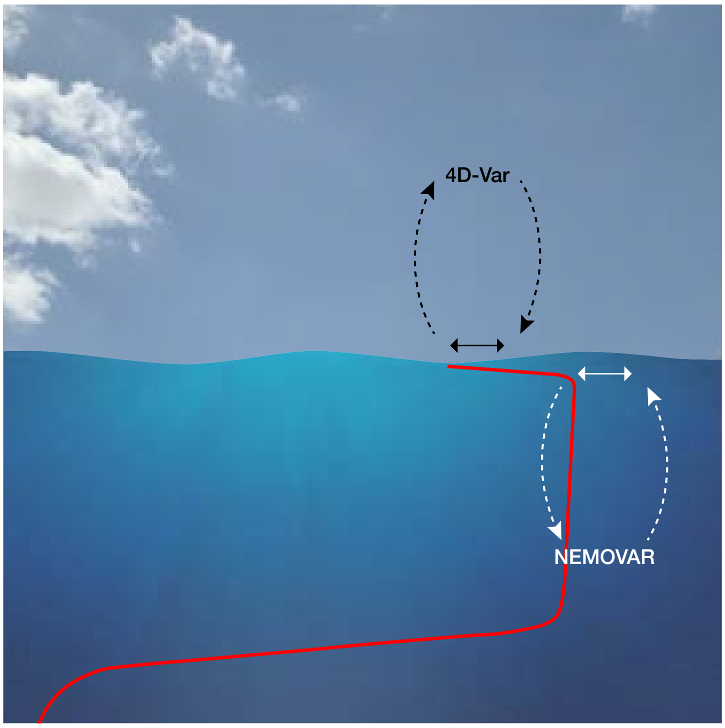
ECMWF Feature article

from Newsletter Number 172 – Summer 2022

COMPUTING

ECMWF's new network and security infrastructure



www.ecmwf.int/en/about/media-centre/media-resources

doi: 10.21957/abc631j7s2

This article appeared in the Computing section of ECMWF Newsletter No. 172 – Summer 2022, pp. 35–40.

ECMWF's new network and security infrastructure

Ahmed Benallegue, Stanislav Burlakov, George Margaritis, Hassan El Ghouizy, Michele Di Mascolo

In September 2021 ECMWF officially opened a new data centre in Bologna (Italy), while its data centre in Reading (UK) is due to close at the end of 2022. As part of the Bologna Our New Datacentre (BOND) programme of activities, the main aim of the Network and Security (N&S) workstream was to deliver the most reliable and performant access to ECMWF's information and communications technology (ICT) infrastructures to ECMWF Member and Co-operating States, other services and applications, and end users. These infrastructures include the high-performance computing facility (HPCF), the Data Handling System (DHS), the Climate Data Store (CDS) and the forthcoming operational European Weather Cloud.

Whilst most of the BOND activities relate to the design, procurement and deployment of a new N&S infrastructure in Bologna, substantial complementary activities also had to be undertaken at the Reading data centre (DC) site to transform the operational network infrastructure. The aim was to introduce the necessary improvements for a smooth and effective interconnection of the DCs in Reading and Bologna through 100 Gbps site-to-site connectivity as well as to prepare for the migration of services and data from Reading to Bologna.

This article details the origins of the N&S infrastructure design and describes the various necessary transformational activities that took place at the Reading DC site.

Design drivers

The first step towards the creation of the N&S design was to ensure compliance with ECMWF's ICT design principles through the definition of more specific N&S design drivers. The table below shows the translation of the design principles to clearer network-level drivers that were followed when producing the N&S design components.

Generic ECMWF design principles	Specific network and security design drivers
Start with user needs	Automation and self service
Strong development and operations Work on working practices	Physical layout considerations Remote access and operations
Focus on the core business	Address the needs of ECMWF's core businesses Focus on ECMWF's new HPCF in Bologna
Defence in depth	Network segmentation Security multi-layer approach Security ecosystem
Fail in place and fix later Be consistent (but not uniform)	Create a future-proof, scalable and flexible design Holistic design approach Reliability and resiliency Fault isolation and simplicity Modularity Reliable and manageable scalability
Minimise technical debt	Follow industry best practice
Enterprise in the cloud	Secure and reliable access to cloud-based services

