Success Story
Poznan Supercomputing and Networking Center Improves I/O Performance to Prepare for Petascale Research Computing

KEY HIGHLIGHTS

Industry
Research

The Challenge
Keep pace with the growing demand for compute power and data storage as Poland’s high-performance computing (HPC) infrastructure takes on more international scientific research projects.

The Solution
Address HPC needs by deploying NetApp® E-Series storage and NetApp EF-Series flash array.

Benefits
• Double capacity for research computing to accommodate new R&D projects
• Improve performance for researchers by serving metadata 10 times faster with SSD
• Simplify storage management with unified interface for HDD and SSD
• Support cloud and virtualization initiatives by offering hybrid storage for VMs

Customer Profile
Supporting academic and scientific research
Formed in 1993, the Poznan Supercomputing and Networking Center (PSNC) works to integrate and develop the e-Infrastructure for the academic community in Poland. Affiliated with the Institute of Bioorganic Chemistry at the Polish Academy of Sciences, PSNC is the leader in implementing innovative technologies for the Polish National Research and Education Network (PIONIER). PSNC is multifunctional, providing computing power and archive systems for national and international scientific research projects and initiatives while also supporting Internet and network services in Poland.

The Challenge
Removing barriers to growth
Many of the high-performance computing (HPC) projects that PSNC supports are now funded by the European Union with the goal of creating a pan-European infrastructure for academic and research computing. The increasingly important role that PSNC plays in international research projects is driving unprecedented growth for the organization. Data has been growing by 2TB every day, and PSNC plans to grow its computing power to petaflop scale in the near future.

As data grew, users began to experience latency in their compute jobs. Backups and other IOPS-intensive administrative activities began to impact production traffic, and parallel compute jobs competed for resources. A full scan of the namespace took several hours to complete, making it difficult to locate files and perform maintenance.

"With the huge number of files our users upload and the increasing complexity of their data structures, we were fast approaching the limits of our previous storage system," explains Maciej Brzeźniak, storage systems specialist at PSNC. "To take the next step in our evolution as a world-class HPC facility, we needed more storage capacity for HPC and better I/O performance to serve file system metadata quickly, or we would have hit a wall."

PSNC needed to enable researchers to store large files in their home directories while also processing hundreds of millions of small files with fast I/O to keep performance high. "We deployed two different storage systems to separate the bulk data traffic from the metadata
traffic, but we wanted a single administrative interface to simplify storage management,” says Brzeźniak.

The organization also needed storage that could efficiently support virtualization and cloud initiatives as its infrastructure evolves to encompass more than 100 physical servers and thousands of virtual machines.

The Solution
Supporting mixed HPC workloads with flash
Following a public tender, PSNC decided to deploy a NetApp E-Series storage system for home directory data and a NetApp EF-Series flash array for home directory metadata. Both would be connected by Fibre Channel to a cluster of Linux®-based servers running IBM General Parallel File System (GPFS).

Another NetApp E-Series system stores data for a separate HPC cluster using an InfiniBand interconnect.

PSNC now has a modular solution to which it can add SSD or HDD as needed to scale for performance, capacity, or both. “It was a good experience collaborating with NetApp to arrive at a storage solution that meets our HPC needs for both capacity and performance,” says Brzeźniak. “With the NetApp EF-Series flash array we can use SSD for workloads that benefit from higher IOPS, while achieving maximum storage density with the NetApp E-Series.”

NetApp SANtricity® Storage Manager provides an intuitive management interface for both E-Series and EF-Series arrays. “The ability to use the same management interface for both our NetApp E-Series storage and EF-Series flash array reduces complexity, which is a huge advantage,” says Brzeźniak.

The NetApp E-Series systems also store VMware® ESXi®, Kernel Virtual Machine (KVM), and Microsoft® Hyper-V® virtual machine images. This capability gives PSNC the capacity and performance it needs to virtualize more physical systems to accelerate testing and development and improve data center efficiency.

Business Benefits
Improving performance for researchers
The ability to serve file system metadata quickly has accelerated research computing at PSNC. A full scan of the namespace no longer has any noticeable impact on production traffic.

“We’ve improved performance for researchers by serving metadata 10 times faster using the NetApp EF-Series flash array,” says Brzeźniak. “We’re getting enough IOPS to eliminate any noticeable latency, which means we can keep researchers productive and engaged in their work.”

Doubling capacity for research computing
PSNC now has enough storage capacity to accommodate new HPC projects, and when the time comes to add computing power at the petalop scale, storage will not be a bottleneck to results. “Adding compute power is useless if storage I/O is slow, or if researchers cannot store their results,” says Brzeźniak. “With the NetApp E-Series and EF-Series flash array, we’ve more than doubled our capacity for research computing.”

By making efficient use of flash storage, PSNC meets performance requirements while conserving power and cooling resources. “To get equivalent performance with spinning drives, we would have needed two more racks of space and a lot more power,” says Brzeźniak. “Having the compact NetApp EF-Series flash array delivering those IOPS is a great thing. We’ll be able to continue serving metadata quickly well into the future without adding more SSD.”

Ready to drive future discoveries
In addition to providing HPC capability, NetApp E-Series systems support cloud initiatives under development at PSNC—including storage as a service with OpenStack—as PSNC continues to expand its service offerings to the international research community.

“We partner with NetApp because the company understands the public sector, has a commitment to HPC, and offers integration with emerging technologies such as OpenStack,” says Brzeźniak. “We can reduce our TCO while accommodating more research projects for our partners in government and academia.”

SOLUTION COMPONENTS

NetApp Products
NetApp E5500 and E5400 storage systems
NetApp EF540 flash array
NetApp SANtricity Storage Manager

Environment
Applications: research-specific solvers and academic applications
Operating systems: Red Hat Enterprise Linux, CentOS, and Ubuntu Server
File System: IBM GPFS
Virtualization: VMware, KVM, and Microsoft Hyper-V
Cloud computing: OpenStack

Protocols
FC-SAN
InfiniBand

Leading organizations worldwide count on NetApp for software, systems and services to manage and store their data. Customers value our teamwork, expertise and passion for helping them succeed now and into the future.

www.netapp.com

© 2015 NetApp, Inc. All rights reserved. No portions of this document may be reproduced without prior written consent of NetApp, Inc. Specifications are subject to change without notice. NetApp, the NetApp logo, and SANtricity are trademarks or registered trademarks of NetApp, Inc. in the United States and/or other countries. Linux is a registered trademark of Linus Torvalds. VMware and VMware ESX are registered trademarks of VMware, Inc. Microsoft and Hyper-V are registered trademarks of Microsoft Corporation. All other brands or products are trademarks or registered trademarks of their respective holders and should be treated as such. CSS-6802-0415