The Challenges Facing HPC Environments in the Cloud (and how to solve them)

Two Use Cases: Oil and Gas and EDA

NetApp®

Contents

Introduction to HPC Environments in the Cloud

In the early days of the industrial revolution, most factories needed to generate their own power to run their machines. Factories were built beside rivers for hydro power or included their own steam-powered generators. Self-generated power came with its own set of problems. Sometimes factories would expand beyond the capacity of their power plants, requiring costly capital upgrades. Power plants were expensive to run and maintain, and if they experienced a failure, the entire factory ground to a halt. When central power generation became available, most industries moved away from private generation stations and joined the grid. The grid was more reliable. The utilities were able to build redundancy in generation and were specialists at building and maintaining power plants. Most importantly, power became a metered commodity and businesses could get as much as the lines would support, paying for it only when the machines were running.

Enterprises looking to move their on-premises data centers to the cloud are in the same situation, particularly those with High Performance Computing (HPC) applications running in their environments. In many ways, on-premises data centers are like personal power plants. You are responsible for all the capital and operating costs. If your business users require more computing power than your current data center supports, you must procure more capacity, and if that need then declines, you are still paying for the new computing power. You are also responsible for consistent performance of your servers and must build in redundancy and monitor your systems to ensure maximum performance, security and uptime for the business. The cloud, like a centralized power generation plant, removes these constraints. Performance, redundancy and availability are your cloud provider's concern, as is capitalizing and configuring new hardware, providing cooling, power and real estate, and maintaining the physical security of the data center. You connect to the cloud and leverage the elastic nature of the service to turn up resources when you need them and turn them down when you don't.

Electronic Design Automation (EDA) and HPC in the Cloud

There are very clear industry examples where the benefits of the cloud for HPC application are evident. One such example is the Electronic Design Automation (EDA) space. Semiconductor design firms, using applications such as Cadence, Mentor Graphics or Synopsys, face a series of challenges that the cloud addresses. The first two phases of the EDA process (Develop and Simulate) create a massive number of files and tend to be very write-intensive while the CAD operators are working on the files. Traditionally, EDA solutions required a large amount of shared storage just to support the performance required for simulation runs. This performance requirement also meant that inexpensive storage was not viable. Imagine trying to manage a dedicated high-speed storage environment that must be available while the designers were developing or simulating and continually looking for more performance. The solutions are also reliant on Linux file systems and, until recently, customers had to build and maintain their own NFS server environment in the cloud to meet their end-users' requirements.

EDA has another challenge that is not storage based. The largest cost in EDA is not infrastructure: it is licensing of the software. The CAD tools used by the designers are often up to eight times the cost of the infrastructure, and the vendors charge licensing fees to the minute for processing. In this scenario, time is, literally, money. Therefore, completing design and processing tasks faster is key, but only for the time that the designers are working.

The elastic nature of the cloud makes it well suited for this workload. Scaling storage is no longer an issue, there is no hardware to capitalize and adding space is instantaneous and metered.

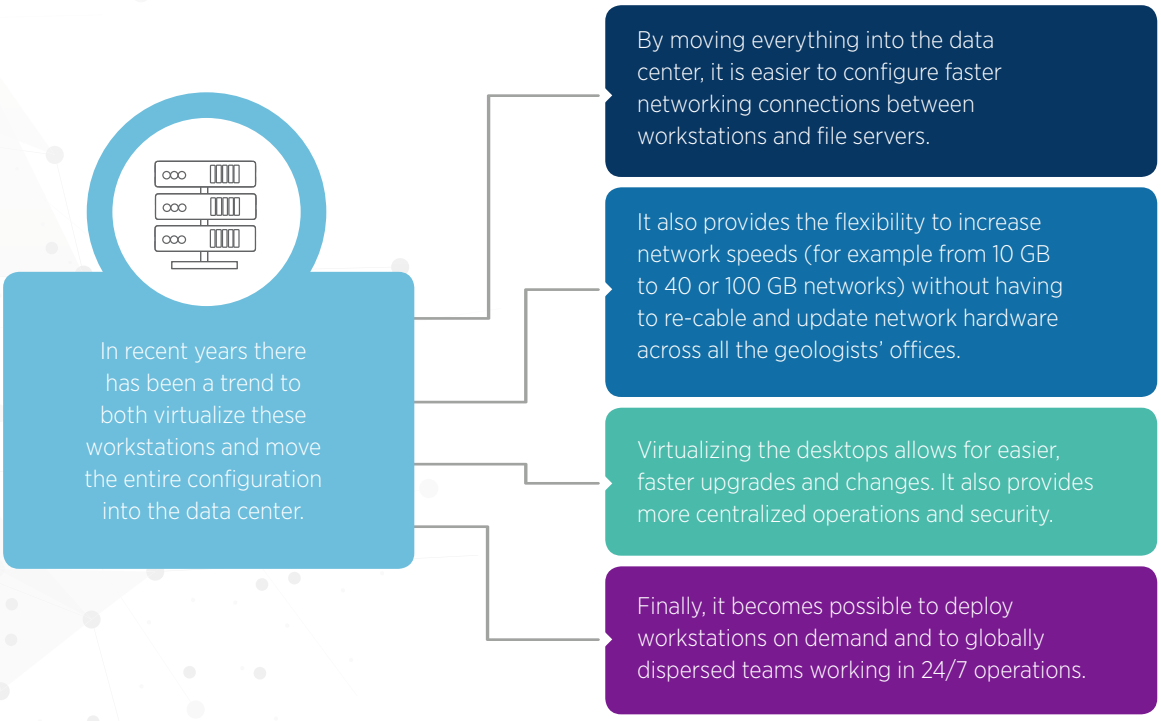
Businesses pay only for the compute and storage resources they use, not what they anticipate needing in the future. Furthermore, development occurs mostly during the workday schedule, allowing resources for development to be shut down outside of work hours.

