Microsoft

Azure Whitepaper

Azure for the Semiconductor Industry

Using Azure for Silicon Development with Electronic Design Automation (EDA) Software

Microsoft

Table of Contents

Summary 4

Silicon Design in Azure – Overview 4

Infrastructure Options 5

- Compute 7
  - General Purpose Virtual Machines 7
  - Compute Optimized Virtual Machines 7
  - Memory Optimized Virtual Machines 8
- Scaling & Scheduling Dynamic Workloads 8
  - Cluster / Resource Deployment 10
  - Policy and Governance 10
  - Azure CycleCloud Features 10
- Storage 11
  - Storage Services 11
  - Page Blob / Disk Storage 11
  - Azure Avere NFS File Systems 12
  - Azure Avere Cloud NAS 12
  - Azure Avere Cloud Bursting 13
  - Azure Avere Blob Storage Gateway 14
  - Parallel Virtual File Systems 14
  - Durability and High Availability 15
  - Encryption 16
  - NetApp ONTAP Cloud 16
  - Microsoft Azure Enterprise Network File System (NFS) service, powered by NetApp 17
- Networking 17
  - Virtual Networks 17
  - On-premises connectivity 18
  - Point-to-site (VPN over SSTP) 18
  - Site-to-site (IPsec/IKE VPN tunnel) 18
  - ExpressRoute (dedicated private connection) 18
  - Accelerated Networking 19

Managing EDA Licensing 19
Security & Governance 19
Azure Security Center 19
Encryption 20
Key Vault and SSH Keys 20
Policies 20
Role-based access control 21
Azure Monitor 21
Azure CLI 21
Azure Infrastructure Automation Tools 21
Azure Cost Management 22
Document Updates and Contributors 23

Summary

Azure provides a globally available, high performance computing (HPC) platform, that is secure, reliable and scalable. Azure is well suited to addressing current and emerging infrastructure needs with the silicon design and development workflow that is based on electronic design automation (EDA) software. This document provides an overview of the infrastructure options and Azure services for silicon design workloads using EDA software. It also summarizes key tools and technologies specifically supported to help optimize silicon design workflows. The options presented are based on Microsoft's internal silicon development experience running EDA tools, as well as deployments and feedback from industry customers and partners.

Silicon Design in Azure – Overview

The general workflow for silicon design, using EDA software, has remained steady over many years, though tools have become more complex and many point tools have been added. Figure 1 shows a high-level overview of the silicon design workflow, consisting of front end design, also referred to as logical design, and back end design, also referred to as physical design.

Figure 1 - Overview of Silicon Design Flow

Front end design spans specification to a logically validated design using software simulation. This phase shares similarities with software R&D and is primarily code (Verilog or VHDL) heavy, with some block level simulation and debug cycles. Azure's DevOps tools and services are applicable in improving the design cycle by fostering team collaboration and agile practices. Following block design, the next step in the logical phase is to verify the design by running it on a large compute grid. With the support of a job scheduler, a design is distributed to available compute resources across the entire gird, up to thousands of cores. This is to bypass the performance limitations of a single computer (node) for running large and complex operations. Azure with its near infinite compute resources can help in improving this verification process by

shortening each job's run time and also allowing more jobs to run in parallel. Design synthesis in software is the next step. Once Synthesis is complete, another set of verifications and validations are performed using point tools. This activity, like the verification done earlier, can also be sped up and multiple runs can be launched at the same time to optimize the step. The project team relies on a configuration management system for source control, and project backups.

In back end design, EDA software tools are used to map the logical design to the specific foundry process that it is targeted to. This involves steps like Static Timing Analysis (STA), Design For Test (DFT), and design validation tools like Design Rule Check (DRC) and a Layout Versus Schematic test (LVS), in multiple iterations to ensure that the design is accurately mapped and conforms to all the elements of the physical fabrication process. Many of these iterations are compute and memory intensive operations, requiring a very large number of cores and using huge data sets. This phase also includes a check on Design For Manufacturing (DFM), which are designed for parallel compute. Such tools are ideally suited to run on Azure, where thousands of cores can be spun up to deliver a result in hours instead of days in an on-prem setup. With Azure, EDA workflows can be targeted to the optimal infrastructure configuration to achieve efficient results. Azure also enables parallel deployment for faster convergence, if license availability permits.

Infrastructure Options

IT infrastructure for silicon design is typically optimized for compute and memory intensive applications, supported with high performance file systems and efficient job scheduling to maximize throughput and performance of EDA software license investments.

Compute: Azure customers can choose from a range of compute- and memory-optimized Linux and Windows Virtual Machines (VMs) to run their workflows. Given a range of VM instances with a Core to Memory ratio ranging from 1:2 to nearly 1:30, Azure has the flexibility needed for optimizing the VM mix depending on the design stage. As an example, an Azure E series 1:8 core to memory ratio VM running at turbo frequency may be the optimal choice for running front end simulation and synthesis workloads, and an Azure M-series machine with 128 cores and 3.8 Terabytes of memory can be utilized for handling intensive back-end timing analysis and physical verification jobs.

