Network Working Group
Internet-Draft
Intended status: Experimental
Expires: September 15, 2011

R. Penno
J. Medved
Juniper Networks
R. Alimi
Google
R. Yang
Yale University
S. Previdi
Cisco Systems
March 14, 2011

ALTO and Content Delivery Networks
draft-penno-alto-cdn-03

Abstract

Networking applications can request through the ALTO protocol information about the underlying network topology from the ISP or Content Provider (henceforth referred as Provider) point of view. In other words, information about what a Provider prefers in terms of traffic optimization -- and a way to distribute it. The ALTO Service provides information such as preferences of network resources with the goal of modifying network resource consumption patterns while maintaining or improving application performance.

One of the main use cases of the ALTO Service is its integration with Content Delivery Networks (CDN). The purpose of this draft is twofold: first, to describe how ALTO can be used in existing and new CDNs, both within an ISP and in separate organizational entities from the ISP; second, to collect requirements for ALTO usage in CDNs and to provide recommendations into the development of the ALTO protocol for better support of CDNs.

Requirements Language

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in RFC 2119 [RFC2119].

Status of this Memo

Table of Contents

1. Introduction

Content Delivery Networks are becoming increasingly important in the Internet [ARBOR] and many CDNs today already use some form of proximity such as latency-based proximity [GoogleCDN]. But in many cases the content provider/distributor and the Internet Service Provider (ISP) are disjoint entities. Consequently, even if content servers are co-located into the ISP's networks, there is not a standardized way to share server location and/or network topology information. Therefore a natural step forward would be to use ALTO to share this information.

Another key aspect of ALTO in the context of CDNs deployments is that it is desirable that no changes to the hosts are needed (or that changes to hosts would be transparent to the user). In other words, a traditional web browser using standard HTTP flow is all there is needed to take advantage of ALTO information. This is a significant difference from the P2P applications where a special client is typically needed and ALTO is normally used as a way to reduce operational expense.

2. Scope

This document discusses how Content Delivery Networks can benefit from ALTO through integration of the ALTO Service with the main request routing techniques. There are two objectives:

o Present basic integration schemes of ALTO into CDNs.

o Provide protocol recommendations to ALTO: Whenever a new requirement on protocol functionality is identified to achieve integration with CDNs, it will be enumerated with 'REQ-<N>'. Each requirement is documented in a section of its own in order to foster parallel discussions and possible adoption.

3. Terminology

We use the following terms defined in ALTO Problem Statement [RFC5693]: Application, ALTO Service, ALTO Server, ALTO Client, ALTO Query, ALTO Reply, ALTO Transaction.

In addition to the above, the following terms are defined:

Content-aware Proximity Request Routing Function: The Request Routing function knows about locations and presence of content & media objects in the network. Therefore the redirection to a CDN

node is made based on both the availability of content and/or content-type in that CDN node and the proximity of the CDN node to the requesting user.

Service-aware Proximity Request Routing Function: The Request Routing function knows about locations of CDN nodes in the network and redirects user to the closest CDN node. A redirection is made irrespective of content presence in the CDN node; if content is not present, the node will be populated with the content while the content is being served to the user.

HTTP Request Routing Function: a Content-aware or Service-aware Proximity Request Routing function for HTTP. It embeds an HTTP Server that performs HTTP Redirects, an ALTO client that retrieves network mapping from the ALTO Server, and a Location Database which stores network mappings received from the ALTO Client. The HTTP Server consults the Location Database when making redirection decisions.

4. Request Routing as an Integration Point of ALTO into CDN

Content Distribution is a rich and evolving field. New architectures and approaches (e.g., a hybrid architecture using both servers and P2P) continue to be developed in the research community and industry. Several CDN architectures are being deployed in production. While we would like to provide a survey of each possible CDN architecture and show how it may be integrated with ALTO, it would be a daunting task to track such a rapidly-changing field.

One scheme that is out of the scope of this document is P2P-only CDNs, where the application tracker takes the role of the ALTO Client, fetching the Network and Cost Maps from the ALTO Server and integrating them with its peer database. The result is a peer database that takes into account both the current peer metrics, such as peer availability or content availability, and network metrics, such as topological localization. This architecture, in the context of file sharing, has been studied extensively and trialed by ISPs such as Comcast [RFC5632] and China Telecom [I-D.lee-alto-chinatelecom-trial] under the ALTO/P4P [P4P] protocol. Thus, P2P-only CDNs are not discussed in this document.

The Request Routing Component of a CDN directs a request to a serving CDN node, and thus is the major integration point to utilize information available through ALTO. Today, multiple request routing schemes have been used even in CDNs with purely server-based infrastructure. The specific schemes include HTTP Redirect, DNS name resolution, and anycast. We focus on HTTP Redirect and DNS name

resolution.

