

# Designing the RAL Tier-1 Network for HL-LHC and Future data lakes

Alastair Dewhurst<sup>1,\*</sup>, Martin Bly<sup>1</sup>, and James Adams<sup>1</sup>

<sup>1</sup>STFC

**Abstract.** The Rutherford Appleton Laboratory (RAL) runs the UK Tier-1 which supports all the LHC experiments, as well as a growing number of others in HEP, Astronomy and Space Science. In September 2020, RAL was provided with funds to upgrade its network. The Tier-1 not only wants to meet the demands of LHC Run 3, it also wants to ensure that it can take an active role in data lake development and the network data challenges in the preparation for HL-LHC. It was therefore decided to completely rebuild the Tier-1 network with a Spine / Leaf architecture. This paper describes the network requirements and design decision that went into building the new Tier-1 network. It also includes a cost analysis, to understand if the ever increasing network requirements are deliverable in a continued flat cash environment and what limitations or opportunities this may place on future data lakes.

## 1 Introduction

The UK Tier-1 is located at the Rutherford Appleton Laboratory (RAL) in Oxfordshire and supports all LHC experiments, as well as a growing number of others in HEP, Astronomy and Space Science. In total it provides 43PB of usable Disk space, 97PB of Tape Capacity and 595k HS06 of CPU. The Tier-1 also host numerous services such as an FTS instance, CVMFS Stratum 0 and 1 and the GOCDB.

In September 2020, the RAL Tier-1 received £400,000 from UK Research and Innovation's (UKRI), World Class Laboratory fund which is intended for maintaining and refreshing existing UK scientific infrastructure, to ensure the scientific community's ability to carry out exceptional science and to retain the country's prominence in scientific research and output. This money has been invested replacing the RAL Tier-1 network. This paper describes how the new RAL Tier-1 network was designed and implemented. It discusses the cost implications as well as possible implications for data lakes.

## 2 Current Network Infrastructure

The RAL Tier-1 is run by the STFC's Scientific Computing Department (SCD) which runs numerous large scale computing systems to support a range of scientific communities. Traditionally, Tier-1 network usage has been approximately 40% of the total RAL traffic (this may well have changed since the global pandemic changed users working patterns). The following subsection describe the key features of the Tier-1 network.

---

\*e-mail: [alastair.dewhurst@stfc.ac.uk](mailto:alastair.dewhurst@stfc.ac.uk)

## 2.1 Tier-1 Network

The Tier-1 network is currently not significantly different in its design philosophy compared to 2006. It is a single Layer 2 segment containing approximately 1200 hosts. There are three subnets:

- Compute 130.246.216.0/21, which contains all the CPU nodes. This accesses the outside world via the RAL Campus network and traverses the site firewall.
- OPN (130.246.176.0/22), which can bypass the firewall to talk to JANET directly, it can also also talk to CERN via a private router. This contains the storage nodes and gateways.
- Services (130.246.180.0/22), which contains all other Tier-1 services. This accesses the outside world via the RAL Campus network and traverses the site firewall.

The routing to deal with this is a little arcane. There are 3 + 1 physical routers and 7 virtual routers. Nodes have 16 IPv4 and 8 IPv6 routing table entries and there are more default route options than subnets. The current design is not suited for current or planned workflows as the CPU cannot easily access external storage. It should be noted that while Castor storage nodes needed to be externally accessible to allow third party transfers to take place this is not the case for Echo. Echo uses gateway machines for external transfers and the Ceph storage nodes are only accessed directly internally.

## 2.2 External Connectivity

At the start of 2020, there were  $3 \times 10\text{Gb/s}$  links connecting CERN and RAL. It had been planned to upgrade these links to a single  $100\text{Gb/s}$  by April 2020, however due to the global pandemic this took until December 2020 to be completed. This link is known as the LHCOPN (LHC Optical Private Network)[1]. The RAL site also has a pair of  $100\text{Gb/s}$  links to JANET (active/passive). These are each being upgraded to  $200\text{Gb/s}$  during 2021. External traffic to/from the Tier-1 enters via a pair of switches, referred to as the OPNR which connects directly to the site border routers, bypassing the site firewall. The current OPNR needs replacing as it is no longer under warranty and only supports  $40\text{Gb/s}$ . Over the next 5 years, the Tier-1 needs to be able to take full advantage of the external links as well as an upgrade to  $200\text{Gb/s}$  for the LHCOPN.

## 2.3 Site Core

The RAL Tier-1 is connected to the site core, which is used to connect to any other department on site. This includes centrally provided services such as the site DNS. Historically very little data has needed to go over the site core, but this is changing as other departments want to use SCD resources. The Tier-1 is connected to the site core via a pair of switches (active/passive) connected at  $40\text{Gb/s}$ . The site core is being upgraded to  $100\text{Gb/s}$  switching in the next year.