TABLE 1 Translating generic computing principles to network-level drivers was essential to guarantee a network and security design and its resulting operational infrastructure that are both innovative and future-proof. For instance, the "Work on working practices" principle translates to "Remote access and operations", meaning remote access to the network equipment's management interfaces, and the "Defence in depth" principle is better understood as "Network segmentation" and "Security ecosystem" when it refers to IT networks and network security, respectively.

doi: 10.21957/abc631j7s2

High-level functional design

ECMWF’s Technical Design Authority (TDA) was set up at the beginning of the BOND programme to manage the overarching technical governance for ECMWF in the context of BOND. One of the first actions of the TDA was to draw up a target high-level functional design that summarises the criticality of and the expectations from the ICT infrastructures and services provided by ECMWF in Bologna (see Figure 1).

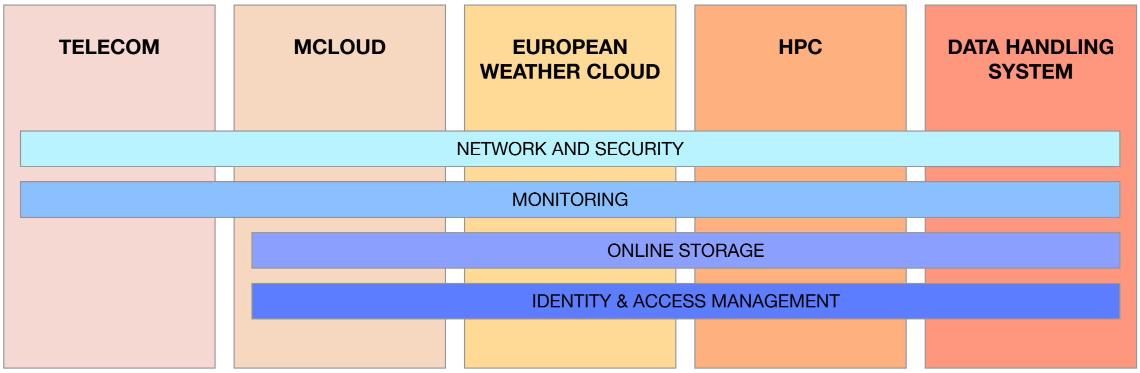


FIGURE 1 The functional design includes the following columns: TELECOM, which includes the Internet and the Regional Meteorological Data Communication Network (RMDCN); MCLOUD, which includes the virtualisation platforms needed to run ECMWF services; the EUROPEAN WEATHER CLOUD, which offers a cloud service to Member States and commercial users; HPC, which integrates the general-purpose batch and interactive cluster; and the DATA HANDLING SYSTEM, which includes the common online storage.

Inspired by and compliant with the TDA’s design, a more specific N&S high-level functional design was put together. The major evolution compared with the Reading DC was the consideration of the two data halls as two separate data centres in which two independent instances of N&S infrastructure are deployed (see Figure 2).

Network and security design overview

With the N&S high-level functional design in mind, the N&S High-Level Design (HLD) activity was initiated. Preliminary workshops were held between the Networks and Security Team and ECMWF’s network equipment suppliers and vendors, with the aim of producing a design that was as standard as possible and that follows industry best practice. In parallel, multiple brainstorming and workshop sessions took place involving all the relevant teams in ECMWF to capture all the requirements for the infrastructures and services to be deployed in or migrated to Bologna.

The result of this substantial piece of work was an HLD document that was submitted to the TDA for peer review and approval. This was achieved in October 2018. This document was consequently used as reference during the next phase of the design, involving the definition of Low-Level Designs (LLDs) for specific infrastructure and services, and the procurement and deployment phases of the N&S workstream.

In summary, the resulting N&S design has a segregated data centre approach, in which two data halls (DHs) in Bologna are considered as two separate data centres as much as applicable. The setup is a simple and scalable solution of two independent data centres that can coexist in any location. This gives ECMWF the flexibility to easily deploy a new infrastructure at any other location. For smaller sites, the design stays the same, but the resulting infrastructure consists of fewer devices.

