps contains redundancy that can be removed. There are at least two ways to reduce the size by removing redundancy. First, adjacent IP prefixes can be merged. When a PID has two adjacent prefix entries it can merge them together to one larger prefix. It is mandatory that both prefixes be in the same PID. However, the large prefix being assigned to another PID cannot be ruled out. This must be checked, and it is up to the grouping function whether or not to merge the prefixes and remove the larger prefix from the other PID. A simple example, when a PID comprises the prefixes 2001:db8:0:0::/64 and 2001:db8:0:1::/64 it can easily merge them to 2001:db8:0:0::/63. Second, a prefix and its next-longest-prefix match may be in the same PID. In this case, the smaller prefix can simply be removed since it is redundant for obvious reasons. A simple example, a PID comprises the prefixes 2001:db8:0:0::/62 and 2001:db8:0:1::/64 and the /62 is the next-longer prefix match of the /64, the /64 prefix can simply be removed. In contrast, if another PID contains the 2001:db8:0:0::/63 prefix, the entry 2001:db8:0:1::/64 cannot be removed since the next-longest prefix is not in the same PID anymore. Operators of an ALTO server thus have to analyze whether their address assignment schemes allows such tuning.

- Cost map calculation: One known implementation challenge with cost map calculations is the vast amount of CPU cycles that may be required to calculate the costs in large networks. This is particular problematic if costs are calculated between the endpoints of each source-destination PID pair. Very often several to many endpoints of a PID are attached to the same node, so the same path cost is calculated several times. This is clearly inefficient. A remedy could be more sophisticated algorithms, such as looking up the routers the endpoints of each PID are connected to in our network graph and calculated cost map based on the costs between the routers. When deploying and configuring ALTO servers, administrators should consider the impact of huge cost maps and possibly ensure that map sizes do not get too large.

In addition, further deployment experiences have been documented. One real example is described in greater detail in reference [CHINA-TRIAL].

Also, experiments have been conducted with ALTO-like deployments in ISP networks. For instance, NTT performed tests with their HINT server implementation and dummy nodes to gain insight on how an ALTO-like service can influence peer-to-peer systems [RFC6875]. The results of an early experiment conducted in the Comcast network are documented in [RFC5632].

4. Using ALTO for P2P Traffic Optimization

4.1. Overview

4.1.1. Usage Scenario

Originally, P2P applications were the main driver for the development of ALTO. In this use case, it is assumed that one party (usually the operator of a "managed" IP network domain) will disclose information about the network through ALTO. The application overlay will query this information and optimize its behavior in order to improve performance or Quality of Experience in the application while reducing the utilization of the underlying network infrastructure. The resulting win-win situation is assumed to be the incentive for both parties to provide or consume the ALTO information, respectively.

P2P systems can be built with or without use of a centralized resource directory ("tracker"). The scope of this section is the interaction of P2P applications with the ALTO service. In this scenario, the resource consumer ("peer") asks the resource directory for a list of candidates that can provide the desired resource. There are different options for how ALTO can be deployed in such use cases with a centralized resource directory.

For efficiency reasons (i.e., message size), only a subset of all resource providers known to the resource directory will be returned to the resource consumer. Some or all of these resource providers, plus further resource providers learned by other means such as direct communication between peers, will be contacted by the resource consumer for accessing the resource. The purpose of ALTO is giving guidance on this peer selection, which should yield better-than-random results. The tracker response as well as the ALTO guidance are most beneficial in the initial phase after the resource consumer has decided to access a resource, as long as only few resource providers are known. Later, when the resource consumer has already exchanged some data with other peers and measured the transmission speed, the relative importance of ALTO may dwindle.

4.1.2. Applicability of ALTO

A tracker-based P2P application can leverage ALTO in different ways. In the following, the different alternatives and their pros and cons are discussed.

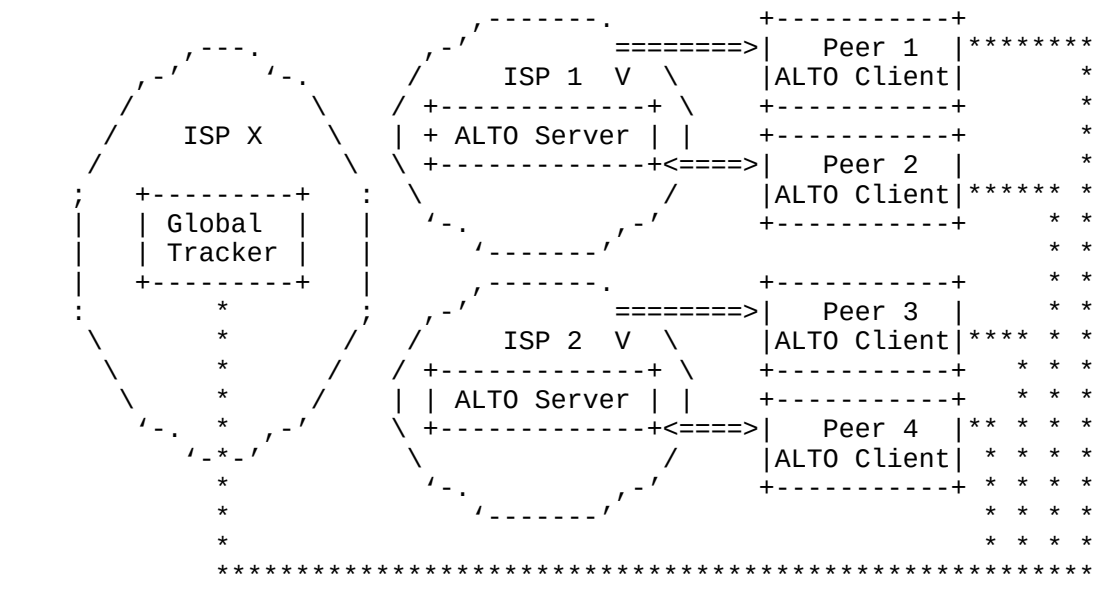

Figure 21: Global Tracker and Local ALTO Servers

Figure 21 depicts a tracker-based P2P system with several peers. The peers (i.e., resource consumers) embed an ALTO client to improve the resource provider selection. The tracker (i.e., resource directory) itself may be hosted and operated by another entity. A tracker external to the ISPs of the peers may be a typical use case. For instance, a tracker like Pirate Bay can serve BitTorrent peers worldwide. The figure only shows one tracker instance, but deployments with several trackers could be possible, too.

The scenario depicted in Figure 21 lets the peers directly communicate with their ISP's ALTO server (i.e., ALTO client embedded in the peers), thus giving the peers the most control on which information they query for, as they can integrate information received from one tracker or several trackers and through direct peer-to-peer knowledge exchange. For instance, the latter approach

is called peer exchange (PEX) in BitTorrent. In this deployment scenarios, the peers have to discover a suitable ALTO server (e.g., offered by their ISP, as described in [RFC7286]).

There are also tracker-less P2P system architectures that do not rely on centralized resource directories, e.g., unstructured P2P networks. Regarding the use of ALTO, their deployment would be similar to Figure 21, since the ALTO client would be embedded in the peers as well. This option is not further considered in this memo.