## 2.4 Super Spine

In 2018, a Super Spine was setup to link the various different services run by SCD. The Super Spine is made from 16 x Mellanox SN2700 switches ( $32 \times 100\text{Gb/s}$  ports each). Currently the Super Spine does not provide a default gateway and is purely to enable the movement of large amount of data between SCD Services. The Tier-1 will have a network pod attached to the Super Spine.

### 3 Requirements

It is currently not possible to measure exactly how the network usage is split between the experiments. Non-LHC experiments make up approximately 10% of the usage, although this fraction is expected to grow in future. Requirements will be estimated using the WLCG predictions, although the ability to upgrade network capacity further, beyond the needs of the LHC should be factored in to the design. From the WLCG’s point of view, the RAL Tier-1 provides CPU, Disk and Tape resources as well as a handful of other services. In any future data lake scenario the RAL Tier-1 is likely to take on a larger role as the primary a source of data for jobs running remotely. Internally the services the Tier-1 provides will be slit across multiple network pods attached to the super spine. The Tier-1 network pod will contain the Echo storage service, the HTCondor batch farm and well as all Grid services.

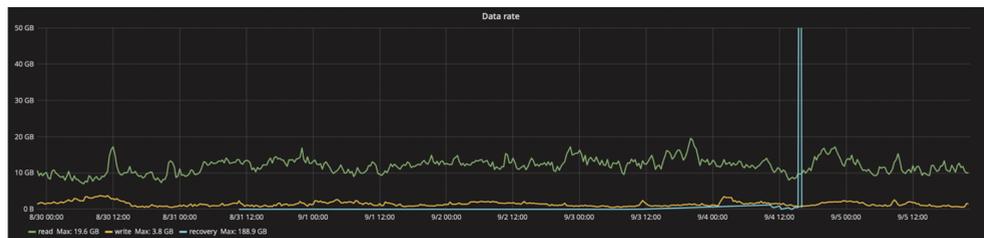
Even though the current network is reliable, the design leads to regular operational problems. For example the FTS service is currently on a different subnet to our storage services. When a VO schedules a transfers via our FTS service to or from either the disk or tape storage, the data is transferred via one route and the control information is sent via another. This can lead to unusual errors. The RAL Tier-1 is also not currently peered with the LHCONE. RAL originally didn’t join as features such as traffic separation weren’t need. Given how ubiquitous a connection to the LHCONE is for larger sites, problems have occur when the other site doesn’t realise the RAL Tier-1 is not connected. As a result of this operational experience, the Tier-1 requires that all externally facing services will be appear equal. e.g. they will have all have access to the LHCOPN, LHCONE and they will all be on dual stack hosts.

#### 3.1 Disk

The RAL disk storage system, known as Echo??, is based on an Erasure Coded Ceph Object Store. It has been designed to provide cost effective high throughput disk storage for the LHC experiment’s data. The cluster has grown each year since 2015 and now contains approximately 70PB of raw storage across 6700 HDD. There are over 220 servers in production.

Data is secured using Erasure Coding with  $k = 8, m = 3$ . The 11 chunks of data are each stored on separate servers. From a network point of view this means that each file that is either written or read will need to go over the network twice. For a write it will initially be written to one storage server which will then send 10 chunks of data to other servers.

Figure 2 shows the throughput in and out of Echo during the first week in September 2020. This is a relatively quiet time as the LHC is not taking data. During previous data taking periods average read rates were 20 – 30GB/s with peaks as high as 50GB/s.



**Figure 1.** Plot showing the shows the throughput in and out of Echo.

There are spikes of rebalancing throughput when hardware is added or removed from the cluster. This can be for operational problems, e.g. we have a few disk fail every week. It is

also not uncommon to remove entire storage nodes from production if they have a hardware problem and return them once they are fixed. This happens once every few months.

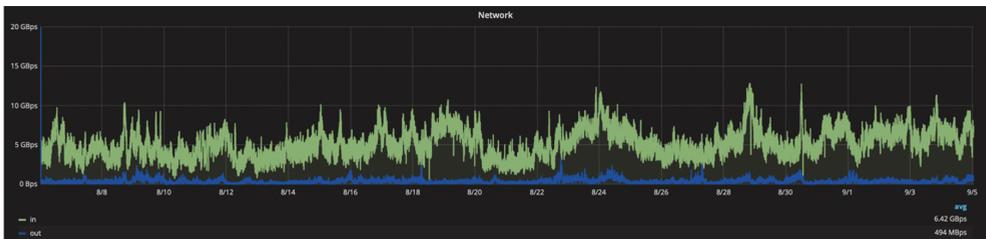
Large scale rebalancing happens when hardware is deployed into Echo for the first time, decommissioned or if we need to change the number of placement groups. These operations could easily saturate our current network and affect production work. We have learnt to control the rate via Ceph config options. These operations do not happen often; around twice a year and are planned.

