Data-Intensive Cloud Service Provision for Research Institutes: the Network Connectivity Problem

CERN, August 2016

Tony Cass & Edoardo Martelli

Draft for Review

1 Abstract

Much effort (and money) has been invested to ensure that academic and research sites are well interconnected with high-capacity networks that, in most cases, span national and continental boundaries. In the last years, these academic and research sites have started using commercial cloud services, which may not be able or allowed to benefit of the high speed network infrastructure put in place by the research and education network operators (RENs.

After a brief summary of the issues involved, we describe three approaches to removing the network connectivity barrier that threatens to limit the ability of academic and research institutions to profit effectively from services offered by Cloud Service Providers (CSPs).

2 Problem statement

The growth of data-intensive science over the past 10-15 years has gone hand-in-hand with a growth in the exploitation of remotely located computing resources, initially as a sharing of publically funded, dedicated resources (the “Grid” model) and more recently through the growing use for scientific purposes of commercially provided resources (the “Cloud” model).

In some cases, for example searching for a match in a genome database, the volume of data exchanged between a client and the remote resource may be relatively small. In others, however, effective exploitation of remote computing resources requires high-speed transfer of high volumes of data. The computing needs of the experiments at CERN’s Large Hadron Collider are perhaps the best-known example of this latter class of data-intensive computing and it is noteworthy that much effort has been devoted to the provision and management of high-bandwidth network connections between the participating institutions.[1]

The high-bandwidth connections serving the needs of the LHC experiments and of other data-intensive science communities, are generally provided by dedicated research and

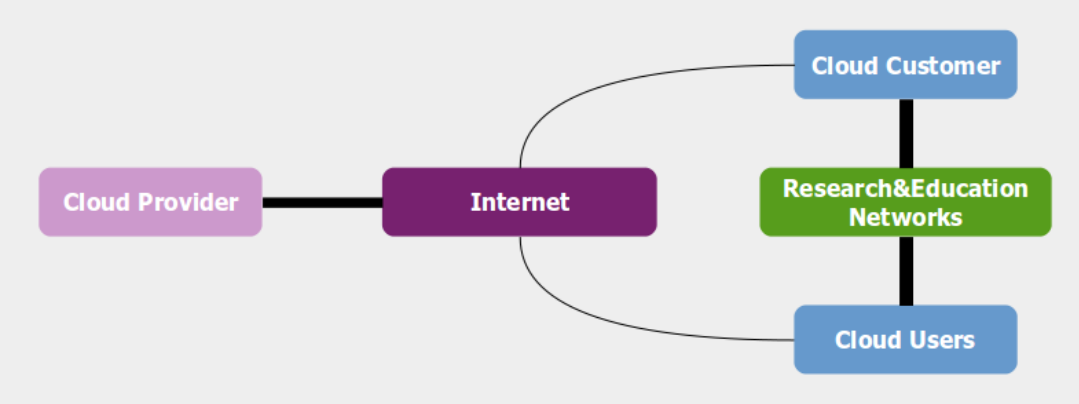
1 See, for example, the architecture document for the private optical network linking the major WLCG centres at http://cern.ch/go/8NFg and the overview of the LHCONE network linking all WLCG sites at http://cern.ch/go/9BnZ.

education networks which, given that they are supported by public funding, generally restrict, if not prohibit, their use by commercial companies. These restrictions affect the use of CSPs for data-intensive science in two ways.

Firstly, the effective bandwidth achievable between an institute with a high bandwidth connection to a REN and a CSP connected to the public internet is limited. CERN, for example, has a 100Gbps connection to GÉANT's trans-European network[2] but only 8Gbps connections to the public internet. Similarly, most of the sites contributing to the Worldwide LHC Computing Grid (WLCG) will have dedicated connections to their NREN of 10Gbps or more but may struggle to achieve transfers above 1Gbps over the public internet. Clearly, the bandwidth differential is a significant obstacle to the effective utilisation of computing resources on the public internet.

Even if the bandwidth achievable over the public internet is acceptable, data outflow from CSPs is often a billable item; were the data transfer to be possible over the academic connection it is easy to imagine this cost could be eliminated or at least significantly reduced.

This leads to the second way in which the restrictions on REN use by commercial companies affect the use of CSPs for data-intensive science: the transfer of research data between CSPs is prohibitively expensive. This is an obstacle on a small scale, e.g. for the use of compute facilities at one CSP to access storage resources at another, and also in a more fundamental way as the potential costs associated with moving data from a CSP may prevent an institution from choosing to use a CSP for data storage purposes even if this would be more cost-effective considering only the storage-related costs.



