Network Working Group
Internet-Draft
Intended status: Experimental
Expires: January 13, 2011

R. Penno
S. Raghunath
J. Medved
M. Bakshi
Juniper Networks
R. Alimi
Yale University
S. Previdi
Cisco Systems
July 12, 2010

ALTO and Content Delivery Networks
draft-penno-alto-cdn-01

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, 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.

A main use case of the ALTO Service is its integration with of Content Delivery Networks (CDN). In this document we describe the deployment scenarios and considerations for a ALTO Service in the case 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

This Internet-Draft is submitted to IETF in full conformance with the provisions of BCP 78 and BCP 79.

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF), its areas, and its working groups. Note that other groups may also distribute working documents as Internet-Drafts.

Internet-Drafts are draft documents valid for a maximum of six months

and may be updated, replaced, or obsoleted by other documents at any time. It is inappropriate to use Internet-Drafts as reference material or to cite them other than as "work in progress."

The list of current Internet-Drafts can be accessed at http://www.ietf.org/ietf/1id-abstracts.txt.

The list of Internet-Draft Shadow Directories can be accessed at http://www.ietf.org/shadow.html.

This Internet-Draft will expire on January 13, 2011.

Copyright Notice

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 through geolocation. But in many cases the content provider/distributor and the Internet Service Provider are disjoint and even if content servers are co-located into the ISP's networks, there is no standardized way to share information. Therefore a natural step forward would be to use ALTO to share 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 would be transparent to the user). In other words, a traditional web browser 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. Whenever a gap in protocol functionality is identified to achieve such integration, it will be enumerated with 'GAP-<N>'. Each gap is documented in a section of its own in order to foster parallel discussions and possible adoption.

3. Terminology

Content-aware Proximity Redirector: The Redirector knows about locations and presence of content & media objects in the network. Therefore the redirection to a CDN node is made based on availability of content or content-type in that CDN node and the proximity of the CDN node to the user.

Service-aware Proximity Redirector: The Redirector 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 before or when the content is served to the user.

HTTP Redirector: a Content-aware or Service-aware Proximity Redirector 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. Content Location Selection

There are multiple mechanisms that ISPs and CDNs can use to select the location from which content is served to a particular host, where information from one or more ALTO Servers can be used to improve or optimize the selection. In particular, two commonly used location selection mechanisms are HTTP Redirect and DNS name resolution. Thus, we focus on these two mechanisms.

4.1. HTTP Redirect

In this case an HTTP GET request from a host is received by an HTTP Redirector which sends back an HTTP responses with Status-Code 302 (Redirect) informing the host of the most optimal 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.

Usage of the ALTO Server with HTTP Redirects is shown in the following figure. Either the Map Service or the Endpoint Cost Service can be used by the ALTO client embedded in the HTTP Redirector entity.