Simulation runs can also be automated to run on-demand; again, paying for the license and resources only when required, reducing costs further when the systems are not in use.

Oil and Gas and HPC in the Cloud

HPC applications in the Oil & Gas industry have challenges that can also be addressed by moving to the cloud. Oil field exploration has specific challenges. To be effective, a project needs to represent an entire oil field, consisting of data collected from tens of thousands of wells. To be the most effective, the data combines all the domains of exploration with all the domains of production. This can include: Interpretation, 3D volume analysis and imaging, well management, drilling management, reservoir simulation, reservoir modeling and drilling modeling. All this data is combined under one application so that multiple groups of geologists and geophysicists can access the same data to create a unified view of seismic workflow and gain an end-to-end view of the field. Such a view is difficult if you have multiple, siloed applications for each data domain. The combined view also saves time and, more importantly, money by eliminating the movement of datasets between disparate systems that would need to be managed independently.

Oil & Gas applications, such as the Petrel Exploration and Production software platform, can pose a challenge for modern IT shops. These applications are older and were designed as two-tier apps running on large 3D visualization workstations having large amounts of memory and large GPUs (such as an NVIDIA M60). Rather than connecting to a database, data is file-based, including large files storing metadata that serve as file locators which are loaded into memory. The files, centralized on SMB 2.x Windows file shares, are connected to the workstations by high-speed networks (primarily 10 GB), allowing the data sets to be shared across multiple workstations.



NetApp®

This trend lends itself well to the cloud.

Oil & Gas customers can gain the elasticity of cloud-based storage. Large numbers of files can be stored in the cloud without worrying about constantly procuring more storage as data sets increase.

Companies pay only for the data space that they use, not the space they anticipate using.

Virtualizing the workstation allows companies to leverage the dynamic compute power of cloud. As user needs change, IT can bring new workstations on- and offline without the need to purchase equipment.

Companies can use automation templates to speed up the time it takes to create a new workstation.

Applications and data are stored in the cloud rather than on the workstations and laptops of multiple scientists, allowing for much greater security.