The main architectural evolutions introduced by the new design in Bologna, which will be detailed in the next paragraphs, are as follows:

- IP Fabric architecture
- Dual-site topology
- Security layer
- Branch office architecture

doi: 10.21957/abc631j7s2

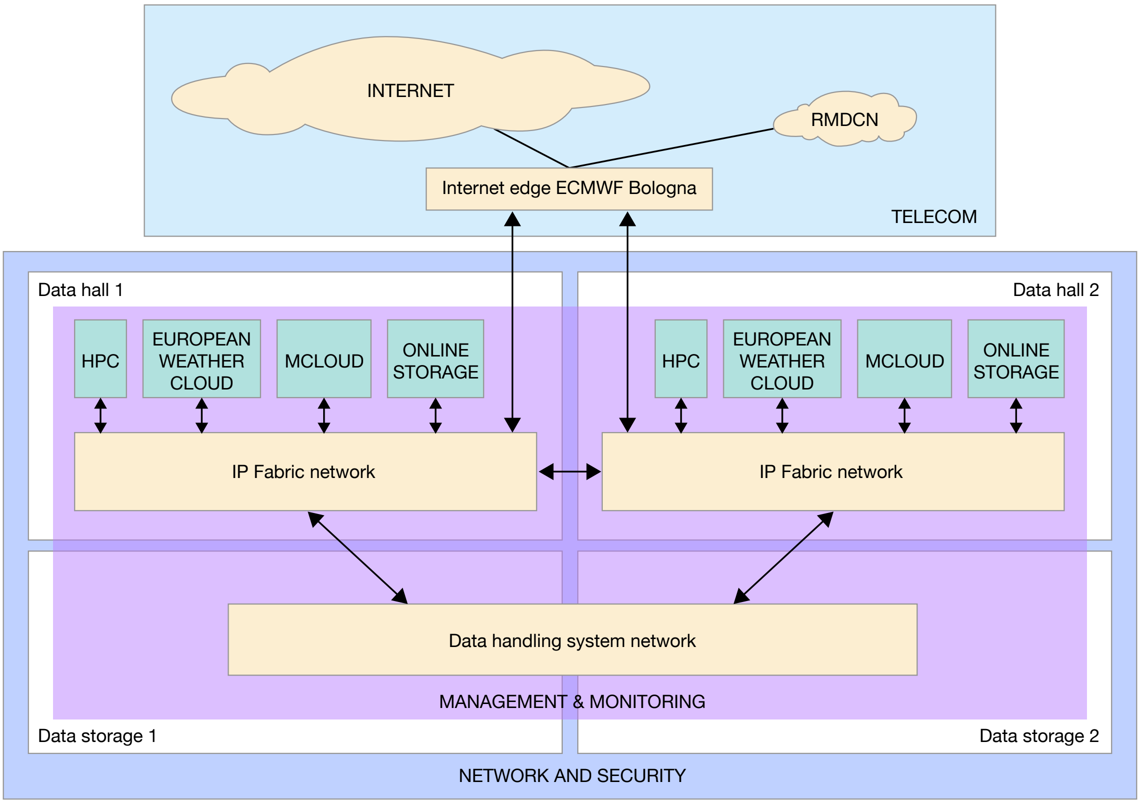


FIGURE 2 Apart from the data handling system network, for which there is no expectation of immediate support, all the other components of the N&S infrastructure are deemed critical for operations and require 24/7 immediate support. The Regional Meteorological Data Communication Network (RMDCN) is a highly available private data network interconnecting all of ECMWF’s Member and Co-operating States used primarily to exchange critical meteorological data.

IP Fabric architecture

An ‘IP Fabric’ architecture, also referred to as ‘Clos’ or ‘Leaf and Spine’, is a state-of-the-art network topology for medium and large-scale data centres. It is a two-layer topology composed of leaf switches and spine switches. Different from the traditional three-layer topology, it minimises latency and bottlenecks whilst offering greater scalability, reliability and performance. It also improves the total available bandwidth, simplifies network configuration and facilitates management and monitoring (see Figure 3).

Dual-site topology

Two physically segregated IP Fabric networks were deployed in the new data centre: one in each data hall. The two data halls are regarded as two separate data centres, which creates two separate fault domains.