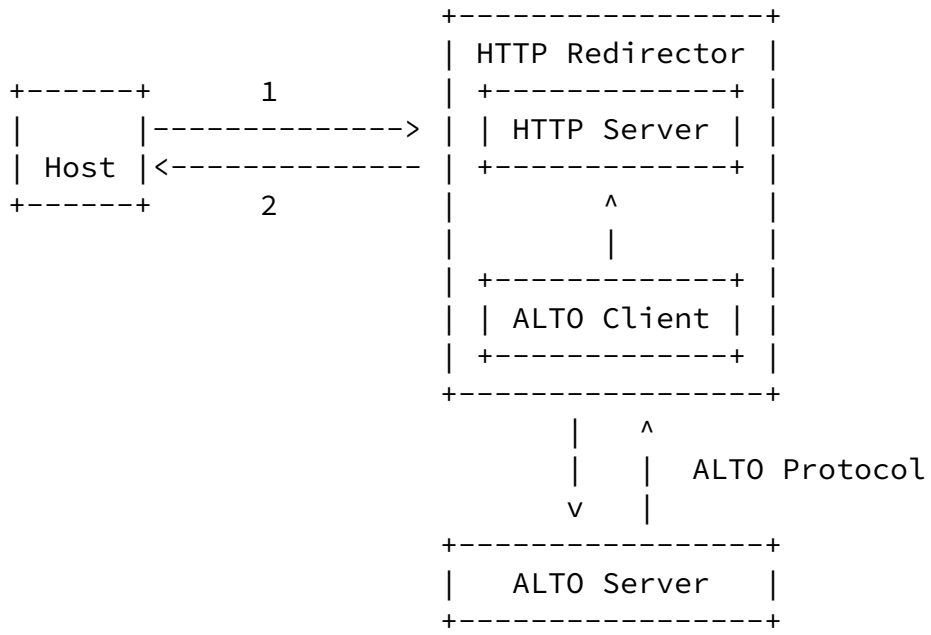


Figure 1: HTTP Redirector

4.1.1. The Map Service

The ALTO client embedded in the HTTP Redirector fetches the Network and Cost Maps from the ALTO Server and provides that information to the HTTP Redirector. As an illustrative example, a simple Redirector may be given (from an external source) the list of available CDN nodes. The Redirector 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 it receives an HTTP GET request (1), it looks up the PID of the source IP address on the GET request, indexes the redirection table using the request PID to select a CDN node, and finally returns an HTTP redirect with the URL of the selected CDN node. In practice, the redirection table may be indexed by both source and content to provide better redirection.

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.

The Network Maps generated by the ALTO server contain Host PIDs and CDN Node PIDs, i.e., Host PIDs contain host subnets; CDN PIDs contain IP addresses of available CDN nodes. Cost Maps contain only cost from each host PID to each CDN PID and not the full matrix across all PIDs. The reason is that the HTTP Redirector can only redirect a host 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 (GAP).

The cost for CDN PID to CDN PID and between host PIDs are assumed to be infinity (GAP). The HTTP Redirector looks up the source address on the HTTP GET request, and uses the cost map to select the best CDN PID and a CDN node from it. The CDN node selection method can be random, round-robin, or the HTTP Redirector can use some level of content awareness (i.e. send requests for the same content (URL) to the same CDN node.

GAP-1 (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. Note that the Alto Server will have to be told which subnets belong to hosts and which subnets belong to CDN Nodes.

GAP-2 (PID Attributes and Query): PID attributes can be used by the ALTO Client to select an appropriate CDN Node and also passed as a constraint in the map filtering service.

GAP-3 (Default Cost): The issue of default cost is one of importance. Without a default PID with endpoint '0.0.0.0/0', what should be the cost between two PIDs? Moreover, is the default PID mandatory in the protocol?

4.1.2. The Endpoint Cost Service

Alternatively, ALTO client embedded in the HTTP Redirector, requests

the endpoint cost service from the ALTO client.

The Redirector knows the IP address of the user (content requester) and the different content locations. It then requests the Endpoint Cost Service in order 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.

Note that the mechanisms through which the CDN acquires the IP addresses of the content locations (i.e.: how to locate the requested content) are part of the CDN implementation and their description is outside the scope of this document.

Once the Redirector 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 location.

4.1.3. CDN Node Discovery and Status

The method of discovering available caches and their locations is outside the scope of this document. We assume the CDN nodes are discovered in some way. It is desirable that not only CDN node locations, but also real-time status (like health, load, cache utilization, CPU, etc.) is communicated either to the HTTP Redirector or to the ALTO Server.

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. The HTTP Redirector or the ALTO Server can implement an SNMP agent and get to know whatever is needed. For greenfield installations, the ALTO Server could also 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.

In addition to the CDN node status, network status can also be retrieved from TE/RP databases.

4.1.3.1. CDN Node Status Updates received by HTTP Redirector

In this use case the HTTP Redirector receives CDN Status updates.