Storage: Azure Storage offers multiple storage options including Blob, to store large unstructured data, Files, as a drive to share files, and Disk Storage, for high I/O performance. Azure Storage offers capabilities for handling high performance NFS scenarios in cloud-only and hybrid environments. For steps like verification that can generate millions and millions of small, (4 to 16K) temp files, Azure Avere provides high performance NFS support.

Networking: Azure supports several features and offers custom networking options to allow for fast, scalable and secure network connectivity between customer premises and global Azure regions. Microsoft has also invested in private optical fiber capacity as well as undersea cabling, which allows low latency access globally, as well as peering between regions for wide footprint customers. For transfer of large data, Azure offers online as well offline options and supports third-party solutions.

Scalability: Azure offers nearly unlimited scalability. Given the cyclical nature of the silicon industry, using Azure enables organizations to rapidly increase and/or decrease the number of cores needed, while only having to pay for the resources that are used. Azure Autoscale enables programmatic scaling to optimize resource use while Azure Monitor provides infrastructure metrics and logs for most services in Microsoft Azure.

Scheduling & Orchestration: Azure supports silicon industry standard tools like LSF (IBM Spectrum LSF®), and provides alternatives like batch, a high-performance cloud native job scheduler. Azure CycleCloud is a site-installed, web-based orchestration tool for HPC clusters and workflows and helps create, manage, use and optimize dynamic clustered compute environments. It can also set policies that govern access and use of Azure, which helps manage resources and costs.

Global Presence: Azure has more regions globally than any other cloud provider, offering the scale needed to bring applications closer to users around the world, preserving data residency, and providing comprehensive compliance and resiliency options for customers. Using Azure's footprint, the cost, the time, and the complexity of operating a global semiconductor infrastructure can be reduced.

Security: As silicon IP is the centerpiece of any Semiconductor company, security is a key concern for the industry. Microsoft has decades of experience building enterprise software and running some of the largest online services in the world. This experience is implemented to continuously improve security-aware software development, operational management, and threat-mitigation practices that are essential to the strong protection of services and data. Azure offers a wide array of security tools and

capabilities, to enable customers to secure their platform, maintain privacy and controls, meet compliance requirements and ensure transparency.

The rest of this document provides more details of Azure services relevant to the Silicon design workflow.

Compute

Figure 2, below, provides a high-level overview of Azure VM size mapping to different stages of the Silicon development workflow, along with key Azure capabilities that support the requirements of the industry.

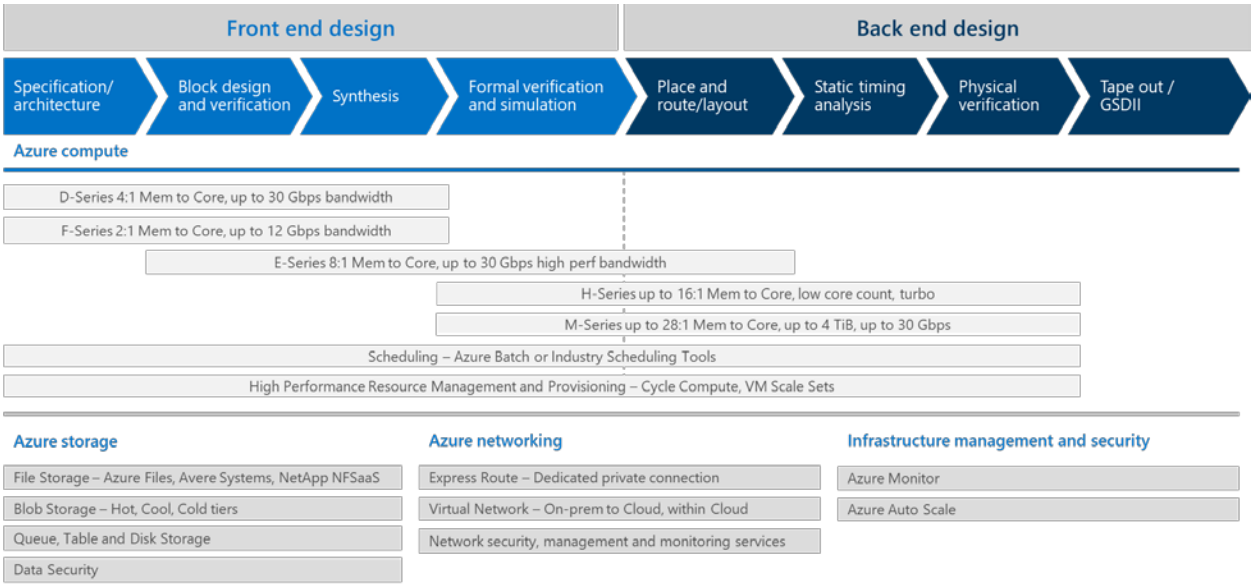


Figure 2 - Optimizing Azure services for Silicon Development

General Purpose Virtual Machines

Dv3-series sizes are based on the 2.4 GHz Intel Xeon® E5-2673 v3 (Haswell) processor or 2.3 GHz Intel XEON ® E5-2673 v4 (Broadwell) processor that can achieve 3.5GHz with Intel Turbo Boost Technology 2.0. The Dv3-series sizes offer a combination of vCPU, memory, and temporary storage for most production workloads.

