riverbed

Think fast.™

WHITE PAPER

Important considerations when optimizing Office 365 in an Azure Express Route (EXP) environment

April 2016

Table of Contents

Introduction 2
Understanding the Exchange Online (EXO) architecture 2
Understanding the SharePoint Online (SPO) architecture 3
Challenges for SteelHead 4
Non-Local Host Optimization Latency 4
Recommendation 6
Asymmetric Routing 7
Recommendation – Customer-to-Microsoft traffic 7
Recommendation – Microsoft-to-customer traffic 8
DNS 9
Recommendation 9
Traffic Redirection 10
Recommendation 10
SSL Certificates 10
Existing SteelHead O365-D customers migrating to the next generation design 10
New O365 Multi-Tenant tenant customers 10
Conclusion 10

Introduction

Azure Express Route (AER) provides the option for customers to peer directly with Microsoft rather than traversing the Internet. With this direct peering, AER promises the following:

- provide better network connectivity for resources located in Azure and Office 365 (O365);
- provide availability SLA; and
- avoid the inherent congestion of the Internet

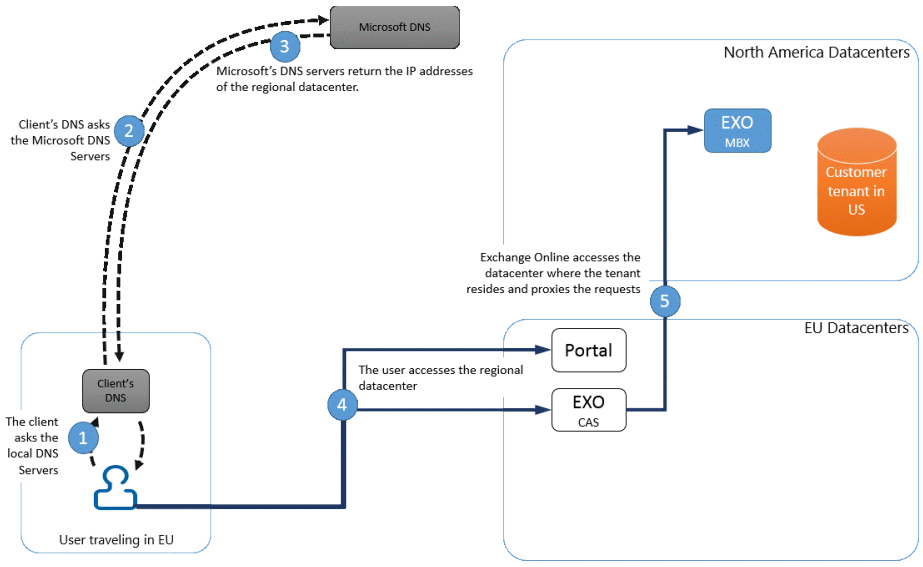
AER poses some unique considerations with SteelHead optimization. This paper highlights some of the important issues customers should take into account prior to deploying AER with Riverbed SteelHead optimization. Please note that this paper is geared towards the Exchange Provider (EXP) model and not the Network Service Provider (NSP) model.

The reader should be familiar with the following products and concepts:

- Azure Express Route and BGP peering;
- SteellHead operation;
- Traffic engineering;
- Exchange Online terminologies;
- DNS; and
- SSL certificates

Understanding the Exchange Online (EXO) architecture

In Figure 1, the O365 tenant is in the US meaning that the locations of the users' mailboxes are typically all located in the US. Furthermore, the users' mailboxes are typically spread across multiple datacenters within the US. The user in this example is logging in while traveling overseas.



Source: Microsoft Technet

Figure 1

When a user accesses their email in Exchange Online (EXO), the client performs a DNS query for outlook.office365.com amongst other hostnames and the DNS server (steps 1 and 2) and Microsoft then returns a list of the closest client access front end (CAFÉ) server (step 3).

Once the client has the list of IP addresses, the user connects to a CAFÉ server (step 4) and the CAFÉ server initiates a separate connection to the actual Exchange server where the mailbox (MBX) is held (step 5). In other words, there are two separate TCP connections here: one between the end-user and the CAFÉ server and another one between the CAFÉ server and the MBX server.

