White Paper

NetApp Engineering Takes to the Cloud

Sanjay Gulabani and Tony Leger, NetApp
August 2020 | WP-7330-0820

Abstract

Data-driven organizations undergoing digital transformation know that public cloud services play an essential role in accelerating DevOps and in optimizing IT operations. Discover how NetApp Engineering's data fabric and NetApp® cloud services in AWS, Microsoft Azure, and Google Cloud help our teams get new services to market quickly while saving millions of dollars on-premises and in the cloud.

NetApp®

TABLE OF CONTENTS

LIST OF FIGURES

ACKNOWLEDGMENTS

The authors of this paper would like to thank the following contributors: Matt Dinsmore, Brad Flanary, Pat Knight, Pete Lanzone, Al Lawlis, Ramanjaneyalu M., Fred Peiffer, Anand Prahlad, and Nagaraju S.

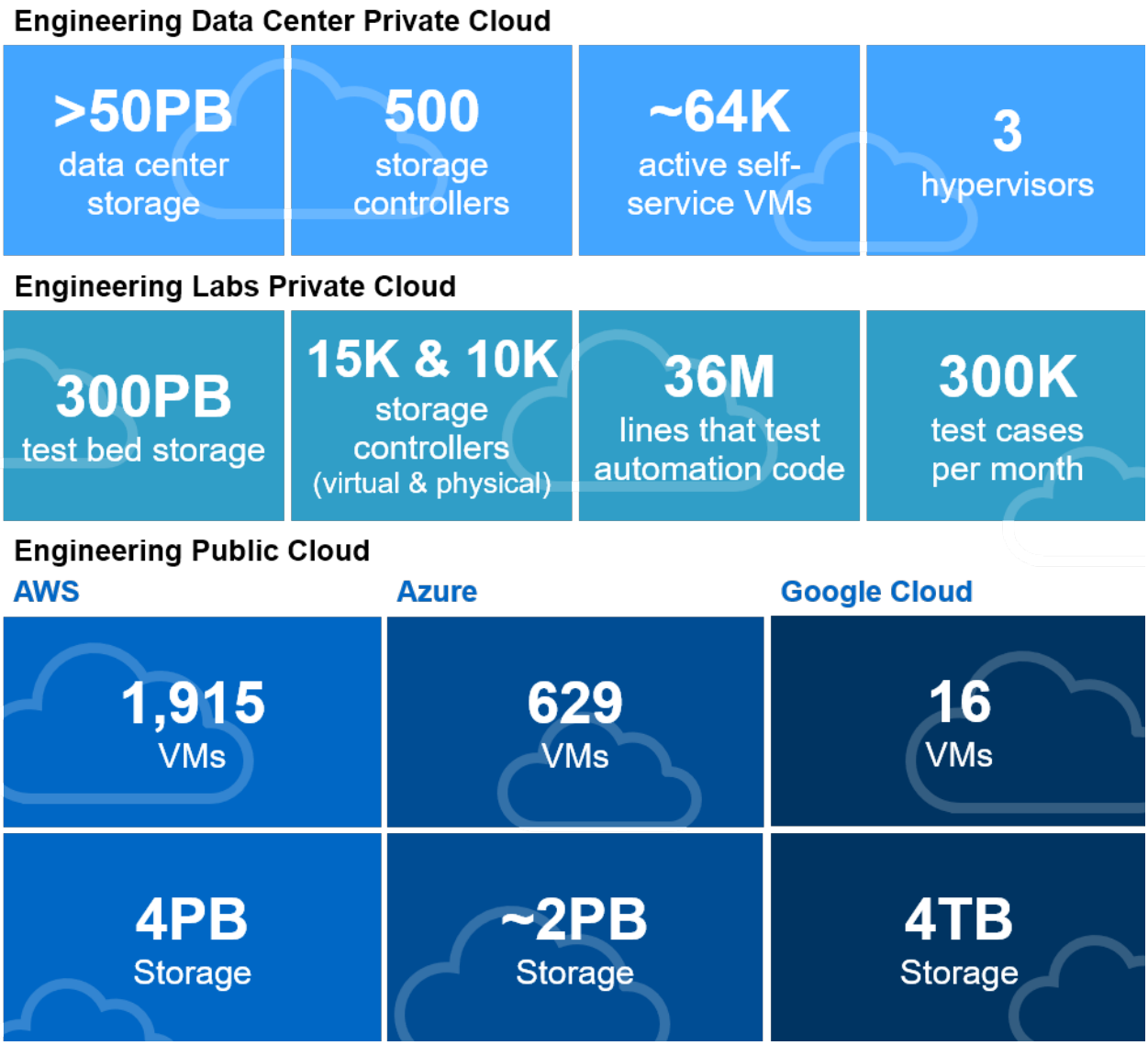
1 Introduction

NetApp's own digital transformation efforts and cloud-based services development have increased significantly over the past several years. Like many organizations, to get new products and services to market faster—while controlling costs—we have been shifting our product strategies, IT environment, and operations to the cloud. This paper describes the shared infrastructure and the NetApp data fabric that the NetApp Engineering team uses to enable secure and cost-effective development (DevOps). It also explains how we accelerate development and use innovation in the cloud with NetApp® cloud services that are hosted by AWS, Microsoft Azure, and Google Cloud.