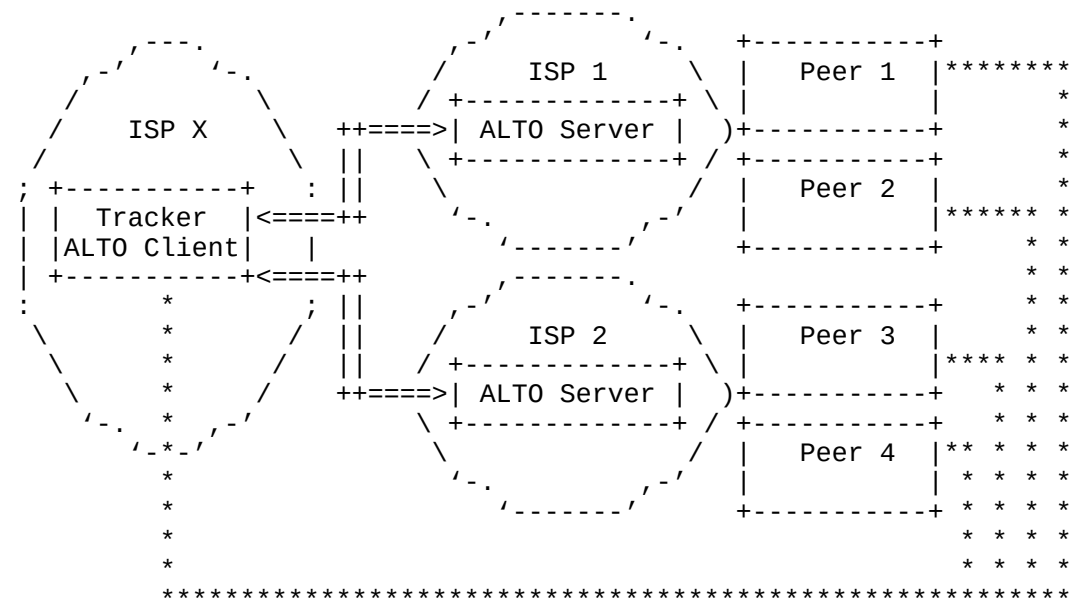

Figure 22: Global Tracker Accessing ALTO Server at Various ISPs

An alternative deployment scenario for a tracker-based system is depicted in Figure 22. Here, the tracker embeds the ALTO client. When the tracker receives a request from a querying peer, it first discovers the ALTO server responsible for the querying peer. This discovery can be done by using various ALTO server discovery mechanisms [RFC7286] [XDOM-DISC]. The ALTO client subsequently sends to the querying peer only those peers that are preferred by the ALTO server responsible for the querying peer. The peers do not query the ALTO servers themselves. This gives the peers a better initial selection of candidates, but does not consider peers learned through direct peer-to-peer knowledge exchange.

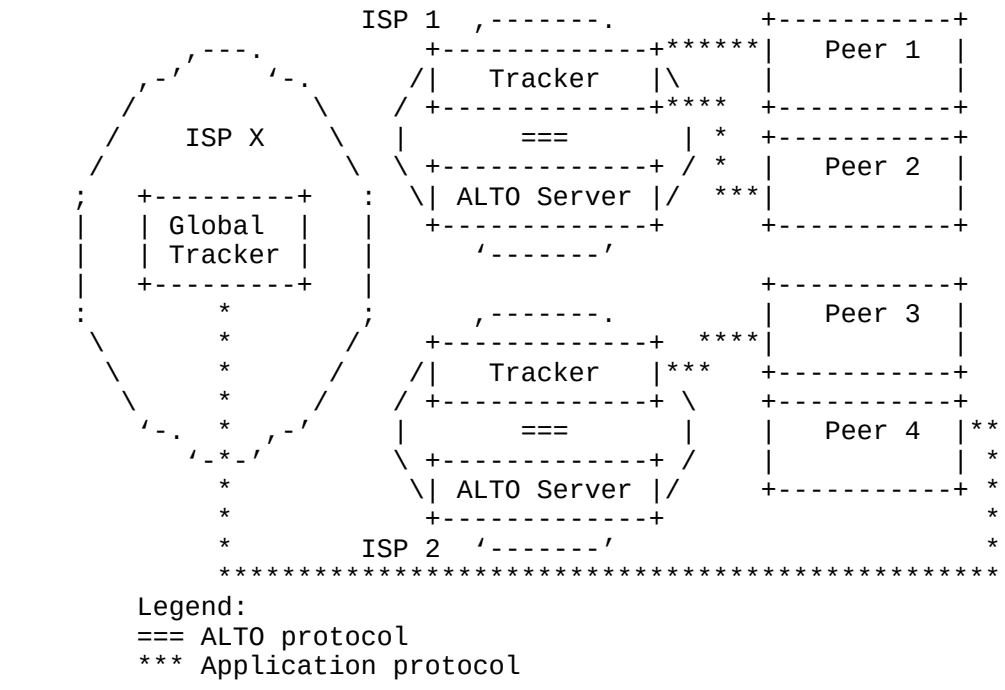

Figure 23: Local Trackers and Local ALTO Servers (P4P Approach)

There are some attempts to let ISPs deploy their own trackers, as shown in Figure 23. In this case, the client cannot get guidance from the ALTO server other than by talking to the ISP's tracker, which in turn communicates with the ALTO server using the ALTO protocol. It should be noted that the peers are still allowed to contact other trackers operated by entities other than the peer's ISP, but in this case they cannot benefit from ALTO guidance.

4.2. Deployment Recommendations

4.2.1. ALTO Services

The ALTO protocol specification [RFC7285] details how an ALTO client can query an ALTO server for guiding information and receive the corresponding replies. In case of peer-to-peer networks, two different ALTO services can be used: the cost map service is often preferred as solution by peer-to-peer software implementors and users, since it avoids disclosing peer IP addresses to a centralized entity. Alternatively, network operators may have a preference for the ECS, since it does not require exposure of the network topology.

For actual use of ALTO in P2P applications, both software vendors and network operators have to agree which ALTO services to use. The ALTO protocol is flexible and supports both services. Note that for other use cases of ALTO, in particular in more controlled environments, both the cost map service and the ECS might be feasible; it is more of an engineering trade-off whether to use a map-based or query-based ALTO service.

4.2.2. Guidance Considerations

As explained in Section 4.1.2, for a tracker-based P2P application, there are two fundamentally different possibilities where to place the ALTO client:

1. ALTO client in the resource consumer ("peer")

2. ALTO client in the resource directory ("tracker")

Both approaches have advantages and drawbacks that have to be considered. If the ALTO client is in the resource consumer (Figure 21), a potentially very large number of clients has to be deployed. Instead, when using an ALTO client in the resource directory (Figures 22 and 23), ostensibly peers do not have to directly query the ALTO server. In this case, an ALTO server could even not permit access to peers.

However, it seems to be beneficial for all participants to let the peers directly query the ALTO server. Considering the plethora of different applications that could use ALTO, e.g., multiple-tracker-based or non-tracker-based P2P systems or other applications searching for relays, this renders the ALTO service more useful. The peers are also the single point having all operational knowledge to decide whether to use the ALTO guidance and how to use the ALTO guidance. For a given peer, one can also expect that an ALTO server of the corresponding ISP provides useful guidance and can be discovered.

Yet, ALTO clients in the resource consumer also have drawbacks compared to use in the resource directory. In the following, both scenarios are compared more in detail in order to explain the impact on ALTO guidance and the need for third-party ALTO queries.