The hardware of the more recent generations of storage nodes has a 24 HDD and a single 25Gb/s connection. We aim to provide 1Gb/s per HDD. We have observed storage nodes sustain throughput of 20 - 25Gb/s each while the Ceph cluster rebalances data.

### 3.2 CPU

The Tier-1 provides a 42000 logical core HTCondor batch farm. These cores are provided by just over 1000 Worker Nodes. The most recent generation in production, purchased in 2019 from Dell is made up of 96 servers. Each server has 2 x AMD EPYC 7542 CPUs, which provide 128 job slots. The servers have SSD storage and Mellanox ConnectX-4 Dual Port 10/25GbE SFP28 Adapters. This generation provides a total of 12288 job slots which are kept on average over 95% occupied.

Figure 2 shows the total network throughput for the Dell19 servers during August and the start of September 2020. The average network usage is 6.42GB/s reads, which means that each job slot averages 0.52MB/s throughput. Each server is currently utilising 500Mb/s out of their 25Gb/s link.



**Figure 2.** Plot showing the total network throughput of the Dell19 generation of Worker Nodes.

The current throughput is much lower than would be expected when the LHC is taking data, probably by a factor 2 – 4. Core count per CPU is also expected to continue growing over the next 5 years (another factor of 2–4) which will increase the throughput per machine. Assuming the worst case scenario (factor of 16 growth) the average throughput per machine won't exceed 8Gb/s. A 25Gb/s network card will remain acceptable and an oversubscription ratio of 3:1 would be fine.

### 3.3 Tape

The RAL Tier-1 Tape system is known as Castor. SCD currently runs 4 different Tape services. Two different instances of Castor, one for the WLCG experiments and one for the experiments based on the RAL campus. There is also an instance of DMF as well as a service called ADS which was written at RAL and provides backups for various systems (primarily database snapshots). SCD is in the process of replacing Castor with CTA and are aiming to consolidate the all users onto a single CTA instance. In order to do that the entire tape archive

system will be placed in its own network pod. In order to get the EOS nodes (the frontends to CTA) to appear in the Tier-1 subnet, VxLANs will be used between the network pods.

The robotics, tape drives, tape servers and disk buffers are expensive certainly when compared to the cost of the network to connect them. CTA has been designed to make maximum use out of the tape drives. The network for the tape system should therefore ensure that data can be written/read from all tape drives at line speed. The RAL Tier-1 primary uses the IBM enterprise JE Media with TS1160 tape drives. These have a line speed of 400MB/s. The next generation of IBM Tape drives are promising to have a line speed of 1000Mb/s. The RAL Tier-1 has 20 tape drives dedicated to LHC use so the network connectivity to the Tier-1 network should be approximately 200Gb/s.

### 3.4 External Connectivity

The WLCG has provided some estimation for the network requirements between CERN and the Tier-1s[2]. The numbers for RAL can be seen in 1. The estimate is based on a calculation of the total amount of data produced at CERN by each experiment that will need exporting multiplied by the pledged size of each Tier-1. There are also factors to account for network overheads. It was possible to refine this estimate further by making the following changes to the assumptions:

- The WLCG predictions over estimated the fraction of each experiments pledge that RAL provides. The largest discrepancy was the estimate that RAL provides 10% of CMS’s Tier-1 requirements when it is actually currently 6%. This accounts for an 80Gb/s reduction in the minimal network requirement for 2027. The overall reduction in pledge means that under the WLCG model a 500Gb/s network link would be required.
- The WLCG model assumes a factor of 2 to take into account burst of activity and a further factor of 2 to ensure that the network link is not regularly running at capacity. The flexible model is an additional factor of 2 above the minimum model. At RAL, the site link to JANET is in general at least twice the capacity of the LHCOPN link to CERN. If RAL is on the LHCONE it is possible to use this capacity if the LHCOPN is saturated / unavailable. It is therefore reasonable to reduce the factor of 2 that was included to ensure that the network link is not regularly running at capacity to 1.5 (i.e. split the overcapacity requirements between the two links). This reduces the minimum WLCG requirement from 500Gb/s to 375Gb/s.