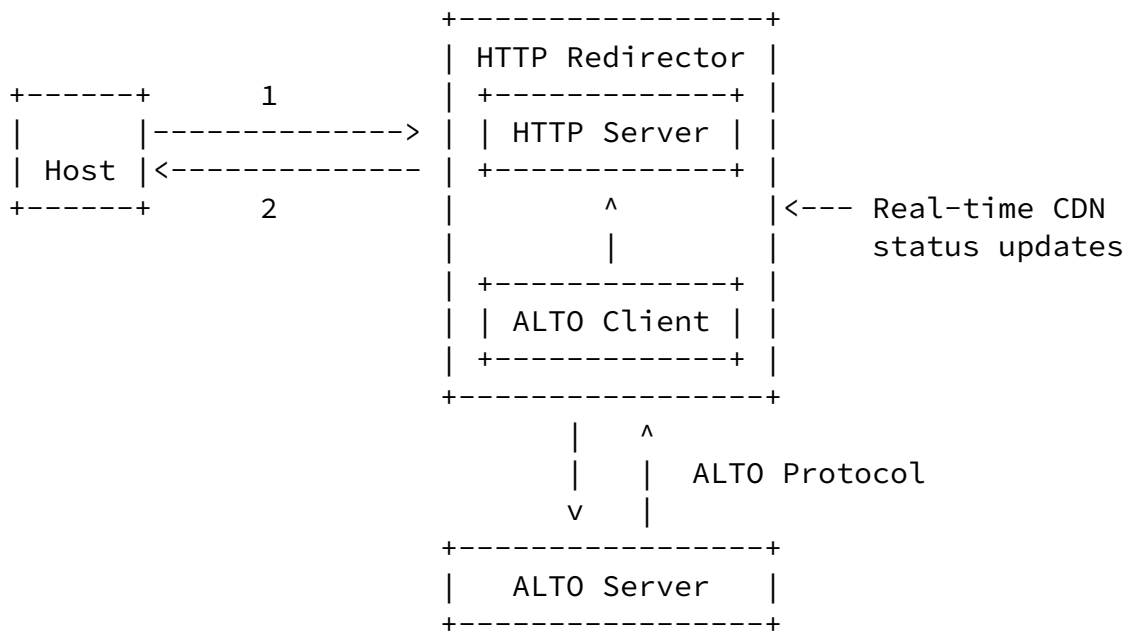


Figure 2: RT CDN Updates to HTTP Redirector

4.1.3.2. CDN Node Status Updates received by ALTO Server

This model generally simplifies the HTTP Redirector. It allows an easier distribution of the HTTP Redirector, 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 HTTP Redirector and the ALTO Server to be in different administrative domains. For example, the HTTP Redirector can be in a Content Provider's domain, the ALTO Server and CDN Nodes in a Network Service Provider's domain.

