Big Compute

El héroe para nuestros problemas de computación masiva

Jose Angel Fernandez - @jangelfdez
Iria Quiroga - @iriaq

Iria Quiroga

Service Delivery Manager @ Microsoft

@iriaq

iriaquiroga

Jose angel Fernandez

Cloud Solution Architect @ Microsoft

@jangelfdez

jangelfdez

Patrocinadores Locales

avanade

Bravent
IT consulting company

capside
architects of the digital society

clevertask

encamina
PIENSA EN COLORES

enimbos
the cloud enabler company

ERNI

everis
an NTT DATA Company

ilitia

Insight

intelequia
SOFTWARE SOLUTIONS

isolutions

KABEL

myCloudDoor

plain concepts

raona

東京 TOKIOTA

Colabora

Microsoft

Agenda

- HPC on Azure, seriously?
- The key elements
  - Storage
  - Networking
  - Compute
  - Management
- Let's play

HPC on Azure Seriously?

What people initially thinks when we discuss about running High Performance Applications on the Cloud in general?

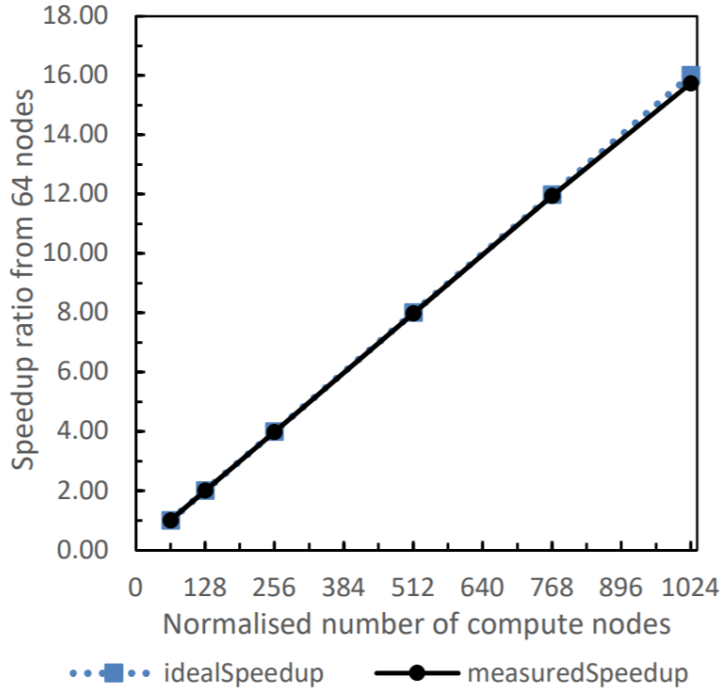
Petascale computing for Terapixel Visualization

Visualization results:

- The first terapixel visualization of an urban digital twin, photo-realistic rendering using Blender Cycles path tracing.
- Tiled image output format makes supercomputer visualization outputs accessible on low cost, thin client devices.
- **Rendering time reduced from an estimated 34 days on one V100 GPU to 48 min**. on 1024 V100 GPU.

Visual supercomputing results:

- Efficiently deployed a 1024 GPU, 14 PFlop cloud supercomputer: faster than any public GPU HPC system in the UK.
- **Cost for one performance scaling graph: <£20,000, providing access to a > £10 million supercomputer**.
- **Azure upgrade** (from K80 to V100) during the project, transparently benefited from a 3x performance increase.
- Enables experiments with future scaled performance approximately 20 years ahead of current workstation systems.