In the first scenario (see Figure 24), the peer (resource consumer) queries the tracker (resource directory) for the desired resource (F1). The resource directory returns a list of potential resource providers without considering ALTO (F2). It is then the duty of the resource consumer to invoke ALTO (F3/F4), in order to solicit guidance regarding this list.

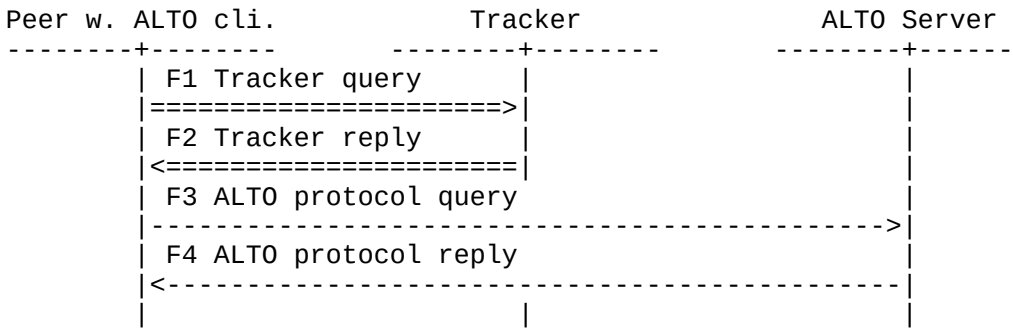

Figure 24: Basic Message Sequence Chart for Resource-Consumer-Initiated ALTO Query

In the second scenario (see Figure 25), the resource directory has an embedded ALTO client, which we will refer to as Resource Directory ALTO Client (RDAC) in this document. After receiving a query for a given resource (F1), the resource directory invokes the RDAC to evaluate all resource providers it knows (F2/F3). Then, it returns a, possibly shortened, list containing the "best" resource providers to the resource consumer (F4).

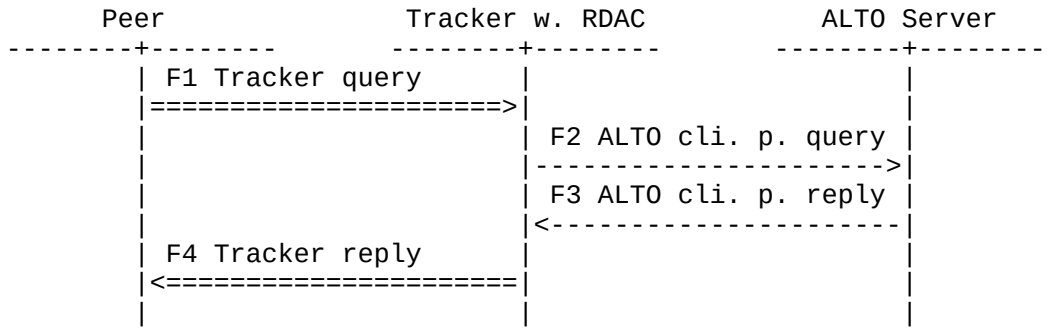

Figure 25: Basic Message Sequence Chart for Third-Party ALTO Query

Note: The message sequences depicted in Figures 24 and 25 may occur both in the target-aware and the target-independent query mode (cf. [RFC6708]). In the target-independent query mode, no message exchange with the ALTO server might be needed after the tracker

query, because the candidate resource providers could be evaluated using a locally cached "map", which has been retrieved from the ALTO server some time ago.

The first approach has the following problem: While the resource directory might know thousands of peers taking part in a swarm, the list returned to the resource consumer is usually shortened for efficiency reasons. Therefore, the "best" (in the sense of ALTO) potential resource providers might not be contained in that list anymore, even before ALTO can consider them.

Much better traffic optimization could be achieved if the tracker would evaluate all known peers using ALTO. This list would then include a significantly higher fraction of "good" peers. If the tracker returned "good" peers only, there might be a risk that the swarm might disconnect and split into several disjunct partitions. However, finding the right mix of ALTO-biased and random peer selection is out of the scope of this document.

Therefore, from an overall optimization perspective, the second scenario with the ALTO client embedded in the resource directory is advantageous, because it is ensured that the addresses of the "best" resource providers are actually delivered to the resource consumer. An architectural implication of this insight is that the ALTO server discovery procedures must support third-party discovery. That is, as the tracker issues ALTO queries on behalf of the peer that contacted the tracker, the tracker must be able to discover an ALTO server that can give guidance suitable for that respective peer (see [XDOM-DISC]).

In principle, a combined approach could also be possible. For instance, a tracker could use a coarse-grained "global" ALTO server to find the peers in the general vicinity of the requesting peer, while peers could use "local" ALTO servers for a more fine-grained guidance. Yet, there is no known deployment experience for such a combined approach.

5. Using ALTO for CDNs

5.1. Overview

5.1.1. Usage Scenario

This section briefly introduces the usage of ALTO for CDNs, as explained in [CDN-USE]. CDNs are used in the delivery of some Internet services (e.g., delivery of websites, software updates, and video delivery) from a location closer to the location of the user. A CDN typically consists of a network of servers often attached to

ISP networks. The point of attachment is often as close to content consumers and peering points as economically or operationally feasible in order to decrease traffic load on the ISP backbone and to provide better user experience measured by reduced latency and higher throughput.

CDNs use several techniques to redirect a client to a server (surrogate). A request-routing function within a CDN is responsible for receiving content requests from user agents, obtaining and maintaining necessary information about a set of candidate surrogates, and selecting and redirecting the user agent to the appropriate surrogate. One common way is relying on the DNS system, but there are many other ways, see [RFC3568].

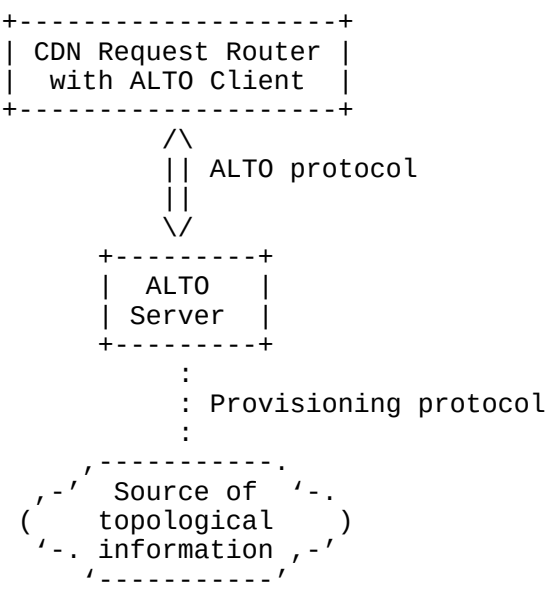

Figure 26: Use of ALTO Information for CDN Request Routing

In order to derive the optimal benefit from a CDN, it is preferable to deliver content from the servers (caches) that are "closest" to the end user requesting the content. The definition of "closest" may be as simple as geographical or IP topology distance, but it may also consider other combinations of metrics and CDN or ISP policies. As illustrated in Figure 26, ALTO could provide this information.