Picture 1: Commercial Cloud providers are well connected to the Internet, but not to the RENs. On the other hand, cloud Customer and Users may have limited connectivity to the Internet

2 See https://www.geant.org/

3 Connectivity options

Putting the problem even more succinctly, if a CSP and an academic customer are not directly connected and cannot be connected via the NRENs in their respective countries, they must find a transit Internet Service Provider (ISP) that can transport the traffic between them.

Three solutions have been proposed to address the connectivity issue.

- Cloud eXchange Points
  provide basic connectivity to CSp and their users (or the NREN of the users); CXPs are essentially open-policy Internet eXchange Points (IXP or OXP) that can be interconnected by RENs with bandwidth sufficient to meet the research traffic loads.
- Cloud Service Virtual Private Networks
  which, in addition to delivering connectivity and bandwidth, are provided by an organiser, such as a REN, that can handle routing relationships between customers and CSPs, thus simplifying the customer and CSP workload.
- On-net VPN termination points
  also simplify the management of network routing for customers and CSPs. Compared to the Cloud Service VPN model, on-net VPN termination points simplify the VPN setup for the CSP at the expense of additional complexity for the provider.

We review these three solutions, highlighting their advantages and drawbacks, in the following sub-sections. It should be noted that we have to consider not only simple "physical" connectivity issues (whether or not data can in principle flow between two points), but also the "logical" connectivity—how the network routing is setup to ensure that the desired data, and only the desired data, does flow. Often, achieving the logical connectivity is the more difficult part given the shared nature of the Internet and the policies that control and limit how traffic is exchanged amongst the different operators of the various domains that the traffic has to cross.

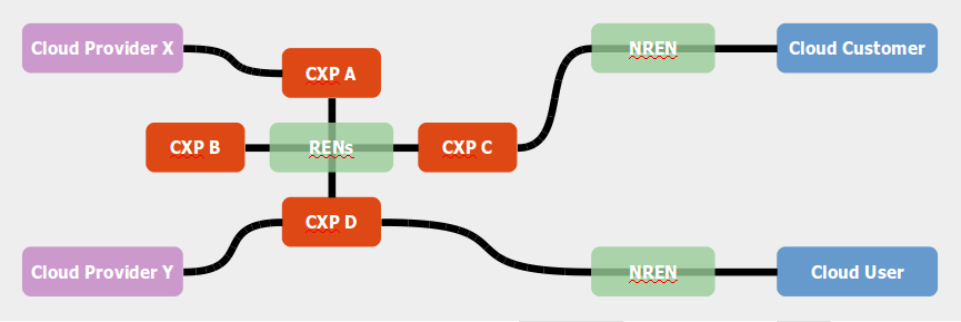
For the sake of completeness, we should mention that there is a fourth solution: having research customers restrict their choice of CSP to a provider that has been selected as a partner or approved provider by their NREN. However, since, firstly, such a restriction of purchasing options may not be acceptable and, secondly, traffic between the customer and the CSP will, effectively, be routed over one of the options above, this solution is not further mentioned here.

3.1 Cloud eXchange Points

In a similar manner to the way in which Internet eXchange Points provide a location to which internet customers can connect and exchange traffic with others, according to their own policies, a Cloud eXchange Point, or CXP, is a facility that provides bare connectivity between CSPs and their customers. In addition, CXPs are connected to other CXPs to provide a distributed network access infrastructure and enough bandwidth to exchange

research traffic. By arranging a connection to a CXP, rather than a "random" IXP, a CSP can be sure of adequate bandwidth to potential research and education clients.

Nevertheless, since many existing Internet eXchange Points (IXPs) today have open policies permitting free interconnection of commercial and academic networks, it is expected that these could act as the required CXPs, simplifying their creation.



Picture 2: CXPs are interconnected by connections provided by RENs. Cloud Customers and Users can connect to a CXP directly or through their NREN

3.1.1 Advantages

- Simplicity, especially as CXPs will develop naturally from existing IXPs
- Reduced costs as compared to having CSPs simply connect to their nearest IXP since concentrating traffic at CXPs minimises the number of high-bandwidth inter-XP trunks that will be needed.
- Routing control is the responsibility of the users of the service, who can thus decide how, and by whom, their bandwidth to the CSP is used.