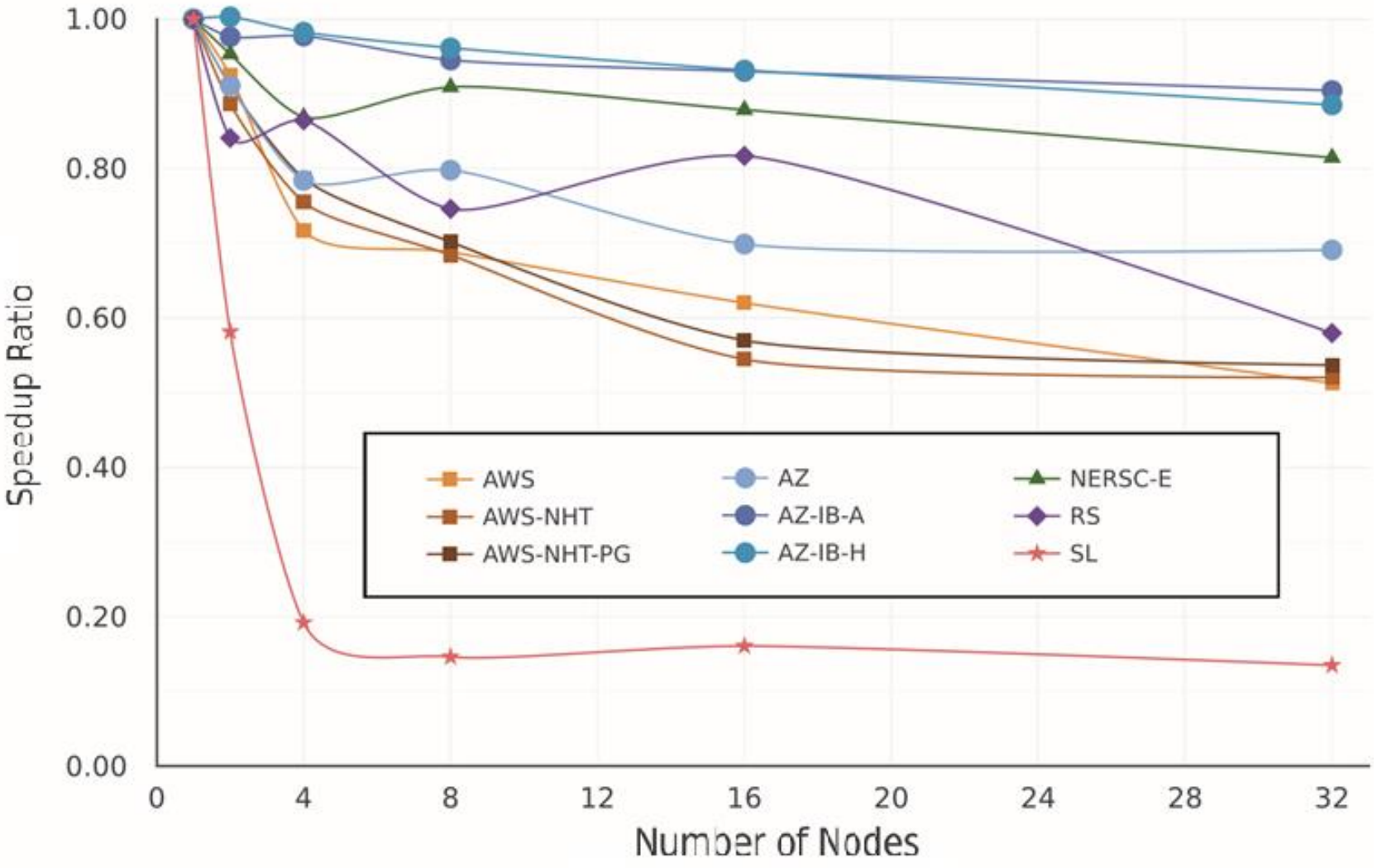


The speedup ratio plotted for the terapixel computation from 64 to 1024 NC6v3 V100 GPU nodes, the speedup shows close to linear scaling achieving 98% efficiency at 1024 nodes

Source: Nicolas S. Holliman, Manu Antony, James Charlton, Stephen Dowsland, Philip James, Mark Turner (Feb 13, 2019) - https://arxiv.org/abs/1902.04820

"Petascale Cloud Supercomputing for Terapixel Visualization of a Digital Twin"

Azure Delivers Best-in-Class Performance

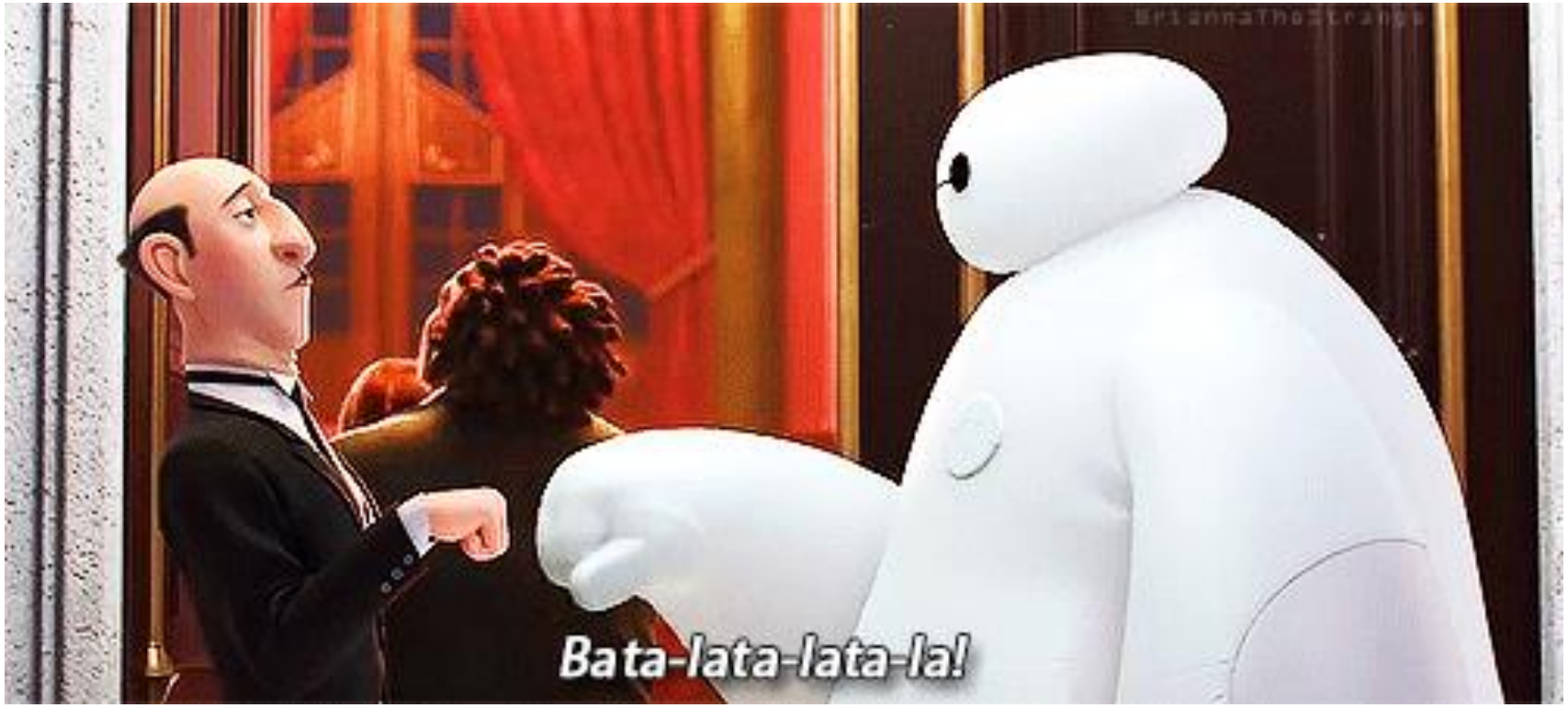


CONCLUSION

"We benchmarked the performance of the best available computing hardware from public cloud providers with high performance Linpack. We optimized the benchmark for each computing environment and evaluated the relative performance for distributed memory calculations. **We found Microsoft Azure to deliver the best results, and demonstrated that the performance per single computing core on public cloud to be comparable to modern traditional supercomputing systems.**"