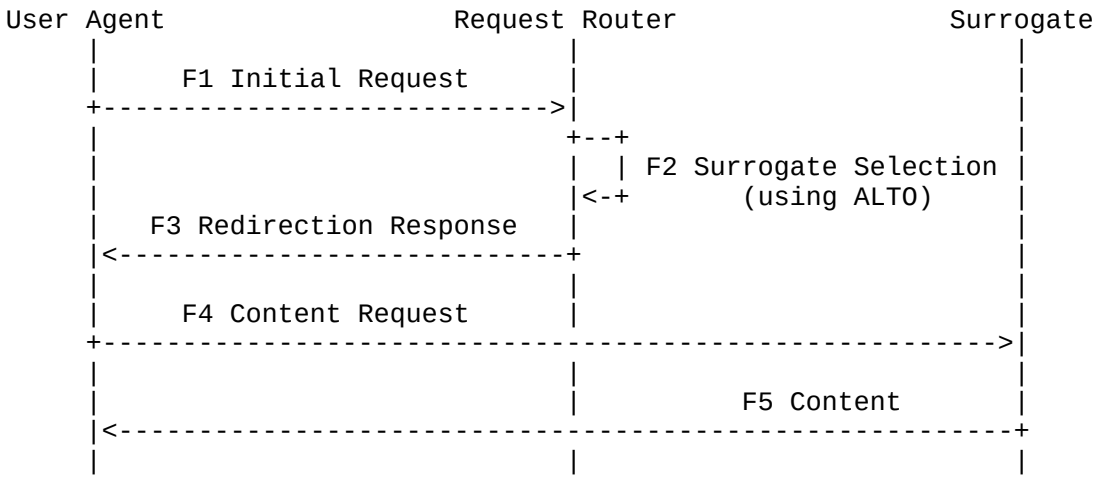

Figure 27: Example of CDN Surrogate Selection

Figure 27 illustrates the interaction between a user agent, a request router, and a surrogate for the delivery of content in a single CDN. As explained in [CDN-USE], the user agent makes an initial request to the CDN (F1). This may be an application-level request (e.g., HTTP) or a DNS request. In the second step (F2), the request router selects an appropriate surrogate (or set of surrogates) based on the user agent's (or its proxy's) IP address, the request router's knowledge of the network topology (which can be obtained by ALTO) and reachability cost between CDN caches and end users, and any additional CDN policies. Then, the request router responds to the initial request with an appropriate response containing a redirection to the selected cache (F3), for example, by returning an appropriate DNS A/AAAA record or an HTTP 302 redirect, etc. The user agent uses this information to connect directly to the surrogate and request the desired content (F4), which is then delivered (F5).

5.1.2. Applicability of ALTO

The most simple use case for ALTO in a CDN context is to improve the selection of a CDN surrogate or origin. In this case, the CDN makes use of an ALTO server to choose a better CDN surrogate or origin than would otherwise be the case. Although it is possible to obtain raw network map and cost information in other ways, for example, passively listening to the ISP's routing protocols or use of active probing, the use of an ALTO service to expose that information may provide additional control to the ISP over how their network map/cost is exposed. Additionally, it may enable the ISP to maintain a

functional separation between their routing plane and network map computation functions. This may be attractive for a number of reasons, for example:

- The ALTO service could provide a filtered view of the network and/or cost map that relates to CDN locations and their proximity to end users, for example, to allow the ISP to control the level of topology detail they are willing to share with the CDN.

- The ALTO service could apply additional policies to the network map and cost information to provide a CDN-specific view of the network map/cost, for example, to allow the ISP to encourage the CDN to use network links that would not ordinarily be preferred by a Shortest Path First routing calculation.

- The routing plane may be operated and controlled by a different operational entity (even within a single ISP) than the CDN. Therefore, the CDN may not be able to passively listen to routing protocols, nor may it have access to other network topology data (e.g., inventory databases).

When CDN servers are deployed outside of an ISP's network or in a small number of central locations within an ISP's network, a simplified view of the ISP's topology or an approximation of proximity is typically sufficient to enable the CDN to serve end users from the optimal server/location. As CDN servers are deployed deeper within ISP networks, it becomes necessary for the CDN to have more detailed knowledge of the underlying network topology and costs between network locations in order to enable the CDN to serve end users from the optimal servers for the ISP.

The request router in a CDN will typically also take into account criteria and constraints that are not related to network topology, such as the current load of CDN surrogates, content owner policies, end user subscriptions, etc. This document only discusses use of ALTO for network information.

A general issue for CDNs is that the CDN logic has to match the client's IP address with the closest CDN surrogate, for approaches that are both DNS or HTTP redirect based (see, for instance, [ALTO-CDN]). This matching is not trivial, for example, in DNS-based approaches, where the IP address of the DNS original requester is unknown (see [RFC7871] for a discussion of this and a solution approach).

In addition to use by a single CDN, ALTO can also be used in scenarios that interconnect several CDNs. This use case is detailed in [CDNI-FCI].

5.2. Deployment Recommendations

5.2.1. ALTO Services

In its simplest form, an ALTO server would provide an ISP with the capability to offer a service to a CDN that provides network map and cost information. The CDN can use that data to enhance its surrogate and/or origin selection. If an ISP offers an ALTO Network and Cost Map Service to expose a cost mapping/ranking between end user IP subnets (within that ISP's network) and CDN surrogate IP subnets/ locations, periodic updates of the maps may be needed. As introduced in Section 3.3), it is common for broadband subscribers to obtain their IP addresses dynamically, and in many deployments, the IP subnets allocated to a particular network region can change relatively frequently, even if the network topology itself is reasonably static.

An alternative would be to use the ALTO ECS: when an end user requests a given content, the CDN request router issues an ECS request with the endpoint address (IPv4/IPv6) of the end user (content requester) and the set of endpoint addresses of the surrogate (content targets). The ALTO server receives the request and ranks the addresses based on their distance from the content requester. Once the request router obtained from the ALTO server the ranked list of locations (for the specific user), it can incorporate this information into its selection mechanisms in order to point the user to the most appropriate surrogate.

Since CDNs operate in a controlled environment, the ALTO Network and Cost Map Service and ECS have a similar level of security and confidentiality of network-internal information. However, the Network and Cost Map Service and ECS differ in the way the ALTO service is delivered and address a different set of requirements in terms of topology information and network operations.

If a CDN already has means to model connectivity policies, the map-based approaches could possibly be integrated into that. If the ECS service is preferred, a request router that uses ECS could cache the results of ECS queries for later usage in order to address the scalability limitations of ECS and to reduce the number of transactions between the CDN and ALTO server. The ALTO server may indicate in the reply message how long the content of the message is to be considered reliable and insert a lifetime value that will be used by the CDN in order to cache (and then flush or refresh) the entry.

5.2.2. Guidance Considerations

The following discusses how a CDN could make use of ALTO services.

In one deployment scenario, ALTO could expose ISP end-user reachability to a CDN. The request router needs to have information about which end-user IP subnets are reachable via which networks or network locations. The network map services offered by ALTO could be used to expose this topology information while avoiding routing-plane peering between the ISP and the CDN. For example, if CDN surrogates are deployed within the access or aggregation network, the ISP is likely to want to utilize the surrogates deployed in the same access/aggregation region in preference to surrogates deployed elsewhere, in order to alleviate the cost and/or improve the user experience.

In addition, CDN surrogates could also use ALTO guidance, e.g., if there is more than one upstream source of content or several origins. In this case, ALTO could help a surrogate with the decision about which upstream source to use. This specific variant of using ALTO is not further detailed in this document.