3.1.2 Drawbacks

- Dedicated inter-CXP trunks will be required and it may be difficult to agree a cost-sharing model for these CXP interconnects. The Cloud Service VPN and On-net VPN termination point options that follow are proposed precisely to allow existing high-capacity networks (such as those provided by GÉANT and ESNET) to, effectively, act as inter-CXP trunks.
- Routing relationships have to be defined between each customer and each CSP. However, such relationships are quick to implement once basic connectivity is in place and if contracts or agreements have already been signed.
- Each CSP/Customer contract requires a peering relationship that has to be configured by the two parties, exposing them to configuration effort and complexity. The effort involved can, however, be reduced if a Route Server service is established by the CXP.
- CXPs and RENs can't easily control the traffic exchanged over the infrastructure they provide

3.1.3 Implementations

The CXP solution has been proposed and is supported by NORDUnet[3]. NORDUnet suggests the use of existing Open Exchange points for the CXP role, such as Netherlight (Amsterdam), CERNlight (Geneva) and Starlight (Chicago).

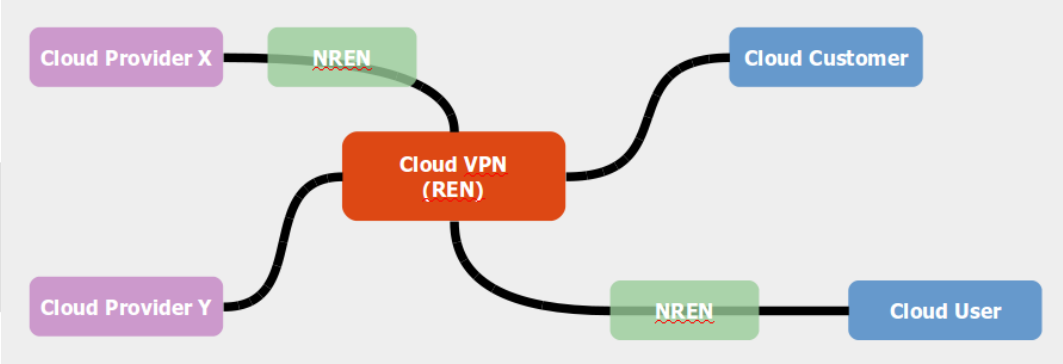
3 http://www.nordu.net/

3.2 Cloud Service VPN

As noted above, the CXP option provides the necessary connectivity but has the drawback that routing relationships have to be defined individually by and between each customer/provider pair, introducing an n^2 level of complexity. This complexity can be addressed through the use of a Layer-3 Virtual Private Network (L3VPN)—a private instance of a routed wide-area network—to interconnect customers and CSPs.[4]

A VPN is normally configured by an individual company or organisation to pass IP packets across an intermediate wide area network. In such circumstances, the company concerned is in control of the networks at either end of the VPN and thus entirely responsible for managing the IP routing. Here, however, the VPN endpoints belong to different entities, and a 3rd party—the organiser of the Cloud Service VPN—takes care of distributing the routing information to the connected entities. Not only does this reduce the complexity involved in the distribution, it also, in the case where a REN is acting as the organiser, allows the REN to ensure compliance with its routing policies, for example preventing the exchange of data between CSPs.

Cloud customers and providers acquire connectivity to a Cloud Service VPN either directly (e.g. if they are already connected to the REN or present in an appropriate IXP) or via an NREN. Once connected, customers and providers establish a routing peering with the Cloud Service VPN organiser in order to exchange the necessary routing information.



Picture 3: Cloud providers, customers and users are interconnected by a continent-wide private network provided by a REN. National RENs provide connectivity from the VPN border to customers/providers

3.2.1 Advantages

- The solution allows a REN acting as a Cloud Service VPN organiser to provision resources dedicated to the Cloud Service VPN whilst keeping control of the traffic transported on their network. In particular, this enables the REN to prevent traffic being exchanged between CSPs.
- Virtual Routing and Forwarding technology (VRF, supported by all modern routers) enables a single router to support multiple virtual routers, just as hypervisors enable the creation of virtual machines. VRF thus enables deployment of a Cloud Service VPN without the need for any additional hardware.

4 Indeed, creation of the LHCONE L3VPN mentioned in footnote 1 was driven by the need to reduce the routing complexity involved in connecting over 140 WLCG sites.