**Table 1.** Table showing the predicted network bandwidth required between CERN and the RAL Tier-1 in the years running up to the HL-LHC

(Gb/s)	2023	2025	2027 Minimum	2027 Flexible
WLCG prediction	90	180	610	1220
RAL prediction	90	180	375	750
Planned LHCOPN Capacity	100	200	400	800

## 4 Design

The new Tier-1 network design will be based on a 3 tier Clos Network[3] also known as a Spine / Leaf network. This design is used by major cloud providers in their data centres [4]. In a Spine / Leaf network, hosts are connected to leaf switches (which in this case are the top of rack switches). Each Leaf is connected to every spine switch. Leaf switches are

not connected to other Leaf switches and neither are Spine switches connected to each other. This leads to a topology where every leaf is attached to every other via the spine. Equal Cost Multi Pathing (ECMP) is used to share traffic across all the links, allowing the full bandwidth to be used. If one of the spine switches were to be removed from production it would simply result in a reduction in the total available throughput.

One leaf will be configured as an exit router. This will be connected to the site border routers which connect directly to JANET, the site core as well as the LHCOPN. The total external throughput will need to be 400Gb/s, it was therefore decided to have four spine switches.

If it is assumed that 50 Worker Nodes on average will be attached to a leaf switch, there will need to be around 20 of them. If it is assumed that 16 storage nodes will be attached to a leaf switch, there are likely to be 15 of them. A few more switches are likely to be needed for additional services (databases, VMWare infrastructure), which means the spine will need to support at least 40 leaf switches. The spine switches choose each have 64Gb/s ports, which will allow up to 60 leaf switches if 4 ports are used to connect to the super spine. Figure 3 show the design of the Tier-1 network.

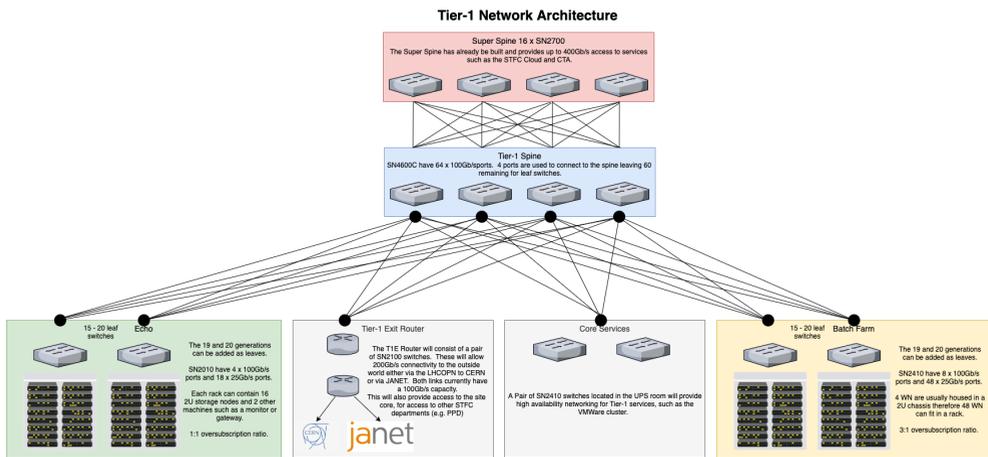


Figure 3. Figuring showing the design of the RAL Tier-1 Network.

The new network equipment arrived in December 2020 and was due for installation in January 2021. Unfortunately due to the global pandemic, a national lockdown was implemented in the UK at the start of January, this delayed the installation by approximately two months. The new network is scheduled to be installed towards the end of March and beginning of April 2021. Configuration, testing and deployment is expected during May and June 2021.