Compute Optimized Virtual Machines

Compute optimized VM sizes have a high CPU to memory ratio and are good for EDA applications. The Fsv2-series is based on the Intel® Xeon® Platinum 8168 (Skylake) processor, featuring a base core frequency of 2.7 GHz and a maximum single-core turbo frequency of 3.7 GHz. Intel® AVX-512 instructions, which are new on Intel Scalable Processors, will provide up to a 2X performance boost to vector processing workloads on both single and double precision floating point operations.

At a lower per hour list price, the Fsv2-series is the best value in price and performance in the Azure portfolio based on the Azure Compute Unit (ACU) per vCPU.

F-series is based on the 2.4 GHz Intel Xeon® E5-2673 v3 (Haswell) processor, which can achieve clock speeds as high as 3.1 GHz with the Intel Turbo Boost Technology 2.0.

F-series VMs are an excellent choice for workloads that demand faster CPUs but do not need as much memory or temporary storage per vCPU such as the initial stages of front end design, including specification and smaller block design and simulation.

Memory Optimized Virtual Machines

Ev3-series instances are based on the 2.3 GHz Intel XEON ® E5-2673 v4 (Broadwell) processor and can achieve 3.5GHz with Intel Turbo Boost Technology 2.0. Ev3-series instances are ideal for memory-intensive enterprise applications.

Azure H-series VMs are the latest in high performance computing VMs aimed at high end computational needs. These 8 and 16 vCPU VMs are built on the Intel Haswell E5-2667 V3 processor technology featuring DDR4 memory and SSD-based temporary storage. With Turbo Boost, they can achieve 3.2 GHz.

The M-Series offers the highest vCPU count (up to 128 vCPUs) and largest memory (up to 3.8 TiB) of any VM in the cloud. It's ideal for applications that benefit from high vCPU counts and large amounts of memory.

Depending on the specific requirements of the silicon development project, E-, H- and/or M-series VMs will be well-suited to address more demanding jobs such as Formal verification, place and route, static timing analysis and physical verification.

A list of Azure endorsed Linux distros can be found here:

https://docs.microsoft.com/en-us/azure/virtual-machines/linux/endorsed-distros

Scaling & Scheduling Dynamic Workloads

EDA workloads can consume anywhere from tens to thousands of cores in a single run. Multiple parallel runs, varying machine configurations, multiple project needs, last minute design changes etc. all add to the complexity of right sizing on-premises IT infrastructure.

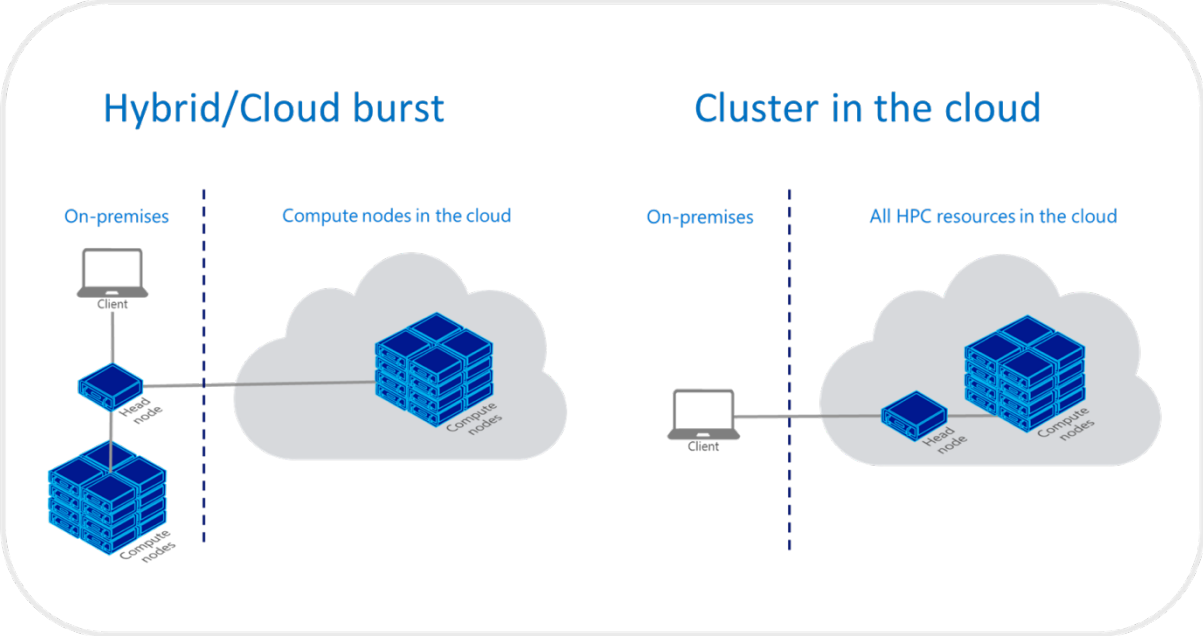
Azure CycleCloud is a site-installed, web-based orchestration tool for HPC clusters and workflows and helps create, manage, use and optimize dynamic clustered compute