If content can be provided by several CDNs, there may be a need to interconnect these CDNs. In this case, ALTO can be used as an interface [CDNI-FCI], in particular, for footprint and capabilities advertisement.

Other, and more advanced, scenarios of deploying ALTO are also listed in [CDN-USE] and [ALTO-CDN].

The granularity of ALTO information required depends on the specific deployment of the CDN. For example, an "over-the-top" CDN whose surrogates are deployed only within the Internet backbone may only require knowledge of which end-user IP subnets are reachable via which ISP’s networks, whereas a CDN deployed within a particular ISP’s network requires a finer granularity of knowledge.

An ALTO server ranks addresses based on topology information it acquires from the network. By default, according to [RFC7285], distance in ALTO represents an abstract "routingcost" that can be computed, for instance, from routing protocol information. But an ALTO server may also take into consideration other criteria or other information sources for policy, state, and performance information (e.g., geolocation), as explained in Section 3.2.2.

The different methods and algorithms through which the ALTO server computes topology information and rankings is out of the scope of this document. If rankings are based on routing protocol information, it is obvious that network events may impact the ranking

computation. Due to internal redundancy and resilience mechanisms inside current networks, most of the network events happening in the infrastructure will be handled internally in the network, and they should have limited impact on a CDN. However, catastrophic events such as main trunks failures or backbone partitioning will have to be taken into account by the ALTO server to redirect traffic away from the impacted area.

An ALTO server implementation may want to keep state about ALTO clients in order to inform and signal to these clients when a major network event happened, e.g., by a notification mechanism. In a CDN/ALTO interworking architecture with few CDN components interacting with the ALTO server, there are less scalability issues in maintaining state about clients in the ALTO server, compared to ALTO guidance to any Internet user.

6. Other Use Cases

This section briefly surveys and references other use cases that have been tested or suggested for ALTO deployments.

6.1. Application Guidance in Virtual Private Networks (VPNs)

Virtual Private Network (VPN) technology is widely used in public and private networks to create groups of users that are separated from other users of the network and allows these users to communicate among themselves as if they are on a private network. Network Service Providers (NSPs) offer different types of VPNs. [RFC4026] distinguishes between Layer 2 VPN (L2VPN) and Layer 3 VPN (L3VPN) using different sub-types. In the following, the term "VPN" is used to refer to provider supplied virtual private networking.

From the perspective of an application at an endpoint, a VPN may not be very different from any other IP connectivity solution, but there are a number of specific applications that could benefit from ALTO topology exposure and guidance in VPNs. As, in the general Internet, one advantage is that applications do not have to perform excessive measurements on their own. For instance, potential use cases for ALTO application guidance in VPN environments are:

- Enterprise application optimization: Enterprise customers often run distributed applications that exchange large amounts of data, e.g., for synchronization of replicated data bases. Network topology information could be useful for placement of replicas as well as for the scheduling of transfers.

- Private cloud computing solution: An enterprise customer could run its own data centers at several sites. The cloud management

system could want to understand the network costs between different sites for intelligent routing and placement decisions of Virtual Machines (VMs) among the VPN sites.

- Cloud-bursting: One or more VPN endpoints could be located in a public cloud. If an enterprise customer needs additional resources, they could be provided by a public cloud, which is accessed through the VPN. Network topology awareness would help to decide in which data center of the public cloud those resources should be allocated.

These examples focus on enterprises, which are typical users of VPNs. VPN customers typically have no insight into the network topology that transports the VPN. Similar to other ALTO use cases, better-than-random application-level decisions would be enabled by an ALTO server offered by the NSP, as illustrated in Figure 28.

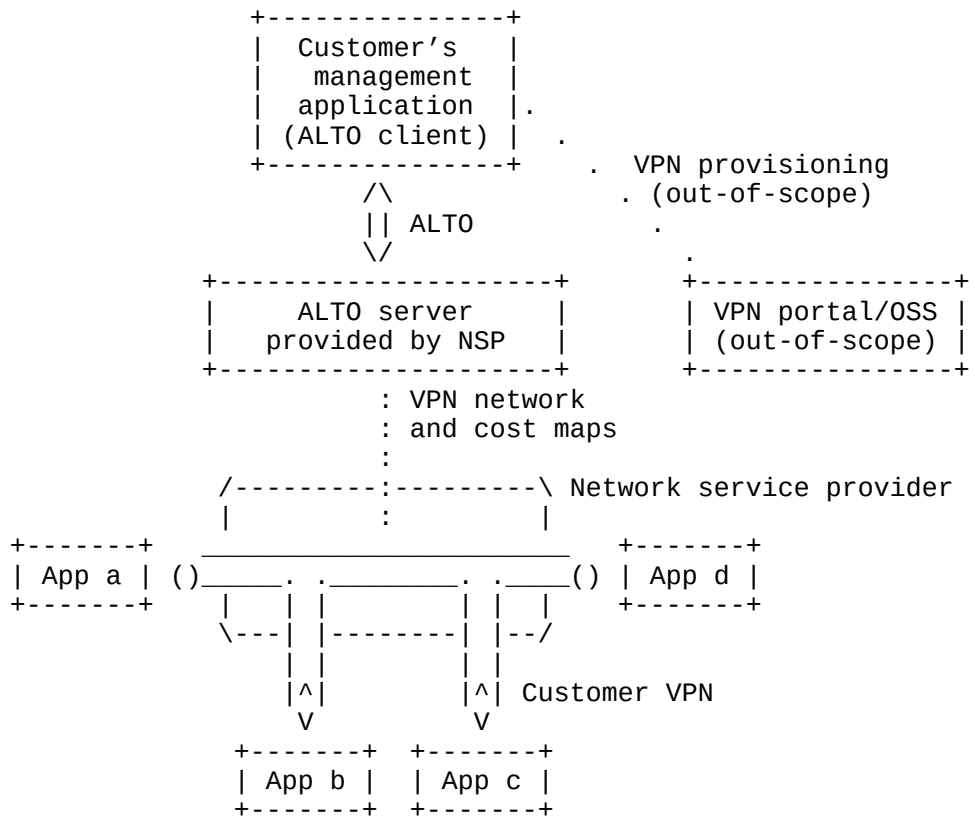

Figure 28: Using ALTO in VPNs

A common characteristic of these use cases is that applications will not necessarily run in the public Internet, and that the relationship between the provider and customer of the VPN is rather well defined. Since VPNs often run in a managed environment, an ALTO server may have access to topology information (e.g., traffic engineering data) that would not be available for the public Internet, and it may expose it to the customer of the VPN only.

Also, a VPN will not necessarily be static. The customer could possibly modify the VPN and add new VPN sites by a Web portal, network management systems, or other OSS solutions. Prior to adding a new VPN site, an application will not have connectivity to that site, i.e., an ALTO server could offer access to information that an application cannot measure on its own (e.g., expected delay to a new VPN site).

The VPN use cases, requirements, and solutions are further detailed in [VPN-SERVICE].

6.2. In-Network Caching

Deployment of intra-domain P2P caches has been proposed for cooperation between the network operator and the P2P service providers, e.g., to reduce the bandwidth consumption in access networks [ALTO-P2PCACHE].

