Arista Selected for ECMWF Critical and Extensible Network Infrastructure

Highlights

Challenge

The European Centre for Medium-Range Weather Forecasts needed to ensure its network layer was not a bottleneck for advanced research and weather prediction on upgraded supercomputer clusters.

Solutions

- Arista 7500-E Family Switches
- Arista EOS®

Results

- Arista Multi-Chassis Link Aggregation, with Virtual ARP, enables highest levels of resiliency for 24/7 operation
- Flexible Arista EOS allows custom on-switch scripting to support supercomputer network node failover
- Multi-process state sharing architecture allows incremental software updates without affecting the state of the system for improved uptime
- Arista network infrastructure provides high performance and low latency access to the largest meteorological data archive in the world

The European Centre for Medium-Range Weather Forecasts selects Arista for critical and extensible network infrastructure for one of Europe's most advanced supercomputing facilities

ECMWF

The European Centre for Medium-Range Weather Forecasts is an independent intergovernmental organisation supported by 34 member states and a world leader in global medium-range numerical weather prediction. As part of an ongoing programme to expand scientific and technical research to improve forecasting, ECMWF continually upgrades its critical supercomputing facility to meet advances in technology. Through switching to Arista, ECMWF has removed potential bottlenecks within its underlying network architecture and delivered unprecedented levels of performance while benefitting from Arista EOS (Extensible Operating System) to improve integration and resiliency between its supercomputing clusters.

Project Background

The European Centre for Medium-Range Weather Forecasts (ECMWF) was established in 1975 and now employs around 300 staff from more than 30 countries. The Centre is both a research institute and a 24/7 operational service, producing and disseminating numerical weather predictions and data to the national meteorological services in its Member and Co-operating States.

ECMWF's core mission is to produce numerical weather forecasts and monitor the Earth-system, which includes scientific and technical research to improve forecast skill and maintain an archive of meteorological data. To deliver this core mission, the Centre provides twice-daily global numerical weather forecasts plus a host of other data useful for scientific and health applications such as air quality analysis, atmospheric composition, and climate monitoring. Headquartered in Reading, in the UK, ECMWF provides advanced training to scientific staff in Member and Co-operating States and assists the World Meteorological Organization with its programmes. The Centre also offers a catalogue of forecast data that can be purchased by commercial customers worldwide.

At the heart of much of the predictions and research are powerful supercomputers which are a necessity when running weather forecast models to a schedule with limited time slots. ECMWF has been using supercomputers since 1977, operating one of the largest supercomputer facilities of its type in Europe together with the largest meteorological data archive in the world.

The upgrade cycle of its core supercomputer takes place on a regular basis to meet the Centre's strategic objectives. This process is often the catalyst for an upgrade to the underlying data centre network. As Oliver Gorwits, Head of Networks and Computer Security for ECMWF explains, "the supercomputer generates a huge amount of data which we need to offload to storage and post-processing as quickly as possible so that it can generate more data – the network cannot be a bottleneck for this process and we need to ensure that the network layer is able to cope with predicted upgrades or new workloads." In 2010, with the Centre's existing IBM based supercomputers due for upgrade, the team began looking at a possible replacement for its current Force 10 based network.

Solution

The selection process was handled as an open tender, and saw bids from all the main players in the networking space, as Oliver continues, "we went through a rigorous scoring method based on the features, performance and roadmap of the competing vendors and Arista scored highly, offering some compelling additional benefits."