Source: Mohammadi, M., Bazhirov, T. (Feb 9, 2017) Exabyte Inc. "Comparative benchmarking of cloud computing vendors with High Performance Linpack" (https://arxiv.org/pdf/1702.02968.pdf)

How we finish meetings about HPC after our presentations?



The key elements

Key points on HPC

STORAGE

NETWORKING

COMPUTING

MANAGEMENT

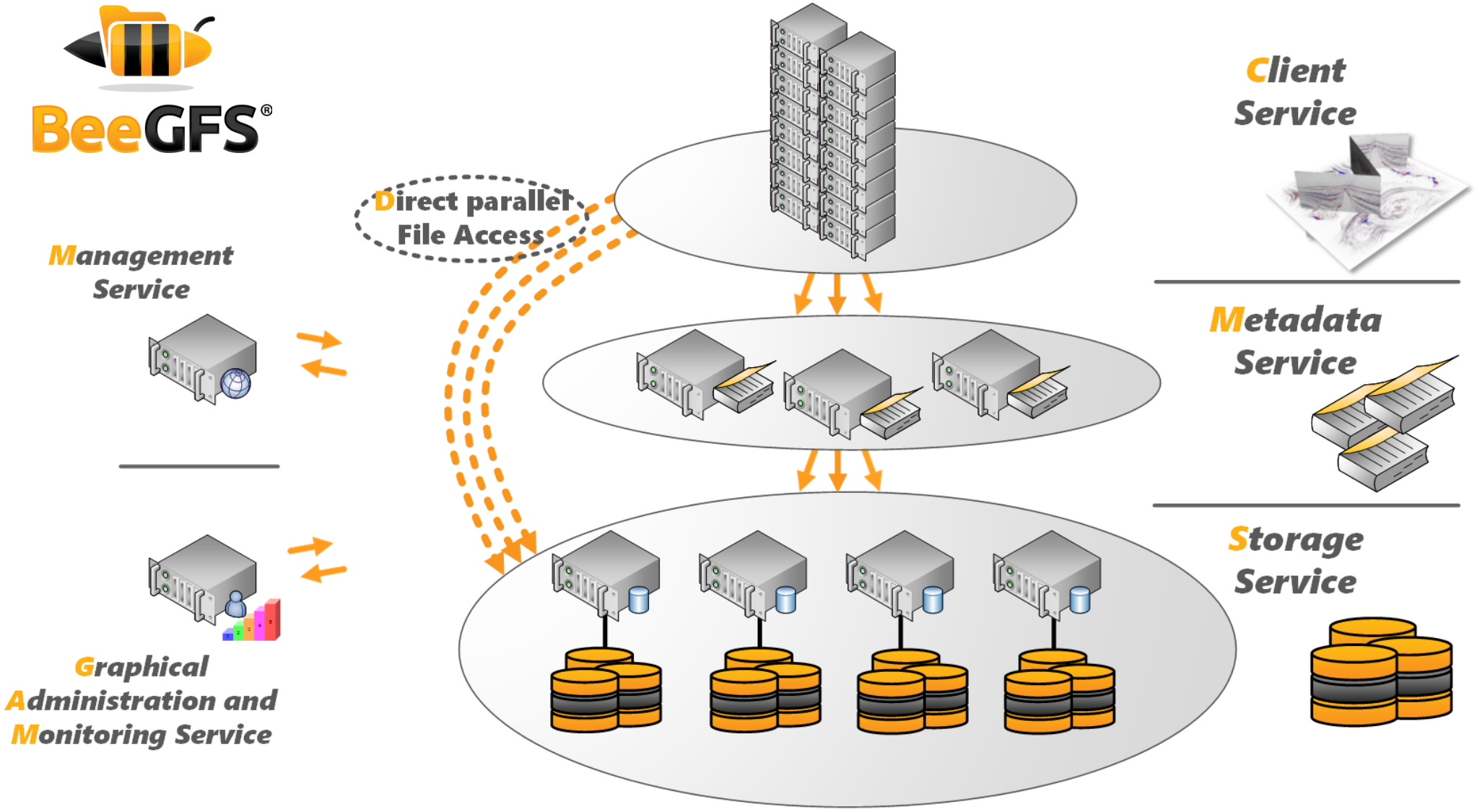
Storage

Distributed network file systems

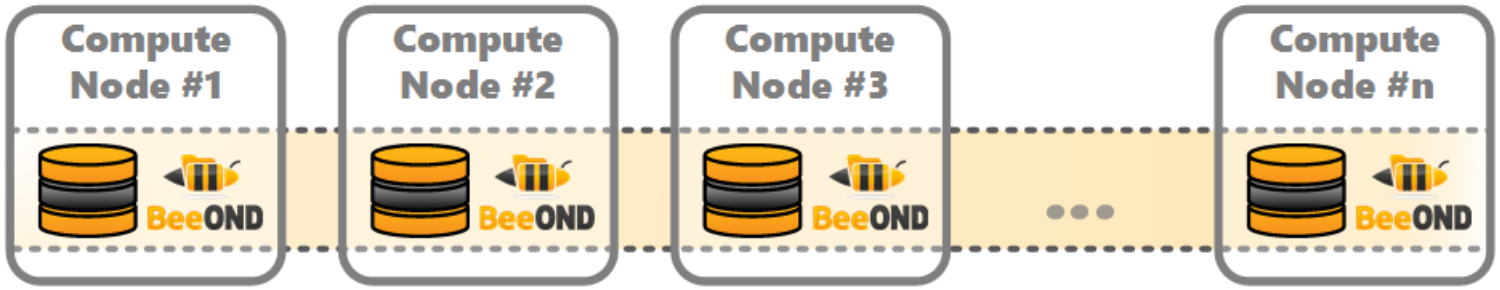
"In computing, a distributed file system (DFS) or network file system is any file system that allows access to files from multiple hosts sharing via a computer network. This makes it possible for multiple users on multiple machines to share files and storage resources"

https://en.wikipedia.org/wiki/Comparison_of_distributed_file_systems