- The Cloud Service VPN organiser can provide tools to enable CSP customers and providers to influence the routing behaviour of the VPN, for example to prefer an existing customer/CSP path rather than using the VPN.
- Different Cloud Service VPNs can be interconnected to allow broader reachability across different domains.

3.2.2 Drawbacks

- As noted just above, a Cloud Service VPN enables a REN to prevent traffic being exchanged between CSPs. Such traffic should be allowed, though, if the exchange is initiated by, or is on behalf of, an academic customer. Ad-hoc routing policies must be negotiated with the Cloud Service VPN organiser should such "third-party" traffic be needed.
- Both CSPs and customers connecting to a Cloud Service VPN must be identifiable Internet "entities" (technically, they must have a public Autonomous System Number). This is unlikely be an issue for CSPs, but some customers—a University Department, for example—may be part of a larger entity and so have to work with third parties to establish the VPN connection.
- Both CSPs and customers must have public IP addresses; this is a consideration addressed by the On-net VPN termination point solution that follows.
- Unless CSPs are able to allocate a clearly identifiable set of resources (and hence IP addresses) to their research customer(s), it is likely to be difficult for CSPs to ensure that only research-related traffic is routed over the VPN.
- Agreeing on a symmetric path through the VPN may be problematic for multi-homed customers and CSPs.

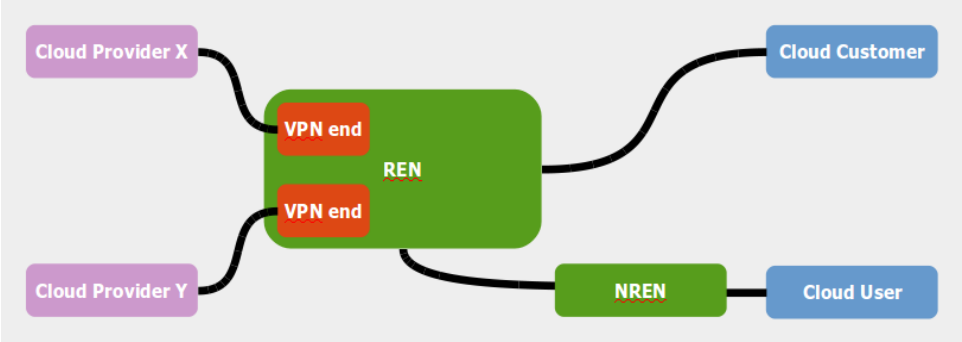
3.2.3 Implementations

After proposing this model, GÉANT has implemented a Cloud Service VPN which, at the time of writing, is being used to provide connectivity between CERN and T-Systems.

3.3 On-net VPN termination points

The third solution that has been proposed also makes use of VPNs. However, instead of creating multiple VPNs between customers and one or more CSPs, this requires each CSP to create a VPN to establish a "point of presence" on the REN network using the credentials of its customer. From this "point of presence" onwards, the CSP traffic is routed normally over the REN network to all the REN customers and peering partners.

To establish such a point of presence, a CSP needs, as previously, to connect to the REN of the customer either directly or through an IXP. The REN provider then allocates a virtual router to terminate the VPN connection and route the CSP traffic into the REN as if it were originated by the customer.



Picture 4: CSPs terminate their VPNs inside the REN network, not at the customer premises

3.3.1 Advantages

- Traffic to other cloud users does not cross the customer's firewall, an advantage compared to classical setups where the VPN endpoint is inside the customer's premises.
- Encapsulation of traffic over a VPN link matches well with standard CSP connectivity options.
- Cloud resources are immediately accessible to all the REN's customers and also to all its peering partners.
- Customers can assign their own IP addresses to their resources at the CSP if the CSP does not have enough public IP addresses and use of Network Address Translation (NAT) is not acceptable or possible.

3.3.2 Drawbacks

- The REN provider will have either to support multiple VPN technologies or to support only a restricted set with the risk of preventing connectivity for certain CSPs.
- Large-scale adoption may lead to scalability problems for the REN's routers given the need to support the encapsulation load.

3.3.3 Implementations

The On-net-VPN model has been proposed by ESnet[5] and tested by them, using the name Collaborative Cloud Services (CCS), to provide connectivity to Amazon Web Services (AWS) for Fermilab. For this trial, ESnet provided public IP addresses to FNAL which were assigned to the AWS cloud resources. Since these IP addresses were already routed in the ESnet network, the resources were immediately reachable by any ESnet customer and by those of Esnet's peering partners.

