

Enabling our HPC Customers

HPC CAT Team



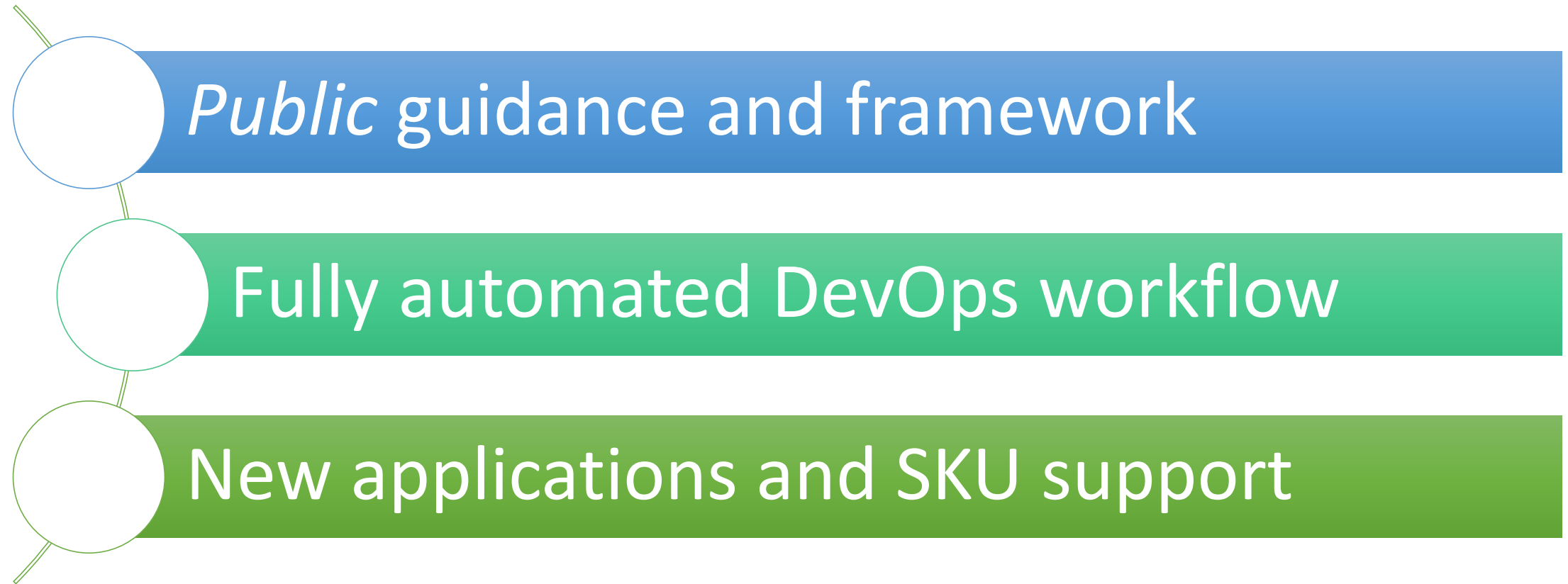
Application catalog: Recap

Goal: Framework we created to accelerate the testing of high-performance computing (HPC) applications using Azure Batch pools or Azure CycleCloud clusters.

Where to find it:

[Public GitHub](#)

