CEA's Military Applications Division selects Arista Networks for major 400Gb/s network upgrade to power next generation of HPC and digital simulation breakthroughs.

Highlights

Challenge

CEA's Military Applications Division, a major European centre for research and innovation, selected Arista Networks to create an 400Gb/s Ethernet solution to overcome HPC limitations between compute partitions and global storage system posed by its legacy InfiniBand infrastructure.

Solution

- Dedicated Leaf and Spine architecture based on Arista DCS-7358X4 and Arista DCS-7060X5 switches
- Arista EOS® and CloudVision® for workload orchestration and workflow automation with enhanced network monitoring and visibility
- 400Gb/s RDMA (Remote Direct Memory Access) over Converged Ethernet workloads

Results

- Improved data throughput and latency with enhanced workflow flexibility
- Proven roadmap from 400Gb/s to 800Gb/s within a more open supplier ecosystem
- EVPN-VxLAN mixed workloads without causing user disruption with better segmentation and Quality of Service
- A rich variety of tools for administration, automation and supervision that have been previously absent from InfiniBand

As one of Europe's largest research institutions, CEA is continually enhancing its High-Performance Computing footprint to deliver breakthrough discoveries. With its InfiniBand based High Performance Network becoming a bottleneck for expansion, CEA's Military Applications Division selected Arista as the core for its new 400Gb/s Ethernet based Leaf and Spine network architecture that delivers enhanced throughput between compute partitions and central storage systems, improved latency, better quality and services, and workload flexibility.

Company Background

CEA is a key player in research, development, and innovation. For over 75 years, CEA has been tasked with guiding public decisions and providing the scientific and technical means that civil society (businesses and local authorities) needs to better manage major societal changes, such as the renewable energy transition, digital transformation, future healthcare, defense and global security. Its mission is supported by 21,000 employees and 9 research centers equipped with major research facilities that provide an innovative environment conducive to academic and industrial partnerships in France, Europe and abroad. For the previous 10 years, CEA has made the Top 100 Global Innovators list, a list that recognises organisations that have demonstrated above-the-bar innovation performance (Clarivate ranking). It is also the leading French organisation that registers the most patent applications in Europe according to the European Patent Office (EPO) 2021.

Across all its projects, CEA employs a vast array of cutting-edge technology, for example, CEA's Military Applications Division has "EXA1", a 36 petaflop scalar supercomputer developed in conjunction with Atos that in 2021 was ranked 14th in the world's 500 most powerful machines.

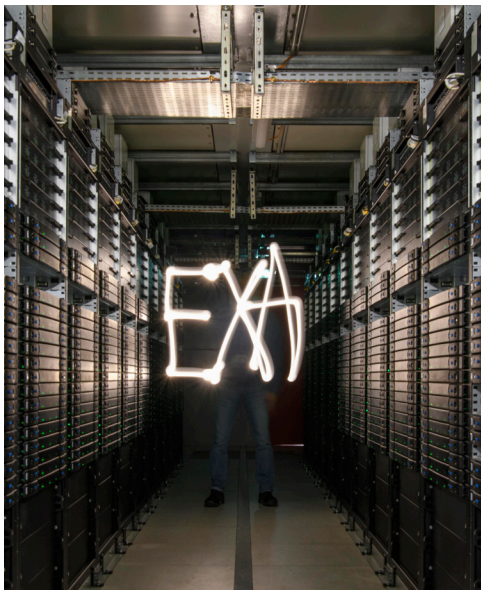
In June 2023, The EuroHPC European Joint Undertaking announced it had selected the project for the future European Exascale supercomputer. This will be led in France by the Jules Verne consortium, which brings together France – represented by the Grand Equipement National de Calcul Intensif (GENCI) as the hosting entity in collaboration with CEA as the hosting site – and the Netherlands, represented by SURF, the Dutch national supercomputing centre.

These supercomputers are used to run advanced simulation and analysis projects that range from simulated nuclear tests, to helping to develop new technologies for health, such as large-scale proteomics and genomics, medical imaging, and structural biology.

Challenge

Much of CEA's high performance computing (HPC) has been developed using a modular and data-centric approach where multiple storage partitions and services are served to several compute partitions via a High-Performance Network (HPN). Traditionally, InfiniBand was the preferred choice for this HPN as it provides high throughput and low latency through RDMA (Remote Direct Memory Access) primitives. However, InfiniBand has some significant technological limitations. Critically, InfiniBand switches have failed to go beyond the 200Gb/s threshold, leading to potential obsolescence in the face of the continued evolution of Ethernet.