Though anycast is a request routing technique that has been used in deployed CDNs, we do not discuss it in this document. Even though one may be able to integrate ALTO with anycast, we do not believe that this is a proper use of ALTO's capabilities. In particular, ALTO has been developed to improve selection amongst multiple content providers at the application level. In contrast, anycast operates by adjusting the routing layer to match content consumers with the desired content providers. Applying ALTO to routing layer decisions introduces additional complexity because it directly adjusts the routing layer from which the ALTO information is typically generated, creating a tight feedback loop. We leave a more detailed study of integrating ALTO with anycast-based CDNs as future work.

We next briefly review the two mechanisms presented in this document, HTTP Redirect and DNS Request Routing.

4.1. HTTP Redirect

In this mechanism, an HTTP GET request from a host is received by an HTTP Request Routing Function which sends back an HTTP response with Status-Code 302 (Redirect) informing the host of the most preferred location to fetch the content. The HTTP Redirection method is already commonly used in production CDNs as described in RFC3568 [RFC3568]. ALTO integration provides localization services where the device that performs the redirection becomes an ALTO client.

4.2. DNS Request Routing

In this mechanism, the DNS server handling host requests provides the Request Routing Component. When the host performs a DNS query/ lookup, the IP address(es) in the DNS response will indicate the selected location to serve the request.

DNS queries can be either iterative or recursive. Iterative queries can be used with ALTO if the host itself queries the DNS Servers, or if the DNS Proxy used by the host is topologically close to the host. If the Host directly queries the DNS Servers, the authoritative DNS Server can see directly the host's IP address. If the DNS Proxy is topologically close to the Host, its IP address is a good approximation for the host's location. In recursive queries, the authoritative DNS Server sees the IP address of the previous DNS Server in the resolution chain, and the IP address of the host is unknown. DNS-based request routing does not work well with recursive DNS queries.

In an iterative DNS lookup with a DNS Proxy (say for cdn.com), the

host queries the Proxy, which in turn first queries one of the root servers to find the server authoritative for the top-level domain (com in our example). The Proxy then queries the obtained top-level-domain DNS server for the address of the DNS server authoritative for the CDN domain. Finally, the Proxy queries the DNS server that is authoritative for the cdn.com domain. The authoritative DNS Server for cdn.com will perform the request routing to the most appropriate CDN node, based on the source IP address of the requestor. The host will then request the content directly from the CDN Node.

Recently, an EDNS0 option in DNS query has been proposed in [I-D.vandergaast-edns-client-subnet] that will provide a mechanism to carry sufficient network information about the client for the authoritative DNS server to tailor DNS response based on the client's subnet. Using this mechanism, the authoritative DNS server can achieve the same request routing accuracy as that of the HTTP Request Routing Function, and both recursive and iterative queuries can be supported.

5. Basic Scheme of CDN/ALTO Integration

Although HTTP Redirect and DNS are quite different mechanisms to direct a request to a serving CDN node, as we will see, the basic structure of integrating ALTO with them can be quite similar. Thus, we first present common structures. We refer to the HTTP Redirect component or the DNS component of a CDN as a CDN Request Routing Function.

5.1. Basic Integration Scheme

Figure 1 shows a general structure to embed an ALTO Client into a CDN Request Routing Function.

```
                                 +-----------------+
                                 | Request Routing |
         +-----------+   1       |     Function    |
         |           |---------> |                 |
         | Requestor |<--------  |                 |
         +-----------+   2       |                 |
                                 | +-------------+ |
                                 | | ALTO Client | |
                                 | +-------------+ |
                                 +-----------------+
                                        |   ^
                                        |   |  ALTO Protocol
                                        v   |
                                 +-----------------+
                                 |   ALTO Server   |
                                 +-----------------+
```

Figure 1: Request Routing Function with ALTO

An ALTO Server may aggregate information from multiple sources, such as routing protocols, traffic engineering policies, and monitoring systems. Thus, ALTO is complementary to existing infrastructure. For further detail, see Figure 1 of [I-D.ietf-alto-protocol].

5.1.1. ALTO for HTTP Redirect

To make the basic scheme more concrete, Figure 2 shows the case that the Request Routing Function is HTTP Redirect.

```
                                 +---------------------+
                                 |     HTTP Request    |
                                 |   Routing Function  |
         +------+      1         |   +-------------+   |
         |      |--------------> |   | HTTP Server |   |
         | Host |<-------------- |   +-------------+   |
         +------+      2         |          ^          |
                                 |          |          |
                                 |   +-------------+   |
                                 |   | ALTO Client |   |
                                 |   +-------------+   |
                                 +---------------------+
                                          |   ^
                                          |   |  ALTO Protocol
                                          v   |
                                 +---------------------+
                                 |     ALTO Server     |
                                 +---------------------+
```