environments. With Azure CycleCloud, customers can choose how to deploy on-premises and on Azure, and dynamically grow or shrink the compute capacity as needed.

Azure CycleCloud integrates with popular job schedulers like Grid Engine, LSF, Slurm, and PBS Pro to automatically provision compute resources based on the requirements of the jobs in the queue.

Target Architectures: Burst & Cloud-Only:

Azure CycleCloud targets both cloud only clusters, composed of Azure's IaaS resources, and Hybrid/Cloud burst architectures. In fully automated hybrid/burst scenarios, Azure CycleCloud acts as a meta-scheduler, providing an interface for job submission with routing to the right cluster, job queue, or scheduler based upon user-defined policy. Policies can also provision additional resources and move data.



In both architectures, Azure CycleCloud primarily offers three capabilities:

1. Dynamic deployment of clusters and other resources: scheduler, compute node, storage, cache, etc.

2. Orchestration of data for jobs and cloud workflow(s.)

3. Policy and governance: what jobs run where, what users can do, etc.

Cluster / Resource Deployment

Azure CycleCloud enables users to define and provision Azure clusters quickly, including networking, storage and compute resources through expressive and parameterizable cluster templates. All Azure IaaS components, including VMs, GPUs, standard and premium disks, etc. are supported. Expressive and parameterizable cluster templates ensure consistent, repeatable, easily extensible cluster configurations

Cluster administrators have the option to use Azure CycleCloud to abstract the cloud from users, providing a controlled and managed interface to scalable compute. Alternatively, CycleCloud can provide users with self-service capabilities, providing the policies and engine for managing access, security, and costs.

Policy and Governance

Azure CycleCloud provides both the orchestration and the policies that govern that orchestration. This includes policies such as:

- Who can access compute, and what types.
- What data can be transferred, and by whom.
- What versions of workflows are executed, on exactly what infrastructure.
- Overall costs and budget, per cluster, user, etc.

Users often find the policy and governance pieces the most unique and valuable. They enable administrators to provide an SLA for compute as opposed to managing compute directly, within the proper limits and restrictions per user. Effectively, they can offer extreme flexibility for compute without the risk of an unlimited bill in a secure, managed way.

Azure CycleCloud Features

- Dynamic scaling of EDA workloads based on work queues, from one to thousands of instances.
- Support for almost any job scheduler, application stack, or cluster configuration.
- Enable existing workflows for the cloud without requiring application rewrites.
- Governance and policy enforcement, including AD integration, cost reporting, cost controls, and extensive error handling capabilities.
- Job submission, monitoring, and administration.
- Data transfer on-demand or on-schedule.
- SubmitOnce™ technology for seamless submission to internal or Azure resources (Hybrid)

More information on CycleCloud can be found here: https://aka.ms/AzureCycleCloud.

Storage

Azure CycleCloud can orchestrate typical HPC storage resources such as NFS file systems, Parallel Virtual File Systems (PVFS), such as Lustre or BeeGFS, and Azure's own Avere Virtual Edge filer (vFXT). Storage components can be defined and managed both separately or as part of compute clusters. This enables administrators and users to manage data lifecycle separately from compute life cycle which optimizes for cost, access, and performance as appropriate.

Storage Services

Azure storage provides the following four services: Blob, Table, Queue, and File storage.

- Blob Storage stores unstructured object data. A Blob can be any type of text or binary data, such as a document, media file, or application installer.
- Block Blob storage is used for streaming and storing documents, videos, pictures, backups, and other unstructured text or binary data.
- Page Blobs should be used when random write operations are required.
- Disk Storage is used by Azure VMs for persistent storage of data (ex: OS disk). Many EDA workflows use OS Disk (local SSD) as a write temp space to reduce the tax on the shared NFS filer. Azure supports the use of multiple Disks with each VM. On the back-end, this storage type uses Page Blobs.
- Azure VM's and cloud services can share file data across application components via mounted shares, and on-premises applications can access file data in a share via the File service REST API.
- Azure Avere provides a POSIX compliant interface to Blob and supports Tcl/Tk, which has been an Semiconductor industry standard language for scripting. This allows users to read and write to Blob storage without having to spend engineering time or money to rewrite existing legacy applications that cannot communicate directly to Blob storage.