In an effort to improve the underlying network connectivity into Office 365, EXO relies on DNS to direct traffic to the closest geographical CAFÉ servers in each region. This feature is known as “GeoLocation” or “GeoDNS”. For performance reasons, users typically obtain their DNS servers via DHCP that belongs to the same region. In Figure 1, a US-based user traveling to Europe utilizes a DNS server that is located in the Europe. As such, the user receives a list of IP addresses for Exchange Online CAFÉ servers based in Europe and the client will pick a CAFÉ server from that list. The idea is to bring the traffic destined to EXO into the Microsoft backbone as soon as possible as to avoid the unpredictability of the Internet.

In Figure 1, step 4) is typically low latency while the latency in step 5) can exceed 100ms even though the connection between the CAFÉ and MBX server is carried over the Microsoft backbone.

This traffic flow of connecting to the closest regional CAFÉ server is not limited to international travelers as the same concept also applies when users are traveling within a region. For example, a US-based user logging in from the west coast could connect to a CAFÉ server in Texas but their mailbox residing in east coast. The only difference is that the latency may be lower for intra-region rather than inter-region connectivity.

Understanding the SharePoint Online (SPO) architecture

With SPO, the architecture is simpler because there is no concept of front-end and back-end servers. When a client connects, it connects directly to the SPO instance.

Challenges for SteelHead

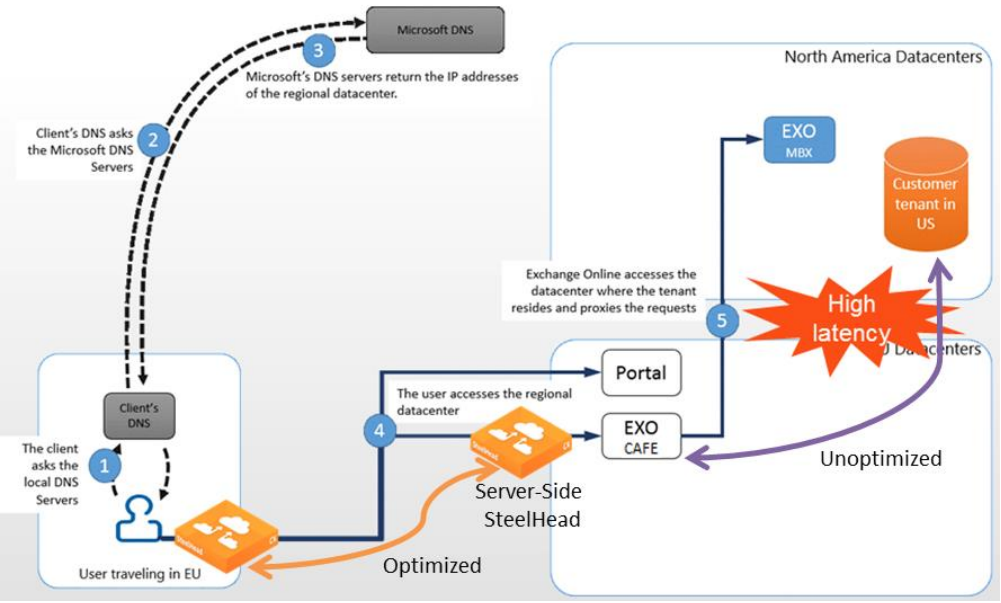


Figure 2

Continuing our example, the customer has an AER peering point in Europe and deploys SteelHead optimization as shown in Figure 2. The connection between the user and Microsoft regional datacenter in Europe will be optimized while the connection between the CAFÉ server and the MBX server will be unoptimized as the servers reside within Microsoft's network.

The fact that the CAFÉ and MBX servers can reside in different Microsoft data centers within a region or across different regions means that the latency most likely would not be LAN-like. The potentially high latency between the CAFÉ and MBX servers has a direct impact on the performance of the SteelHead optimization. Riverbed calls this type of situation common to cloud scenarios as Non-Local Host Optimization (NLHO) latency. The impact NLHO latency has on SteelHead optimization is further discussed in the next section.

While SPO has a different architecture, it also suffers from the NLHO latency as the server-side SteelHead could be far away from the SPO instance. In the example above, the server-side SteelHead could be in Europe while the SharePoint instance is in the US.

