Dell Technologies Cloud PowerScale: Microsoft Azure

Efficiently run compute intensive file workloads in Microsoft Azure Cloud with Dell EMC PowerScale OneFS

Customer Benefits

- High bandwidth, low latency connection from PowerScale and Isilon nodes to Microsoft Azure using Azure ExpressRoute Local

- No outbound data traffic costs

- Scale up compute and storage on demand without complexity

- Durable, persistent cloud-attached storage

- No secondary data center or infrastructure to manage

- End-to-end managed service

- Simple and predictable subscription-based pricing

- 24x7 support

- Ideal for technical workloads in industries such as Life Sciences and Media and Entertainment

Business Challenges

As we enter the data decade, cloud adoption has never been higher, but data centers aren't going away either. In fact, 91% of cloud strategies include on-premises data centers[1]. In this new data era, cloud is not a destination, it's an operating model that requires a Data First approach to balancing workload requirements across on-premises and cloud and ensuring that data is secured, protected, available where and when it is needed, and stored and managed within compliance policies.

Unstructured data such as file faces its own set of challenges in the cloud such as maintaining user experience, scale and performance limitations, and the difficulty of applying cloud data analysis tools to on-premises data. In addition, compute intensive workloads are also very demanding on storage and many require high read/write throughput and symmetric data access to storage while producing hundreds of gigabytes of data.

Dell Technologies has partnered with Microsoft to offer an Azure-based cloud service that addresses these challenges through cloud-attached NAS storage that delivers high performance at scale and has a flexible design to optimize costs and keep business in control of their data.

Solution Overview

Dell Technologies Cloud PowerScale: Microsoft Azure combines Dell EMC PowerScale, the #1 scale-out NAS platform powered by PowerScale OneFS, with the Microsoft Azure public cloud, which offers enterprise-grade compute for operational flexibility. This solution provides a high bandwidth (up to 200Gbps), low latency ([as low as 1.2ms)](#) connection from Isilon to Azure using Azure ExpressRoute Local. It also eliminates outbound data traffic costs for data written to PowerScale storage nodes from within Azure.

Dell Technologies Cloud PowerScale: Microsoft Azure allows for the right combination of compute and storage for data-intensive, high I/O throughput, file-based workloads that require high compute performance on a periodic and/or unpredictable basis. This makes them suitable for a cloud consumption model. Eliminating egress charges enables workloads that require a lot of temporary writes to the PowerScale storage nodes to cost-effectively take advantage of Azure's application services. This is ideal for industries such as Life Sciences and Media and Entertainment, which can require on-demand computing power tied to a large file system.

---

[1] ESG Research Hybrid Cloud Tipping Point, June 2018

For compute, Azure offers the choice of dozens of VMs with a wide variety of CPUs, some optimized for HPC workloads, memory capacity and network options. When combined with PowerScale's unmatched performance, reliability and scalability, and a single multi-petabyte namespace which supports symmetric data access across its nodes, organizations get a fully managed cloud that can address the most demanding requirements.

Life Sciences

Genome analysis is one of the key use cases for life sciences. The raw data generated by a genomic sequencer for the complete genome of a single human is approximately 100GB. This dictates a requirement for a massively scalable file system to which capacity and performance could be added. Genome alignment and sorting, which are both part of the secondary analysis stage, are the most compute and storage demanding and can require network throughput of 10Gbps or even 100Gbps and beyond. Dell Technologies and Azure testing has demonstrated that the performance of Isilon scales out linearly with the increasing IO demands and growing number of Azure VMs supporting the genome alignment stage. High bandwidth ExpressRoute Local connections between Isilon and Azure enable both the compute performance in Azure and the storage performance in Isilon to scale up in order to process real-world genome analysis.
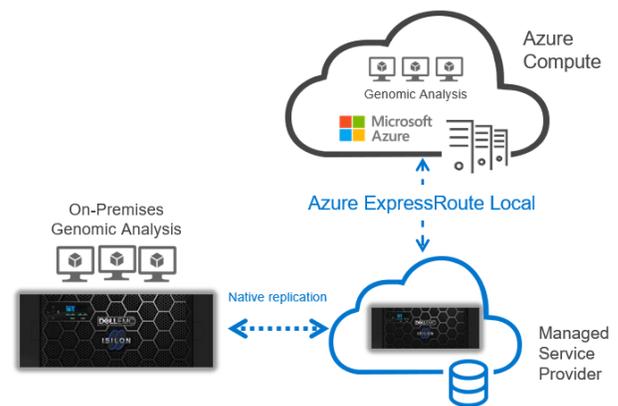
Large research facilities processing hundreds of thousands of genomes per year, generate petabytes of very large file data (typically 500GB per file set) to be stored, and have a demand for computing power that is bursty by nature – a perfect application for on-demand, easily scalable cloud computing. In addition, since genomic processing is, at its core, a pattern-matching application, there are writes to temporary files on the PowerScale or Isilon storage node during a large part of the analysis workflows.

Media and Entertainment

Video rendering, a critical process for Media and Entertainment (M&E) organizations, is another key use case for Dell Technologies Cloud Storage with Azure and PowerScale storage. A typical movie may require rendering of several hundred thousand frames at 50MB each, generating hundreds of TBs per movie. Plus, studios' archives grow to multiple PBs of raw and processed footage that are retained forever. Time to market may be a great consideration, as production often needs to be contained to a few months. The entire film may be as large as several PBs and on-premises compute infrastructure may not be available to produce the film, which benefits from very highly-parallelized rendering processes. Rendering video in the cloud allows M&E organizations to combine flexible Azure compute resources with the unmatched scalability, performance and efficiency of the PowerScale storage platform.

Why Microsoft Azure

Microsoft Azure provides a range of compute options and software services for on-demand, cost-effective processing of high-throughput, filed-based workloads. Enterprise-grade reserve compute enables operational flexibility. Azure has also made investments in infrastructure and analytics services. The Microsoft Azure H-series family of high performance computing virtual machines (VMs) deliver the throughput required for genomic analysis workloads and deploying one to thousands of VMs is simple and flexible. Plus, customers only need to pay for the compute and services used.

Why Dell Technologies Cloud Storage

The Dell EMC PowerScale family scale-out NAS all-flash and hybrid storage connected to Microsoft Azure cloud compute capacity and scale on demand, with no outbound data traffic costs, gives users with data-intensive workloads the best of both worlds – cost-effective Isilon storage performance at scale and Microsoft Azure scalable compute performance.

With Dell Technologies Cloud Storage, organizations gain advantages of Dell EMC Storage such as high-availability for business continuity, data resiliency, and flexible scalability coupled with the economic benefits of public cloud delivered as an end-to-end managed service. Plus, native replication capabilities allow businesses to move their data from on-premises to workloads in the cloud.

DELLTechnologies