Page Blob / Disk Storage

Page blobs are optimized for representing IaaS disks and supporting random writes and can be up to 1 TB in size. An Azure VM network attached IaaS disk is a VHD stored as a page blob.

Azure Avere NFS File Systems

EDA workflows can involve well over 10,000 cores on premise, driving up to a million metadata NFS ops. If a single node VM is not sufficient to provide enough performance, Azure has a scale-out, caching, tiered file system to offer. Azure Avere is that file system. Each Avere vFXT node in the cluster, provides 70,000 random NFS ops. A ten-node cluster would provide 700,000 NFS ops and it can scale to 24 nodes in a single cluster.

Azure Avere file system provides tiering for EDA applications. There are three levels of storage with Azure Avere. Hot data and metadata is stored and served by DRAM. Warm data is stored in SSDs. All other data (cold) is stored on Azure Blob.

The Azure cold storage tier is optimized for storing data that is rarely accessed and long lived. Data in the cold storage tier can tolerate a slightly lower availability, but still requires high durability and similar but less performance. For cold data, slightly lower availability and higher access costs are acceptable tradeoffs for much lower storage costs. In typical application environments, more than 90 percent of the total data is cold, so Blob storage provides a dense, very cost-effective tier.

The archive storage tier is optimized for storing data that is rarely accessed and stored for at least six months.

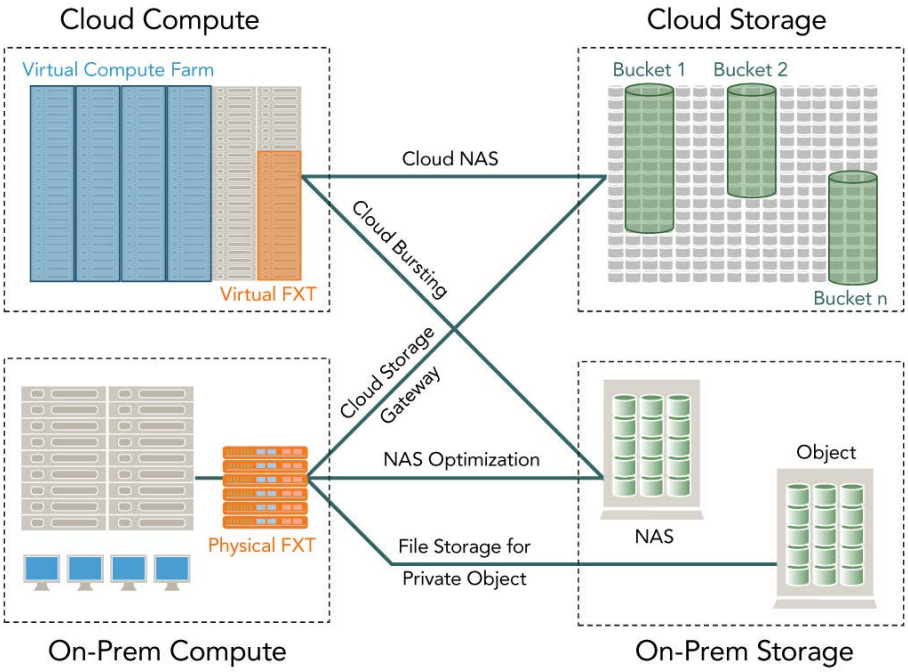
At a high level, Azure Avere offers four solutions to EDA customers:

- All-cloud, a good option for new companies with no investment in on-premises infrastructure.
- Cloud bursting to cloud resources to handle peak or unexpected demand.
- Seamless migration of data and file-based enterprise applications with a goal of using Azure as permanently provisioned IT infrastructure.
- Cloud Storage Gateway that provides cost effective storage for rarely access data.

Azure Avere Cloud NAS

Cloud NAS is powered by the Avere Virtual Edge filer (vFXT), a software only file server that bridges the storage cloud and Azure compute. The Avere vFXT provides a fully functional NAS to support both NFS and SMB applications running in Azure. Clustering enables this Cloud NAS architecture to scale compute cloud resources to massive performance and capacity levels while ensuring high availability for data.

For high availability (HA), horizontal clustering provides N+1 redundancy with active/active node failover. At the node-level, self-healing protects against outages.



Azure Avere Cloud Bursting

In Cloud Bursting use cases, customers also use the Avere vFXT to access the public cloud compute, but access data in on-premises storage. The near infinite supply of Azure compute nodes can serve as a supplement to on-premises compute for legacy applications. For example, customers can burst EDA or rendering, genomics, and financial batch jobs to the cloud, rapidly applying tens or many tens of thousands of compute cores to accommodate peak or unexpected demand. Data for these applications can remain in place on an enterprise's existing NAS or private object storage. The Avere vFXT automatically caches active data in the compute cloud and hides the latency to the customer's on-premises datacenter.