Non-Local Host Optimization Latency

Figure 3a illustrates the impact NLHO latency has on performance. As the NLHO latency increases, the amount of time it takes to complete an operation in Outlook increases. The tests were performed under the following parameters:

- Fixed WAN latency of 60ms and unlimited bandwidth;
- AER with 200Mbps of bandwidth;
- Outlook 2013;
- Office 365 tenant based in the US; and
- Sending a 13MB Word document

The NLHO latency in the graph includes the latency between the server-side SteelHead to the CAFE server and between the CAFÉ server to the MBX server. 3ms NLHO latency is the equivalent of having the CAFÉ server and MBX server at the same location (i.e. there is no NLHO latency).

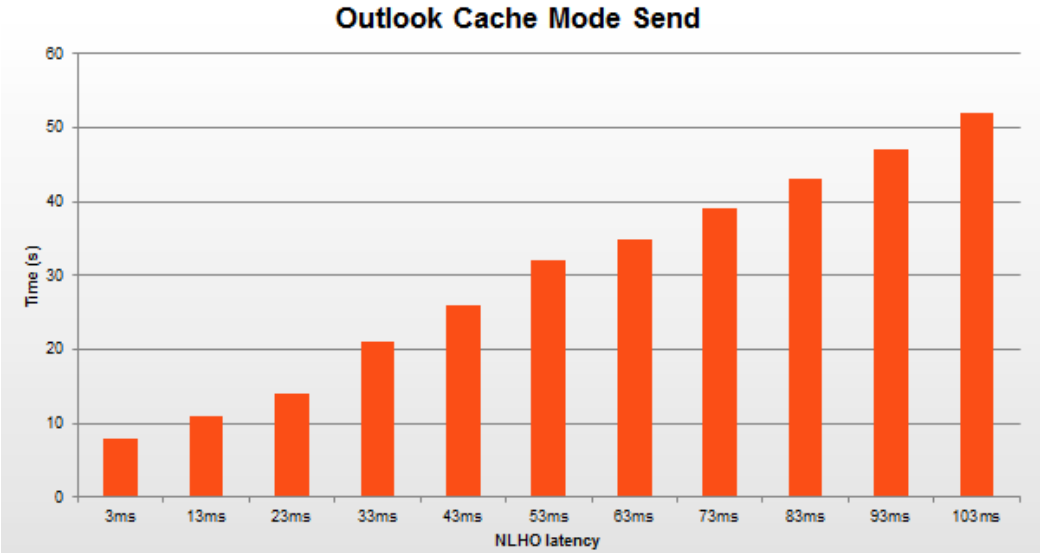


Figure 3a

Figure 3b compares the performance between no SteelHead optimization and SteelHead optimization with the presence of NLHO latency. The orange bars in Figure 3b corresponds with the orange bars in Figure 3a. "Latency" in Figure 3b includes the WAN latency and the NLHO latency. For example, 83ms = 60ms WAN latency + 23ms of NLHO latency.