Figure 2: ALTO for HTTP Request Routing Function

5.1.2. ALTO for DNS Resolution

Figure 3 shows the case that the Request Routing Function uses DNS Resolution.

```
                     2      +----------------+
         +------------------> |      root      |
         | +----------------- |  Name Server   |      +----------+
         | |          3      +----------------+      | Content  |
         | |                                         | Provider |
         | |          4      +----------------+      +----------+
         | |   +------------> |      com       |
         | |   | +----------- |  Name Server   |
         | |   | |      5      +----------------+
         | |   | |
         | V   | V
        +----------+    6      +----------------+
        |   DNS    |---------> |    cdn.com     |
        |  Proxy   |<--------- |  Name Server   |
        +----------+    7      |                |
            ^ |                | +------------+ |
          1 | | 8              | |ALTO Client | |
            | V                | +------------+ |
        +----------+           +----------------+
        |  Host    |                  |   ^
        +----------+                  |   |  ALTO Protocol
             |                        |   |
             |                        V   |
             V                 +----------------+
          CDN Node             |   ALTO Server  |
                               +----------------+
```

Figure 3: ALTO for DNS Resolution.

5.2. Multi-hop Redirection

The preceding examples show the logical flow for redirection. It is important to state that there maybe multiple redirection hops.

For HTTP Redirect, the requestor may be redirected again by the first CDN node. For DNS, the first DNS server may direct, using aggregated ALTO information (e.g., from multiple ALTO Servers of multiple ISPs), the DNS resolution to a second level DNS server, which then may use more specific ALTO information as well as CDN node status.

6. Request Routing using ALTO Services

Either the Map Service or the Endpoint Cost Service of ALTO can be used by the Request Routing Function. We first discuss two common issues: how to configure ALTO topology at ALTO servers; and how to achieve CDN node discovery and status notification. Then we give specific details on using the Map Service or the Endpoint Cost Service.

6.1. ALTO Topology vs. Network Topology

To answer queries from CDN Request Routing Functions, the ALTO server builds a ALTO-specific network topology that represents the network as it should be understood and utilized by the application layer (the CDN). 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 visibility of the network, and at which level of granularity, is required by the CDN and more in general by the application layer.

The ALTO server builds topology (for either Map and ECS services) based on multiple sources that may include routing protocols, network policies, state and performance information, geo-location, etc. In all cases, the ALTO topology will not contain any details that would endanger the network integrity and security (for example, there will be no leaking of OSPF/ISIS/BGP databases to ALTO clients).

6.2. CDN Node Discovery and Status Notification

A design issue of integrating ALTO into Request Routing is how CDN Request Routing discovers the available CDN nodes and their locations. The exact mechanism is outside the scope of this document.

It is desirable that not only CDN node locations, but also real-time CDN node status (like health, load, cache utilization, CPU, etc.) is communicated to the Request Routing Function.

Specifically, CDN node status can be retrieved from the existing Load Balancer infrastructure. Most Load Balancers today have mechanisms to poll caches/servers via ping, HTTP Get, traceroute, etc. Most LBs have SNMP trap capabilities to let other devices know about these thresholds. Specification of a particular mechanism or API used to fetch load status information into an ALTO Server is out of scope of this document.

Note that in addition to the CDN node status, network status can also be retrieved from TE/RP databases. The Request Routing Function may

also need to be configured with a proper set of policies and business rules that control routing of requests. For example, it may be desirable to set up a rule that within a CDN certain requests have higher priority.

We see two approaches that CDN node status can be communicated to the Request Routing Function.

6.2.1. CDN Node Status Updates received by Request Routing Function

In the first approach, the Request Routing Function receives CDN Status updates directly.

For example, the Request Routing Function can implement an SNMP agent and get to know whatever is needed.

```
                              +-----------------+
                              | Request Routing |
                              |     Function    |
         +-----------+   1    |                 |
         |           |---------> |              |<--- Real-time CDN
         | Requestor |<--------  |              |     status updates
         +-----------+   2    |                 |
                              | +-------------+ |<--- Business rules
                              | | ALTO Client | |     and Policies
                              | +-------------+ |
                              +-----------------+
                                     |   ^
                                     |   |  ALTO Protocol
                                     v   |
                              +-----------------+
                              |   ALTO Server   |
                              +-----------------+
```

Figure 4: CDN Node Status to Request Routing Function

6.2.2. CDN Node Status Updates received by ALTO

In the second approach, the Request Routing Function receives CDN Status from ALTO instead of CDN nodes.

This model generally simplifies the Request Routing Function. It allows an easier distribution of the Request Routing Function, and to keep real time CDN status data updates in a logically centralized ALTO Server or in an ALTO Server Cluster. It allows for the Request Routing Function and the ALTO Server to be in different administrative domains. For example, the Request Routing Function can be in a Content Provider's domain; the ALTO Server and CDN Nodes

in a Network Service Provider's domain.