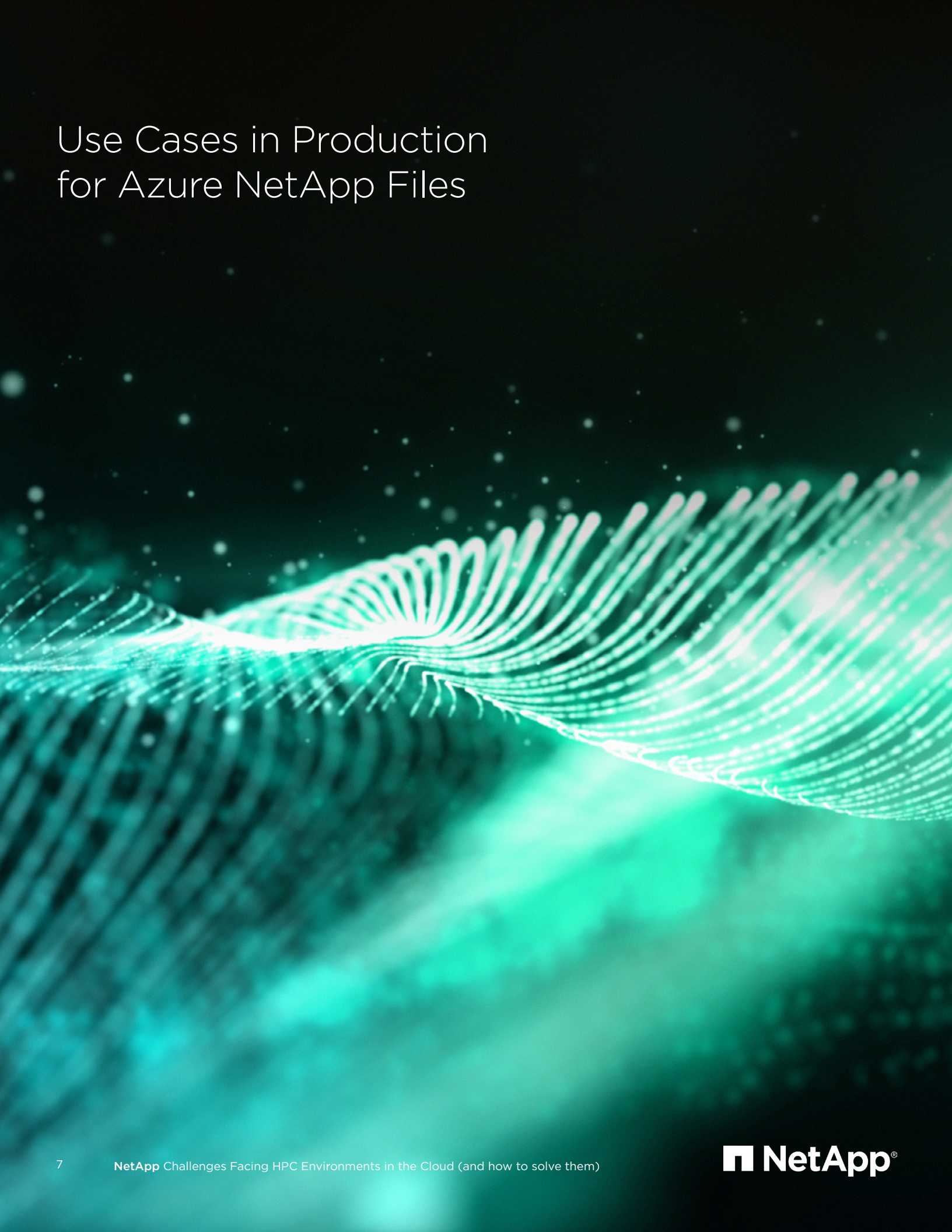
3-10 years

$1.2B risk

It has been estimated that an oil company could spend 3-10 years using the current technology at a cost of 1-2 billion dollars of risk before the first well is drilled.[1]

With the cloud they could cut that time by 33-66%.

Use Cases in Production for Azure NetApp Files



NetApp®

"We have a long history of applying technology," says the CIO. "This partnership will allow us to digitally transform and leverage the scale and capabilities of Microsoft to ensure we harness the value of our data. We have an incredible amount of data from our production facilities, drill ships and fiber optic cables inside well casings. The cables measure pressure, temperature and other metrics, and can generate up to 1 TB of data a day."

Industry: Oil and Gas
Use Case: SAP

Azure NetApp Files Chosen for SAP Workloads in Oil and Gas Customer's Cloud Transformation