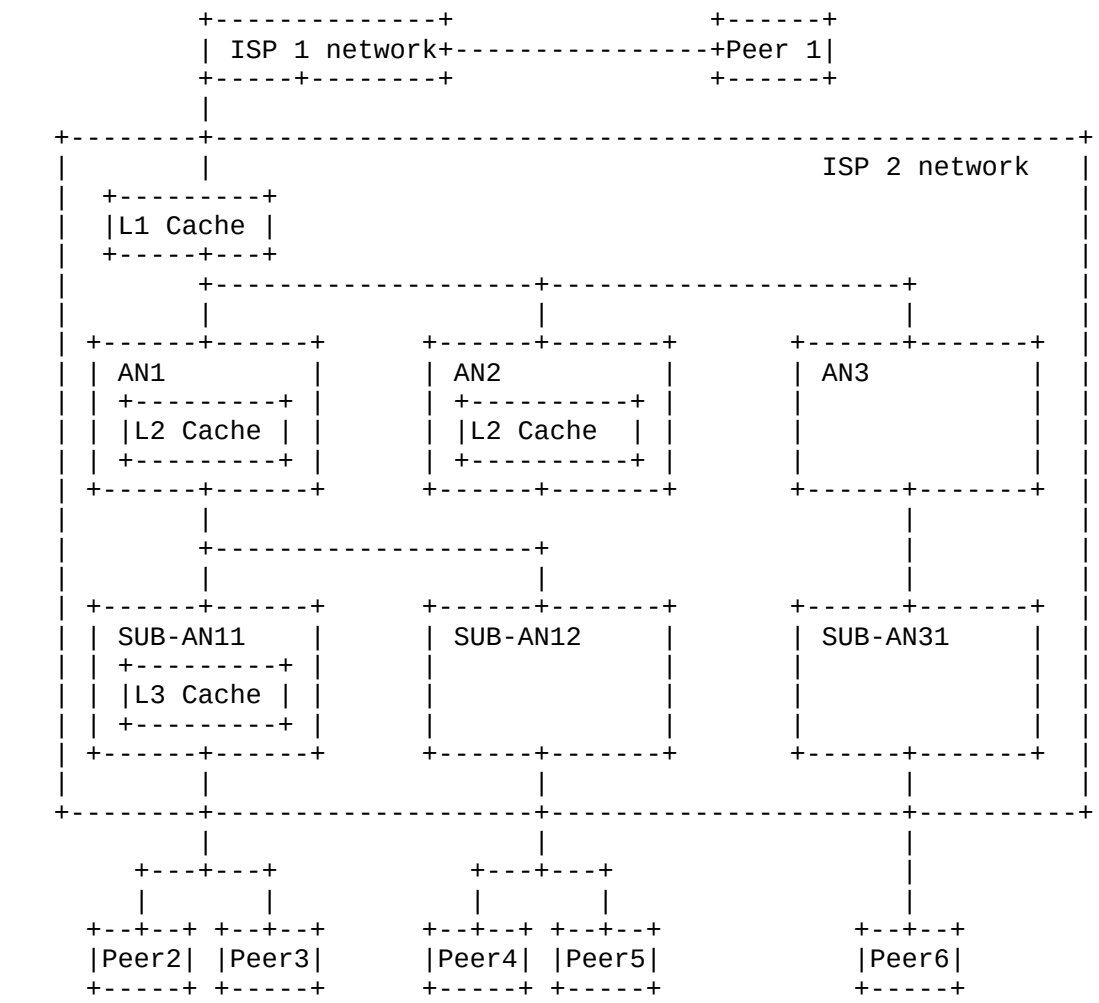

Figure 29: General Architecture of Intra-ISP Caches

Figure 29 depicts the overall architecture of potential P2P cache deployments inside an ISP 2 with various access network types. As shown in the figure, P2P caches may be deployed at various levels, including the interworking gateway linking with other ISPs, internal access network gateways linking with different types of accessing networks (e.g., WLAN, cellular, and wired), and even within an accessing network at the entries of individual WLAN subnetworks. Moreover, depending on the network context and the operator's policy, each cache can be a Forwarding Cache or a Bidirectional Cache [ALTO-P2PCACHE].

In such a cache architecture, the locations of caches could be used as dividers of different PIDs to guide intra-ISP network abstraction and mark costs among them according to the location and type of relevant caches.

Further details and deployment considerations can be found in [ALTO-P2PCACHE].

6.3. Other Application-Based Network Operations

An ALTO server can be part of an overall framework for Application-Based Network Operations (ABNO) [RFC7491] that brings together different technologies. Such an architecture may include additional components such as a PCE for on-demand and application-specific reservation of network connectivity, reliability, and resources (such as bandwidth). Some use cases how to leverage ALTO for joint network and application-layer optimization are explained in [RFC7491].

7. Security Considerations

Security concerns were extensively discussed from the very beginning of the development of the ALTO protocol, and they have been considered in detail in the ALTO requirements document [RFC6708] as well as in the ALTO protocol specification document [RFC7285]. The two main security concerns are related to the unwanted disclosure of information through ALTO and the negative impact of specially crafted, wrong ("faked") guidance presented to an ALTO client. In addition to this, the usual concerns related to the operation of any networked application apply.

This section focuses on the peer-to-peer use case, which is -- from a security perspective -- probably the most difficult ALTO use case that has been considered. Special attention is given to the two main security concerns.

7.1. ALTO as a Protocol Crossing Trust Boundaries

The optimization of peer-to-peer applications was the first use case and the impetus for the development of the ALTO protocol, in particular, file sharing applications such as BitTorrent [RFC5594].

As explained in Section 4.1.1, for the publisher of the ALTO information (i.e., the ALTO server operator), it may not be apparent who is in charge of the P2P application overlay. Some P2P applications do not have any central control entity and the whole overlay consists only of the peers, which are under control of the individual users. Other P2P applications may have some control entities such as super peers or trackers, but these may be located in

foreign countries and under the control of unknown organizations. As outlined in Section 4.2.2, in some scenarios, it may be very beneficial to forward ALTO information to such trackers, super peers, etc., located in remote networks. This situation is aggravated by the vast number of different P2P applications that are evolving quickly and often without any coordination with the network operators.

In summary, it can be said that in many instances of the P2P use case, the ALTO protocol bridges the border between the "managed" IP network infrastructure under strict administrative control and one or more "unmanaged" application overlays, i.e., overlays for which it is hard to tell who is in charge of them. This differs from more-controlled environments (e.g., in the CDN use case), in which bilateral agreements between the producer and consumer of guidance are possible.

7.2. Information Leakage from the ALTO Server

An ALTO server will be provisioned with information about the ISP's network and possibly also with information about neighboring ISPs. This information (e.g., network topology, business relations, etc.) is often considered to be confidential to the ISP and can include very sensitive information. ALTO does not require any particular level of details of information disclosure; hence, the provider should evaluate how much information is revealed and the associated risks.

Furthermore, if the ALTO information is very fine grained, it may also be considered sensitive with respect to user privacy. For example, consider a hypothetical endpoint property "provisioned access link bandwidth" or "access technology (ADSL, VDSL, FTTH, etc.)" and an ALTO service that publishes this property for individual IP addresses. This information could not only be used for traffic optimization but, for example, also for targeted advertising to residential users with exceptionally good (or bad) connectivity, such as special banner ads. For an advertisement system, it would be more complex to obtain such information otherwise, e.g., by bandwidth probing.