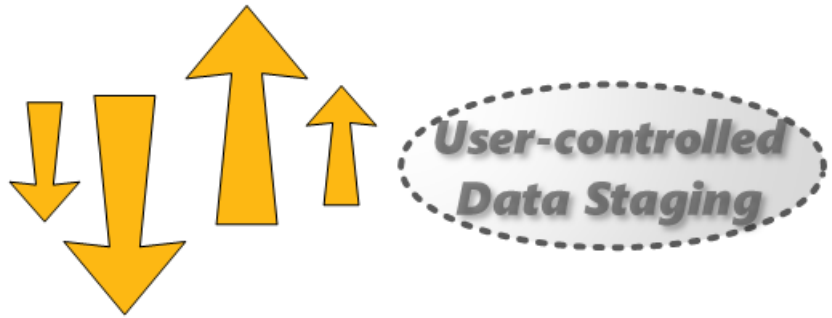
Distributed network file systems: an example



On-demand file system per job



BeeOND



BeeGFS

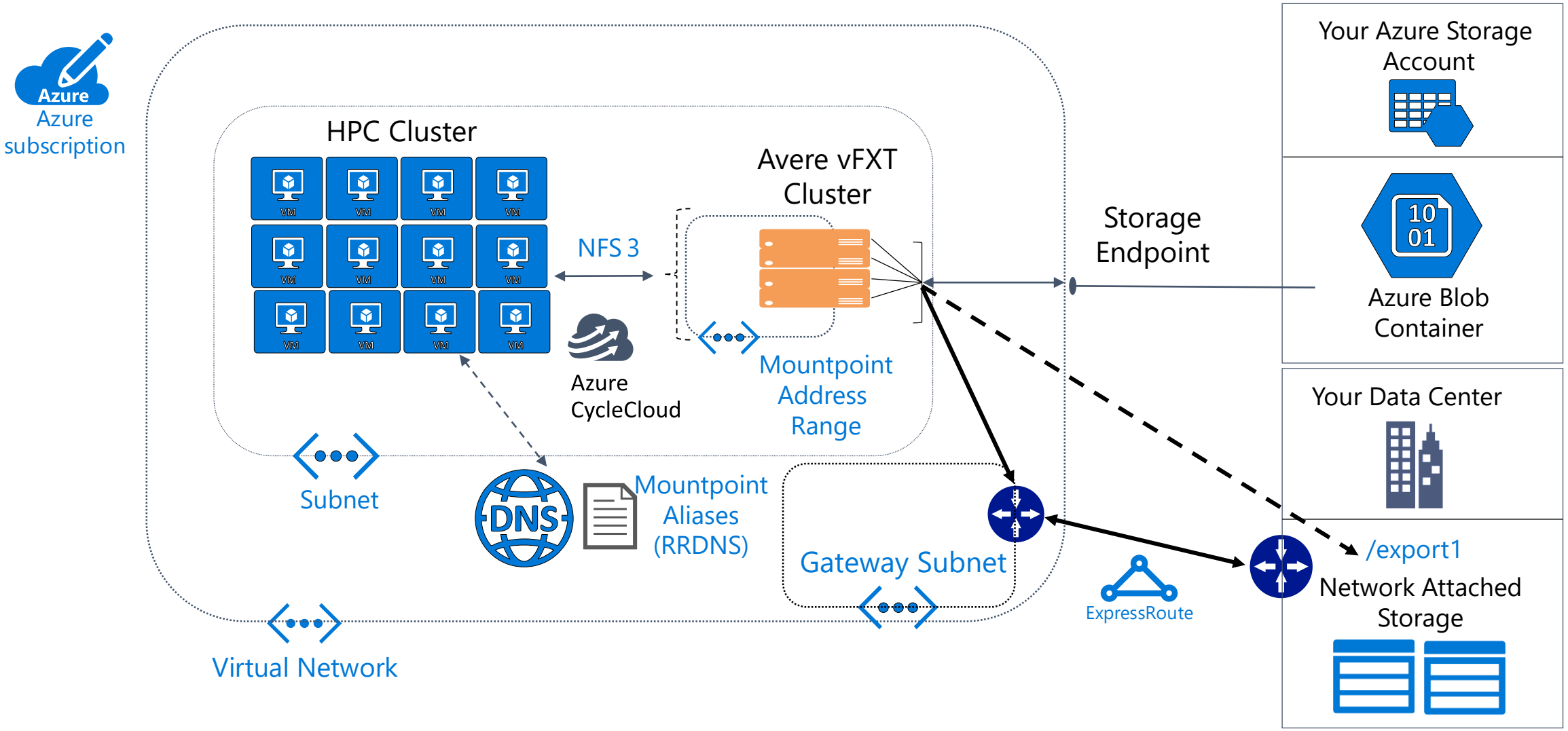
Cloud-based cache layer, Avere

Avere vFXT for Azure provides a ***cloud-based cache layer*** *for data-intensive high performance computing (HPC) tasks. By caching files close to your compute nodes, Avere vFXT speeds read times and lets you work more smoothly even at peak load. Avere vFXT works best with systems that have* ***between 1,000 and 40,000 client cores.***

Avere vFXT can work with your existing on-premises data storage to provide Azure-based computing resources local access to active files that are stored long-term in your datacenter.