1.1 NetApp Engineering: A Proving Ground for New Technologies

NetApp Engineering is Customer Zero: a proving ground for new technologies and a showcase for their value. Our Customer Zero team supports 3,500 engineers across the NetApp hybrid multicloud portfolio. The numbers in Figure 1 illustrate the complexity of our engineering environment.

Figure 1) The Global Engineering data fabric integrates our private and public clouds.

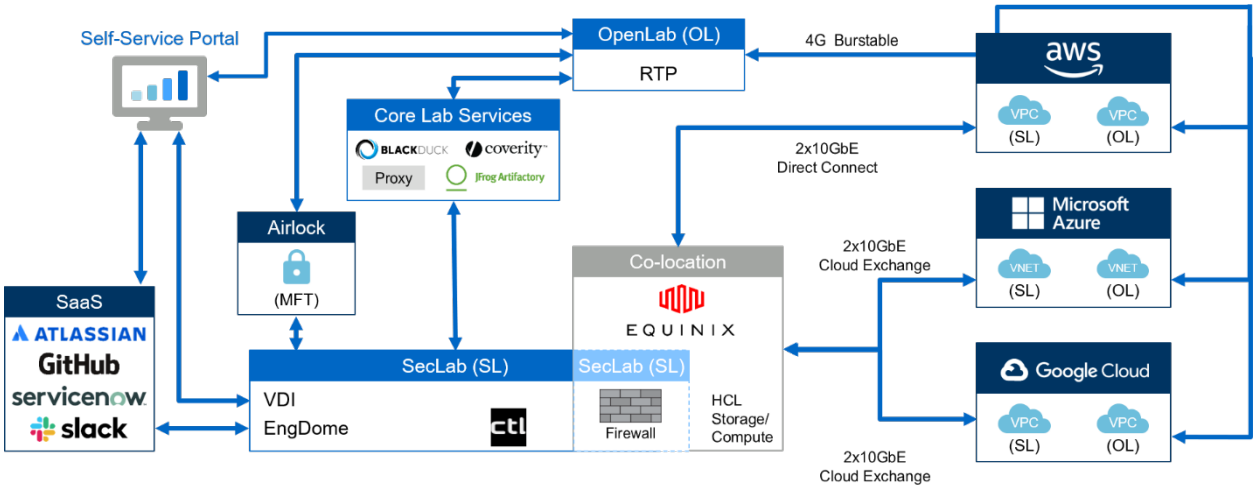


2 A Shared Infrastructure for Secure and Cost-Effective Development

To support taking product development to the next level, our first step was to create a shared infrastructure that would enable secure, reliable access to all engineering resources. Our resulting hybrid multicloud infrastructure is based on a data fabric that allows us to streamline development while providing predictable costs and performance.

All our engineering services are available through a self-service portal and through APIs that enable automation. Self-service is an important factor in our ability to accelerate development. The hybrid cloud lab portal (Figure 2) automatically performs key management functions and enables self-service by engineers. Engineers can create accounts, reset passwords, and manage resources that are allocated to them. They can perform multifactor authentication (MFA) tests, prevent automatic shutdowns, or shut down resources as necessary.

Figure 2) NetApp Engineering's hybrid multicloud infrastructure provides secure access to all resources.



2.1 Providing Secure Access to All Resources

Development of intellectual property takes place in the SecLab. It is firewalled and proxied from our open corporate network—no public internet routing is enabled for any device in that environment. To disallow any public internet access from our secure cloud accounts, we use policies on Virtual Private Cloud (VPC) in AWS and Google Cloud. We use similar policies on Azure Virtual Networks (VNets).

An Airlock managed file transfer (MFT) connects the SecLab to our OpenLab to support development where access to the internet is required. Services that are common to our secure and open environments are hosted by our Core Lab Services. Equinix and dedicated network connections extend these networks to the three hyperscalers.

Public internet traffic is routed back from cloud accounts by way of Equinix through firewalls and proxies that are hosted on premises and at Equinix. We connect securely to all three hyperscalers by using two 10GbE Cloud Exchange interconnects. We use encryption for data at rest in the public cloud. To encrypt data during transport from remote sites, we use dedicated networks through Equinix Cloud Exchange or through Cisco Intelligent WAN (IWAN) connections.

2.2 Governing the Use of Cloud Services

The success of our business depends on managing the cost of services in the cloud. Our Engineering and Cloud Operations team is responsible for managing up to 600 individual accounts across AWS, Microsoft Azure, and Google Cloud. Because each cloud provider limits the number of accounts that can

be paid for by purchase order, we have three accounts at each cloud provider. Those three accounts make it possible to roll up charges for NetApp Engineering, revenue-generating products hosted in the cloud, and NetApp IT. To manage costs across the clouds, we rely on a proprietary tool that automates watchdog governance, tagging, MFA enforcement, and weekend test bed shutdown, as well as other functions.