Different scenarios related to the unwanted disclosure of an ALTO server's information have been itemized and categorized in RFC 6708, Section 5.2.1., cases (1)-(3) [RFC6708].

In some use cases, it is not possible to use access control (see Section 7.3) to limit the distribution of ALTO knowledge to a small set of trusted clients. In these scenarios, it seems tempting not to use network maps and cost maps at all, and instead completely rely on

ECS and endpoint ranking in the ALTO server. While this practice may indeed reduce the amount of information that is disclosed to an individual ALTO client, some issues should be considered: first, when using the map-based approach, it is trivial to analyze the maximum amount of information that could be disclosed to a client -- the full maps. In contrast, when providing endpoint-cost service only, the ALTO server operator could be prone to a false feeling of security, while clients use repeated queries and/or collaboration to gather more information than they are expected to get (see Section 5.2.1., case (3) in [RFC6708]). Second, the ECS reveals more information about the user or application behavior to the ALTO server, e.g., which other hosts are considered as peers for the exchange of a significant amount of data (see Section 5.2.1., cases (4)-(6) in [RFC6708]).

Consequently, users may be more reluctant to use the ALTO service at all if it is based on the ECS instead of providing network and cost maps. Given that some popular P2P applications are sometimes used for purposes such as distribution of files without the explicit permission from the copyright owner, it may also be in the interest of the ALTO server operator that an ALTO server cannot infer the behavior of the application to be optimized. One possible conclusion could be to publish network and cost maps through ALTO that are so coarse grained that they do not violate the network operator's or the user's interests.

In other use cases, in more-controlled environments (e.g., in the CDN use case) bilateral agreements, access control (see Section 7.3), and encryption could be used to reduce the risk of information leakage.

7.3. ALTO Server Access

Depending on the use case of ALTO, it may be desired to apply access restrictions to an ALTO server, i.e., by requiring client authentication. According to [RFC7285], ALTO requires that HTTP Digest Authentication be supported, in order to achieve client authentication and possibly to limit the number of parties with whom ALTO information is directly shared. TLS Client Authentication may also be supported.

In general, well-known security management techniques and best current practices [RFC4778] for operational ISP infrastructure also apply to an ALTO service, including functions to protect the system from unauthorized access, key management, reporting security-relevant events, and authorizing user access and privileges.

For peer-to-peer applications, a potential deployment scenario is that an ALTO server is solely accessible by peers from the ISP network (as shown in Figure 21). For instance, the source IP address can be used to grant only access from that ISP network to the server. This will "limit" the number of peers able to attack the server to the user's of the ISP (however, including compromised computers that are part of a botnet).

If the ALTO server has to be accessible by parties not located in the ISP's network (see Figure 22), e.g., by a third-party tracker or by a CDN system outside the ISP's network, the access restrictions have to be looser. In the extreme case, i.e., no access restrictions, each and every host in the Internet can access the ALTO server. This might not be the intention of the ISP, as the server is not only subject to more possible attacks, but also the server load could increase, since possibly more ALTO clients have to be served.

There are also use cases where the access to the ALTO server has to be much more strictly controlled, i.e., where an authentication and authorization of the ALTO client to the server may be needed. For instance, in case of CDN optimization, the provider of an ALTO service as well as potential users are possibly well-known. Only CDN entities may need ALTO access; access to the ALTO servers by residential users may neither be necessary nor be desired.

Access control can also help to prevent Denial-of-Service (DoS) attacks by arbitrary hosts from the Internet. DoS can both affect an ALTO server and an ALTO client. A server can get overloaded if too many requests hit the server, or if the query load of the server surpasses the maximum computing capacity. An ALTO client can get overloaded if the responses from the sever are, either intentionally or due to an implementation mistake, too large to be handled by that particular client.

7.4. Faking ALTO Guidance

The ALTO services enables an ALTO service provider to influence the behavior of network applications. An attacker who is able to generate false replies, or e.g. an attacker who can intercept the ALTO server discovery procedure, can provide faked ALTO guidance.

Here is a list of examples of how the ALTO guidance could be faked and what possible consequences may arise:

Sorting: An attacker could change the sorting order of the ALTO guidance (given that the order is of importance; otherwise, the ranking mechanism is of interest), i.e., declaring peers located outside the ISP as peers to be preferred. This will not pose a

big risk to the network or peers, as it would mimic the "regular" peer operation without traffic localization, apart from the communication/processing overhead for ALTO. However, it could mean that ALTO is reaching the opposite goal of shuffling more data across ISP boundaries, incurring more costs for the ISP. In another example, fake guidance could give unrealistically low costs to devices in an ISP's mobile network, thus encouraging other devices to contact them, thereby degrading the ISP's mobile network and causing customer dissatisfaction.

Preference of a single peer: A single IP address (thus a peer) could be marked as to be preferred over all other peers. This peer can be located within the local ISP or also in other parts of the Internet (e.g., a web server). This could lead to the case that quite a number of peers to trying to contact this IP address, possibly causing a DoS attack.

The ALTO protocol protects the authenticity and integrity of ALTO information while in transit by leveraging the authenticity and integrity protection mechanisms in TLS (see Section 8.3.5 of [RFC7285]). It has not yet been investigated how wrong ALTO guidance given by an authenticated ALTO server can impact the operation of the network and the applications.

8. References

8.1. Normative References

[ALTO-REG] IANA, "Application-Layer Traffic Optimization (ALTO) Protocol", <http://www.iana.org/assignments/alto-protocol>.

[RFC5693] Seedorf, J. and E. Burger, "Application-Layer Traffic Optimization (ALTO) Problem Statement", RFC 5693, DOI 10.17487/RFC5693, October 2009, <http://www.rfc-editor.org/info/rfc5693>.

[RFC6708] Kiesel, S., Ed., Previdi, S., Stiemerling, M., Woundy, R., and Y. Yang, "Application-Layer Traffic Optimization (ALTO) Requirements", RFC 6708, DOI 10.17487/RFC6708, September 2012, <http://www.rfc-editor.org/info/rfc6708>.

[RFC7285] Alimi, R., Ed., Penno, R., Ed., Yang, Y., Ed., Kiesel, S., Previdi, S., Roome, W., Shalunov, S., and R. Woundy, "Application-Layer Traffic Optimization (ALTO) Protocol", RFC 7285, DOI 10.17487/RFC7285, September 2014, <http://www.rfc-editor.org/info/rfc7285>.

[RFC7286] Kiesel, S., Stiemerling, M., Schwan, N., Scharf, M., and H. Song, "Application-Layer Traffic Optimization (ALTO) Server Discovery", RFC 7286, DOI 10.17487/RFC7286, November 2014, <http://www.rfc-editor.org/info/rfc7286>.

8.2. Informative References

[ALTO-CDN]
Penno, R., Medved, J., Alimi, R., Yang, R., and S. Previdi, "ALTO and Content Delivery Networks", Work in Progress, draft-penno-alto-cdn-03, March 2011.

[ALTO-H12]
Kiesel, S. and M. Stiemerling, "ALTO H12", Work in Progress, draft-kiesel-alto-h12-02, March 2010.

[ALTO-P2PCACHE]
Lingli, D., Chen, W., Yi, Q., and Y. Zhang, "Considerations for ALTO with network-deployed P2P caches", Work in Progress, draft-deng-alto-p2pcache-03, February 2014.