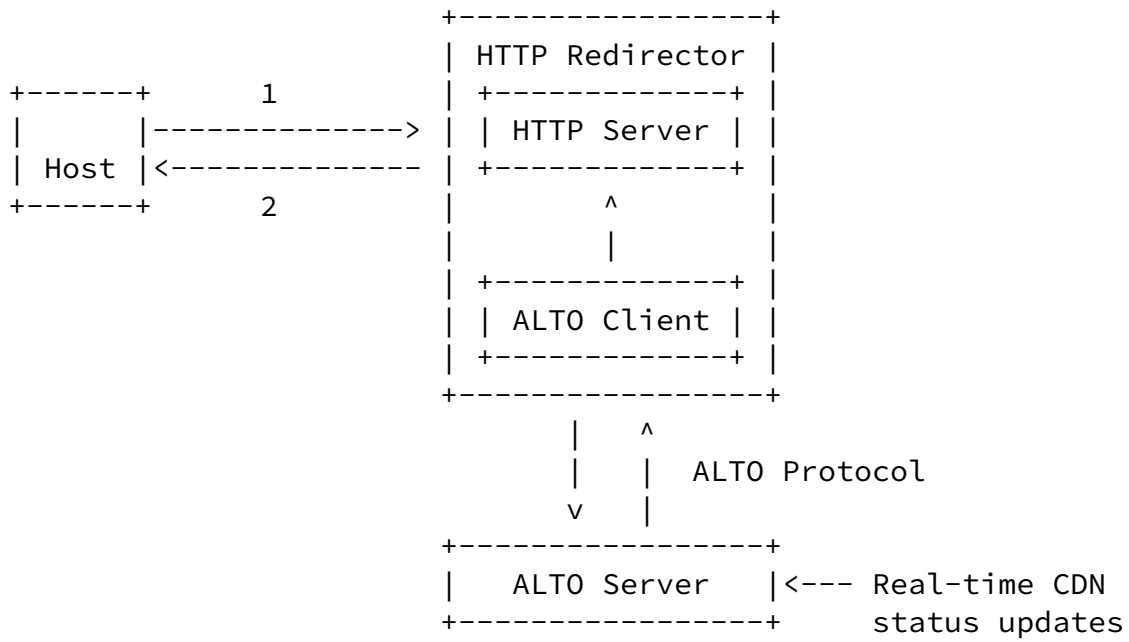


Figure 3: RT CDN Updates to ALTO Server

In this model it is recommended that a given HTTP Redirector may be designated as being responsible for a fixed set of Host PIDs. This information can be made available to the HTTP Redirector 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, an HTTP Redirector 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 maps for relevant PID 2-tuples.

An HTTP Redirector that uses the Endpoint Cost Service may query the ALTO Server for rankings of CDN Node IP addresses for each interesting Host and cache the results for later usage.

The HTTP Redirector 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. In order to better deal with outages of caches or changes to CDN PIDs, a push mechanism from ALTO server to ALTO client would be needed (GAP-4). In the general P2P scenario this may not make sense, but with content delivery this may be important from a service continuity perspective.

If the maps are large and change often a natural extension to the protocol is to allow incremental Map Updates (GAP-5). This

requirement becomes more emphasized when the ALTO Server is the recipient of CDN nodes' status updates, because their load/status

changes are typically more frequent than topology changes in the network.

GAP-4 (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.

GAP-5 (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.

4.2. DNS Integration

In the case of DNS request routing, the DNS server handling host requests is integrated with an ALTO client. When the host performs a DNS query/lookup, the IP address contained in the response is already optimal for that query. As in the previous example, no changes in the host are needed.

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 queries the DNS Servers, the authoritative DNS Server can see directly the host's IP address. If the the DNS Proxy's 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 with recursive DNS queries.

In an iterative DNS lookup with DNS Proxy, as shown in examples in the next section, 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 the 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.