Security layer

The ‘Defence in Depth’, one of ECMWF’s ICT design principles, is an information assurance concept based on the application of a multi-layered defence approach. It is intended to ensure the confidentiality, integrity, and availability of information assets. Defence in Depth entails multiple layers of protection so that, if one defensive measure fails, there are more behind it to continue protecting the assets. The aim of such an approach is to use security controls to reduce the risk and effects of vulnerabilities, attacks and intrusions. In the context of the BOND N&S design work, this resulted in the construction of a security ecosystem and network segmentation.

doi: 10.21957/abc631j7s2

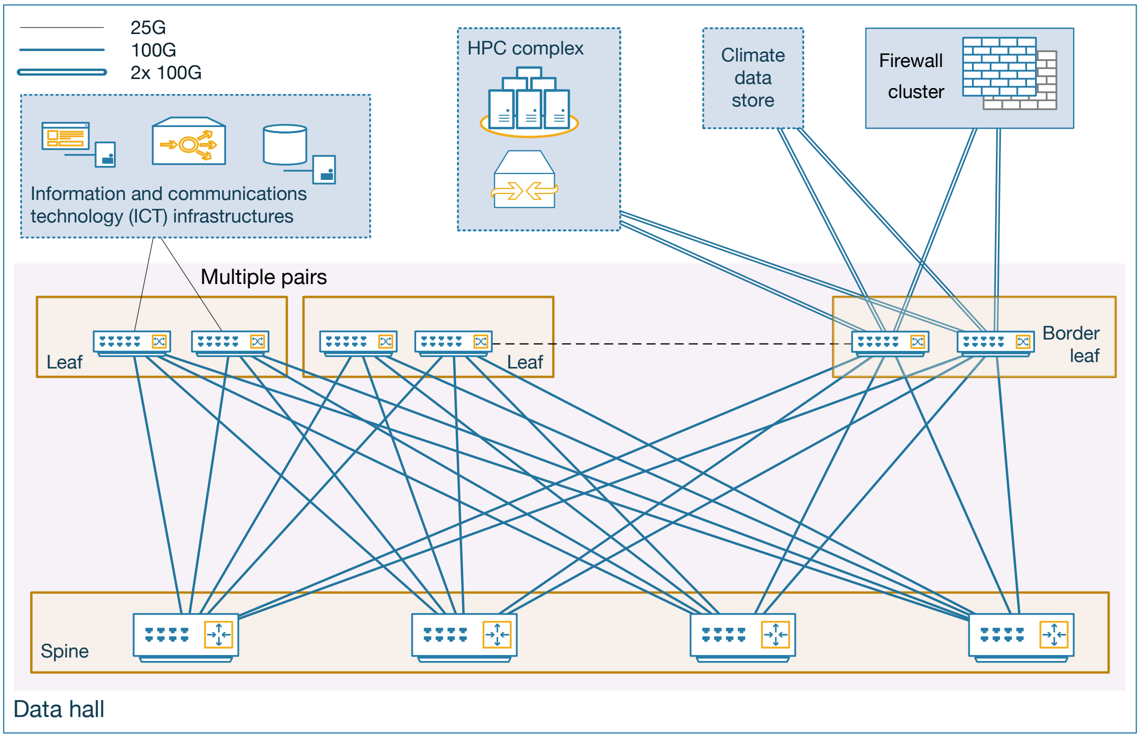


Figure 3 In an IP Fabric network, the total capacity is restricted by the number of ports available in the spine switches, which makes it very scalable. Furthermore, as long as a single spine switch is fully operational, the capacity of the network is degraded but uninterrupted.

In particular, the segmentation of the data centre network into different security zones offers higher control and visibility of the traffic. In addition, it makes it possible to introduce new security defence controls to improve operational security. It thus boosts ECMWF's ability to prevent and react to internal and external threats. See Figure 4 for a representation of key security layers in Bologna.

