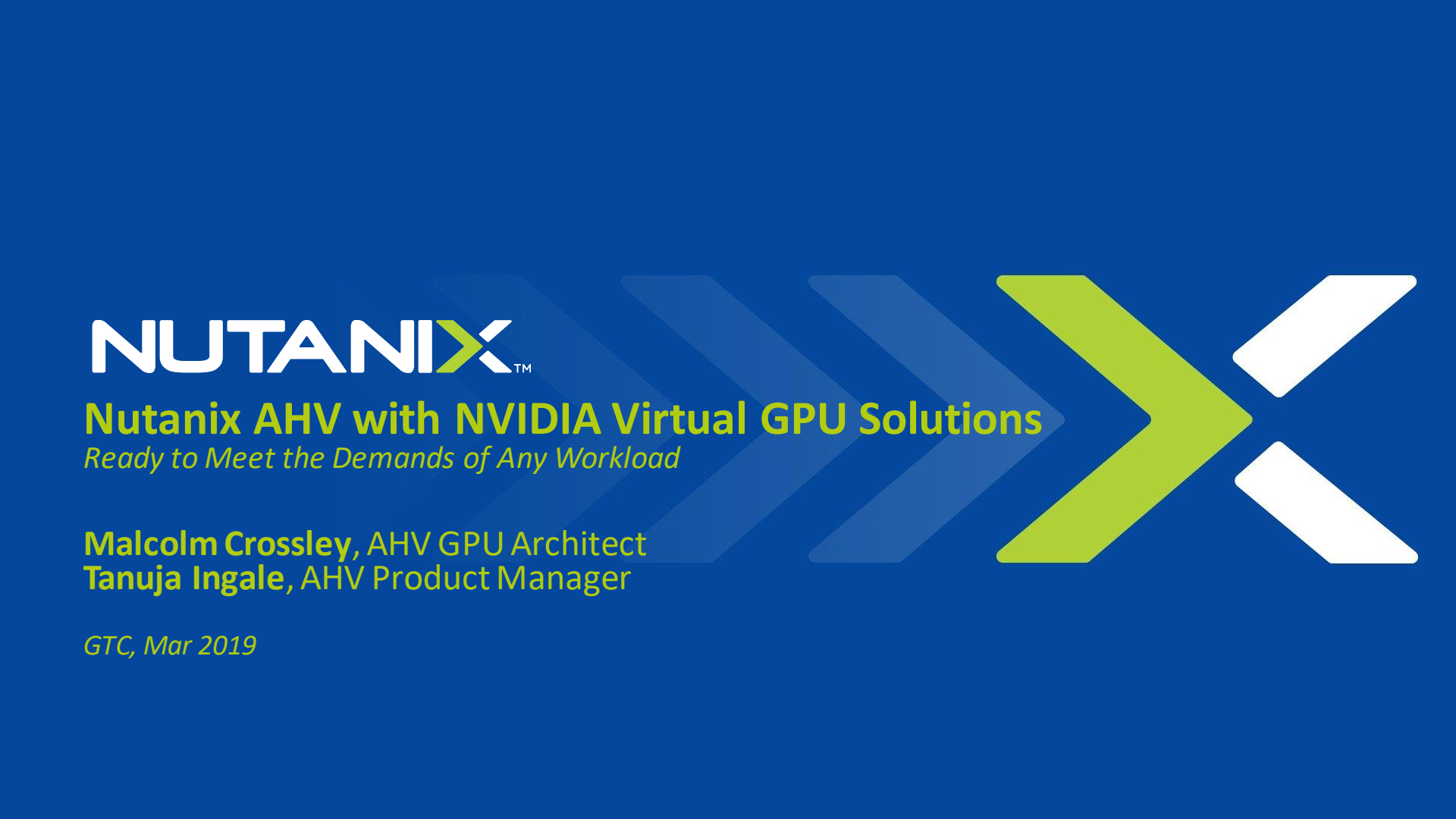
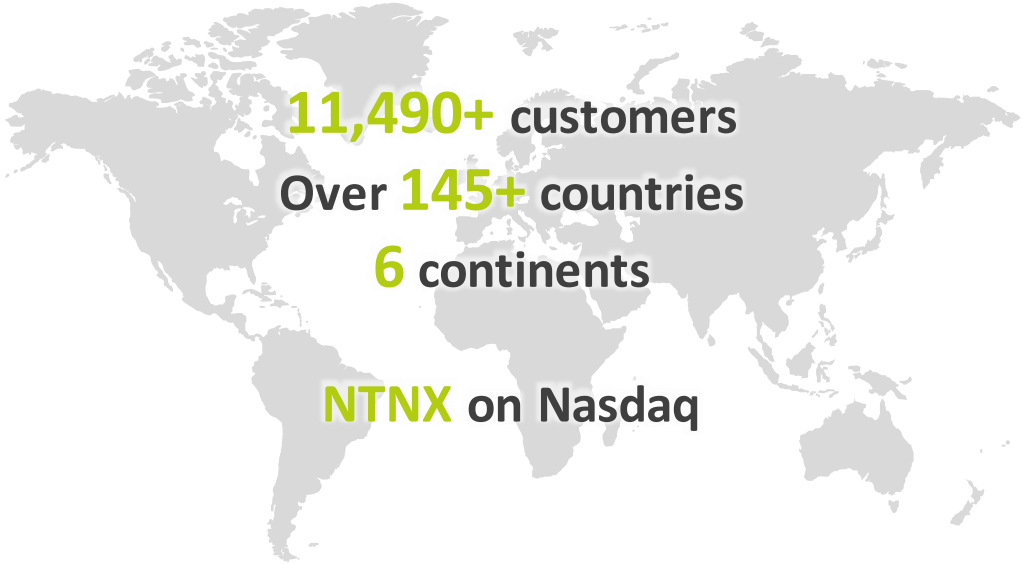