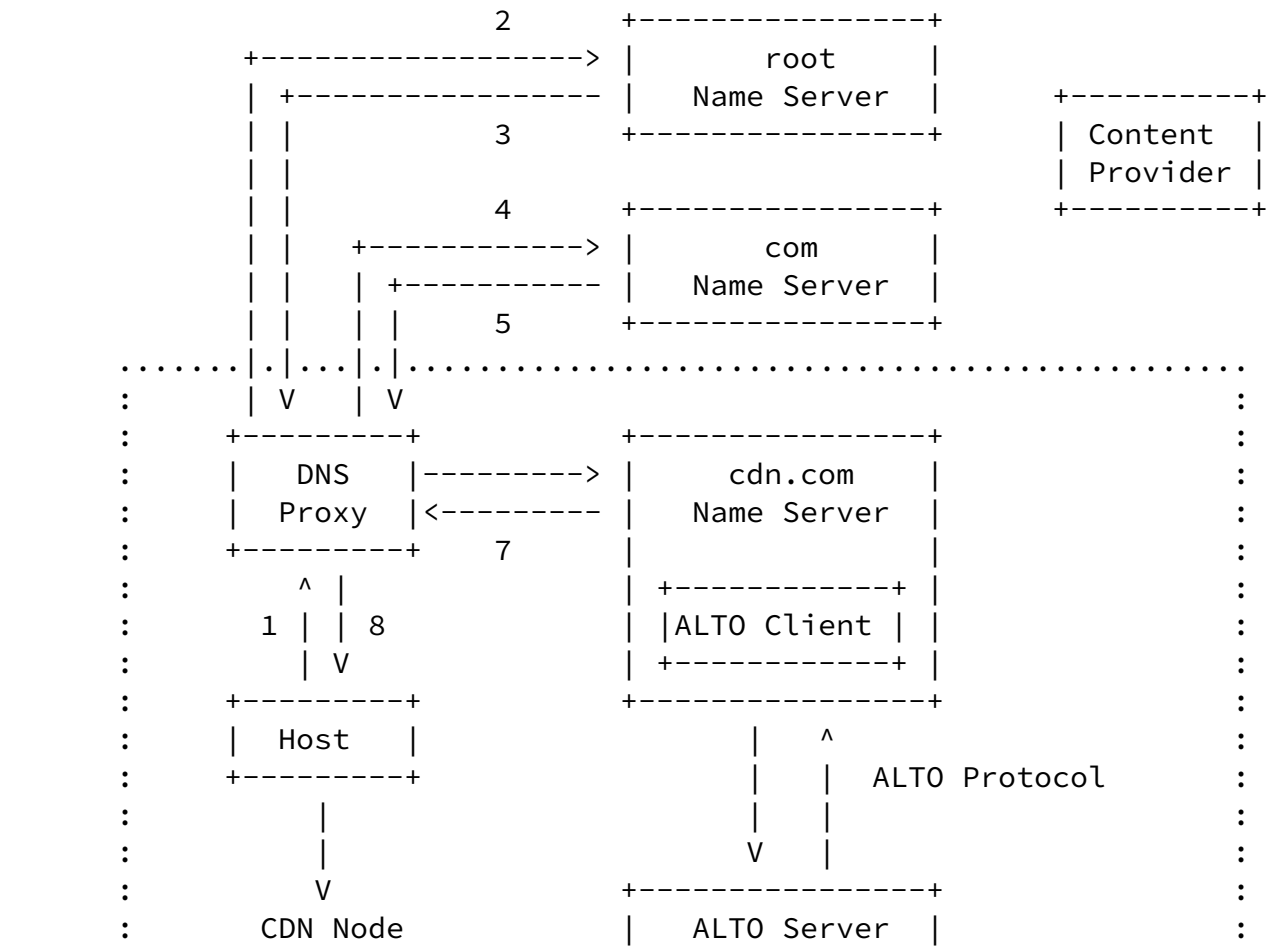
4.3. ALTO Server Discovery

5. Administrative domains and ALTO

With DNS-based redirection, among others, there are two models that are worth further study - one, where the CDN nodes are in the administrative domain of the ISP and two, where CDN nodes are part of a separate domain from that of the ISP. In the first use case, the Host, the CDN Nodes, the ALTO Server and the Authoritative DNS Server for the CDN domain are in the same administrative domain. In the second use case, Hosts and CDN Nodes are in different administrative domains.

5.1. CDN nodes in the ISP administrative domain

When the CDN nodes are within the ISP's administrative domain, the DNS server with the ALTO client is under the ISP's management. A best CDN nodes can be picked by performing ALTO query on the source IP address (and the target domain name) of the DNS request.


2
root
Name Server
3
Content
Provider
4
com
Name Server
5
DNS
Proxy
7
cdn.com
Name Server
1
8
ALTO Client
Host
ALTO Protocol
CDN Node
ALTO Server