3 Accelerate Development and Leverage Innovation in the Cloud

With our shared infrastructure in place, we are integrating NetApp cloud services that are hosted by AWS, Microsoft Azure, and Google Cloud into our environment. Our cloud integration efforts that are described in this paper have focused on three challenges: monitor and optimize resources, simplify data protection, and deliver high-performance file services. Read on to learn how we use NetApp Cloud Insights, Cloud Volumes ONTAP®, and Cloud Volumes Service solutions to meet these challenges.

3.1 Monitor and Optimize Resources: NetApp Cloud Insights

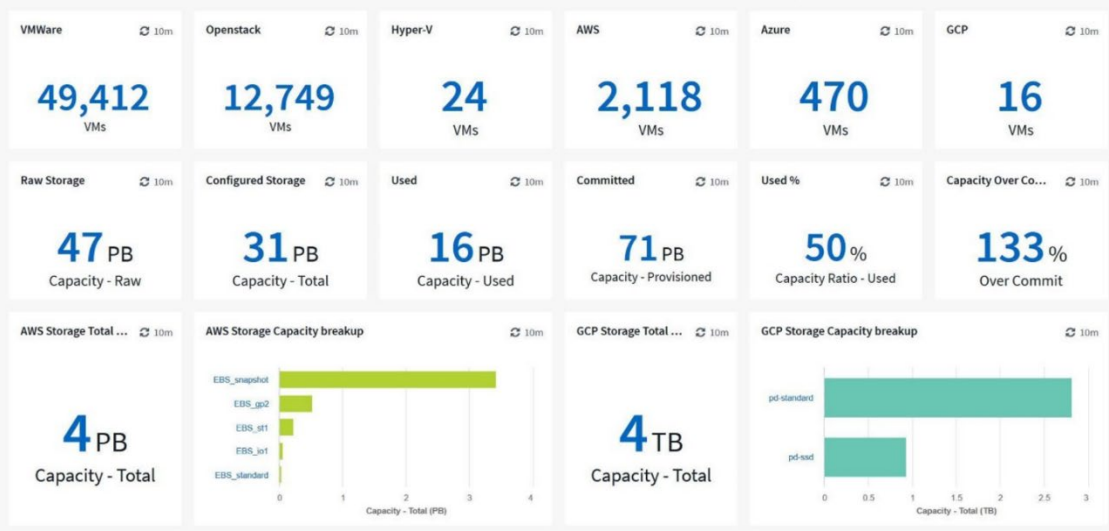
For years, we have used NetApp OnCommand® Insight and an ever-expanding portfolio of custom tools to manage our Global Engineering cloud. With each expansion of our environment and each new tool, we would create a dedicated environment to host it. With the cloud, the days of patching, upgrading, and monitoring our OnCommand Insight environment are over. Today, we rely on NetApp Cloud Insights to proactively manage and optimize our hybrid cloud environment. With Cloud Insights, we can easily visualize the capacity, performance, and usage of our complete IT stack.

Cloud Insights makes it simple to manage metadata through annotations, to search assets through a visual search engine, to customize alerts, and to create custom dashboards. With the custom dashboards that we have created to answer our most common questions, Cloud Insights has saved us innumerable work hours.

What Is the Inventory of Components Across Our Entire Environment?

NetApp Cloud Insights provides visibility into our on-premises, hybrid cloud, and cloud environments—including our Cloud Volumes ONTAP and Cloud Volumes Service volumes. We were able to quickly create a home page dashboard (Figure 3) that gives us visibility into servers, hypervisors, storage, switches, and interconnects. It also monitors third-party storage and services such as Kubernetes, Docker, MongoDB, MySQL, and PostgreSQL.

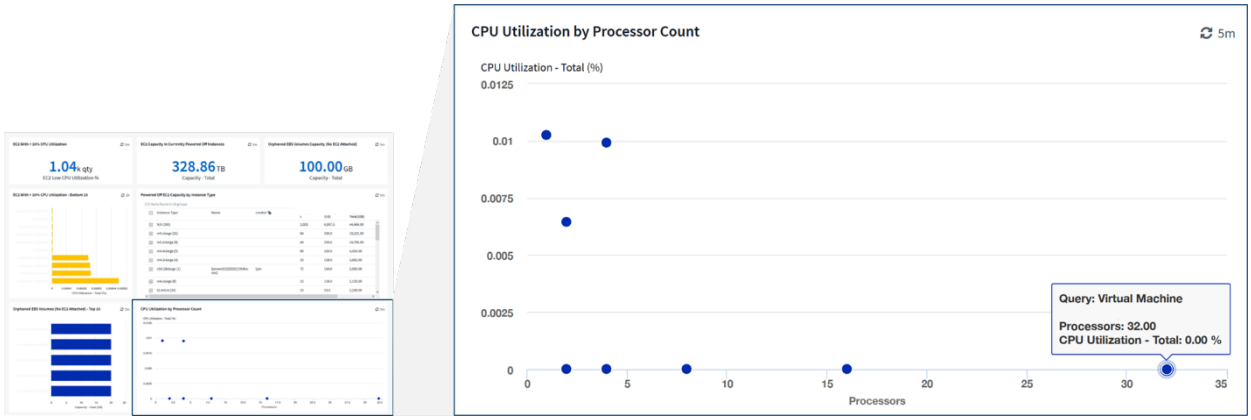
Figure 3) NetApp Cloud Insights home page.