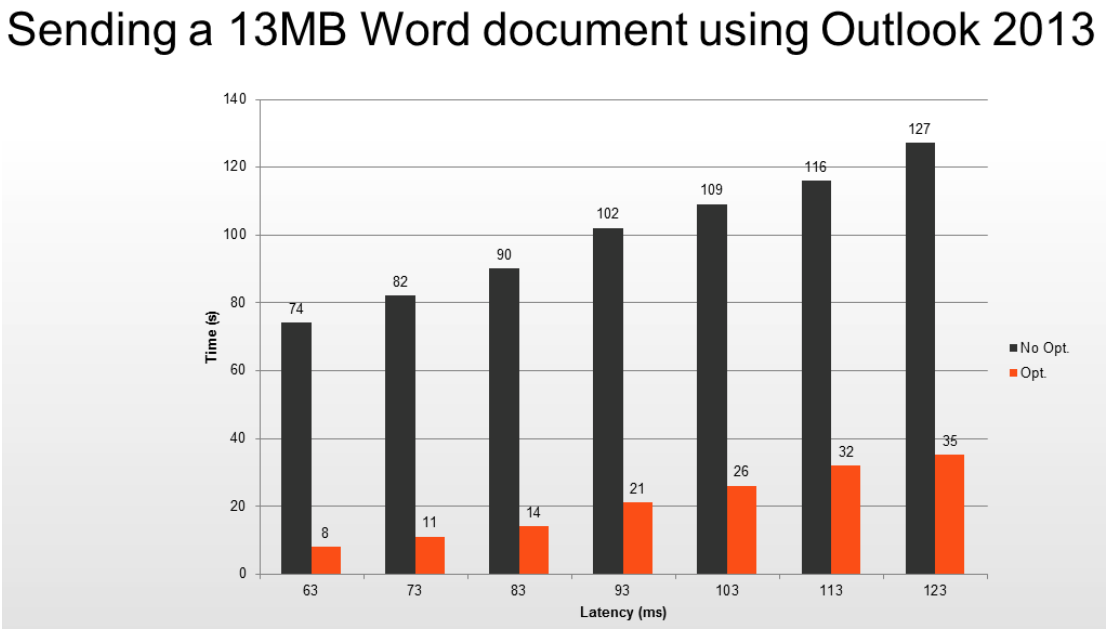


Figure 3b

When deploying SteelHead in an AER environment, NLHO will almost always be present and therefore the solution is not to eliminate it but rather to try and minimize its impact.

Recommendation

Microsoft Azure offers peering locations at different locations within the US. In order to minimize the impact NLHO latency has on EXO, the recommended peering locations are Dallas and/or Chicago. The reason why these two locations are ideal for optimizing EXO is because they are centrally located within the US (refer to Figure 4).

By peering in Dallas and/or Chicago, the NLHO latency could be reduced. For example, if the client is directed by DNS to use a CAFÉ server in VA but with the MBX server located in CA, then the NLHO latency will be 36ms across the AER network (Dallas-to-VA) and 60ms across the Microsoft backbone network (VA-to-CA) for a total latency of 96ms. Compare this to a peering location in CA in which case it could incur a "full boomerang" of 60ms across the AER network (CA-to-VA) and then 60ms across the Microsoft backbone network (VA-to-CA) for a total of 120ms in the worst case scenario.

Indeed, it is possible that the peering location, the CAFÉ server, and the MBX server are all located at one location and therefore there is no NLHO latency. However, given the architecture of EXO, this scenario is likely to be an exception rather than the norm.

Furthermore, to minimize the impact NLHO latency has on SPO, a peering location should be chosen that is close to the SPO instance.

The recommended AER peering points for the rest of the world are:

- Europe: Amsterdam and London;
- Asia: Hong Kong and Singapore;
- Australia: Sydney and Melbourne

Figure 4

Asymmetric Routing

When SteelHeads are deployed to optimize the Office 365 traffic, it is imperative that the Office 365 traffic flows symmetrically. In other words, traffic going out of a one peering location must return back via the same peering location. There two reasons why this is important:

1) Customers who are using RFC1918 address must NAT the source address to the publicly registered IP addresses. As NAT is involved, the router performing the NAT must be able to process both inbound and outbound traffic; and
2) In a traditional datacenter environment, the SteelHead appliances can support asymmetric routing by leveraging the Connection Forwarding feature. However, Connection Forwarding was designed to work within the same datacenter with LAN-like latency. Connection Forwarding across different Azure peering points (e.g. Dallas and Chicago) is not supported.

Asymmetric traffic could take place in either direction: customer-to-Microsoft and Microsoft-to-customer.

Recommendation – Customer-to-Microsoft traffic

When peering with Microsoft, all peering points will receive the same routes from Microsoft. Therefore, depending on the interior gateway protocol (IGP) being used and how the routes are redistributed from BGP into the IGP, it's conceivable that there may be equal cost routes in reaching the Microsoft network.

From a network symmetry perspective, having equal cost routes to a destination is not an issue as traffic for a certain flow should traverse the same path (i.e. no per-packet load balancing). However, it has been observed that by minimizing the latency between the SteelHead appliance and the SPO instance, there is a noticeable performance increase.

For example, if the SPO instance is located in Virginia, then it would make sense to ensure that the traffic flow through the Washington DC peering point as the latency is typically less than 10ms versus about 36ms from Dallas. A common way to influence the traffic flow is via the well-known BGP attributes such as local preference.