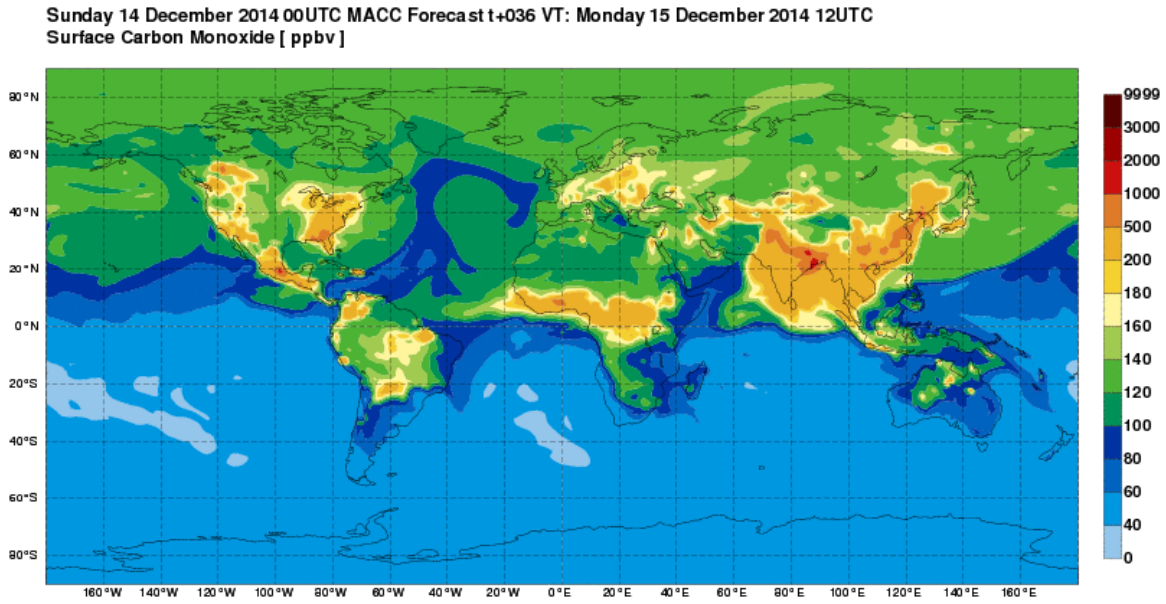
One of these was Arista EOS - a multi-process architecture that separates networking state from the processing itself. This enables fault recovery and incremental software updates on a fine-grain process basis without affecting the state of the system, as well as security patches behind the scenes. In addition, protocol processing, management functions, and even device drivers run in user address space, not in the kernel itself. This greatly increases the stability of the kernel, which is a standard Linux kernel, making it safe to extend the operating system with additional functionality.

Since the selection of Arista in 2010 ECMWF has transitioned through three major supercomputer cluster upgrades, in 2011, 2014 and most recently in 2016. Whilst the 2011 installation was almost a like for like change, in 2012 ECMWF deployed a custom EOS extension that really unlocked the full potential of the Arista switches and powerful operating systems design.

The supercomputer system is not without its idiosyncrasies, one of which is how it manages transparent failover and recovery. "This was an area in which we utilised the power of the extensible Arista EOS," explains Oliver, whose team develop custom scripts for its core Arista 7508 switches. "We could have deployed another hardware element in between the supercomputer and the network or run routing software on the supercomputers as we had done in the past, but neither is an optimal or desirable method. The extensibility features allow us to provide transparent failover to supercomputer network nodes without compromising on any design, latency or operational processes."

Again through an open tender, the IBM supercomputers were replaced with new Cray XC30 supercomputers which went live in September 2014, comprising two self-sufficient clusters with their own storage, but with equal access to the high performance working storage of the other cluster. This cross-connection allows most of the benefits of having one very large system but adds significantly to the resiliency of the system, allowing flexibility in performing maintenance and upgrades. Along with this two Arista 7504-E switches were deployed, interconnected via 40-Gigabit Ethernet based aggregated links, to provide connectivity between the supercomputers, tape archive, and Linux batch farms. The 16 supercomputer network nodes are each connected via dual 40-Gigabit Ethernet and all other systems via dual 10-Gigabit Ethernet.

Most recently in 2016 the two supercomputers were upgraded to Cray's XC40 model, bringing next-generation Intel Xeon processors for a combined peak performance of 15.7 petaflops across 260,000 processing cores in two clusters.



Conclusion

Oliver describes the deployment of Arista as "one of the most reliable network technologies that we have ever deployed at ECMWF," but as a cutting edge facility, there have been a few bumps. "The architecture is designed to cope well with unanticipated issues and we have worked with a fantastically responsive team at Arista to do things which would normally be incredibly challenging with monolithic network operating system architecture – and crucially, while delivering the highest levels of performance and availability for the network."

With more use of virtualisation within its server farms, ECMWF is now examining some of the orchestration and virtualisation features with EOS to improve operation flexibility. "It's fair to say that this has been a very successful network transformation for ECMWF, built on a strong foundation of Arista for our future evolution," Oliver concludes.

Santa Clara—Corporate Headquarters
5453 Great America Parkway,
Santa Clara, CA 95054

Phone: +1-408-547-5500
Fax: +1-408-538-8920
Email: info@arista.com

Ireland—International Headquarters
3130 Atlantic Avenue
Westpark Business Campus
Shannon, Co. Clare
Ireland

Vancouver—R&D Office
9200 Glenlyon Pkwy, Unit 300
Burnaby, British Columbia
Canada V5J 5J8

San Francisco—R&D and Sales Office 1390
Market Street, Suite 800
San Francisco, CA 94102

India—R&D Office
Global Tech Park, Tower A & B, 11th Floor
Marathahalli Outer Ring Road
Devarabeesanahalli Village, Varthur Hobli
Bangalore, India 560103

Singapore—APAC Administrative Office
9 Temasek Boulevard
#29-01, Suntec Tower Two
Singapore 038989

Nashua—R&D Office
10 Tara Boulevard
Nashua, NH 03062