How Can We Optimize Costs and Run More Efficiently?

We created a waste reclamation dashboard (Figure 4) that allows us to identify provisioned resources—storage capacity and VMs—that are unused so that we can reclaim them for active workloads. It also identifies candidates for reducing the number of CPU cores.

Figure 4) Cloud Insights waste reclamation dashboard reveals opportunities for cost savings.



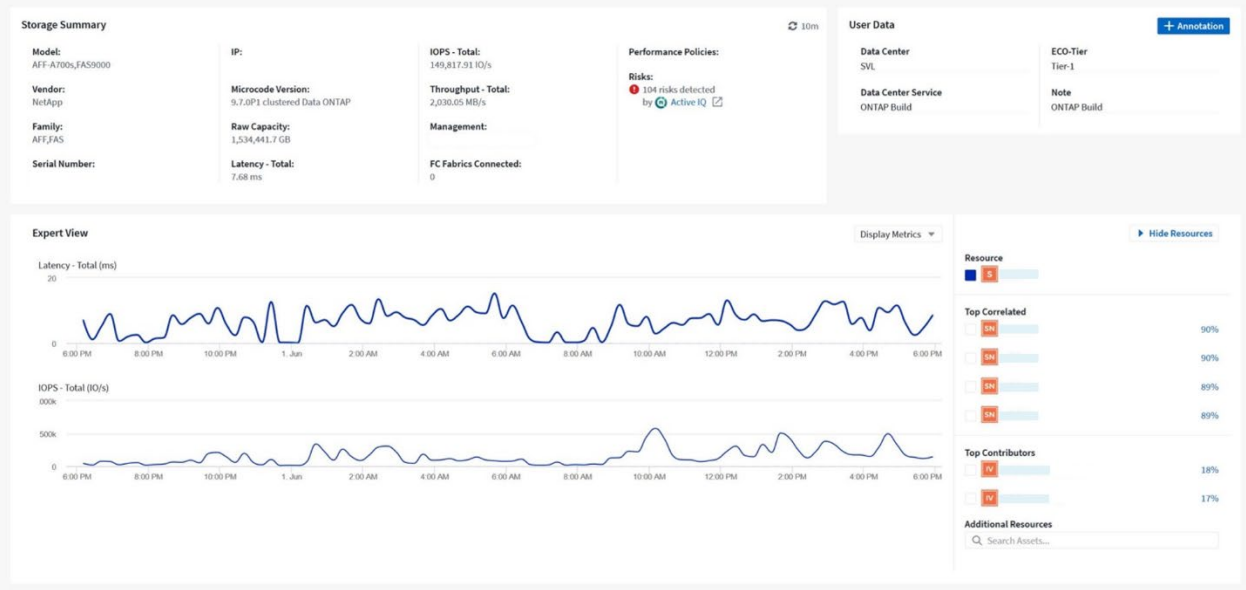
How Is Our Environment Performing and Where Are Bottlenecks?

Our performance dashboard (Figure 5) enables us to monitor performance across all architectures and to easily identify the root cause of issues. For example, we can identify bully workloads in servers and fabric interconnects, or in storage.

Figure 5) Cloud Insights performance dashboard provides insights into system performance.

After we identify underperforming systems, we can drill down for details by clicking the system name. Figure 6 shows an example.

Figure 6) Expert View provides detailed information about storage virtual machines (SVMs) that comprise bully workloads.