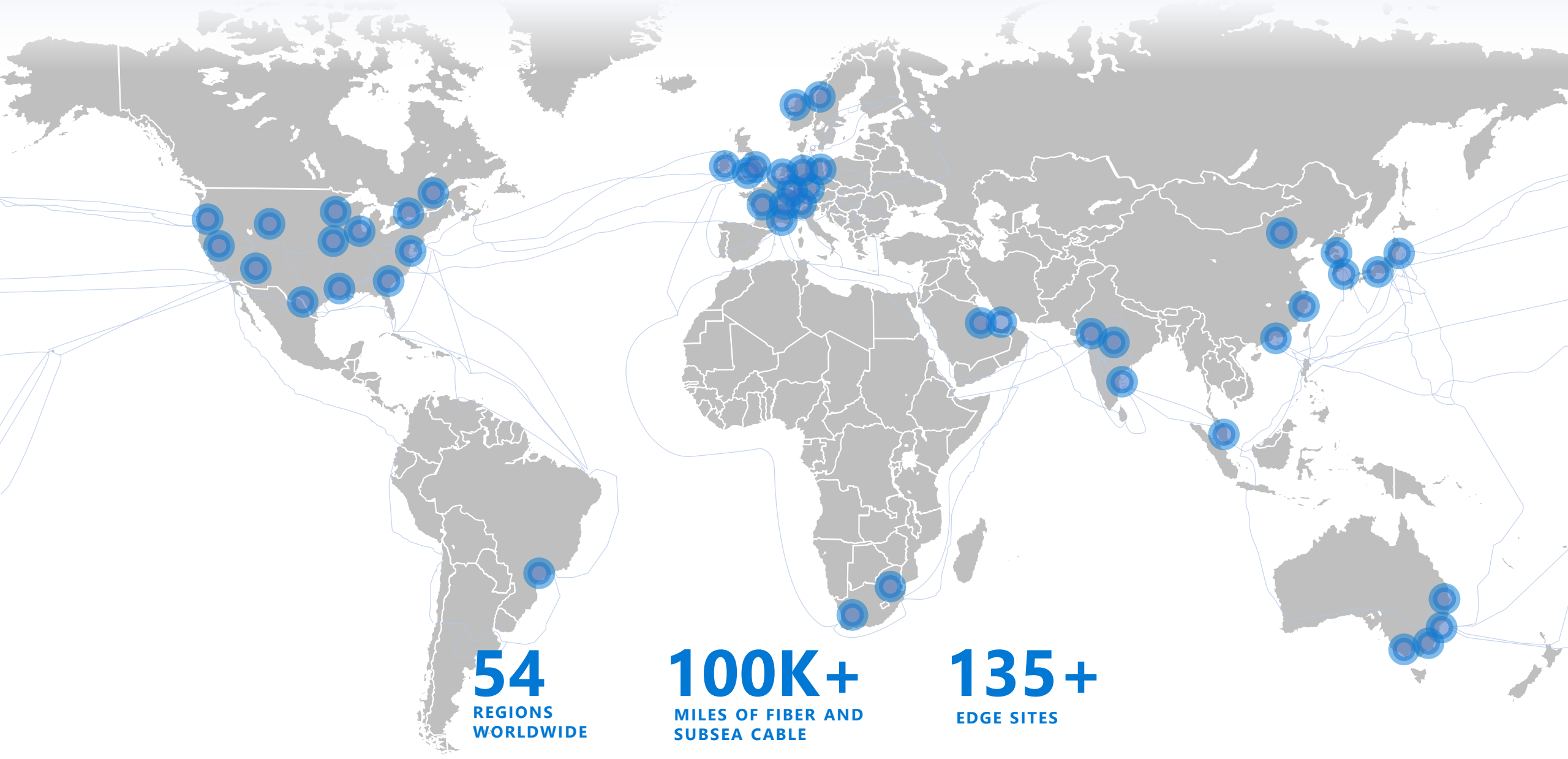
https://docs.microsoft.com/en-us/azure/avere-vfxt/