With Azure Avere's cloud bursting architecture, data is automatically moved between compute and storage resources, whether it is in your datacenter or in Azure. Using Avere's software only technology, complexity of data movement and synchronization is eliminated. Storage cluster setup, administration, and teardown are simpler and data is

not locked in to any location. Avere's global namespace ensures simple NAS access with a single mount point.

Azure Avere Blob Storage Gateway

The Blob Storage Gateway technology is similar to the on-premises NAS environment with the added benefit that customers can take advantage of Azure's Blob to reduce requirements for on-premises capacity (and datacenter real estate). The Avere FXT Edge Filer automatically retrieves data from the cloud and caches it on premises so that all subsequent accesses occur at low latency.

Parallel Virtual File Systems

For those workloads where a single NFS server is not sufficient from either a capacity or performance perspective, several open-source Parallel Virtual File Systems (PVFSs) such as Lustre, GlusterFS, and BeeGFS are available in the Microsoft Azure Marketplace. These solutions distribute file data across multiple servers and provide concurrent access by multiple tasks spread across multiple client machines. Used in high performance computing (HPC) environments, a PVFS delivers high-performance access to large data sets.

PVFS clusters include nodes designated as management, storage, and metadata servers. Storage servers hold file data, while metadata servers store statistics, attributes, data file-handles, directory entries, and other metadata. The clients access the file system across a network.

Lustre

Lustre is currently the most widely used PVFS in HPC solutions. Lustre file systems can scale to tens of thousands of client nodes and tens of petabytes of storage.

A Lustre file system has three major functional units:

- One or more metadata servers (MDS) nodes that has one or more metadata target (MDT) devices per Lustre filesystem that stores namespace metadata, such as filenames, directories, access permissions, and file layout.
- One or more object storage server (OSS) nodes that store file data on one or more object storage target (OST) devices. Depending on the server's hardware, an OSS typically serves between two and eight OSTs, with each OST managing a single local disk filesystem. The capacity of a Lustre file system is the sum of the capacities provided by the OSTs.
- Client(s) that access and use the data. Lustre presents all clients with a unified namespace for all the files and data in the filesystem, using standard POSIX

semantics, and allows concurrent and coherent read and write access to the files in the filesystem.

GlusterFS

The GlusterFS file system is a free, scalable, open source distributed file system specifically optimized for cloud storage, and it works well for applications hosted on Azure. Unlike Lustre, it does not have separate metadata servers - metadata is integrated into the file storage.

The GlusterFS architecture includes a server node and a client node in a cluster. A single node can be designated for both storage and metadata eliminating the need for separate servers. GlusterFS stores and locates data using an elastic hash algorithm and doesn't need any metadata servers so I/O bottlenecks are removed, which improves performance and parallel access. Servers can be added or removed whenever required.

BeeGFS

BeeGFS is a Linux-based, hardware-independent parallel file system designed for high performance and high throughput environments. BeeGFS is free to use and offers commercial support.

One advantage of the BeeGFS client service is that it provides a normal mount point that your applications can use to directly access the BeeGFS storage system. Compared to Lustre, BeeGFS offers more flexibility, greater robustness, and ease of use on top of enhanced performance and scalability.

Durability and High Availability

The data in your Microsoft Azure storage account is always replicated to ensure durability and high availability. Azure Storage replication copies your data so that it is protected from planned and unplanned events ranging from transient hardware failures, network or power outages, massive natural disasters, and so on. You can choose to replicate your data within the same data center, across zonal data centers within the same region, and even across regions. When you create a storage account, you must select one of the following replication options:

- Locally redundant storage (LRS). Locally redundant storage maintains three copies of your data. LRS is replicated three times within a single facility in a single region. LRS protects your data from normal hardware failures, but not from the failure of a single facility.
- Zone redundant storage (ZRS). Zone-redundant storage maintains three copies of your data. ZRS is replicated three times across two to three facilities, either within a

single region or across two regions, providing higher durability than LRS. ZRS ensures that your data is durable within a single region.

- Geo redundant storage (GRS). Geo redundant storage maintains six copies of your data. With GRS, your data is replicated three times within the primary region and is also replicated three times in a secondary region hundreds of miles away from the primary region, providing the highest level of durability. In the event of a failure at the primary region, Azure Storage will failover to the secondary region. GRS ensures that your data is durable in two separate regions.
- Read access geo redundant storage (RA-GRS). Read access geo redundant storage replicates your data to a secondary geographic location and provides read access to your data in the secondary location. Read-access geo redundant storage allows you to access your data from either the primary or the secondary location, if one location becomes unavailable.

Encryption

Azure Storage Service Encryption (SSE) for Data at Rest helps you protect and safeguard your data to meet your organizational security and compliance commitments. With this feature, Azure Storage automatically encrypts your data prior to persisting to storage and decrypts prior to retrieval. The encryption, decryption, and key management are totally transparent to users.