Microsoft has indicated that routes will be tagged with a certain BGP community value based on the service. Riverbed suggests matching the route based on the BGP community value (e.g. SharePoint) and then redistribute those routes into the IGP. However, as of April 2016, the BGP community values are not part of the route updates.

Service	BGP community value
Exchange	12076:5010
SharePoint	12076:5020
Skype For Business	12076:5030
CRM Online	12076:5040
Other Office 365 Services	12076:5100

Source: https://azure.microsoft.com/en-us/documentation/articles/expressroute-routing/

Figure 5

In the absence of BGP community values to match the routes that are related to SPO, the alternative is to simply match using the subnet. In general, both SPO and OneDrive traffic use the URL <domain>.sharepoint.com or <domain>-my.sharepoint.com and performing a "ping" or DNS lookup on the name will reveal the IP address. Once the IP address is known, execute the command "show ip route x.x.x.x" on the router to determine which route matches that IP address. Note that it's possible Microsoft may be blocking ICMP in their network and therefore the "ping" command may not receive a response from the server. This is not an issue as the hostname is should still resolve.

```
$ ping rvbdaer.sharepoint.com
PING prodnet320-281ipv4a0001.sharepointonline.com.akadns.net (104.146.156.34): 56 data
bytes
64 bytes from 104.146.156.34: icmp_seq=0 ttl=236 time=227.037 ms
64 bytes from 104.146.156.34: icmp_seq=1 ttl=236 time=228.120 ms

rtr#sh ip route 104.146.156.34
Routing entry for 104.146.0.0/15
  Known via "bgp 18597", distance 20, metric 0
  Tag 12076, type external
```

Once the route has been determined, the Washington DC router can then redistribute the route into the IGP by matching the 104.146/15 prefix using a more favorable metric than the Dallas router and therefore attracting the traffic to the Washington DC peering point.

Recommendation – Microsoft-to-customer traffic

Instead of advertising a large summarized route to Microsoft, it is recommended to allocate smaller blocks of IP addresses to each peering location and advertise the more granular routes to Microsoft. For example, Riverbed owns the range 208.70.196.0/22 and therefore it is possible for Riverbed to advertise this single prefix at both Dallas and Chicago peering points. However, by only advertising this 208.70.196.0/22 prefix, Microsoft will have two equal cost routes in their routing table and could send the traffic back via Dallas and/or Chicago.

On the other hand, by allocating smaller ranges to Dallas and Chicago, then Microsoft will have more specific routes in its routing table and traffic will always flow through corresponding peering point. For example, Dallas would be allocated 208.70.196.0/25 and the Dallas router will only advertise this prefix to Microsoft while Chicago would be allocated 208.70.196.128/25 and the Chicago router will only advertise this prefix to Microsoft via BGP. From

Microsoft's perspective, the only way to reach 208.70.196.0/25 is via Dallas and the only way to reach 208.70.196.128/25 is via Chicago.

Outbound traffic will then be source NAT to these ranges depending on which peering point the traffic flows through and the return traffic from Microsoft will traverse the same router on the way back.

DNS

As mentioned earlier in the "Understanding Exchange Online Architecture" section, the purpose of the "GeoLocation" feature is to bring the traffic destined for EXO into the Microsoft backbone as soon as possible as to avoid the unpredictability of the Internet. This makes sense when there is local Internet breakout but likely to create an undesirable effect when used in conjunction with AER.

Consider the scenario in Figure 6 whereby the customer has an AER peering point in the US but with local Internet breakout in Europe and using local DNS servers in Europe.