NUTANIX
Nutanix AHV with NVIDIA Virtual GPU Solutions
Ready to Meet the Demands of Any Workload
Malcolm Crossley, AHV GPU Architect
Tanuja Ingale, AHV Product Manager
GTC, Mar 2019

Nutanix: The Enterprise Cloud Company

Our Mission: Nutanix makes IT infrastructure invisible with an enterprise cloud platform that delivers the agility and economics of the public cloud, without sacrificing the security and control of on-premises infrastructure.



11,490+ customers

Over 145+ countries

6 continents

NTNX on Nasdaq

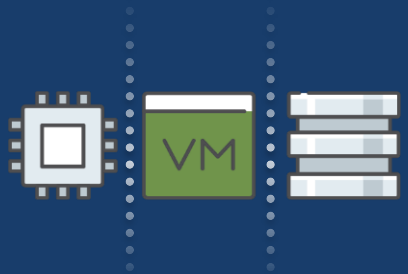
Founded in 2009

4,380+ employees

NUTANIX

IT Complexity is Hurting Business

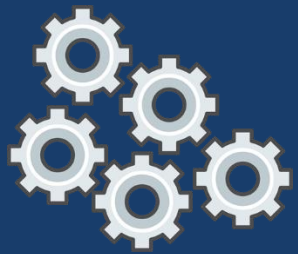
Infrastructure



Time consuming to provision

Multiple points of failure

Process



Difficult to scale and upgrade

Large upfront CapEx

People

Little time for innovation

Requires IT specialists

The Need for Desktop Virtualization (VDI)

Increase employee productivity with anywhere access

Simplify support and enable choice of BYO devices

Centralized security to protect sensitive information

Reduce cost and complexity of app and desktop management

Comprehensive desktop virtualization for any use case

VDI and GPU workload evolution

VDI, Client/Server & business apps

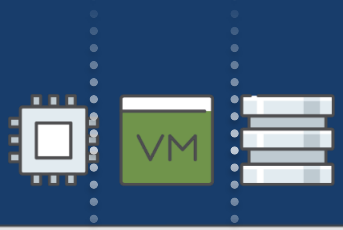
- Nutanix Comprehensive desktop virtualization
- 100s of general knowledge & task workers per Compute + Expansion module
- ~2000 VDI users per X rack units w/o GPU acceleration
- 200+ VDI sessions with accelerated virtualized graphics