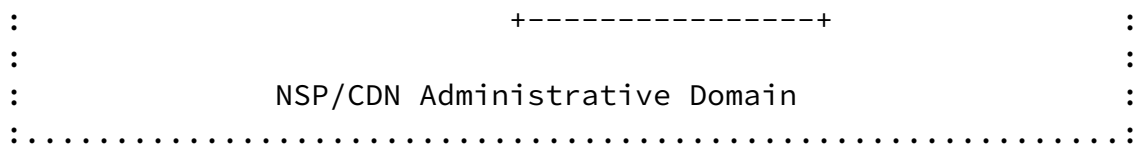
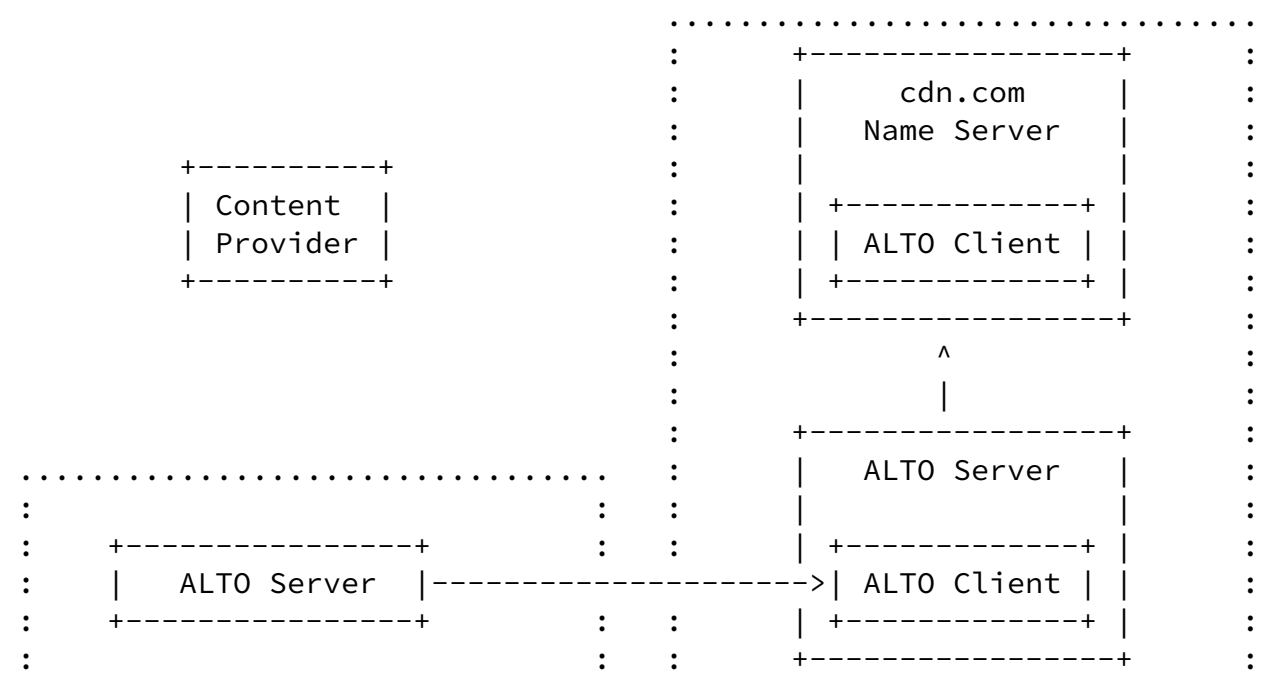


Figure 4: DNS Resolution with single admin domain

5.2. CDN nodes in a separate administrative domain from that of ISP

In many situations, the CDN nodes 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 subscriber belongs.