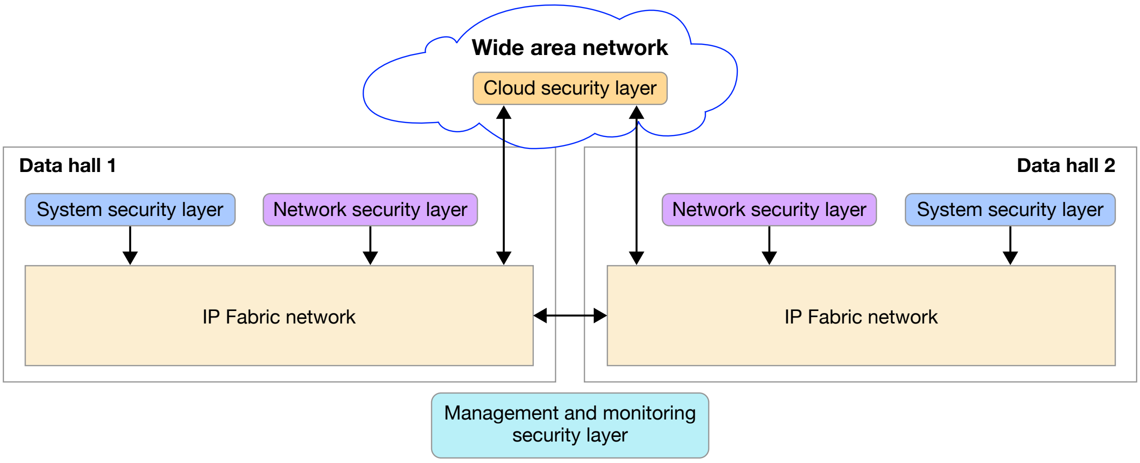


Figure 4 The Bologna data centre includes various security layers, which are complementary.

doi: 10.21957/abc631j7s2

Network segmentation: the case of the CDS

A practical use of network segmentation is illustrated by the deployment of the CDS infrastructure. As this infrastructure is managed by a third-party partner, CloudFerro, it is important to properly control traffic to and from it. This way the CDS, and other Bologna ICT infrastructures it exchanges data with, can be protected from external and internal network-level threats. This is done through the configuration of network segments and security zones, through which routing and packet filtering controls can be applied.

Branch office architecture

The new Bonn and Bologna offices network architecture is based on a branch office architecture with local survivability and total independence from the data centre network. In case of a complete data centre isolation, the branch office network can provide basic Internet connectivity and services to staff and visiting guests. The same architecture and its associated topology will be applied to the Reading offices network during the decommissioning phase of the Reading DC. For an overview, see Figure 5.

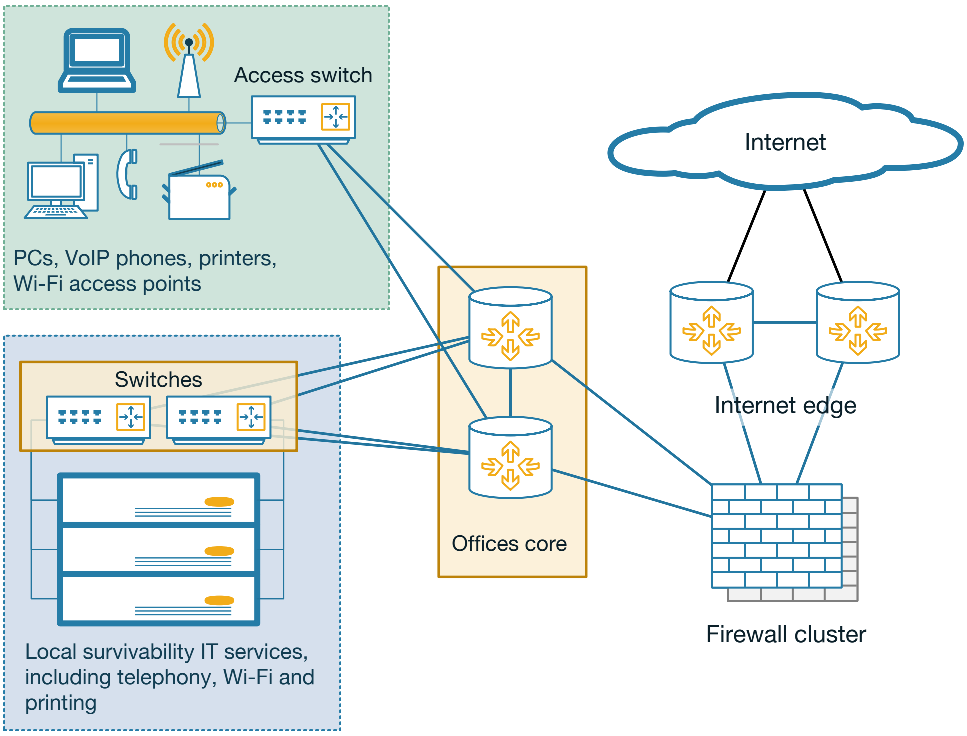

Access switch
PCs, VoIP phones, printers, Wi-Fi access points
Switches
Local survivability IT services, including telephony, Wi-Fi and printing
Offices core
Internet
Internet edge
Firewall cluster