Specifically, ALTO Server could provide an API (for example, a Web Service or XMPP-based API) that could be used by CDN nodes to communicate their status to the ALTO server directly.

```
                              +-----------------+
                              | Request Routing |
                              |    Function     |
     +-----------+   1        |                 |
     |           |---------> |                 |
     | Requestor |<--------   |                 |
     +-----------+   2        |                 |
                              | +-------------+ |
                              | | ALTO Client | |
                              | +-------------+ |
                              +-----------------+
                                     |   ^
                                     |   |  ALTO Protocol
                                     v   |
                              +-----------------+
                              |                 |<--- Real-time CDN
                              |   ALTO Server   |     status updates
                              |                 |
                              |                 |<--- Business rules
                              +-----------------+     and policies
```

Figure 5: CDN Node Status to ALTO

6.3. Request Routing using the Map Service

The ALTO client embedded in the Request Routing Function fetches the Network and Cost Maps from the ALTO Server and provides that information to the Request Router.

As an illustrative example, we consider the case of HTTP Redirect. A simple Request Router may be given (from an external source) the list of available CDN nodes. The Request Router precomputes a redirection table indexed by source PID with values being the closest CDN nodes. This redirection table can be built based on information from Network and Cost Maps. Then when the Request Router receives an HTTP GET request, it looks up the PID of the source IP address on the request, indexes the redirection table using the request PID to select a CDN node, and finally returns a response that is an HTTP redirect with the URL of the selected CDN node. The URL in 302 Redirect may contain the IP address of the selected CDN node or a domain name instead of IP address due to virtual hosting. Therefore the IP addresses contained in the cost maps may need to be correlated to

domain names a priori. In practice, the redirection table may be indexed by both source and content to provide better redirection.

The illustrative example can also be extended to DNS.

The Network Maps generated by the ALTO Server will contain both Host PIDs and CDN Node PIDs, i.e., Host PIDs contain host subnets; CDN PIDs contain IP addresses of available CDN nodes. Cost Maps may contain only cost from each host PID to each CDN PID and not the full matrix across all PIDs. The reason is that the Request Router may redirect a host only to a CDN node, not to another host as in the P2P case. Moreover, there is no generic way to disambiguate PIDs containing only hosts from PIDs containing CDN nodes.

It is possible that a Request Router may be designated as being responsible only for a fixed set of Host PIDs. This information can be made available to the Request Router before it receives requests from hosts. If the set of Host PIDs is not known ahead of time, the latency for serving requests will be impacted by the capabilities of the ALTO server.

With such information ahead of time, a Request Router that uses the Network Maps Service may pre-download the Network Map for the interesting Host PIDs and the CDN PIDs. It can also start periodically pulling Cost Map for relevant PID 2-tuples.

The Request Router can rely on the ALTO Server generated Cache-Control headers to decide how often to fetch CDN PID network map and Host PID network maps.

For Alto protocol requirements related to request routing with the Map Service see Section 8.1.1 and Section 8.1.2.

6.4. Request Routing using the Endpoint Cost Service

Alternatively, the Request Router may request the Endpoint service from the ALTO client.

Specifically, the Request Router requests the Endpoint Cost Service to rank/rate the content locations (i.e., IP addresses of CDN nodes) based on their distance/cost (by default the Endpoint Cost Service operates based on Routing Distance) from/to the user address.

Once the Request Router obtains 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 location.

A Request Router that uses the Endpoint Cost Service may query the ALTO Server for rankings of CDN Node IP addresses for each requesting host and cache the results for later usage.

Maps Services and ECS deliver similar ALTO service by allowing the Request Routing Functino to optimize internal selection mechanisms. Both services deliver similar level of security, confidentiality of layer-specific information (i.e.: application and network) however, Maps 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.

6.4.1. Topology Computation and ECS Delivery

ECS allows the Request Routing Function to 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. Thus ECS is used in order to implement a lightweight integration of ALTO services in the CDN layer. ECS implies an ALTO and CDN implementation with the necessary scalability in order to cope with the amount of transactions that CDN and ALTO server will have to handle (knowing that the CDN is able to cache ALTO ECS results for further use).

6.4.2. Ranking Service

When a user requests a given content, the Request Routing Function locates the content in one or more caches and executes a selection algorithm to redirect the user to the 'best' cache. In order to achieve that, the CDN issues an ECS request with the endpoint address (IPv4/IPv6) of the user (content requester) and the set of endpoint addresses of the content caches (content targets). The ALTO server, receives the request and ranks the list of content targets addresses based on their distance from the content requester. By default, according to [I-D.ietf-alto-protocol], the distance represents the routing cost as computed by the routing layer (OSPF, ISIS, BGP) and may take into consideration other routing criteria such as MPLS-VPN (MP-BGP) and MPLS-TE (RSVP), policy and state & performance information.

Once the ALTO server has computed the distance it replies with the ranked list of content target addresses. The list being ranked by distance, the CDN is capable of integrating the rankings into its selection process (that will also incorporate other criteria) and redirect the user accordingly.