SSE works by encrypting the data when it is written to Azure Storage and can be used for Azure Blob Storage and File Storage. It works for the following:

- General purpose storage accounts and Blob storage accounts.
- Standard storage and Premium storage.
- All redundancy levels (LRS, ZRS, GRS, RA-GRS.)
- Azure Resource Manager storage accounts but not classic.
- All regions for Blob Storage.

NetApp ONTAP Cloud

NetApp ONTAP Cloud is deployed using OnCommand Cloud Manager to deliver secure, proven NFS, SMB data management for Azure cloud storage. A software only storage service running the ONTAP storage operating system, ONTAP Cloud combines data control with enterprise-class storage features such as data deduplication and compression to minimize your Azure storage footprint. It also enables you to take snapshots of your data without requiring additional storage or impacting your

application's performance. ONTAP Cloud can tie your cloud storage to your data center using the industry leading NetApp replication protocol, SnapMirror technology. OnCommand Cloud Manager handles deployment and management of ONTAP Cloud, giving you a simple point-and-click environment to manage your storage and ease control of your data.

Microsoft Azure Enterprise Network File System (NFS) service, powered by NetApp

The new, industry first enterprise Network File System (NFS) service in the cloud is delivered natively in Azure and powered by NetApp.

More information on Azure Storage can be found at https://azure.microsoft.com/en-us/services/storage/

Networking

Virtual Networks

Azure resources such as VMs and VM Scale Sets can communicate privately with each other through an Azure Virtual Network (VNet). A VNet is a logical isolation of the Azure cloud dedicated to your subscription. You can implement multiple VNets within each Azure subscription and Azure region. Each VNet is isolated from other VNets. For each VNet you can:

- Specify a custom private IP address space using public and private (RFC 1918) addresses. Azure assigns resources connected to the VNet a private IP address from the address space you assign.
- Segment the VNet into one or more subnets and allocate a portion of the VNet address space to each subnet.
- Use Azure provided name resolution or specify your own DNS server for use by resources connected to a VNet.

You can connect VNets to each other, enabling resources connected to either VNet to communicate with each other across VNets. You can use either or both of the following options to connect VNets to each other:

- **Peering:** Enables resources connected to different Azure VNets within the same Azure region to communicate with each other. The bandwidth and latency across the VNets is the same as if the resources were connected to the same VNet.

- **VPN Gateway:** Enables resources connected to different Azure VNets within different Azure regions to communicate with each other. Traffic between VNets flows through an Azure VPN Gateway. Bandwidth between VNets is limited to the bandwidth of the gateway.

On-premises connectivity

You can access resources in your VNet securely over either a VPN connection or a direct private connection. To send network traffic between your Azure virtual network and your on-premises network, you must create a virtual network gateway. For the gateway, settings can be configured to create the type of connection that you want, either VPN or ExpressRoute.

You can connect your on-premises network to a VNet using any combination of the following options:

Point-to-site (VPN over SSTP)

This connection is established between a single computer and a VNet. This connection type is great for designers who work remotely and are not connected to the Corporate network, because it requires little or no changes to your existing network. Point-to-site connections are often coupled with a site-to-site connection through the same virtual network gateway. The connection uses the SSTP protocol to provide encrypted communication over the Internet between the computer and the VNet. The latency for a point-to-site VPN is unpredictable, since the traffic traverses the Internet.

Site-to-site (IPsec/IKE VPN tunnel)

This connection is established between your on-premises VPN device and an Azure VPN Gateway. This connection type enables any on-premises resource that you authorize to access the VNet. The connection is an IPSec/IKE VPN that provides encrypted communication over the Internet between your on-premises device and the Azure VPN gateway. You can connect multiple on-premises sites to the same VPN gateway. The latency for a site-to-site connection is unpredictable, since the traffic traverses the Internet.

ExpressRoute (dedicated private connection)

This type of connection is established between your network and Azure, through an ExpressRoute partner. This connection is private. Traffic does not traverse the Internet. The latency for an ExpressRoute connection is predictable, since traffic doesn't traverse the Internet. ExpressRoute can be combined with a site-to-site connection.

Accelerated Networking

Accelerated Networking provides consistent ultra- low network latency via Azure's in-house programmable hardware and technologies such as SR-IOV. By moving much of Azure's software-defined networking stack off the CPUs and into FPGA-based SmartNICs, compute cycles are reclaimed by end user applications, putting less load on the VM, decreasing jitter and inconsistency in latency. Benefits include:

- **Lower Latency / Higher packets per second (PPS):** Removing the virtual switch from the datapath removes the time packets spend in the host for policy processing and increases the number of packets that can be processed inside the VM.
- **Reduced jitter:** Virtual switch processing depends on the amount of policy that needs to be applied and the workload of the CPU that is doing the processing. Offloading the policy enforcement to the hardware removes that variability by delivering packets directly to the VM, removing the host to VM communication and all software interrupts and context switches.
- **Decreased CPU utilization:** Bypassing the virtual switch in the host leads to less CPU utilization for processing network traffic.