CHALLENGE

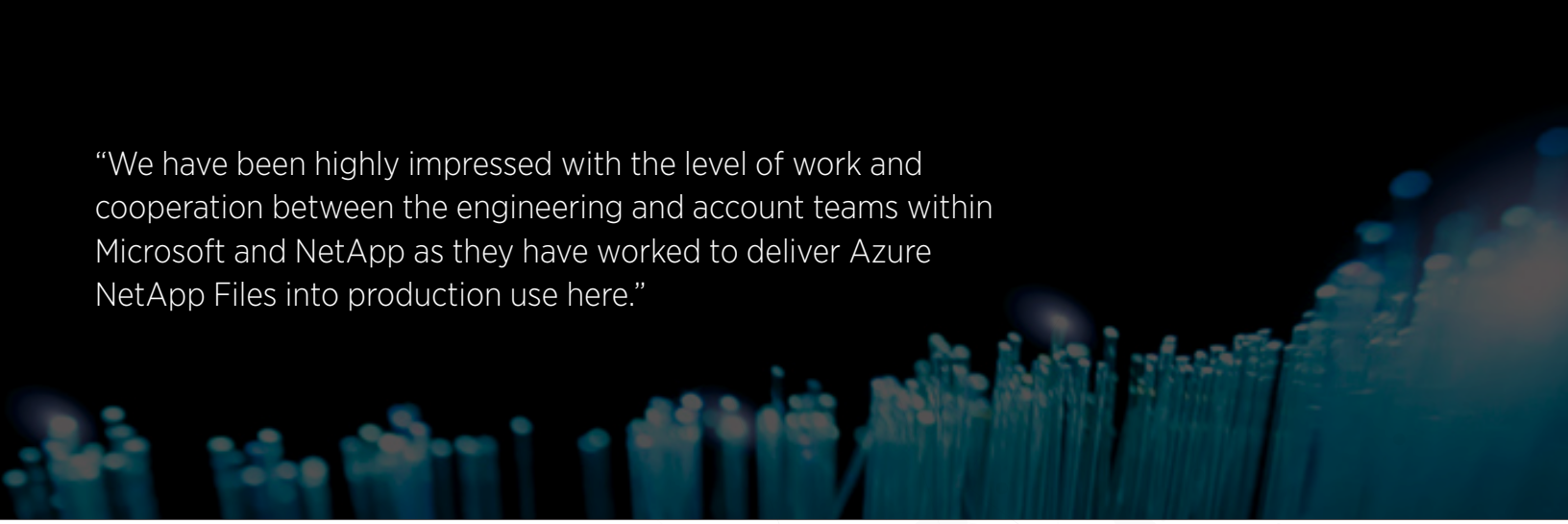
The company is working through a large-scale cloud-first initiative with Microsoft to move applications & workloads to Microsoft Azure. The goal is to save them time and resources and to accelerate the speed at which they can deploy applications. With many applications including their many SAP landscapes requiring shared file access, they cannot achieve the necessary scale by building their own file servers.

SOLUTION

Azure NetApp Files is being used as the file share environment for SAP deployments on Azure, along with many other file share applications.

BENEFIT

Ease of use and performance: the company was surprised by the performance capabilities Azure NetApp Files offers natively in the Azure Portal.



"We have been highly impressed with the level of work and cooperation between the engineering and account teams within Microsoft and NetApp as they have worked to deliver Azure NetApp Files into production use here."

Industry: Oil and Gas
Use Case: Cloud-first Strategy

Petroleum Company Selects Azure NetApp Files to Support its Cloud-first Strategy

CHALLENGE

With the closure of the customer's Core Data Centers and the adoption of a cloud-first strategy, the customer needed to move their NFS and SMB application sets to the cloud while still achieving the same levels of enterprise-class performance. The customer is using Azure NetApp Files for a broad set of workloads, including production systems, staging environments, Disaster Recovery and more.

SOLUTION

The Microsoft Azure Platform Services (MAPS) team are migrating data directly into Azure NetApp Files, and saw the solution as a "fully managed service" which, once integrated into their cloud setup, could be consumed by various customer projects. The customer has chosen Azure NetApp Files as the strategic solution for NFS and SMB environments in Azure, has been an early adopter on each of the feature sets as released, and will continue with this into the Cross-Region Replication and Backup services.

BENEFIT

The customer benefits from the ease of use with Microsoft applications, which helps them to fulfill their Microsoft Azure Monetary Commit (AMC) by ensuring that the right workloads are using the right levels of service in the Azure cloud.

"We're a long-time customer and huge fan of NetApp for many of our enterprise applications requiring file-based storage. Azure NetApp Files is both straightforward to use and cost effective."

Industry: Oil and Gas
Use Case: CAD and HPC

Azure NetApp Files Helps Manufacturer Create HPC in the Azure Cloud

CHALLENGE

The company has a file environment in Azure today with NetApp, but they've grown their footprint and it's become more cumbersome to manage the required Azure VMs and managed disk to support the environments which require massive numbers of files running applications that are very write-intensive and need to run in a high performing environment.

SOLUTION

The customer chose Azure NetApp Files as their primary file storage option in Azure for all their CAD design work because of speed and simplicity. Azure NetApp Files allows the customer to flatten their storage architecture, remove dozens of individual storage elements, migrate from an IaaS-like experience with CVO and experience on-premises-like speed in Azure.

BENEFIT

The performance and cost of Azure NetApp Files vs. multiple Azure Marketplace offerings is a huge benefit.