6.5. Update, Redirection of ALTO Info to CDN Request Routing

The information provided by an ALTO server to Request Routing is based on topology information of the network. The different methods and algorithms through which the ALTO server computes topology information and rankings is out of the scope of this document. However, update and rediction of such information may have an impact on the integration of ALTO into CDN Request Routing.

6.5.1. ALTO Update and Network Events

In the case that ALTO information is based on routing (IP/MPLS) topology, it is obvious that network events may impact the ALTO computation. The scope of the ALTO information delivered to Request Routing is not to maintain the CDN aware of any possible network topology changes since, due to redundancy of current networks, most of the network events happening in the infrastructure will have limited impact on the CDN. However, catastrophic events such as main trunks failures or backbone partition will have to take into account by the ALTO server so to redirect traffic away from the failure impacted area.

6.5.2. Caching and Lifetime

Each reply sent back by the ALTO server to the ALTO client running in the Request Routing Function has a validity in time so that the CDN can cache the results in order to re-use it and hence reducing the number of transactions between 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 Request Routing Function in order to cache (and then flush or refresh) the entry.

An ALTO server implementation may want to keep state about ALTO clients so to inform and signal to these clients when a major network event happened so to clear the ALTO cache in the client. In a CDN/ ALTO interworking architecture, where there are only a few CDN components interacting with the ALTO server, there are no scalability issues in maintaining state about clients in the ALTO server.

6.5.3. ALTO Redirection

When ALTO server receives a request from a CDN Request Routing Function, it may not have the most appropriate topology information to reply. In such case, the ALTO server, may want to adopt the following strategies:

- o Reply with available information (best effort).

- o Redirect the request to another ALTO server presumed to have better topology information (redirection).

- o Doing both (best effort and redirection). In this case, the reply message contains both the rankings and the indication of another ALTO server where more accurate information may be delivered.

The decision process that is used to determine if redirection is necessary (and which mode to use) is out of the scope of this document. As an example, an ALTO server may decide to redirect any request having addresses that are located into a remote Autonomous System. In such case the redirection message includes the ALTO server to be used and that resides in the remote AS. Redirection implies communication between ALTO servers so to be able to signal their identity, location and type of visibility (AS number).

6.5.4. Groups and Costs

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. IP/MPLS networks make use of a common mechanism to aggregate and group prefixes that is called BGP Communities. BGP is the protocol all ISP networks use in order to exchange information about their prefix reachability. BGP Community us an attribute used to tag a prefix so to group prefixes based on mostly any criteria (as an example, most SP networks originate BGP prefixes with communities identifying the Point of Presence (PoP) where the prefix has been originated).

The ALTO server may leverage the BGP information that is available in the ISP network layer and compute group of prefixes. By policy, the ALTO server operator may decide an arbitrary cost to set between groups. Alternatively, there are algorithms that allow dynamic computation of cost between groups.

7. Multiple Administrative Domains

The preceding discussion works well in a single administrative domain setting: the CDN nodes are in the administrative domain of the ISP. However, the CDN nodes, the ISP, and the Request Router can be in different administrative domains. In this section, we consider a few such deployment cases. We use DNS as an example.

7.1. **CDN nodes/Request Router in a separate administrative domain** from that of ISP

In many situations, the CDN nodes and the Request Router are in a separate network managed by an entity that is distinct from the ISP. Consequently, the CDN nodes belong to a network with its own ALTO server that is distinct from the ALTO server of the ISP where the subscribers belong to.

```
                                       .................................
                                       :      +-----------------+      :
                                       :      | Request Routing |      :
                                       :      |     Function    |      :
            +----------+               :      |                 |      :
            | Content  |               :      | +-------------+ |      :
            | Provider |               :      | | ALTO Client | |      :
            +----------+               :      | +-------------+ |      :
                                       :      +-----------------+      :
                                       :              ^                :
                                       :              |                :
                                       :      +-----------------+      :
   .................................   :      |   ALTO Server   |      :
   :                               :   :      |                 |      :
   :    +----------------+         :   :      | +-------------+ |      :
   :    |   ALTO Server  |--------------------->| ALTO Client | |      :
   :    +----------------+         :   :      | +-------------+ |      :
   :                               :   :      +-----------------+      :
   :                               :   :                               :
   : +------+ C(1-4)    +--------+ :   : +--------+    C(6-8) +------+ :
   : | Host |<--------->| Border |: c6  :| Border |<--------->| CDN  | :
   : | PID1 |       +-->| Router |-------| Router |<--+       | PID8 | :
   : +------+       |+->|  PID4  | :    :|  PID6  |<-+|       +------+ :
   :                ||  +--------+ :   : +--------+  ||                :
   :                ||             :   :             ||                :
   : +------+ C(2-4)||             :   :             ||C(6-9) +------+ :
   : | Host |<------+|             :   :             |+------>| CDN  | :
   : | PID2 |        |             :   :             |        | PID9 | :
   : +------+        |             :   :             |        +------+ :
   :                 |             :   :             |                 :
   :                 |             :   :             |                 :
   : +------+ C(3-4) |  +--------+ :   : +--------+  | C(6-10)+------+ :
   : | Host |<-------+  | Border |: c7  :| Border |  +------->|  CDN | :
   : | PID3 |           | Router |-------| Router |           | PID10| :
   : +------+           |  PID5  | :   : |  PID7  |           +------+ :
   :                    +--------+ :   : +--------+                    :
   :                               :   :                               :
   :  ISP Administrative Domain    :   :   CDN Administrative Domain   :
   :...............................:   :...............................:
```