<https://github.com/az-cat/az-hpcapps>



Current Status

Chemistry

- NAMD
- LAMMPS
- GAMESS
- Quantum Espresso
- NWChem

Bioscience

- Blast
- MrBlayes
- AutoDock
- HMMER
- GATK
- Gromacs

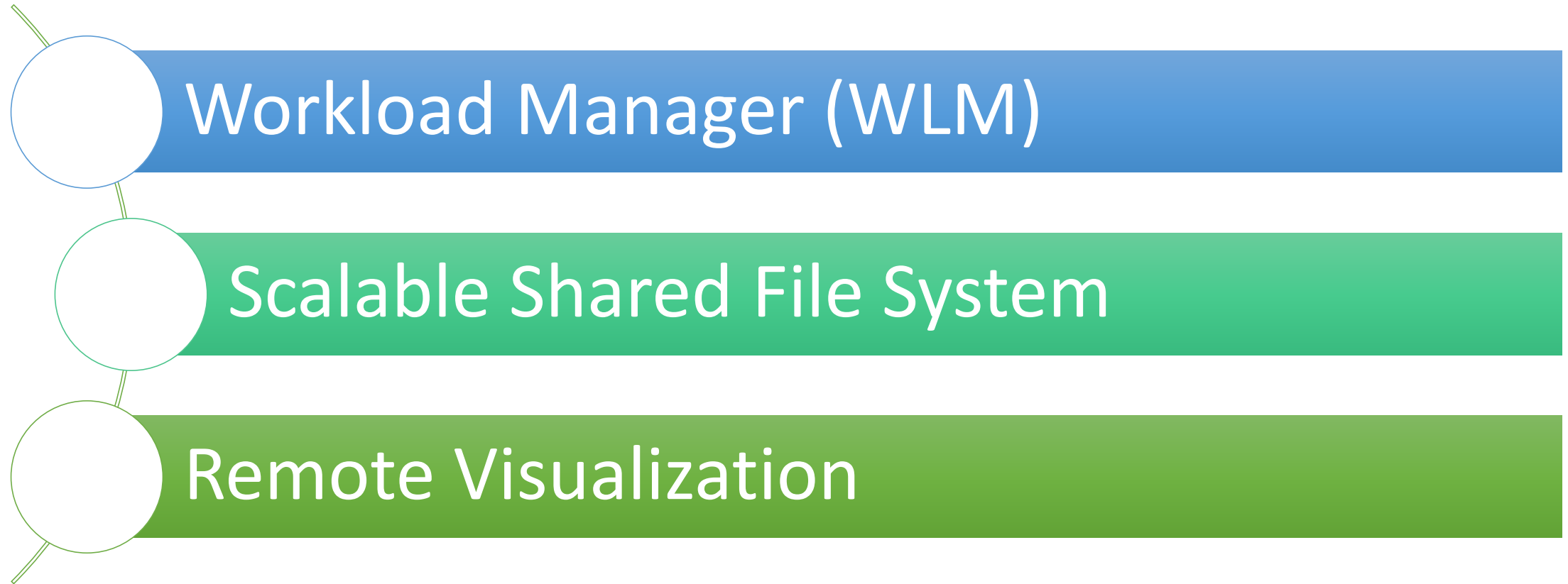
Fluid Dynamics

- Fluent
- OpenFOAM
- CFX
- StarCCM+
- ConvergeCFD

Structural Analysis

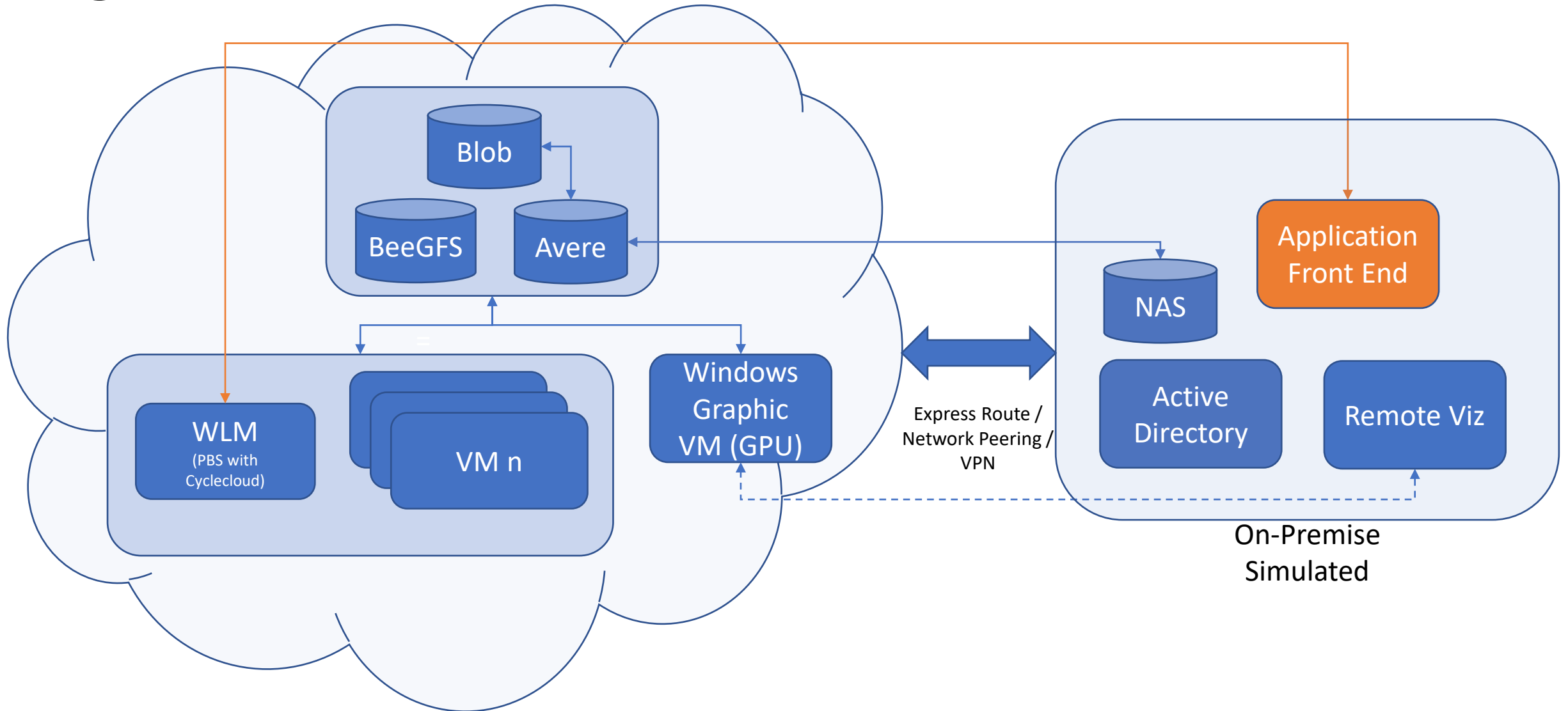
- Abaqus
- Ansys Mechanical
- PamCrash
- Isdyna