APIs for Easy Integration and Automation

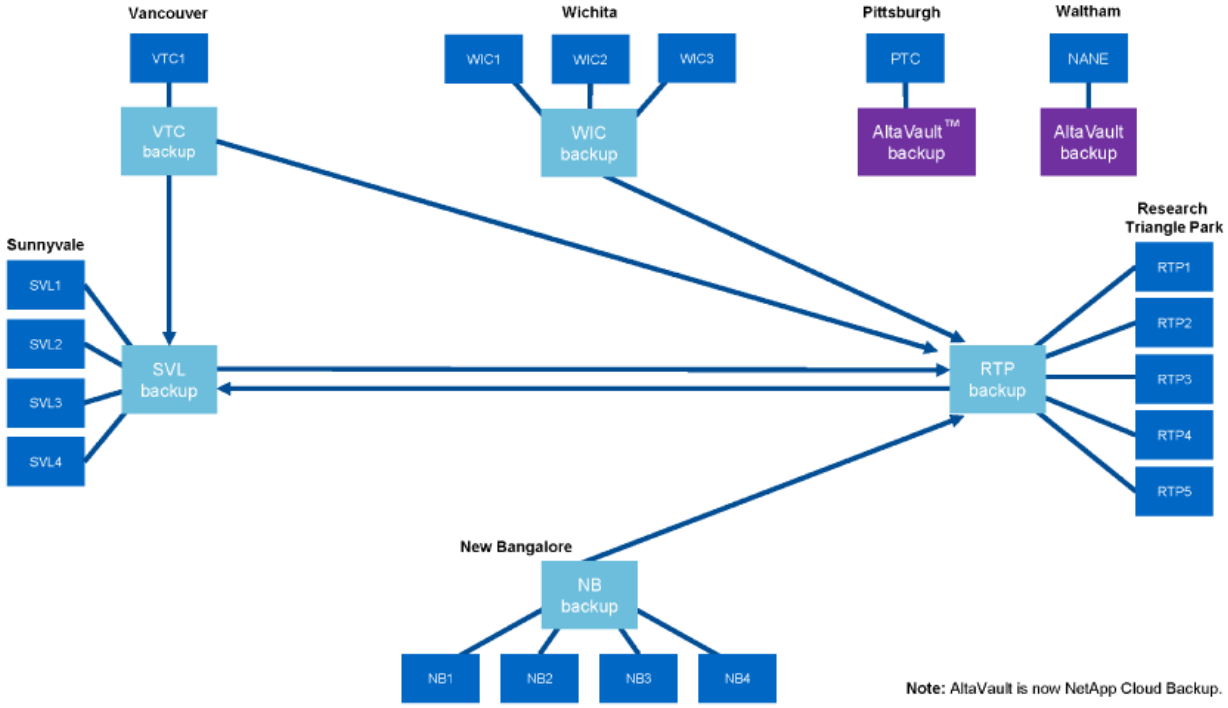
By using Cloud Insights REST APIs, we can automate integration of asset data into our engineering workflows that are managed by ServiceNow. Twice a day, ServiceNow makes calls to Cloud Insights and imports data for all our storage and virtual infrastructure—including data sources and object relationships—into our configuration management database. The integration of Cloud Insights with ServiceNow simplifies service management, accelerates all our management processes, and improves planning.

3.2 Simplify Data Protection: NetApp Cloud Volumes ONTAP

As our engineering environment was expanding, protecting the increasing volumes of data became more challenging and costly. With 14 on-premises systems that were due for a refresh, the time was right to move backup to the cloud. AWS, Azure, and Google Cloud all offer cost-effective block storage in the cloud that eliminates the costs for on-premises infrastructure and management labor. However, none of their as-delivered solutions offer the premium data management efficiencies of our NetApp SnapMirror® architecture.

Before moving to the cloud, we had SnapMirror infrastructure in seven data centers across the globe backing up data to on-site clusters (Figure 7). For business continuity, we used SnapMirror to mirror data across geographies. As a result, we were managing 280 SnapMirror relationships to back up data from 78 tier 1 production clusters (~300 nodes) to 7 backup tier 2 clusters (14 nodes).