Figure 6: Map advertising between ISP and CDN domains

The ALTO server in the CDN provider network is assumed to be initialized with information about the ISP networks it serves. For every such ISP network, it consults the routing plane to find the set of Border routers. The CDN network ALTO server computes the cost of reaching each Border router from every CDN node (say, C_cdn).

Next, the CDN ALTO server contacts the ISP network's ALTO server and downloads the network map. In order to help the CDN ALTO server compute the cost from a CDN node to a subscriber's PID, we break it down into two parts - the cost from the CDN node to the Border Router (C_cdn) and the cost from the Border Router to the subscriber's PID (say, C_isp). Note that for any chosen exit point, C_cdn may be computed locally by the CDN ALTO Server. However, the fundamental issue is that C_isp depends on the exit point (Border outer) chosen by the CDN. There are multiple ways for the CDN ALTO Server to compute C_isp given the Network Map and Cost Map from the ISP's ALTO Server.

One possibility is for the ISP ALTO Server to define a special Border Router PID (denoted by a PID attribute) which also indicates the corresponding Border Router PID in the CDN. The attributes and values may be agreed-upon by the ISP and CDN when the ALTO Services are configured. For example, in the example shown in Figure 5, the ISP ALTO Server indicates that its PID4 and PID5 are Border PIDs, with corresponding PIDs in the CDN as PID6, and PID7, respectively. Then, CDN ALTO Server can locally compute C_isp = cost(ISP Border Router PID, Subscriber PID).

A second possibility for computing C_isp is to make use of Border Router IP addresses. The CDN's Border Router can locally determine the IP address of the connected border router in the ISP. In this approach, neither the CDN ALTO Server nor the ISP ALTO Server define PID attributes. The ISP ALTO Server is not required to define special PIDs for Border Routers - it only needs to ensure that Border Router IP addresses are aggregated appropriately in its Network Map.

Specifically, we identify two scenarios for the CDN ALTO Server to compute C_isp and C_cdn.

In the first scenario, the CDN does not conduct CDN-level multi-path routing from the CDN nodes to the subscriber hosts. Thus, the routing path from a CDN IP address to a subscriber host IP address is typically uniquely (if no ECMP) determined by the network routing system. In this scenario, for a given CDN node IP address to a subscriber host IP address, the CDN ALTO Server uses the routing system to compute the Border Egress router inside the CDN, and the corresponding Border Ingress router inside the ISP. Then the CDN ALTO Server has C_cdn(CDN node IP, Border Egress router IP inside the CDN), and C_isp(Border Ingress router IP inside the ISP, Subscriber IP). The computation of C_cdn and C_isp can be done using ALTO in the traditional way through either the Network Map and Cost Map or the Endpoint Cost Service.

In the second scenario, the CDN may support CDN-level multi-path

routing from the CDN nodes to the subscriber hosts. In particular, from each CDN node, the CDN has a capability (e.g., through tunneling) to send to a subscriber host IP through multiple Border Egress routers (e.g., through any Egress router that receives an announcement from the ISP of the subscriber host IP). In this case, the cost of reaching a host PID from a given CDN node is then determined as the minimum cost among all possible intermediate Border Routers.

If the network is homogeneous, then a good approximation of the cost between each host PID and a given CDN node can be given as: C_cdn(CDN Node, Border router) + C_isp(Border router, Subscriber PID). In this computation, the Border Router is the one that is on the best path from the CDN node to the Subscriber PID.

The CDN ALTO server now has a cost map that provides the cost from each CDN node to all known Subscriber PIDs. The ALTO client in the CDN DNS server downloads this cost map in preparation for subscriber DNS requests.

When a subscriber DNS request arrives at the CDN provider's DNS server, it looks up the network map and maps the source IP address to a Subscriber PID. It then uses the cost map to pick the best CDN node for this Subscriber PID.

7.2. Managed DNS Domain with Three Administrative Domains

Many organizations / content providers outsource DNS management to the external vendors for various reasons like reliability, performance improvement, DNS security etc. Managed DNS service could be used either with caches owned by the organization itself (section 6.3.1) OR with external CDNs (section 6.3.2)

7.2.1. **Managed DNS** Redirect to Local CDN

One of the common functions offered by managed DNS service vendor is DNS traffic management where DNS resolver can load balance traffic dynamically across CDN servers.