Applications



Requirements

High level view of the architecture



CycleCloud

- Autoscaling capabilities are still a work in progress and you will encounter some issues depending on how sophisticated the job submission scripts are
- Does a good job of standing up the infrastructure needed for our reference architecture

Storage

- Avere performed well serving up software stack and data (read only)
- Avere not recommended for heavy write I/O (and parallel I/O to shared file)
- Avere automated deployment was not repeatable when configuring core filers.
- Use local SSD's if possible.
- BeeGFS good for large I/O, not ideal for small I/O.

Visualization

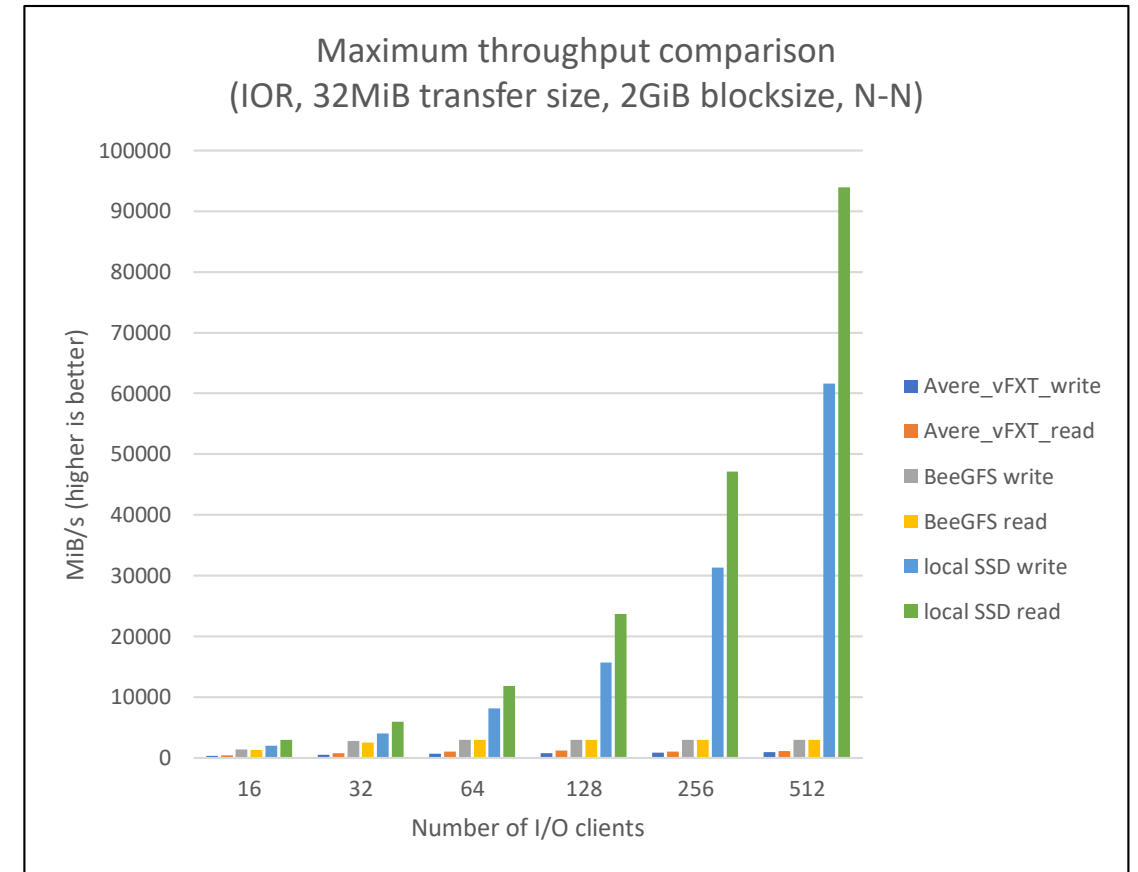
- Provided a good remote visualization experience for large data sets
- Lack of Accelerated Networking limits remote IO performances