Figure 5 The offices network blueprint approach intends to harmonise and simplify the configuration and management of the resulting infrastructures wherever it is used, in Reading, Bologna or Bonn. It provides the same user experience for staff and visiting guests in terms of accessing key services, such as Wi-Fi eduroam or printing.

The transformation of the Reading network

To ensure an effective implementation of site-to-site connectivity, as well as to guarantee minimum disruption during the migration of services and data from Reading to Bologna, it was important to plan for the transformation of the operational N&S infrastructure in Reading. This will be referred to as the Reading Local Area Network (LAN) or the High-Performance Network (HPN) in the rest of the article.

The Reading LAN in June 2018

The Reading LAN has evolved organically over many generations of supercomputers and ever-changing ECMWF needs. Since it is constantly used for operational data transfers, hardware refreshes had to be limited in scope to minimise risk and impact on operations. Over the years, this resulted in a network topology that was suboptimal, with manually configured devices that relied on static routing. While redundancy was always a key aspect in ECMWF network designs, the static nature of the network meant that any failover procedures were generally manual and labour intensive. It was clear that the network as it stood would not be able to support the fast rate of change and dynamic adjustments that would be necessary during the period of migration from Reading to Bologna.

doi: 10.21957/abc631j7s2

The Reading LAN transformation plan

The Reading LAN transformation took place during the period March 2019 to December 2020. It aimed to achieve the following high-level goals:

- Definition and enforcement of a new IP addressing scheme for all ECMWF sites
- General simplification of deployed network services
- Removal of inter-service dependencies, where possible
- Definition and adoption of standardised configuration and naming conventions
- Harmonisation and clean-up of key elements of the infrastructures and services
- Rationalisation of routing and introduction of dynamic routing protocols
- Increased network traffic segregation and adoption of a zone-based firewalling approach
- Improvements to the network's high availability and self-healing capabilities
- Collapse of the firewalling capabilities into a single cluster.

With these high-level goals in mind, an analysis of the network and security solutions implemented in Reading was carried out, concentrating on current service design and any apparent deficiencies. Designs for equivalent services to be implemented in Bologna were then considered, together with any known service migration requirements. Since the number of possible enhancements was quite high, it was clear that only a certain proportion of them could be implemented within the timeframe. Therefore, in the next stage of the process, a list of tasks was compiled together with their impact ratings and perceived implementation difficulty, while taking note of any inter-task dependencies. This produced a final list of tasks to be carried out within the scope of the Reading LAN transformation.

The Reading LAN transformation – making the change

Due to project length, complexity, and the need for continued use of the network for operational traffic, the process of transformation was gradual with many intermediate designs having to be implemented before the final design goals could be achieved. Several platform limitations were discovered, which necessitated some scope changes. However, they resulted in a network infrastructure that is more resilient and capable of supporting rapid adjustments reliably.

One of the core changes enabling network agility was the replacement of static routing with standards-based dynamic routing protocols, namely Border Gateway Protocol (BGP) and Open Short Path First (OSPF). With static routing, making a change to the path of traffic required either a manual intervention or a complex configuration tracking the state and the reachability of interfaces. With dynamic routing protocols, all participating routers exchange messages to check the link between them, monitor network routes they are responsible for, and verify various network path attributes. This means that the network can detect the failure of a router and automatically reconfigure itself to route the traffic along an alternative path. The speed at which this is done depends on protocol and its configuration, but supporting protocols, such as Bidirectional Forwarding Detection (BFD), enable reconfiguration times of under one second. Furthermore, in cases where multiple paths of equal quality are present between devices, all of them could be used simultaneously for routing traffic, meaning that a failure would result in a decrease in network capacity but no user-noticeable impact.