The Case for Azure NetApp Files – Learning More

The case for moving HPC workloads such as Electronic Design Automation and Oil and Gas applications to the cloud is clear. Organizations are already centralizing their HPC applications into the data center, and the cloud enhances this trend with the benefits of ease, scale and cost control. The final question for businesses is which cloud provider will offer the most advantage with the least disruption. The answer to this question is Microsoft Azure and the Azure NetApp Files service. Azure provides all the advantages of a modern enterprise cloud provider: security, performance, redundancy and the ability to elastically scale resources up to meet variable demands and down to save money when resources are not needed. The Azure NetApp Files service allows businesses to meet all the specific requirements of an HPC application's file system storage without requiring expensive and difficult code changes.

Microsoft provides Azure NetApp Files as a native service, under your existing Azure Enterprise Agreement (or pay as you go) and does not require you to purchase any additional licenses. Azure compute services support both Azure and Windows natively, so there are no barriers to moving from on-premises to Azure. Once Azure provisions the service in your tenant, you can seamlessly move your HPC application (whether Linux or Windows based) to Azure, with performance that mirrors your on-premises solution.

"Azure provides all the advantages of a modern enterprise cloud provider: security, performance, redundancy and the ability to elastically scale resources..."

NetApp®

Azure, through their partnership with NetApp, have built Azure NetApp Files to support the needs of modern HPC applications.

Since the Azure NetApp Files service is built on NetApp technology and supports Linux NFS v3, v4.1 and Windows SMB file system, there is no need to re-platform Windows or Linux based applications to move them to the cloud.

The storage capacity of this service is designed for massive workloads, supporting volumes of up to 100 TB for up to 100 million files in the standard offering. Microsoft or NetApp can even increase the file count limit in special circumstances.

Azure NetApp Files also provides three service levels so you can right size your storage performance to the application:

- Standard provides 16 MB of throughput per TB of storage
- Premium provides 64 MB of throughput per TB
- Ultra provides 128 MB of throughput per TB

This allows administrators to have flexible and agile control of performance–and effectively costs–by allocating faster storage only when necessary for the task.

The HPC architecture provided by Azure NetApp Files is highly parallelized and can reduce the computational time for a given task by dynamically scaling resources, saving overall costs to the business and creating shorter development cycles and quicker data response times.

The Ultra service level provides consistent performance that will meet the most demanding application requirements. This storage provides sub-millisecond latencies–speeds unparalleled in any cloud file service.

For organizations such as Oil and Gas that are looking to virtualize their HPC applications in the cloud, this performance allows cloud-based HPC workloads to respond as if they were on-premises (or possibly better). In the case of EDA applications this increase in speed also saves time per simulation run, resulting in less time using the CAD license and lower licensing costs.

As a native service, Azure with Azure NetApp Files provides a high level of management and security. For both Oil and Gas and EDA applications security and uptime are essential. Azure NetApp Files supports native encryption and is integrated into the security of the other Azure services, giving administrators easier control of access, and giving end-users a single sign-on experience.

Equally important, Azure NetApp Files is built directly on highly available NetApp storage systems in the Azure data center, offering true enterprise class storage for your file environments. This provides fault tolerant services that minimize downtime without the need to add additional hardware, ensuring that your most critical applications always have access to the required data.

Finally, as a native service the full solution can be scripted and automated to allow your administrators to quickly spin up user workstations as needed, in any region of the world, and bring them down just as fast when not needed.

HPC application workloads are built for the cloud. While speed, flexibility and simplicity are challenges that are difficult to meet on-premises, Microsoft Azure can alleviate all of these issues. With Microsoft Azure and Azure NetApp Files, businesses can easily bring their existing HPC applications, whether on Linux or Windows, to the cloud and benefit from a higher level of performance and flexibility–a level that on-premises simply cannot provide. A service is only as good as the infrastructure upon which it operates. With Azure and Azure NetApp Files, you can be confident that the service will be available when your HPC applications require it most.

Want to try Azure NetApp Files yourself? Visit Microsoft's web page to learn more and to register for the service.

References

[1] https://www.ft.com/content/19234982-0cbb-11e8-8eb7-42f857ea9f09

About the Author – Jeff Whitaker

As a member of the Cloud Data Services team at NetApp, Jeff is focused on making sure NetApp customers know all about the cool things we are doing in Microsoft Azure. He is a perennial disruptor with a passion for emerging technologies that let start-ups grow up fast, and traditional companies re-discover their entrepreneurial essence. Cool ideas that Jeff considered (but didn't invent) include machine learning, virtual reality and a drone that fetches tapas on demand. You can reach Jeff directly at **jeff.whitaker@netapp.com**