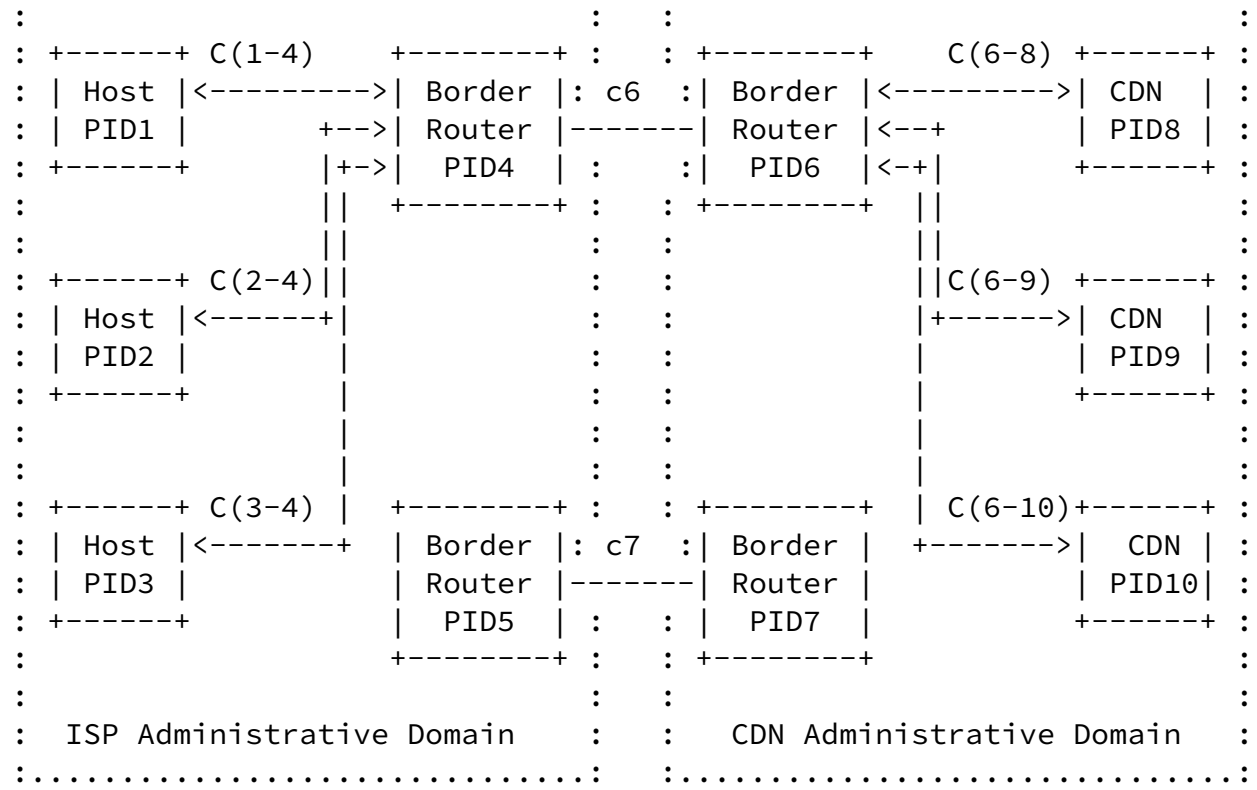


Figure 5: Map advertising between ISP and CDN domains

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.

GAP-6: Federation of ALTO servers: There is a need to define how ALTO servers may communicate with each other in a federated model.

GAP-7: 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.