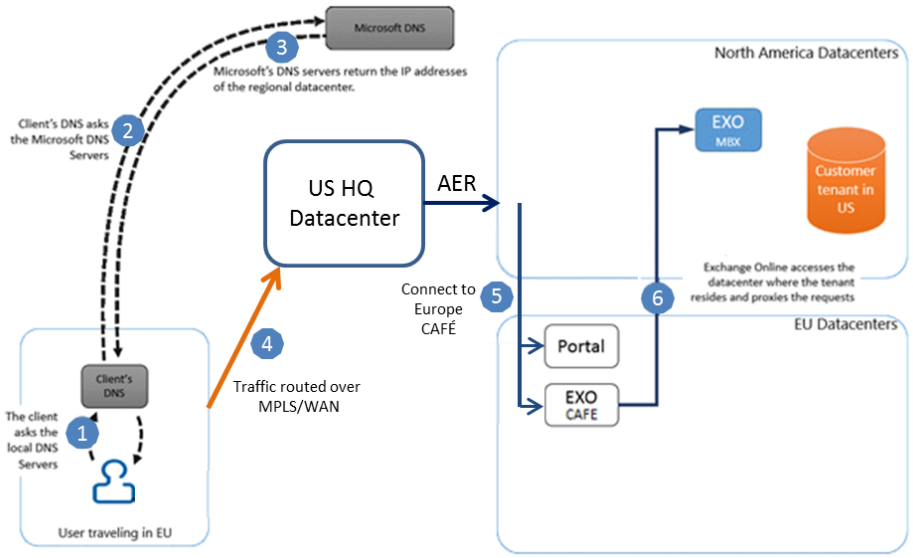


Figure 6

After the DNS lookup, the user receives the list of CAFÉ servers that are based in Europe. When the client initiates a connection, the traffic will traverse the MPLS/WAN rather than going directly to the Internet. This is the expected and desirable behavior as the traffic should be attracted towards the AER peering point in the US. However, as the traffic enters the Microsoft backbone network in the US, the destination IP address remains that of the CAFÉ server in Europe. The traffic will traverse the Microsoft backbone network in the US and connect to the CAFÉ server in Europe. The CAFÉ server in Europe then initiates its connection across the Microsoft backbone network to the MBX server in the US. Depending on the location of the user, AER peering point, and the MBX server, the total latency in this traffic flow could exceed 400ms.

Recommendation

While this is not a SteelHead-specific issue, the recommended solution to this problem is to configure conditional forwarding on the local DNS server for domains related to EXO. The IP address of the conditional forwarder should be in the same region as the O365 tenant. For more information the relevant domain names that should be conditionally forwarded, refer to this article.

Traffic Redirection

EXO and SPO typically experience the most benefit from WAN optimization. As such, only traffic related to EXO and SPO should be optimized while the rest of the traffic should be passed through.

Recommendation

There are various ways to selectively redirect the traffic for optimization depending on the RiOS versions. As such, please consult with the Riverbed account team on the best way to redirect traffic only for EXO and SPO for optimization.

SSL Certificates

Existing SteelHead O365-D customers migrating to the next generation design

Existing SteelHead customers who are migrating to the next generation O365-D environment should check and ensure their existing SSL certificates can be used in the next generation O365-D environment.

New O365 Multi-Tenant tenant customers

New customers who are looking to deploying SteelHead appliances in an AER environment should contact their local Riverbed account team for details on configuring SSL optimization.

Conclusion

Deploying SteelHead appliances in an AER environment is supported although it does introduce challenges not found in traditional on-premise deployments. However, many of these challenges can be mitigated through proper planning and traffic engineering resulting in significant benefits when SteelHead appliances are deployed in an AER environment.

About Riverbed

Riverbed Technology is the IT infrastructure performance company. The Riverbed family of wide area network (WAN) optimization solutions liberates businesses from common IT constraints by increasing application performance, enabling consolidation, and providing enterprise-wide network and application visibility – all while eliminating the need to increase bandwidth, storage or servers. Thousands of companies with distributed operations use Riverbed to make their IT infrastructure faster, less expensive and more responsive. Additional information about Riverbed is available at www.riverbed.com.

riverbed
Think fast.™

Riverbed Technology, Inc.
680 Folsom Street
San Francisco, CA 94107
Tel: (415) 247-8800
www.riverbed.com

Riverbed Technology Ltd.
Farley Hall, London Rd., Level 2
Binfield
Bracknell. Berks RG424EU
Tel: +44 1344 354910

Riverbed Technology Pte. Ltd.
391A Orchard Road #22-06/10
Ngee Ann City Tower A
Singapore 238873
Tel: +65 6508-7400

Riverbed Technology K.K.
Shiba-Koen Plaza Building 9F
3-6-9, Shiba, Minato-ku
Tokyo, Japan 105-0014
Tel: +81 3 5419 1990