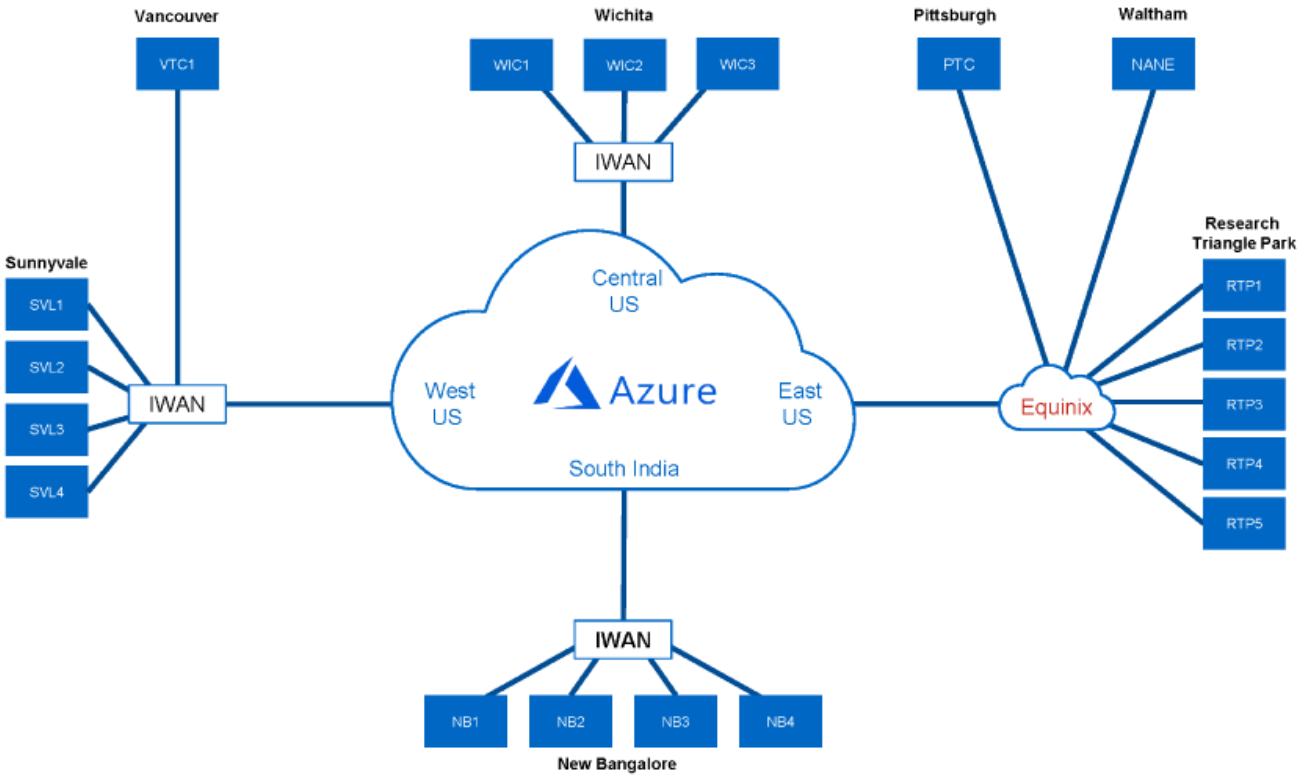
Figure 7) Backup and business continuity architecture before moving to the cloud.



NetApp Cloud Volumes ONTAP enables us to use any cloud provider's compute and block storage more cost-efficiently. We now use cloud regions and availability zones to simplify SnapMirror relationships and to protect our data (Figure 8). With ONTAP data efficiencies, we can reduce the amount of storage that we need in the cloud for our backups. These efficiencies include deduplication, data compression, data compaction, and thin provisioning. To monitor the benefits and cost savings of these technologies, along with the metrics related to CPU and memory utilization of cloud resources, we use Cloud Insights.

When replicating data to the cloud, being able to use Cloud Volumes ONTAP with other familiar processes and tools is a big advantage. Easier processes give our IT staff more time to focus on more strategic tasks. Cloud Volumes ONTAP adds to the capabilities of NetApp Snapshot™ technology. In our Azure cloud, for example, we use cloud regions and availability zones across geographies to protect our Research Triangle Park (RTP) data. It is easy to move data in and out of any cloud. It is also easy to move application workloads to the cloud, with no significant changes required in processes or staffing.

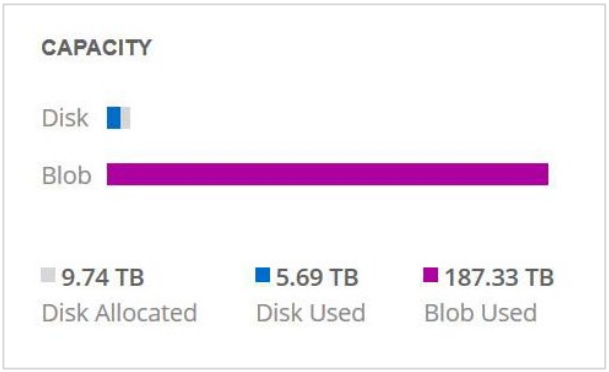
Figure 8) We simplified SnapMirror relationships by eliminating cross-regional relationships.



Use of FabricPool Tiering for Extra Savings

Cloud Volumes ONTAP enables us to optimize storage costs by using storage tiers in the cloud. Using FabricPool to configure aggregates, we can limit hot data in our performance tier on high-performing disk and use Azure Blob storage in the cloud for infrequently used data. Because up to 95% of our backup data is cold, we save about 80% in operational expenditures (opex) in Azure. While flash storage costs as much as $107 per terabyte, Azure Blob storage costs just $10 per terabyte. Figure 9 shows how a 196TB backup uses 9.74TB of disk with 187.33TB of lower-cost Blob storage.

Figure 9) NetApp Cloud Manager shows the effect of using aggregates to tier cold data to Blob storage.



Advantages of Cloud Volumes ONTAP for Data Protection

Cloud Volumes ONTAP delivers many advantages that have enabled us to:

- Decommission seven clusters and eliminate the associated costs of on-premises management, capital expenditures (capex), power, space, and cooling.
- Automate a third tier of cold object storage, further reducing opex by up to 80%.
- Eliminate duplicate backups across geographies by backing up everything to the cloud.
- Use APIs to automate storage management workflows.
- Scale storage up and down as necessary.
- Save 50% on cloud storage by paying only for the storage that we use rather than for allocated storage, thanks to thin provisioning.