3.4 Interconnecting the different solutions

Interconnecting these different solutions is not straightforward given their different nature. A Cloud Service VPN is a closed environment, the On-Net-VPN merges cloud services into the shared network of the REN and CXPs provide a fabric rather than a routed network.

5 http://www.es.net/

If a CXP solution could be supported by a route server service that handles the distribution of routes between the interested parties, then coupling with a Cloud Service VPN would be possible. Arranging for a border router of the Cloud Service VPN to peer with the CXP route server would be sufficient to allow the exchange of traffic between the two worlds.

Similarly, as an On-Net-VPN solution makes the cloud resources concerned available to the general-purpose research Internet, it could be possible to exchange only the necessary routes with CXPs and Cloud Service VPNs provided the CSPs IP addresses are in well-defined prefixes. However, it is likely to be difficult to avoid unwanted CSP-to-CSP traffic in such cases.

As noted in above, however, establishment of peering agreements is all that is needed to interconnect Cloud Service VPNs.

4 Experience to date

CERN, as many other organisations, has been actively investigating the feasibility of working with CSPs, notably as part of the HelixNebula partnership[6] that has included a sequence of cloud-service procurements at increasing scale. Not surprisingly, these activities rapidly identified networking connectivity as a significant obstacle to effective use of cloud services. Problems encountered included

- CSPs with off-the-shelf solutions relying on private IP addresses and NAT or VPNs, the use of which set severe limitations on the network performance that could be achieved;
- CSPs with limited flexibility in their Internet connectivity (for example when managed by a third party) leading to difficulties in setting up ad-hoc routing relationships; and
- Cloud resources with public IPv4 addresses scattered across a large address space shared with other customers making it difficult to define the IP prefixes to be exchanged over the routing relationship.

As a particular example, the most recent exercise involved the use of cloud resources provided by T-Systems at their datacentre near Magdeburg in Germany[7]. The initially available paths to CERN were through one of CERN's commercial Internet providers (expensive and limited to 7Gbps) or via a low capacity peering over the GÉANT Internet service.

Fortunately, the T-Systems cloud service could connect to GÉANT's Cloud Service VPN via a connection provided by DFN, the German NREN, and so achieve the required 10Gbps throughput to CERN. It would not have been practical to use the On-net VPN solution for this connection since the VPN offered by T-Systems that would have been used for this had a bandwidth limitation of just 50Mbps.

As could be expected, however, the addresses of the 1,000 virtual machines provided by T-Systems for CERN were scattered across two /16 IPv4 prefixes (i.e. amongst 131,072 other addresses) thus making the Cloud Service VPN accessible to many unwanted machines. To limit the extent of the problem, T-Systems narrowed down the range of machines that

6 http://www.helix-nebula.eu/
7 See http://www.telekom.com/media/enterprise-solutions/307124

could be used by CERN but this required them to announce 16 IP address ranges, increasing complexity yet still leaving many unwanted machines accessible.

To address this, GÉANT tags routes that come from the CSPs and those that come from customers. These two tags are used to implement a policy by which CSPs learn only customers' routes, not other CSPs' routes, thus preventing the exchange of non-research traffic over the GÉANT backbone.

Elsewhere, although we commented above that the On-Net VPN would not have been helpful to connect CERN to T-Systems, this solution has been used very effectively by ESnet in a trial with Amazon Web Services. AWS provide a service called Virtual Private Cloud (VPC) that matches well the On-Net VPN model, allowing the creation of a high-speed IPsec tunnel and the assignment of public IP addresses provided by the customer to the AWS cloud resources.

5 Final Remarks

Unfortunately, we cannot today conclude that any one of the proposed solutions is better than the others. It should be clear from our analysis above that each option has significant drawbacks to balance its advantages and both the advantages and the disadvantages have been observed in practical tests. Indeed, since the options that are practically applicable in individual cases depend strongly on the connectivity options available to the CSP involved, it may be wise for RENs and NREN's to foresee that all three will be required and hence to concentrate efforts to improve interconnectivity between the three options.

Note that, in all three cases, CSPs are required to have a connection to a REN, either directly or through a CXP or an NREN. Since the procurement of such connections often involves significant delays when the underlying physical infrastructure is not in place, CSPs wishing to provide services to the academic community have every interest in establishing such a connection in a permanent way.