Learnings and Outcomes

Key storage performance outcomes

If I/O is critical use local SSD's (if possible)

- When to use local SSD's
 - Do not need a global shared filesystem
 - Have sufficient capacity (Each H16r has 2TB local SSD)
 - temporary/scratch disk is acceptable
- 60GB/s BeeGFS (you would need >75 VM's and >300 P30 disks)



Note:

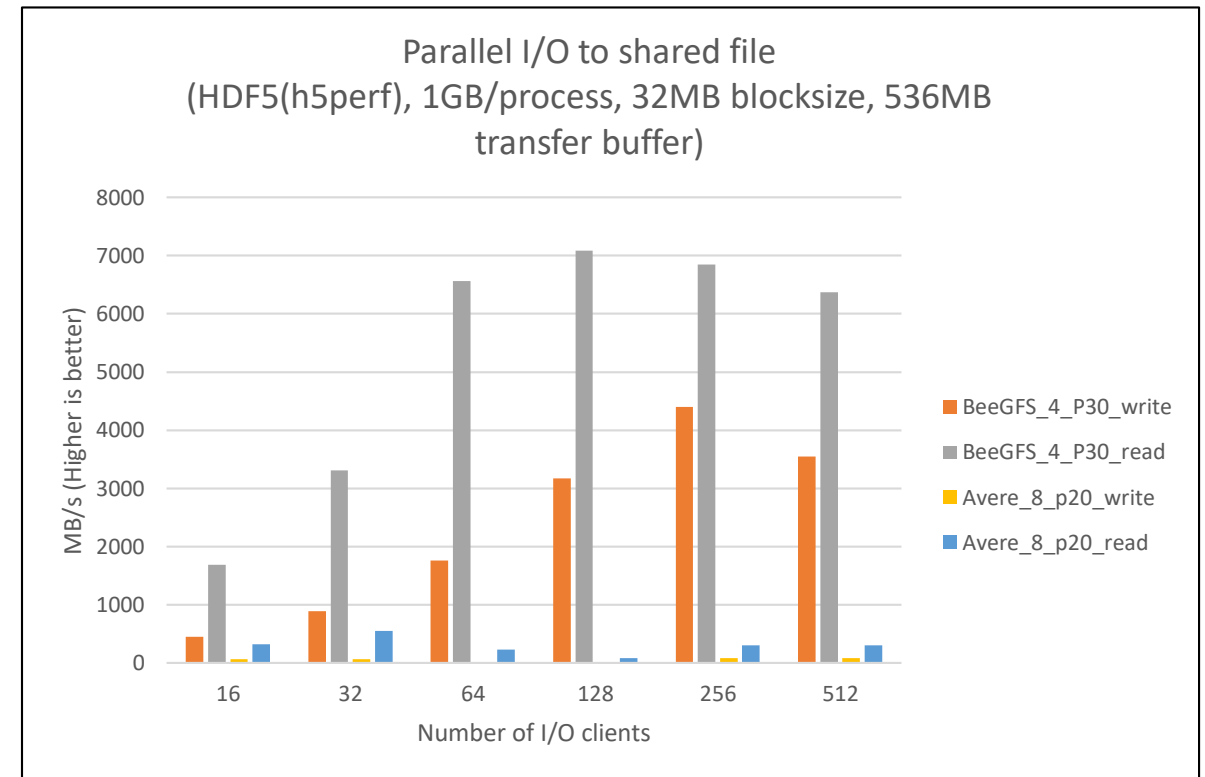
- 16 clients = 1 node (H16r)

Avere vFXT positives

- Works well syncing software stack and datasets with on-prem storage (read only).
- Software stack available via modules (Lmod).
 - Easy to deploy new software (No modifications to CycleCloud necessary).