5.3. Integrating with managed DNS service

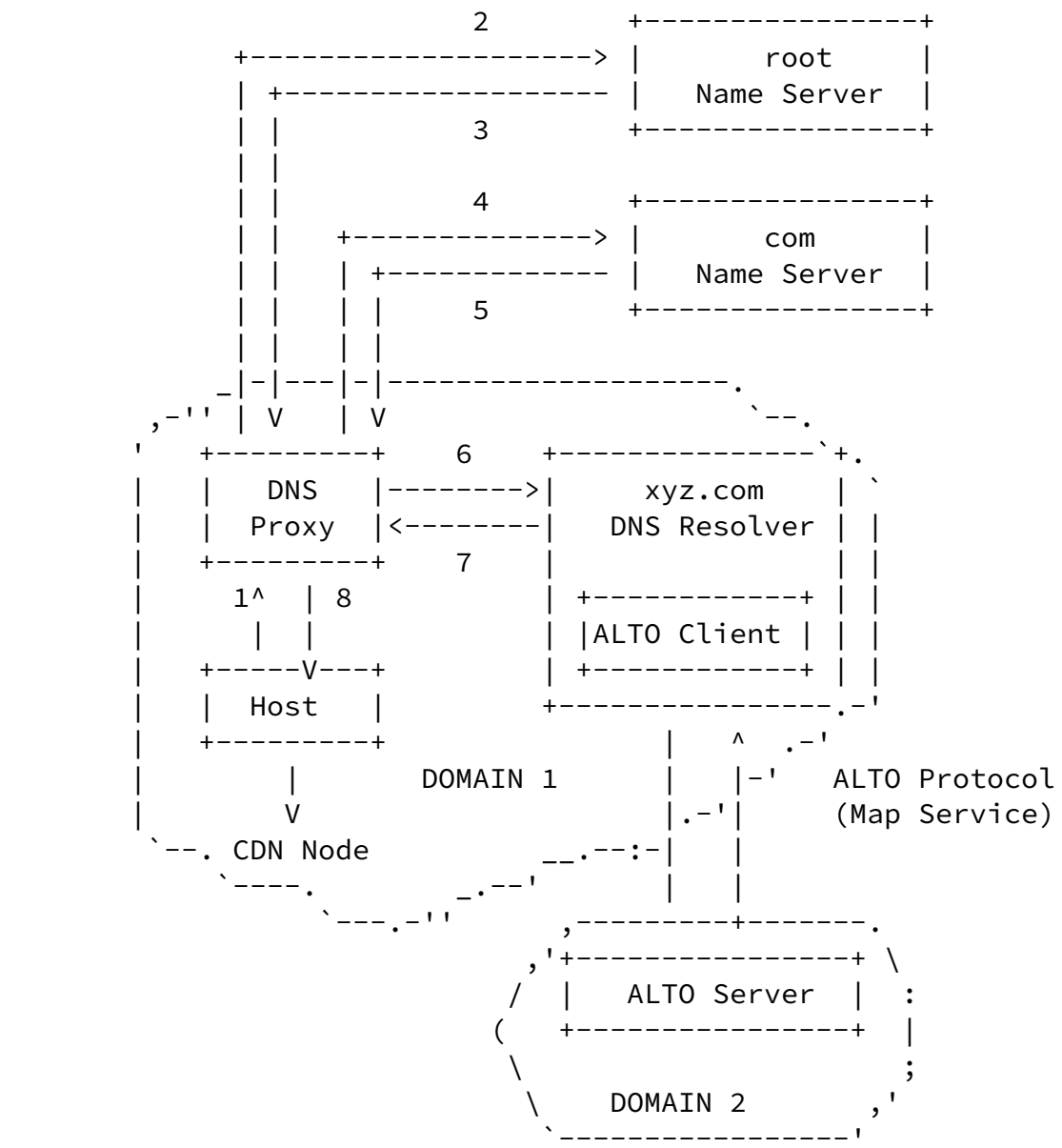
Many organizations / content providers outsource DNS management to the external vendors for various reasons like reliability, performance improvement, DNS security etc. Managed DNS service could be used either with caches owned by the organization itself (section 6.3.1) OR with external CDNs (section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

5.3.1. Managed DNS resolver used to redirect to local cache

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



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

5.3.2. Managed DNS resolver used with multiple CDN vendors

In this Model, Managed DNS service can be used along with multiple CDN vendors where DNS resolver can redirect to different caches depending on the subdomain e.g. DNS resolver could have below records

subdomain1.xyz.com CNAME cdn1.com

subdomain2.xyz.com CNAME cdn2.com

In this case CDN DNS resolver needs to be an ALTO client. This deployment will be similar to ones described in section 6.1 and section 6.2 earlier.

6. Tracker Integration

In the case of P2P CDNs, the application tracker takes the role of the ALTO Client, fetching the map from the ALTO Server and integrating it its peer database. The result is a peer database taking into account current metrics such as peer availability, content availability and also localization. This architecture in the context of file sharing was extensively studied and trialed by ISPs such as Comcast [RFC5632] under the P4P [P4P] protocol.

7. IANA Considerations

This document makes no request of IANA.

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

8. 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.

9. Acknowledgements

We would like to thank Richard Yang 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.

10. References

10.1. Normative References

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

10.2. Informative References

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

[I-D.ietf-alto-protocol] Alimi, R., Penno, R., and Y. Yang, "ALTO Protocol", draft-ietf-alto-protocol-03 (work in progress), March 2010.

[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

Satish Raghunath
Juniper Networks

Email: satishr@juniper.net

Jan Medved
Juniper Networks

Email: jmedved@juniper.net

Mayuresh Bakshi
Juniper Networks

Email: mbakshi@juniper.net

Richard Alimi
Yale University

Email: richard.alimi@yale.edu

Stefano Previdi
Cisco Systems

Email: sprevidi@cisco.com