In addition, the ability to run RDMA over Converged Ethernet (RoCEv2) offers a compelling technical alternative to InfiniBand along with a much more diverse community of silicon, switching and Network Interface Cards (NIC) suppliers. 400Gb/s Ethernet switches are already deployed at scale with 800Gb/s now reaching the market and 1.6 Tb/s Ethernet in development.



Solution

Around the launch of 400G, in 2020 CEA's Military Applications Division began a major project to benchmark Ethernet with RoCEv2 as an alternative to InfiniBand for building its underlying network to connect storage resources to compute partitions. The evaluation, including proof-of-concept testing, took one year and proved successful so that in 2022, they were able to begin a formal tender to move towards an RoCEv2 based architecture.

CEA's Military Applications Division looked at multiple vendors and carried out extensive performance, feature and reliability testing ahead of selection. The criticality of its work is of the utmost importance to the French State. Based on the results of this evaluation phase, the organisation selected Arista Networks for its new Ethernet based core network.

Arista Networks' equipment was tested with RoCEv2, especially for Lustre filesystem RDMA payloads and CEA's Military Applications Division validated that its Ethernet offering matched the throughput of InfiniBand to effectively utilise the storage infrastructure.

CEA's Military Applications Division's High Performance Network, being used to federate storage partitions, services modules, compute partitions and outer world connectivity, must also handle a variety of traffic alongside Lustre traffic. In testing, Arista handled these EVPN-VxLAN mixed workloads successfully without causing disruption to users' access while rich segmentation and Quality of Service (QoS) features allowed for better control and isolation of flows from different Lustre file systems and standard services over a single fabric.

Another critical benefit offered by the Arista based solution is deep support for Data Center Quantized Congestion Notification (DCQCN). Arista's Extensible Operating System (EOS®) provides an end-to-end congestion control scheme using a combination of Priority Flow Control (PFC) and Explicit Congestion Notification (ECN) to support RDMA over Ethernet.

Arista EOS also takes real time traffic utilization of the network links into account and balances flows uniformly across them. The result is less congestion in the network, fewer ECN marked packets, fewer pause frames, and higher aggregate throughput across nodes resulting in shorter completion times for workloads.

With tools like QoS classification, scheduling and adjustable buffer allocation schemes, Arista EOS allows CEA's Military Applications Division to gain a more granular control of the network so it can tailor network performance to meet the requirements of its diverse range of applications.

Conclusion

CEA's Military Applications Division is deploying a dedicated Leaf and Spine architecture based on Arista DCS-7358X4 and Arista DCS-7060X5 switches, for its central network connecting compute partitions and global storage system. This network uses the principles of Arista Universal Cloud Network design that allows the network to scale to multiple racks while keeping the latency predictive and low. This approach uses Arista EOS' intelligent load-balancing capabilities to ensure traffic flows are uniformly distributed across the Leaf and Spine links to avoid flow collisions.

The Arista 7358X4 series is used as Leafs and is built on a single 12.8 Tb/s high-capacity packet processor in an extremely compact and power efficient 4RU modular form factor that simplifies network design while providing flexibility to deploy a range of interface types. Line rate performance with up to 128 ports of 100G or 32 ports of 400G is combined with broad support for enterprise grade features including EVPN-VxLAN for multi-tenancy.

The Arista 7060X5 series, used as Spines, offers high density 400G and 800G in a compact form factor that deliver the highest performance of 25.6 Tb/s combined with layer 2/3 forwarding. These switches can also be configured as Leafs, or fixed configuration Spine deployment, depending on the future needs of the organisation.

CEA's Military Applications Division has also adopted Arista CloudVision®, a network-wide approach for workload orchestration and workflow automation, as a turnkey solution for Cloud Networking. Arista CloudVision extends the Arista EOS publish/subscribe architectural approach across the network for state, topology, monitoring and visibility. This enables CEA's Military Applications Division to move to cloud-class automation and leverage tools such as Ansible without needing any significant internal development.

Although the deployment is still ongoing, early analysis has proven the benefits initially identified in 2020 to have been correct. The move to Ethernet has met the performance and latency criterion while also adding a rich variety of tools for administration, automation and supervision that have been previously absent from InfiniBand, along with the inherent flexibility of being able to run EVPN-VxLAN mixed workloads.

Perhaps most importantly, CEA's Military Applications Division now has a high-performance network with 400G capability which is 800G ready, that can keep pace with the next generation of computing center architecture.

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