Typically managed DNS service provider has DNS resolvers spread across geographical locations to improve performance. This also makes easier for DNS resolver to redirect host to the nearest cache. Such a DNS resolver would be an ideal candidate to implement ALTO client where it can fetch network map and cost map from ALTO servers located in the same geographical area only. Load balancing implemented with the knowledge of network and cost map would be more efficient than other mechanisms like round robin.

```
                         2        +----------------+
           +--------------------> |      root      |
           | +------------------- |   Name Server  |
           | |           3        +----------------+
           | |
           | |           4        +----------------+
           | |   +--------------> |      com       |
           | |   | +------------- |   Name Server  |
           | |   | |     5        +----------------+
           | |   | |
          _|-|---|-|--------------------.
      ,-'' | V   | V                     `--.
     '   +---------+    6    +---------------`+.
     |   |   DNS   |-------->|     xyz.com    | `
     |   |  Proxy  |<--------|   DNS Resolver | |
     |   +---------+    7    |                | |
     |     1^  | 8           | +------------+ | |
     |      |  |             | |ALTO Client | | |
     |   +-----V---+         | +------------+ | |
     |   |  Host   |         +----------------.-'
     |   +---------+                |   ^  .-'
     |        |       DOMAIN 1      |   |-'   ALTO Protocol
     |        V                     |.-'|     (Map Service)
      `--. CDN Node          __.--:-|   |
          `----.        _.--'       |   |
                `---.-''      ,---------+-------.
                            ,'+----------------+ \
                           /  |   ALTO Server  |  :
                          (   +----------------+  |
                           \                      ;
                            \    DOMAIN 2       ,'
                             `-----------------'
```

In the figure above, there exists 2 possibilities:

Case 1: Domain 1 and Domain 2 are connected to the same service provider network. This case is similar to section 6.1

Case 2: Domain 1 and Domain 2 are connected to different service provider network. This case is similar to section 6.2

7.2.2. Managed DNS with CDN-Provided Request Routing

It is also possible to utilize a Managed DNS service and still rely on a CDN's request routing. For example, this could be done if a network provider wishes to utilize a Managed DNS provider, but also wishes to integrate its own CDN using ALTO with DNS-based request routing.

To support this, the network provider may submit any necessary configuration files (e.g., indicating necessary CNAME records) to redirect CDN requests to the CDN's DNS Request Routing mechanism. Requests for the CDN (e.g., 'cdn.isp.com') will then be directed by DNS request routing, while requests for other hosts are handled by the Managed DNS solution.

8. Protocol Recommendations

In the previous sections, this document has taken the approach of providing information on existing CDN approaches and possible benefits of utilizing ALTO. However, in developing the taxonomy, use cases, and deployment scenarios, we have identified cases where the ALTO Protocol [I-D.ietf-alto-protocol] and Server Discovery [I-D.kiesel-alto-3pdisc] [I-D.song-alto-server-discovery] [I-D.stiemerling-alto-dns-discovery] may be lacking capabilities that may be helpful and/or necessary for usage with CDNs. We now focus on detailing these gaps with the goal of providing feedback and recommendations. Note that some protocol changes may be necessary in the core protocol, while others may be implemented as extensions.

This section will be updated to track changes in the ALTO Protocol, ALTO Server Discovery, and accompanying protocols.

8.1. Necessary Additions

This section details changes to the ALTO protocols that would be necessary to make use of ALTO within CDN infrastructures. We classify a change as "necessary" if there is a core feature of a CDN/ ALTO integration that is not possible to implement with the existing protocols.

8.1.1. NA1: PID Attributes

In order to disambiguate between PIDs that contain endpoints of a specific class, a PID property is needed. A PID can be classified as containing "CDN nodes", "Mobile Hosts", "Wireline Hosts", etc. This mechanism can be used to provide an ALTO Client a list of nodes of a particular type, along with the ALTO Costs to each node. In the context of CDNs, the attributes could describe a type of CDN node. For example an Origin would have one type of attribute while an edge cache would have another. This would allow for more intelligent routing.

8.1.2. NA2: PID Attributes and Query

PID attributes can be used by the ALTO Client to select a appropriate host and also passed as a constraint in the map filtering service.

8.2. Helpful Additions

This section details changes to the ALTO Protocol that would be helpful to make use of ALTO within CDN infrastructures. We classify a change as "helpful" if there is a compelling extension to existing CDNs that would be possible with additional functionality within ALTO, or if there is a component of CDN/ALTO integration that could be made more efficient or otherwised improved with additional ALTO functionality.

8.2.1. HA1: Push Mechanism

It is important for the ALTO Service through the ALTO protocol or a companion protocol to provide a push mechanism from server to client. The push mechanism can be a notification that new data is available or the data itself.

8.2.2. HA2: Incremental Map Updates

A natural evolution to the protocol if maps are large and change often is to allow for incremental map updates. In this sense the map contained in the reply would be considered the delta from the previous version.