More information on Azure Networking can be found at https://azure.microsoft.com/en-us/product-categories/networking/

Managing EDA Licensing

Most EDA applications require a license from the vendor whether the application is running on-premises or on Azure. By leveraging either a VPN or ExpressRoute connection, the VMs running in Azure can obtain license tokens from the existing on-premises license server. A separate license server could also be provisioned on a VM in Azure.

Security & Governance

Securing Azure VMs can include one or more Azure services and features that secure access to VMs and secure storage of data.

Azure Security Center

Azure Security Center helps prevent, detect, and respond to threats to VMs. Security Center provides integrated security monitoring and policy management across customer Azure subscriptions, helps detect threats that might otherwise go unnoticed, and works with a broad ecosystem of security solutions.

Encryption

For enhanced Windows VM and Linux VM security and compliance, virtual disks in Azure can be encrypted. Virtual disks on Windows VMs are encrypted at rest using Bitlocker. Virtual disks on Linux VMs are encrypted at rest using dm-crypt.

There is no charge for encrypting virtual disks in Azure. Cryptographic keys are stored in Azure Key Vault using software-protection, or customers can import or generate keys in Hardware Security Modules (HSMs) certified to FIPS 140-2 level 2 standards. These cryptographic keys are used to encrypt, and decrypt virtual disks attached to a VM. The customer retains control of these cryptographic keys and can audit their use. An Azure Active Directory service principal provides a secure mechanism for issuing these cryptographic keys as VMs are powered on and off.

Key Vault and SSH Keys

Secrets and certificates can be modeled as resources and provided by Key Vault. Azure PowerShell can be used to create key vaults for Windows VMs and the Azure CLI for Linux VMs. Keys for encryption can also be created.

Key vault access policies grant permissions to keys, secrets, and certificates separately. For example, it is possible to grant user access to only keys, but no permissions for secrets. However, permissions to access keys or secrets or certificates are at the vault level. In other words, key vault access policy does not support object level permissions.

When connecting to VMs, public-key cryptography should be used to provide a more secure way to log in to them. This process involves a public and private key exchange using the secure shell (SSH) command to authenticate the user rather than a username and password. Passwords are vulnerable to brute-force attacks, especially on Internet-facing VMs such as web servers. With a secure shell (SSH) key pair, it is possible to create a Linux VM that uses SSH keys for authentication, eliminating the need for passwords to log in. SSH keys can also be used to connect from a Windows VM to a Linux VM.

Policies

Azure policies can be used to define the desired behavior for the organization's Windows VMs and Linux VMs. By using policies, an organization can enforce various conventions and rules throughout the enterprise. Enforcement of the desired behavior can help mitigate risk while contributing to the success of the organization.

Role-based access control

Using role-based access control (RBAC), it is possible to segregate duties within your team and grant only the amount of access to users of the VM that they need to perform their jobs. Instead of giving everybody unrestricted permissions on the VM, it is possible to allow only certain actions. Access control can be configured for the VM in the Azure portal, using the Azure CLI, or Azure PowerShell.

Azure Monitor

Azure Monitor is part of Microsoft Azure's overall monitoring solution. Azure Monitor helps track performance, maintain security, and identify trends.

Azure CLI

The Azure CLI 2.0 is Microsoft's cross-platform command line experience for managing Azure resources. It can be used in browser with Azure Cloud Shell, or install it on macOS, Linux, or Windows and run it from the command line.

Azure CLI 2.0 is optimized for managing and administering Azure resources from the command line, and for building automation scripts that work against the Azure Resource Manager.

Azure Infrastructure Automation Tools

To create and manage Azure virtual machines (VMs) in a consistent manner at scale that EDA workflows require, some form of automation is typically desired. Azure supports many tools and solutions that automate the complete Azure infrastructure deployment and management lifecycle.

Automate the configuration of VMs:

- Tools include Ansible, Chef, and Puppet.
- Tools specific to VM customization include cloud-init for Linux VMs, PowerShell Desired State Configuration (DSC), and the Azure Custom Script Extension for all Azure VMs.

Automate infrastructure management:

- Tools include Packer to automate custom VM image builds, and Terraform to automate the infrastructure build process.
- Azure Automation can perform actions across your Azure and on-premises infrastructure.

Automate application deployment and delivery:

- Examples include Visual Studio Team Services and Jenkins.

Azure Cost Management

Azure Cost Management by Cloudyn is a multi-cloud cost management solution that helps utilize and manage Azure and other cloud resources.

Document Updates and Contributors

Version 1 – May 2018

Authors:

- Mujtaba Hamid, Principal PM Manager, Industry Verticals, Azure Global Engineering
- Steve Roach, Azure Global Black Belt, Intelligent Cloud

Contributors:

- Rob Futrick, Principal Program Manager, Azure High-Performance Compute
- Andy Chan, EDA specialist, Azure Storage
- Preeth Chengappa, Director, Semiconductors & EDA, Azure Global Engineering