Avere architecture deployment



Networking

Azure Networking Infrastructure

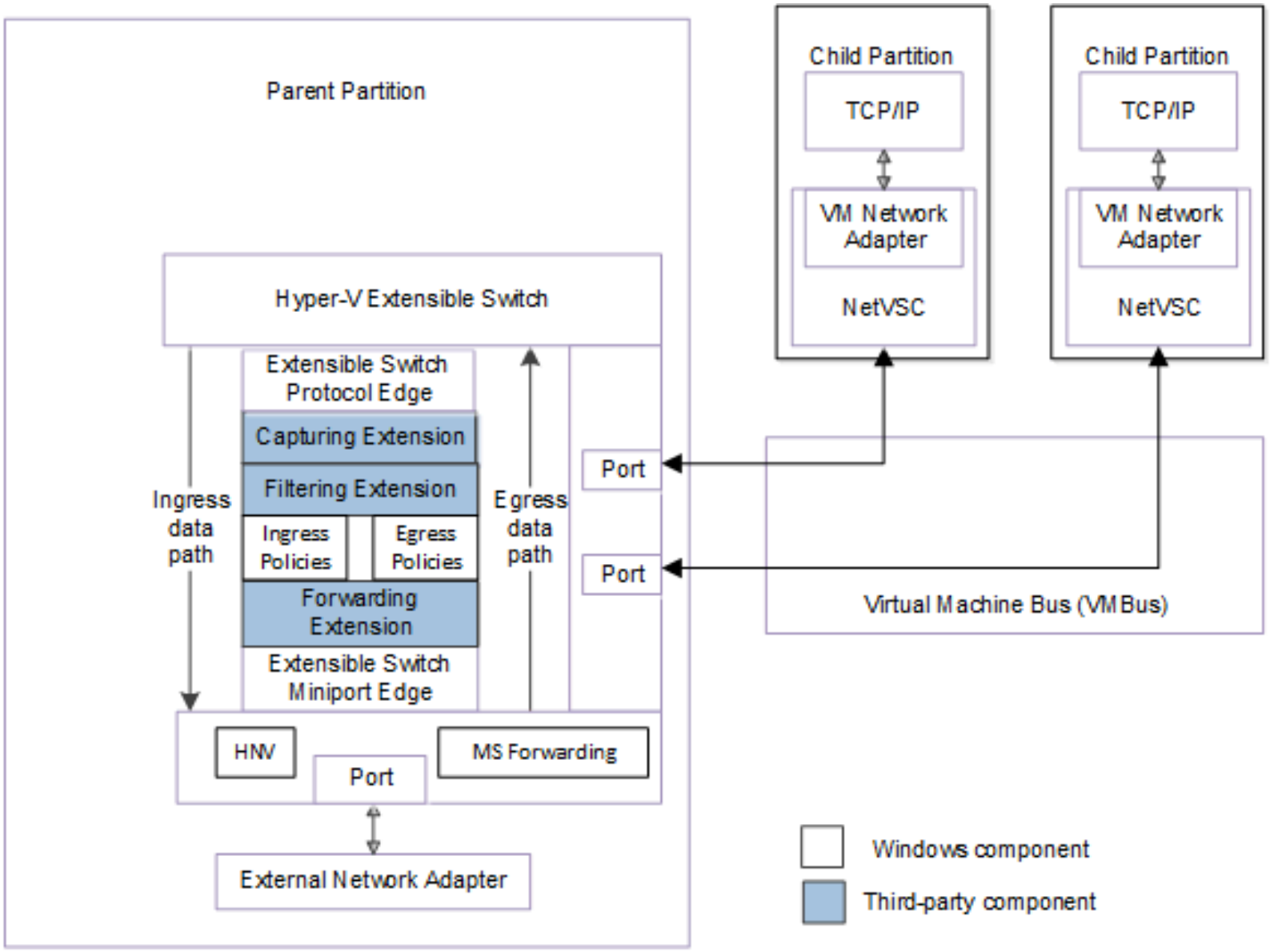


54
REGIONS WORLDWIDE

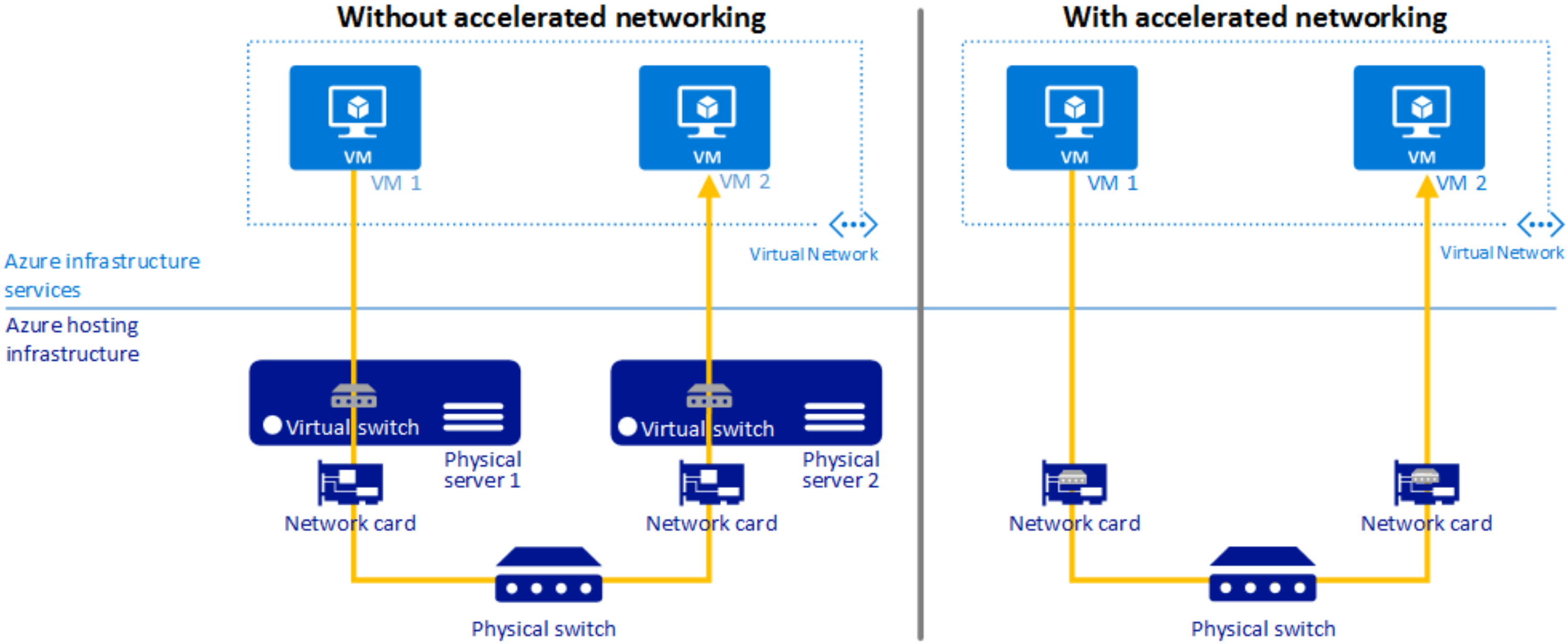
100K+
MILES OF FIBER AND SUBSEA CABLE