Avere vFXT does not work well with parallel shared file I/O (HDF5)

- Intersect runs were 7-8x slower on Avere vFXT compared to BeeGFS.
- Intersect uses HDF5 (file format) for some of its I/O.



Note:

- May include some linux I/O buffering
- 16 clients = 1 node (H16r)

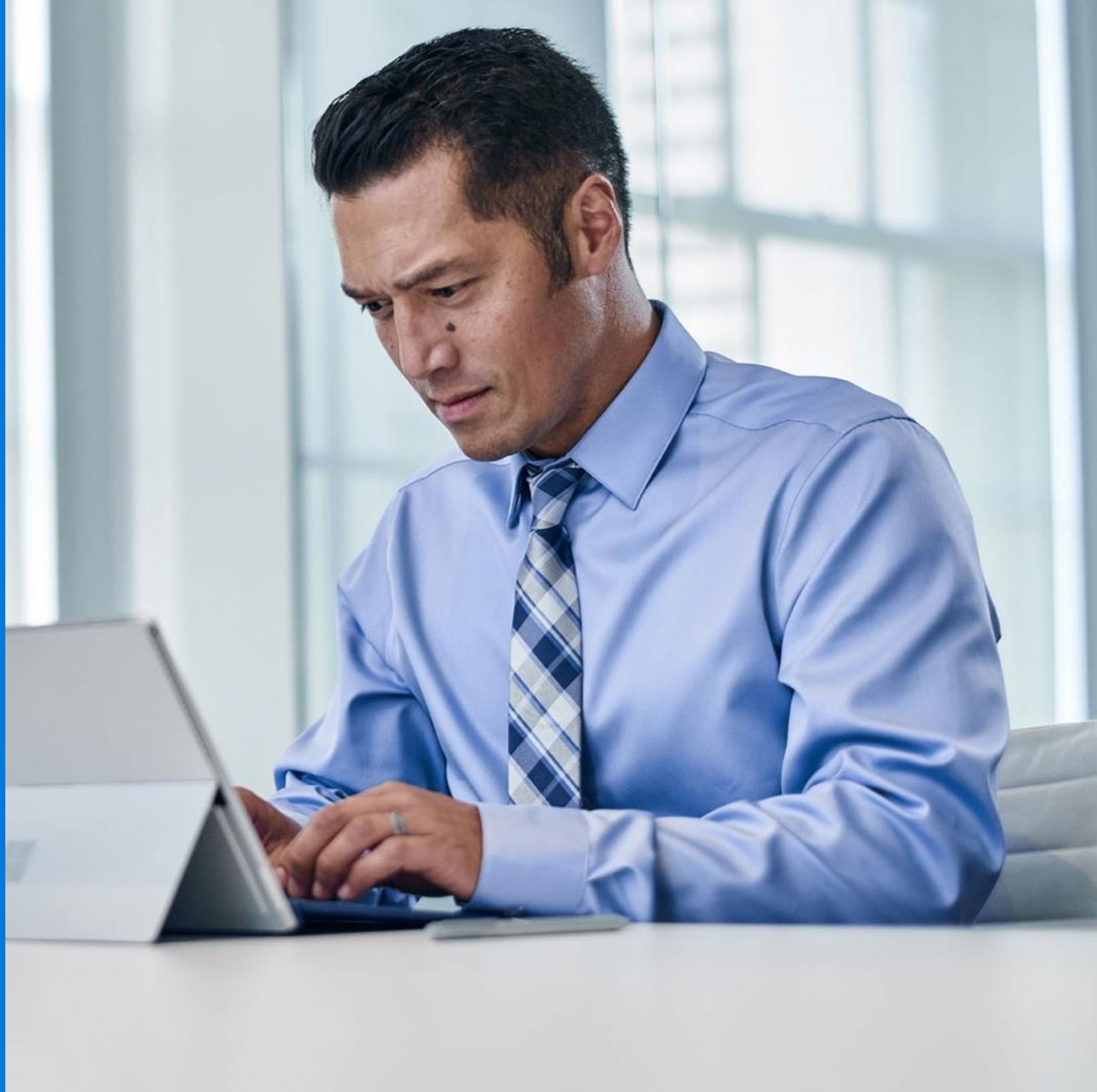
Future Work for Storage/Filesystems in HPC