[CDN-USE] Niven-Jenkins, B., Watson, G., Bitar, N., Medved, J., and S. Previdi, "Use Cases for ALTO within CDNs", Work in Progress, draft-jenkins-alto-cdn-use-cases-03, June 2012.

[CDNI-FCI]
Seedorf, J., Yang, Y., and J. Peterson, "CDNI Footprint and Capabilities Advertisement using ALTO", Work in Progress, draft-seedorf-cdni-request-routing-alto-08, March 2015.

[CHINA-TRIAL]
Li, K. and G. Jian, "ALTO and DECADE service trial within China Telecom", Work in Progress, draft-lee-alto-chinatelecom-trial-04, March 2012.

[MAP-CALC]
Seidel, H., "ALTO map calculation from live network data", Work in Progress, draft-seidel-alto-map-calculation-00, October 2015.

[NETWORK-TOPO]
Clemm, A., Medved, J., Varga, R., Tkacik, T., Bahadur, N., Ananthakrishnan, H., and X. Liu, "A Data Model for Network Topologies", Work in Progress, draft-ietf-i2rs-yang-network-topo-06, September 2016.

[RFC3411] Harrington, D., Presuhn, R., and B. Wijnen, "An Architecture for Describing Simple Network Management Protocol (SNMP) Management Frameworks", STD 62, RFC 3411, DOI 10.17487/RFC3411, December 2002, <http://www.rfc-editor.org/info/rfc3411>.

[RFC3568] Barbir, A., Cain, B., Nair, R., and O. Spatscheck, "Known Content Network (CN) Request-Routing Mechanisms", RFC 3568, DOI 10.17487/RFC3568, July 2003, <http://www.rfc-editor.org/info/rfc3568>.

[RFC4026] Andersson, L. and T. Madsen, "Provider Provisioned Virtual Private Network (VPN) Terminology", RFC 4026, DOI 10.17487/RFC4026, March 2005, <http://www.rfc-editor.org/info/rfc4026>.

[RFC4655] Farrel, A., Vasseur, J., and J. Ash, "A Path Computation Element (PCE)-Based Architecture", RFC 4655, DOI 10.17487/RFC4655, August 2006, <http://www.rfc-editor.org/info/rfc4655>.

[RFC4778] Kaeo, M., "Operational Security Current Practices in Internet Service Provider Environments", RFC 4778, DOI 10.17487/RFC4778, January 2007, <http://www.rfc-editor.org/info/rfc4778>.

[RFC5594] Peterson, J. and A. Cooper, "Report from the IETF Workshop on Peer-to-Peer (P2P) Infrastructure, May 28, 2008", RFC 5594, DOI 10.17487/RFC5594, July 2009, <http://www.rfc-editor.org/info/rfc5594>.

[RFC5632] Griffiths, C., Livingood, J., Popkin, L., Woundy, R., and Y. Yang, "Comcast’s ISP Experiences in a Proactive Network Provider Participation for P2P (P4P) Technical Trial", RFC 5632, DOI 10.17487/RFC5632, September 2009, <http://www.rfc-editor.org/info/rfc5632>.

[RFC6020] Bjorklund, M., Ed., "YANG - A Data Modeling Language for the Network Configuration Protocol (NETCONF)", RFC 6020, DOI 10.17487/RFC6020, October 2010, <http://www.rfc-editor.org/info/rfc6020>.

[RFC6241] Enns, R., Ed., Bjorklund, M., Ed., Schoenwaelder, J., Ed., and A. Bierman, Ed., "Network Configuration Protocol (NETCONF)", RFC 6241, DOI 10.17487/RFC6241, June 2011, <http://www.rfc-editor.org/info/rfc6241>.

[RFC6875] Kamei, S., Momose, T., Inoue, T., and T. Nishitani, "The P2P Network Experiment Council's Activities and Experiments with Application-Layer Traffic Optimization (ALTO) in Japan", RFC 6875, DOI 10.17487/RFC6875, February 2013, <http://www.rfc-editor.org/info/rfc6875>.

[RFC7491] King, D. and A. Farrel, "A PCE-Based Architecture for Application-Based Network Operations", RFC 7491, DOI 10.17487/RFC7491, March 2015, <http://www.rfc-editor.org/info/rfc7491>.

[RFC7752] Gredler, H., Ed., Medved, J., Previdi, S., Farrel, A., and S. Ray, "North-Bound Distribution of Link-State and Traffic Engineering (TE) Information Using BGP", RFC 7752, DOI 10.17487/RFC7752, March 2016, <http://www.rfc-editor.org/info/rfc7752>.

[RFC7871] Contavalli, C., van der Gaast, W., Lawrence, D., and W. Kumari, "Client Subnet in DNS Queries", RFC 7871, DOI 10.17487/RFC7871, May 2016, <http://www.rfc-editor.org/info/rfc7871>.

[RFC7921] Atlas, A., Halpern, J., Hares, S., Ward, D., and T. Nadeau, "An Architecture for the Interface to the Routing System", RFC 7921, DOI 10.17487/RFC7921, June 2016, <http://www.rfc-editor.org/info/rfc7921>.

[RFC7922] Clarke, J., Salgueiro, G., and C. Pignataro, "Interface to the Routing System (I2RS) Traceability: Framework and Information Model", RFC 7922, DOI 10.17487/RFC7922, June 2016, <http://www.rfc-editor.org/info/rfc7922>.

[UPDATE-SSE] Roome, W. and Y. Yang, "ALTO Incremental Updates Using Server-Sent Events (SSE)", Work in Progress, draft-ietf-alto-incr-update-sse-03, September 2016.

[VPN-SERVICE] Scharf, M., Gurbani, V., Soprovich, G., and V. Hilt, "The Virtual Private Network (VPN) Service in ALTO: Use Cases, Requirements and Extensions", Work in Progress, draft-scharf-alto-vpn-service-02, February 2014.

[XDOM-DISC] Kiesel, S. and M. Stiemerling, "Application Layer Traffic Optimization (ALTO) Cross-Domain Server Discovery", Work in Progress, draft-kiesel-alto-xdom-disc-02, July 2016.

Acknowledgments

This memo is the result of contributions made by several people:

- Xianghue Sun, Lee Kai, and Richard Yang contributed text on ISP deployment requirements and monitoring.
- Rich Woundy contributed text to Section 3.3.
- Lingli Deng, Wei Chen, Qiuchao Yi, and Yan Zhang contributed Section 6.2.

Thomas-Rolf Banniza, Vinayak Hegde, Qin Wu, Wendy Roome, and Sabine Randriamasy provided very useful comments and reviewed the document.

Authors' Addresses

Martin Stiemerling
Hochschule Darmstadt

Email: mls.ietf@gmail.com
URI: http://ietf.stiemerling.org

Sebastian Kiesel
University of Stuttgart Information Center
Networks and Communication Systems Department
Allmandring 30
Stuttgart 70550
Germany

Email: ietf-alto@skiesel.de

Michael Scharf
Nokia
Lorenzstrasse 10
Stuttgart 70435
Germany

Email: michael.scharf@nokia.com

Hans Seidel
BENOCS GmbH
Winterfeldtstrasse 21
Berlin 10781
Germany

Email: hseidel@benocs.com

Stefano Previdi
Cisco Systems, Inc.
Via Del Serafico 200
Rome 00191
Italy

Email: sprevidi@cisco.com