135+
EDGE SITES

How Hyper-v net adapter works for “dummies”?



Accelerated Networking

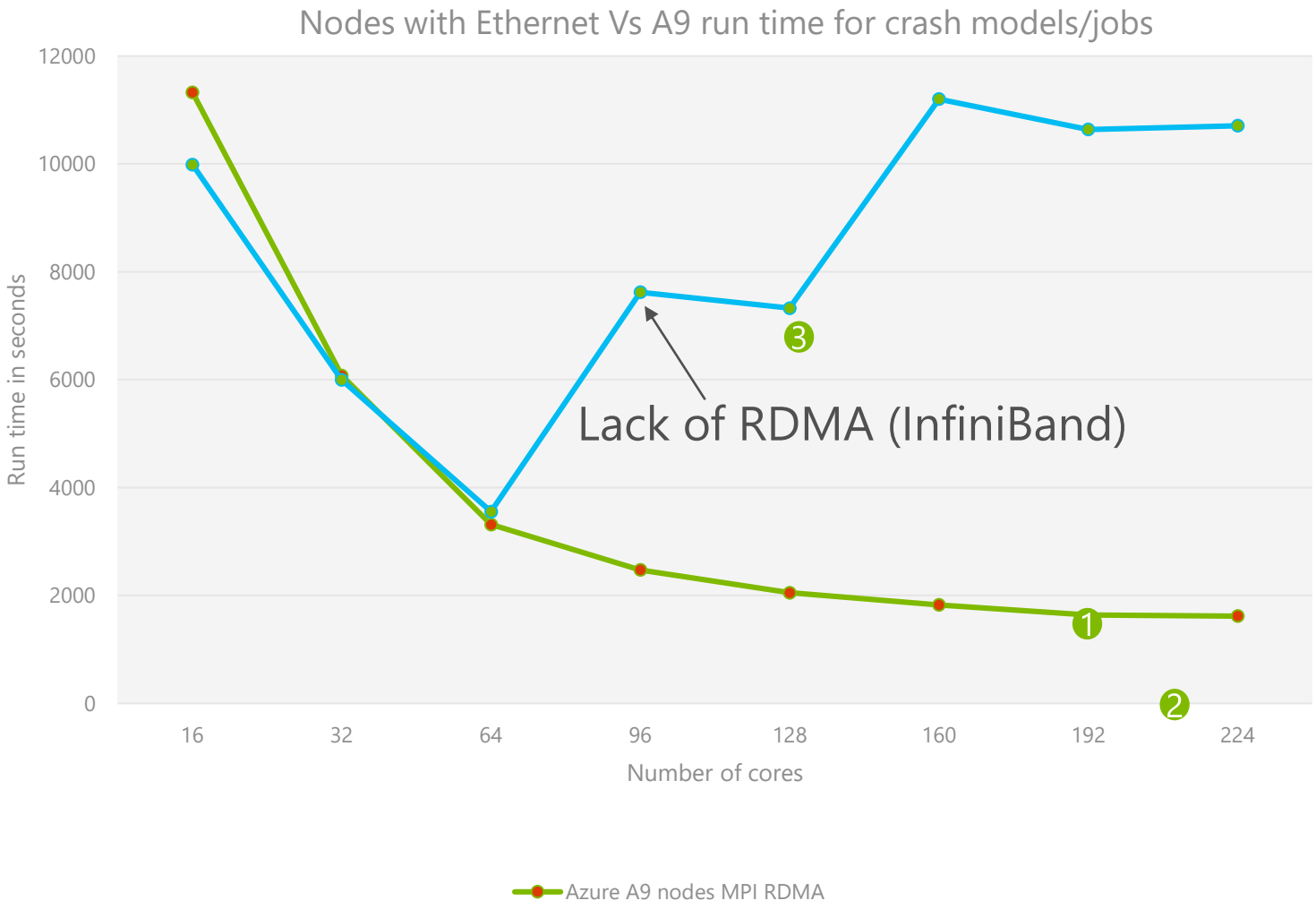


https://www.usenix.org/system/files/conference/nsdi18/nsdi18-firestone.pdf

Infiniband....