- eBook [“Parallel Virtual FileSystems on Microsoft Azure”](#)
- More to come.
 - Define a standard set of comprehensive I/O benchmarks.
 - Detailed performance evaluation of different storage/filesystem options
 - BeeGFS, Lustre, BeeOND, Avere vFXT, Azure NetApp Files (ANF), ultra SSD, premium Files, local SSD, Lv2 VM's, Cray ClusterStor
 - Build guidance to what storage option will work best for different Customer I/O requirements.

Setup Steps

- Configure
 - Add your specific Azure information to the config.json file
- Deploy
 - Run deploy_all.sh from the same directory as the configuration script
- Avere
 - Login to the Avere controller and configure your storage
- Cycle Cloud
 - Login to portal to review setup and add/remove nodes
- PBS headnode
 - Login to cyclecloud server, then to the PBS headnode and submit jobs

Demo



- Terminate
- ✎ Edit
- ↗ Share
- ↺ Refresh

State **Started** at 11/19/18 1:50 PM (up 56d 19h 50m) - [View in Portal](#)

Nodes **1** ready, **8** acquiring

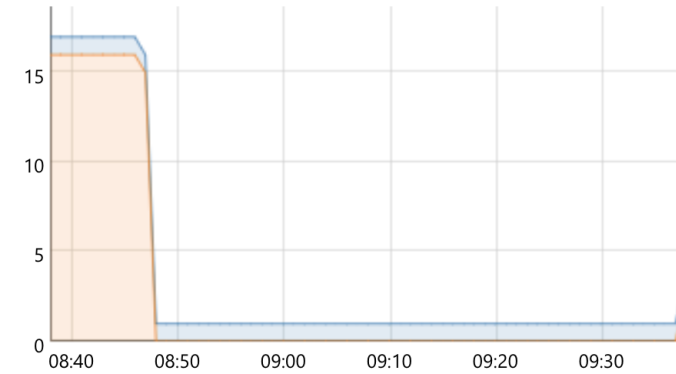
Scalesets **1** configuring




















Size **9** instances, **136** cores (**\$21.7** per hour)

Usage **4.7k core-hours (~\$766)** in the last 24 hours

Alerts [🔔 Create new alert](#)

Volumes **0** volumes, **0 B**

[illegible]

 Show Detail		 Search
Time	Message	
9:38 AM	 Added 8 nodes of type execute to cluster p	
1/14/19 5:37 PM	 Node execute-19 in cluster pbspro-1 finish	
1/14/19 5:21 PM	 Node execute-20 in cluster pbspro-1 finish	
1/14/19 5:20 PM	 Node execute-32 in cluster pbspro-1 finish	
1/14/19 5:20 PM	 Node execute-23 in cluster pbspro-1 finish	
1/14/19 5:20 PM	 Node execute-25 in cluster pbspro-1 finish	
1/14/19 5:19 PM	 Node execute-24 in cluster pbspro-1 finish	
1/14/19 5:19 PM	 Node execute-31 in cluster pbspro-1 finish	
1/14/19 5:19 PM	 2 nodes in cluster pbspro-1 finished startin	
1/14/19 5:19 PM	 Node execute-30 in cluster pbspro-1 finish	
1/14/19 5:18 PM	 Node execute-17 in cluster pbspro-1 finish	
1/14/19 5:18 PM	 Node execute-18 in cluster pbspro-1 finish	
1/14/19 5:18 PM	 2 nodes in cluster pbspro-1 finished startin	
1/14/19 5:18 PM	 Node execute-28 in cluster pbspro-1 finish	
1/14/19 5:17 PM	 Node execute-21 in cluster pbspro-1 finish	
1/14/19 5:07 PM	 Added 16 nodes of type execute to cluster	
1/14/19 4:28 PM	 Added 16 nodes of type execute to cluster	

Things to look for

- Documentation will be available in the March timeframe
- HPC Reference Arch Repo: To be shared when the documentation is released
- Application catalog: <https://github.com/az-cat/az-hpcapps>
- Email: hpccattd@microsoft.com