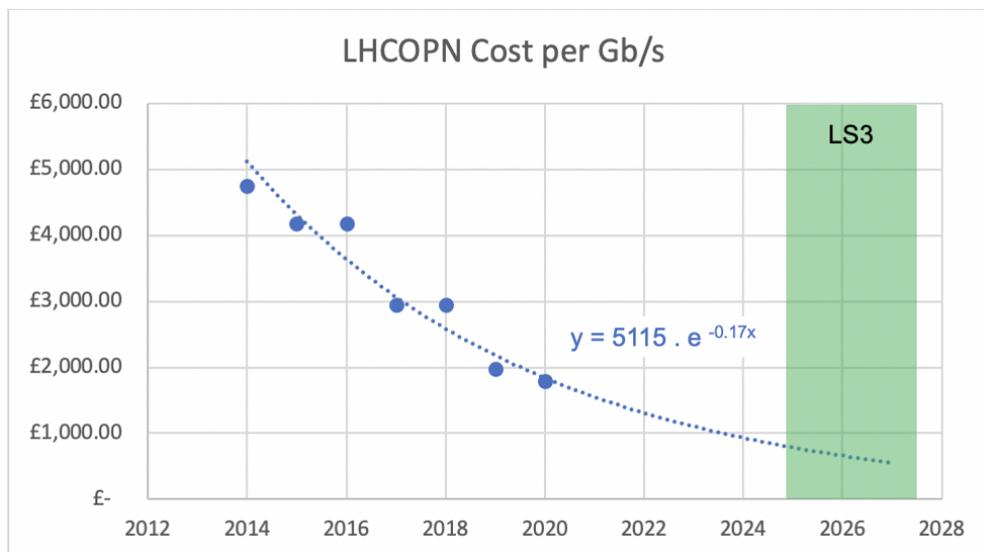
### 4.1 Future Upgrades

During Run 3 it is expected that RAL will upgrade the LHCOPN link from 100Gb/s to 200Gb/s. In the current planned configuration, the Tier-1 Exit Router should be able to handle 400Gb/s. The hardware all comes with a 5 year warranty and given the resilient design can be expected to stay in production for up to 6 years. It would therefore be replaced in 2026 - 2027 towards the end of LS3. At this time the Spine switches as well as the Tier-1 Exit Router would be replaced with their 400Gb/s equivalents, which should be significantly

cheaper than they are now and very much standard technology. Given that ECMP will be used, it should be possible to easily replace each spine switch in turn without impacting production work.

## 5 Cost Analysis

When designing any computing system it is important to take into account cost. For several years now the WLCG have been working on the assumption that they will get flat cash. Figure 4 shows the maintenance cost per Gb/s for the LHCOPN connection between RAL and CERN. Normally the cost is negotiated every two years, which is why the cost doesn't drop every year. If an exponential curve is fitted to these points the cost per Gb/s halves every 4.08 years. This should mean that RAL is able to upgrade to a 200Gb/s link in 2024, towards the end of Run 3 and a 400Gb/s by 2028. It should be possible to move this upgrade forward slightly to before HL-LHC starts.



**Figure 4.** Plot showing the cost of the LHCOPN Link from RAL to CERN per Gb/s

In addition to the ongoing costs, there is also the upfront costs to install the network. The cost to install the 100Gb/s OPN link was £240,000. The cost to install the spin/leaf network was £374,000. If a more traditional "fat tree" network design had been followed it would have cost about £80,000 less. This setup would have been slightly less resilient and had half the internal bandwidth. If costs are considered over the 5 year lifetime of the hardware, the Tier-1 spends approximately 3 times more to maintain the external connection to CERN compared to the internal network. The internal network is effectively infinite (the devices will hit their line speed before the network is saturated), while the external connectivity will just about keep up with the expected growth.

## 6 Conclusion

The RAL Tier-1 has embarked on an ambitious project to completely rebuild its network to provide the throughput and features required for Run 3 and prototype data lakes. With

suitable upgrades the network design should also meet the needs of HL-LHC. To keep within the cost projections data lakes must not add a significant amount to the transfers required between sites. The most data intensive jobs should continue to be sent to the site that have the data. For remote access, even if sites are well connected with low latency (e.g. RAL Tier-1 and Oxford Tier-2), caching of data that is used multiple times at sites would still be beneficial.

## References

- [1] E Martelli and S Stancu, “LHCOPN and LHCONE: Status and Future Evolution.“, *Journal of Physics: Conference Series*. 664. 052025, doi: <https://doi.org/10.1088/1742-6596/664/5/052025> (2015)
- [2] *HL-LHC network needs and data transfer challenges*, [https://indico.cern.ch/event/983436/contributions/4226012/attachments/2213578/3746893/HL-LHC\\_network\\_challenges\\_-\\_V2.pdf](https://indico.cern.ch/event/983436/contributions/4226012/attachments/2213578/3746893/HL-LHC_network_challenges_-_V2.pdf) (2021), accessed: 2021-06-21
- [3] *Clos Networks: What's Old Is New Again*, <https://www.networkworld.com/article/2226122/clos-networks-what-s-old-is-new-again.html> (2021), accessed: 2021-02-28
- [4] *The next generation Facebook data center network*, <https://engineering.fb.com/2014/11/14/production-engineering/introducing-data-center-fabric-the-next-generation-facebook-data-center-network/> (2014), accessed: 2021-02-28
- [5] A Dewhurst et al 2017, “The deployment of a large scale object store at the RAL Tier-1.“, *Journal of Physics: Conference Series*. 898 062051, doi: <https://doi.org/10.1088/1742-6596/898/6/062051> (2017)