3.3 Deliver High-Performance File Services: NetApp Cloud Volumes Service

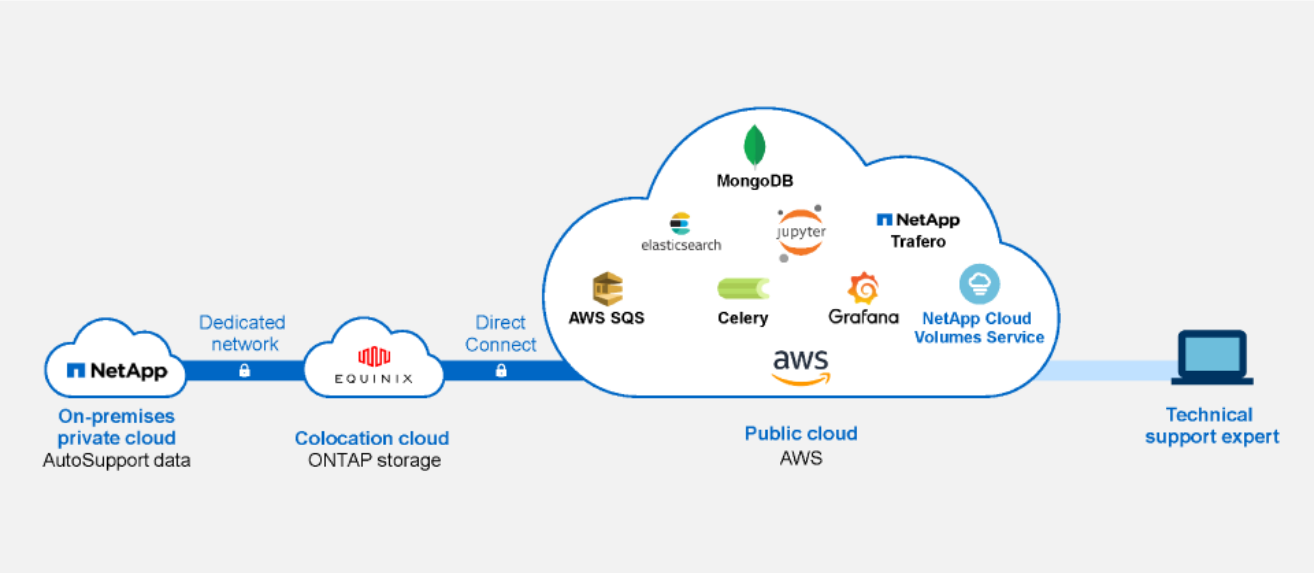
We deliver NetApp Active IQ® Performance Analytics Services (PAS) as a cloud service to NetApp Support Services, Performance Engineering, and select partners. Active IQ PAS provides on-demand access to deeper forensic NetApp AutoSupport® telemetry data on controller performance than is typically available through Active IQ. As demand for our service increased, the burden of scaling and maintaining our high-performing on-premises environment—as well as the expense--caused us to look to the cloud.

To use cloud capabilities for this data analytics service, and to obtain scaling and cost-efficiency, our team turned to Amazon Elastic Container Service (Amazon ECS) to ingest AutoSupport performance data and to visualize that data in dashboards. See Figure 10.

With NetApp Cloud Volumes Service, we can provide consistent high-performance file services, including NFSv3 for Linux and for UNIX and SMB 3 for Windows in the cloud. We chose Cloud Volumes Service for its integrated quality-of-service (QoS) features. It is now easy to use high-performance I/O only where it is really needed. We just select a service level for each storage volume. We can also use Cloud Insights to manage our Active IQ PAS architecture.