Increased traffic segregation was achieved by implementing Virtual Routing and Forwarding instances (VRFs), which are used to separate production traffic from other traffic types. This resulted in an improvement to ECMWF's security posture and enabled risk-free proof-of-concept work required for both BOND designs and LAN transformation work itself.

In 2020 it became clear that some LAN capacity issues would have to be addressed in Reading, prior to the migration of services to Bologna. For example, a simplification of the HPN firewall infrastructure by collapsing three firewall clusters into one was supplemented with an additional requirement for expanded firewall capacity. While this change was very complex due to its sheer size as well as the amount of configuration refactoring required, it was eased by recent implementations of dynamic routing and network segmentation. A new cluster was built and tested inside a separate VRF, and then the routing path was switched over from using the old clusters to the new cluster, during a maintenance window. Traffic interruption was kept to under a minute, and the new cluster doubled the LAN capacity. This can be doubled once again, should it be required in the future. It is notable that this major piece of work was carried out remotely, with only a small group of supporting technicians present on site.

Alongside these major changes, many smaller but no less significant pieces of work were carried out to improve network services and facilitate fast-approaching migration activities. These included the commissioning of site-to-site data connectivity between Reading and Bologna; and a large amount of IP renumbering activities to homogenise usage of our IP space, so that large contiguous blocks of IP space were available for use at ECMWF's new sites in Bologna and Bonn.

The Reading LAN in April 2022

In April 2022, the LAN in Reading is radically different to what it was four years before. While it is by no means as modern and flexible a design as the one deployed in Bologna, the activities carried out as part of LAN transformation were, where appropriate, inspired by designs for the Bologna infrastructure. The result is that the network can support rapid change as part of day-to-date operations, thereby easing migration of services from Reading to Bologna. It has also shown its ability to tolerate failure and self-repair under failure conditions. Crucially, the transformed network can be effectively managed remotely, much in the same way as is possible for Bologna.

Conclusion

The importance and priority given to the design phase of the new N&S infrastructure for Bologna at the start of the BOND Programme proved to be key to producing an innovative and future-proof architecture that will be fit-for-purpose for ECMWF's needs for years to come. However, this was only achieved through an iterative process involving all key ECMWF stakeholders and a validation/approval mechanism required from key bodies such as the BOND Programme board, the Computing Department management team and the TDA. In particular, the gathering of detailed requirements was a necessary step that must never be shortcut or rushed as it is the only way to ultimately ensure the appropriateness of the design. In addition, using external third parties, such as suppliers, vendors, and partners, as a sounding board was extremely beneficial as it gave the resulting design a stamp of approval from industry and allowed for a smoother transition from the design to the procurement phases.

The transformation of the Reading LAN proved to be extremely challenging for two main reasons: the size of the task to make the network BOND-compatible, and the fact that this was done in a 24/7 operational environment. This was further compounded by the need to increase capacity and performing most of the work remotely because of the COVID-19 pandemic. However, the task was successfully completed thanks to a careful and methodical approach.

Although the bulk of the deployment of the N&S infrastructure in Bologna has been completed, the configuration of the network is still on-going to accommodate the deployment and migration of ICT infrastructures and services. However, the benefits of the new design and its resulting implementation are already visible since making incremental changes and fulfilling change requests is much easier than doing the same exercise at the Reading DC, especially before the LAN transformation. The relatively pain-free successful migration of the CDS infrastructure in the first quarter of 2022 is testimony to that.

Further reading

Benallegue, **A.**, 2020: ECMWF's new IT network and security infrastructure in Bologna, *ECMWF Newsletter* **No. 162**, 12–13.

Benallegue, **A.**, **S. Burlakov** & **L. Sorth**, 2021: Testing the Reading–Bologna site-to-site connectivity, *ECMWF Newsletter* **No. 168**, 4–5.

Dell'Acqua, **M.**, **J. Thomas**, **M. Toni** & **A. Gundry**, 2022: ECMWF's new data centre in Italy, *ECMWF Newsletter* **No. 170**, 23–25.

European Centre for Medium-Range Weather Forecasts, Shinfield Park, Reading, RG2 9AX, UK

The information within this publication is given in good faith and considered to be true, but ECMWF accepts no liability for error or omission or for loss or damage arising from its use.

doi: 10.21957/abc631j7s2