Ethernet vs Infiniband

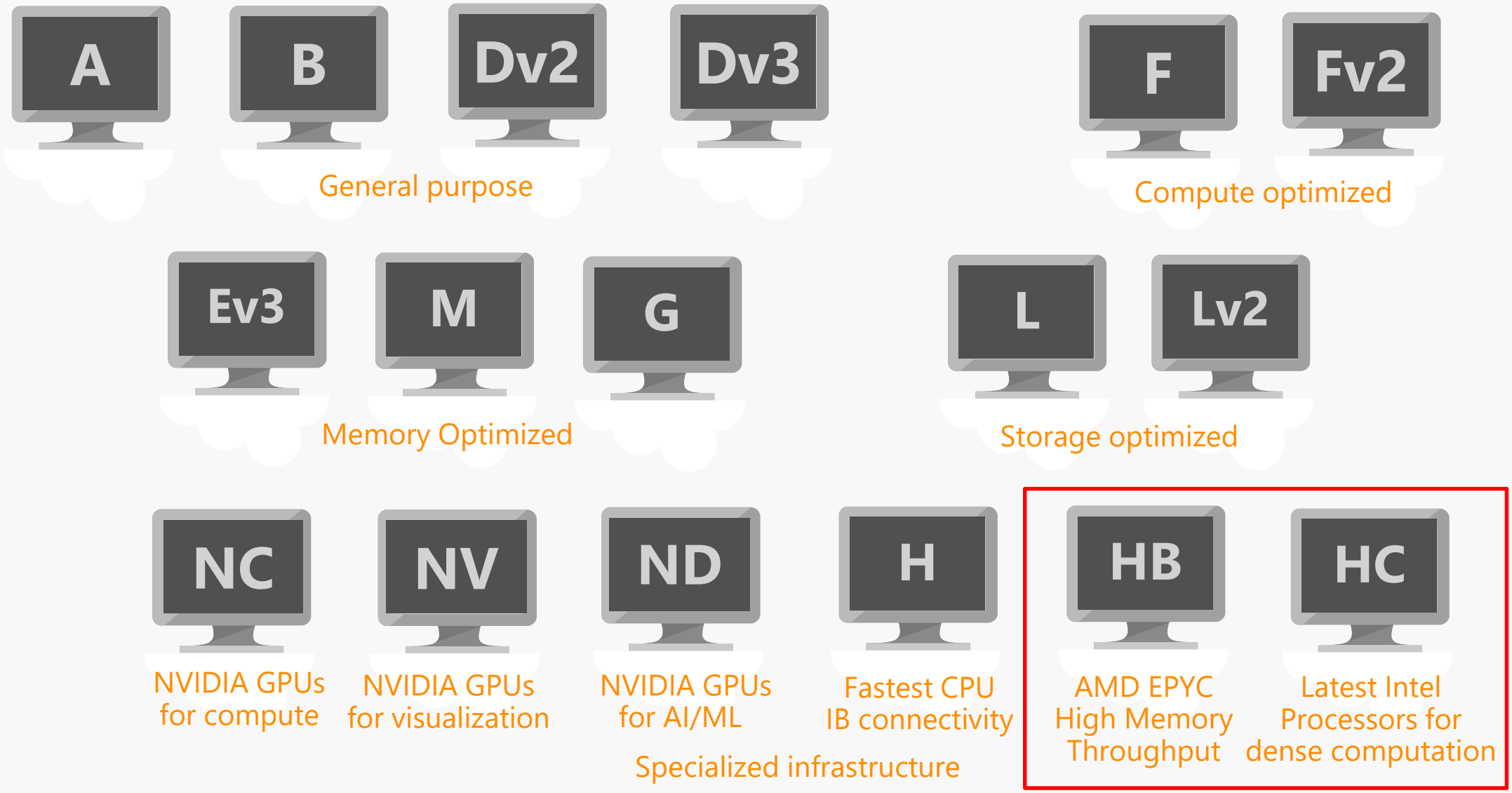
Why Infiniband matters?



1. Lower run time is better
2. As you add more cores the job should run faster
3. Because of lack of RDMA InfiniBand low latency networking jobs run slower after 64 cores on other clouds
1. Azure continues to scale well after adding more cores

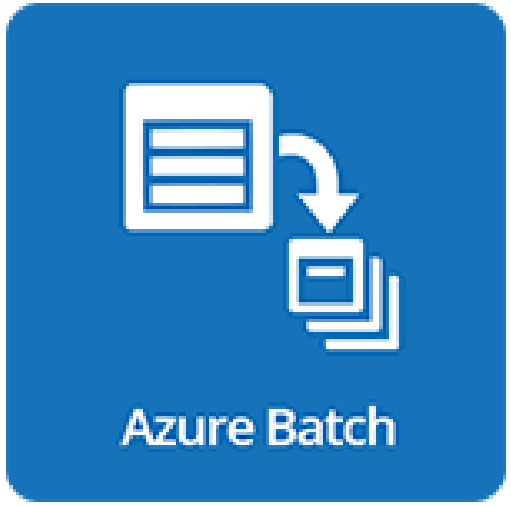
Computing

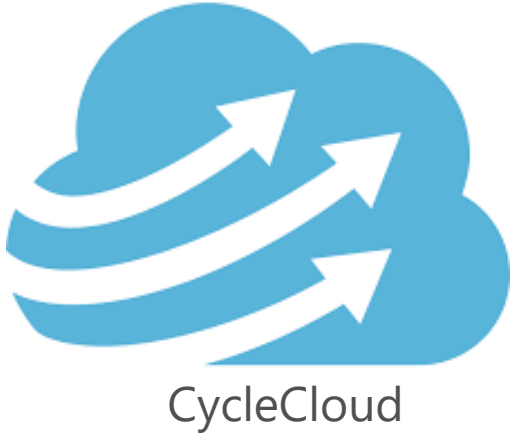
Computing options in Azure



Management

Management





Let’s Play

¡Gracias

Patrocinadores Locales

Colabora

2019

Global Azure BOOTCAMP

2019
Global Azure
BOOTCAMP