8.2.3. HA3: ALTO Border Router PID attribute

In order for administrative domains to collate costs across domain boundaries, the border routers may be placed in their own PIDs. Such PIDs may be identified by a Border Router attribute.

8.2.4. HA4: CDN ALTO Server Discovery

In certain deployment scenarios, it may be beneficial for an ALTO client to directly query a CDN's ALTO Server (instead of the CDN's ALTO Server only being consulted as a backend process). For example, this can provide more accurate guidance than DNS Request Routing since the client's IP address may be directly used by the CDN in order to select a cache node. This would require an ALTO Client (e.g., an ISP subscriber) to be able to discover an ALTO Server owned and/or managed by a CDN. This could be done by an extension to the discovery protocol, or it could be done by allowing an ISP's ALTO Server to redirect certain queries to a CDN ALTO Server.

8.2.5. HA5: Extensible ALTO Cost Maps

Certain deployment scenarios may benefit from additional information being carried within ALTO information. For example, a trusted neighboring ISP B may be able to help ISP A optimize multihoming costs. To provide an extensible way to communicate additional data, the ALTO Protocol could be extended to include opaque data strings (in addition to numeric and ordinal values) in an ALTO Cost Map.

8.2.6. NA4: Federated Deployment of ALTO Servers

There is a need to define how ALTO servers may communicate with each other in a federated model.

9. IANA Considerations

This document makes no request of IANA.

Note to RFC Editor: this section may be removed on publication as an RFC.

10. Security Considerations

When the ALTO Server and Client are operated by different entities the issue of trust and security comes forward. The exchange of information could be done using the encryption methods already present in HTTP but preventing unauthorized redistribution comes into play. A further issue is if the ALTO information information is transitive, which modifications are allowed.

11. Acknowledgements

We would like to thank Satish Raghunath and Mayuresh Bakshi for valuable input and contributions to this draft. We would also like to thank Nabil Bitar, Manish Bhardwaj, Michael Korolyov, Steven Luong and Ferry Sutanto for their comments.

12. References

12.1. Normative References

[I-D.ietf-alto-protocol]
Alimi, R., Penno, R., and Y. Yang, "ALTO Protocol", draft-ietf-alto-protocol-06 (work in progress),

October 2010.

[RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, March 1997.

[RFC5693] Seedorf, J. and E. Burger, "Application-Layer Traffic Optimization (ALTO) Problem Statement", RFC 5693, October 2009.

12.2. Informative References

[ARBOR] Labovitz, "Internet Traffic and Content Consolidation", 2009, <http://www.ietf.org/proceedings/10mar/slides/plenaryt-4.pdf>.

[GoogleCDN] Madhyastha, H., Jain, S., Srinivasan, S., Krishnamurthy, A., Anderson, T., and J. Gao, "Moving Beyond End-to-End Path Information to Optimize CDN Performance", 2009, <http://research.google.com/pubs/pub35590.html>.

[I-D.kiesel-alto-3pdisc] Kiesel, S., Tomsu, M., Schwan, N., Scharf, M., and M. Stiemerling, "Third-party ALTO server discovery", draft-kiesel-alto-3pdisc-03 (work in progress), July 2010.

[I-D.lee-alto-chinatelecom-trial] Li, K., Wang, A., and K. Zhou, "ALTO and DECADE service trial within China Telecom", draft-lee-alto-chinatelecom-trial-00 (work in progress), July 2010.

[I-D.song-alto-server-discovery] Yongchao, S., Tomsu, M., Garcia, G., Wang, Y., and V. Avila, "ALTO Service Discovery", draft-song-alto-server-discovery-03 (work in progress), July 2010.

[I-D.stiemerling-alto-dns-discovery] Stiemerling, M. and H. Tschofenig, "A DNS-based ALTO Server Discovery Procedure", draft-stiemerling-alto-dns-discovery-00 (work in progress), July 2010.

[I-D.vandergaast-edns-client-subnet] Contavalli, C., Gaast, W., Leach, S., and D. Rodden, "Client subnet in DNS requests", draft-vandergaast-edns-client-subnet-00 (work in

progress), January 2011.

[P4P] Xie, H., Yang, YR., Krishnamurthy, A., Liu, Y., and A. Silberschatz, "P4P: Provider Portal for (P2P) Applications", March 2009.

[RFC3568] Barbir, A., Cain, B., Nair, R., and O. Spatscheck, "Known Content Network (CN) Request-Routing Mechanisms", RFC 3568, July 2003.

[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, September 2009.

Authors' Addresses

Reinaldo Penno
Juniper Networks

Email: rpenno@juniper.net

Jan Medved
Juniper Networks

Email: jmedved@juniper.net

Richard Alimi
Google

Email: ralimi@google.com

Richard Yang
Yale University

Email: yry@yale.edu

Stefano Previdi
Cisco Systems

Email: sprevidi@cisco.com