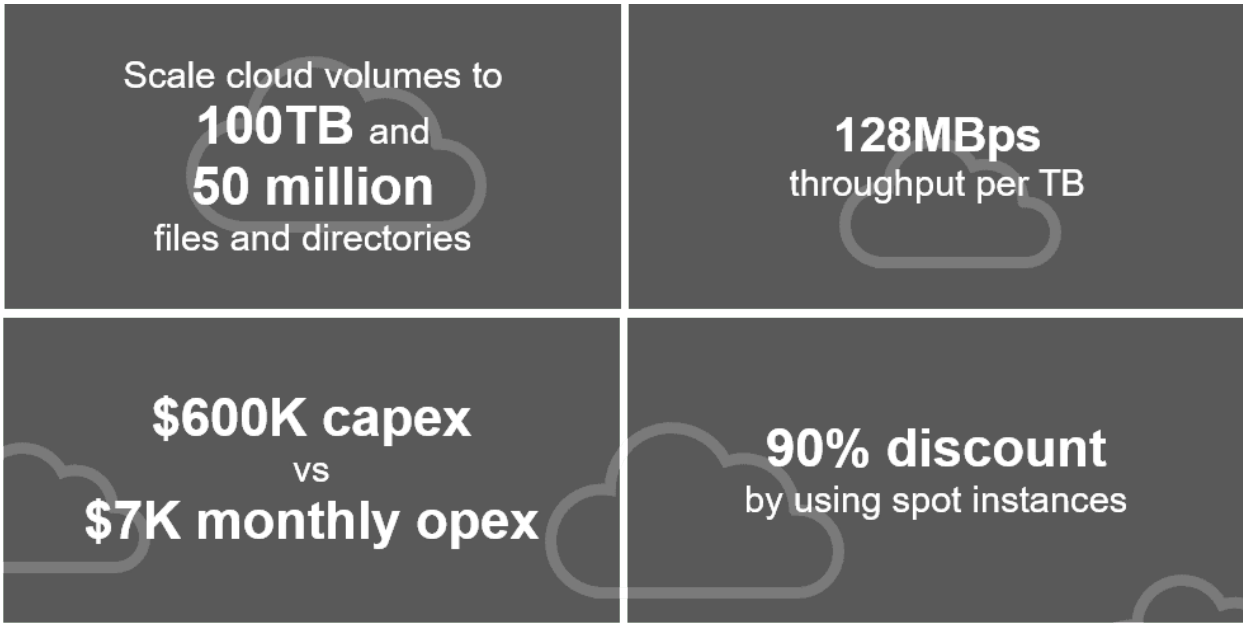
We use Snapshot and instant copies to get fast file-storage performance as a service. We can scale each cloud volume up to 100TB in capacity, up to 50 million files and directories, with up to 128MBps throughput per terabyte—all while maintaining enterprise-level data protection and security.

Figure 10) Active IQ Performance Analytics Services (SaaS architecture).



By designing Active IQ PAS containers to use interruptible and restartable containers in Amazon ECS, we use preemptible instances (Spot instances) for analytical processing and for performing get and put operations. This approach reduces on-demand pricing and lowers cloud costs. The numbers in Figure 11 represent our real-world advantages and savings from the use of NetApp Cloud Volumes Service.

Figure 11) Cloud Volumes Service delivers scalability, performance, and cost-efficiency.



3.4 Fabric View: a Launchpad for NetApp Data Services

With our adoption of NetApp cloud services, we have a growing appreciation of Fabric View (Figure 12), the launchpad for NetApp data services. Fabric View provides an executive view of our hybrid multicloud environment. We can monitor our Cloud Volumes environments, go into our Cloud Insights dashboard, or view our NetApp Cloud Backup Service. It gives us visibility into our footprint in each hyperscaler, along with our on-premises environment. The numerous dashboards provide deeper visibility into each component and help us diagnose performance issues.

Figure 12) NetApp Cloud Central Fabric View provides an executive view of our hybrid multicloud environment.

	Microsoft Azure	Amazon Web Services	Google Cloud Platform	On-Premises
Cloud Volumes Service Get Started	The industry's leading Network File System (NFS/SMB) service in the cloud			
Cloud Volumes ONTAP Start Free Trial	Simple & Fast Enterprise Cloud Storage			
Cloud Backup Service	Fully integrated service for Cloud Volumes ONTAP that delivers backup and restore capabilities for protection and long-term archive of your cloud data. Available automatically via Cloud Manager			
Cloud Insights Go to Cloud Insights	Innovate faster with insights across your application infrastructure stack			
Cloud Compliance View Cloud Managers	Always-on Privacy and Compliance Controls for Cloud Volumes ONTAP, Azure NetApp Files and Amazon S3 – for CCPA, GDPR, HIPAA, PCI DSS and other regulations.			
Cloud Sync Start Free Trial	Secure, Fast, Automated Data Synchronization			
Cloud Tiering Go to Cloud Tiering	Free up high-performance AFF storage space by automatically tiering infrequently-used data to the cloud			
SaaS Backup Go to SaaS Backup	A secure, encrypted cloud-native offering that safeguards your business-critical Microsoft Office 365 and Salesforce data from corruption, malicious or accidental deletion			
NATIVE SERVICES	Cloud Provider native storage in the cloud.			

4 Top Takeaways

There is **no reason to fear the cloud**. It really doesn't matter where a data service resides. With public cloud services, you still have control. You still have security. In fact, we learned that cloud providers deliver better security capabilities than our engineers had expected—but it is a shared responsibility. You still have high performance when you need it. You still protect your data. You can still do it all cost-effectively.

With NetApp Cloud Volumes ONTAP and Cloud Volumes Service, you can reduce your data center footprint and move to an opex model with an architecture that is still familiar. Being in the **cloud has no impact on how your users perform their jobs**. They use the same processes and technologies that they have been using in your on-premises NetApp ONTAP environment.

Businesses demand speed and scale; with NetApp Cloud Volumes Service and Cloud Insights, you can **scale up and down at will**. Our team really appreciates this flexibility.

NetApp makes it easy to get started with cloud technologies. You can start today. To learn about free trials of NetApp cloud services, visit cloud.netapp.com. Be sure to check out NetApp Cloud Manager for centrally managing your hybrid cloud infrastructure. Help your organization get new services to market quickly while saving money in the process.

NetApp