Complex Design & Visualization

- High-end design, CAD, Rendering, Ray tracing (M&E, Manufacturing)
- Support 10s of specialized users (Design) per Compute + Expansion module
- Support AI, Deep ML, Compute workloads for Improved VR perf
- HPC workloads, Monte Carlo analysis

How IT complexity translates to GPU environments

GPU Infrastructure



Operational & workflow complexity

Silos for graphics & compute

Demand for GPU accelerated workloads

Process

Need predictable scaling

VDI CapEx Underutilization

Underutilization of True pre-emptive capabilities of GPU

People

Requires experts for GPU workflows/sizing

Innovating in Three Fronts for VDI/GPU accelerated env

Architecture

Scale out, software-defined architecture

Technology

Powerful data fabric and control fabric capabilities

Infra CapEx Preservation

Purchase based on user needs & provide 100% utilization

Innovating in Three Fronts for VDI/GPU accelerated env

Architecture

Scale out, software-defined architecture

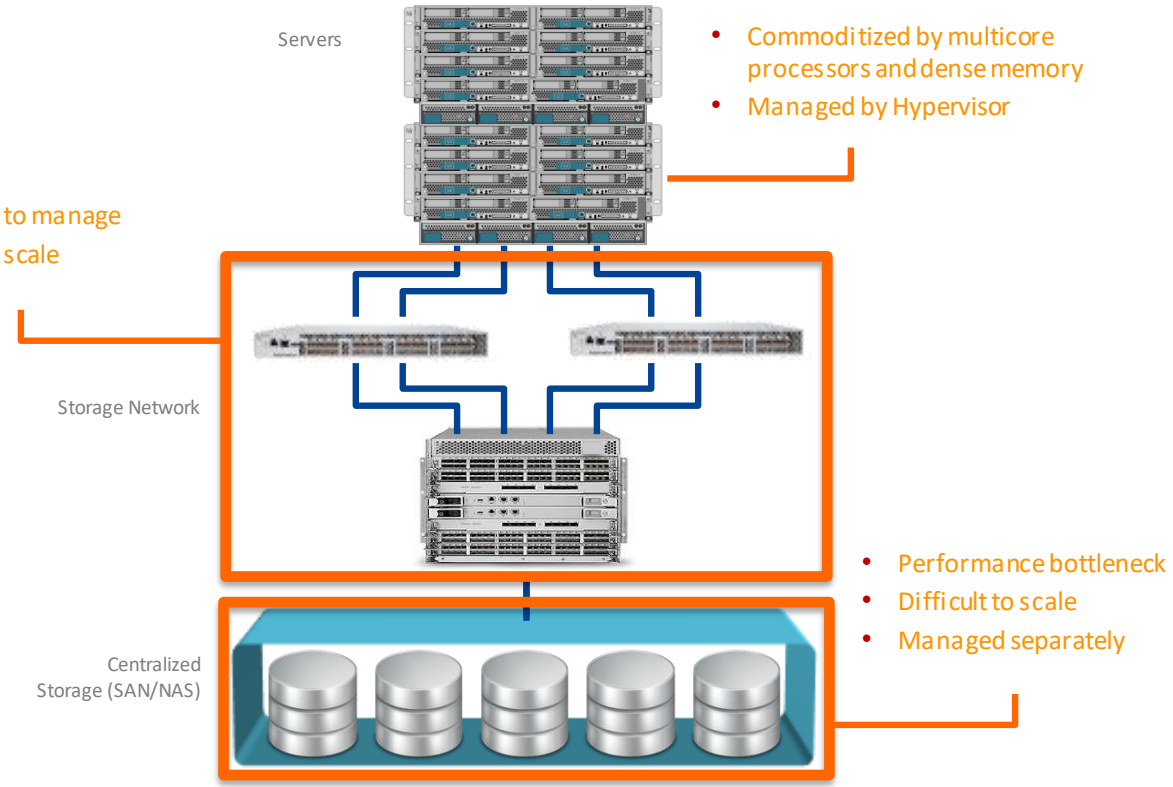
Technology

Powerful data fabric and control fabric capabilities

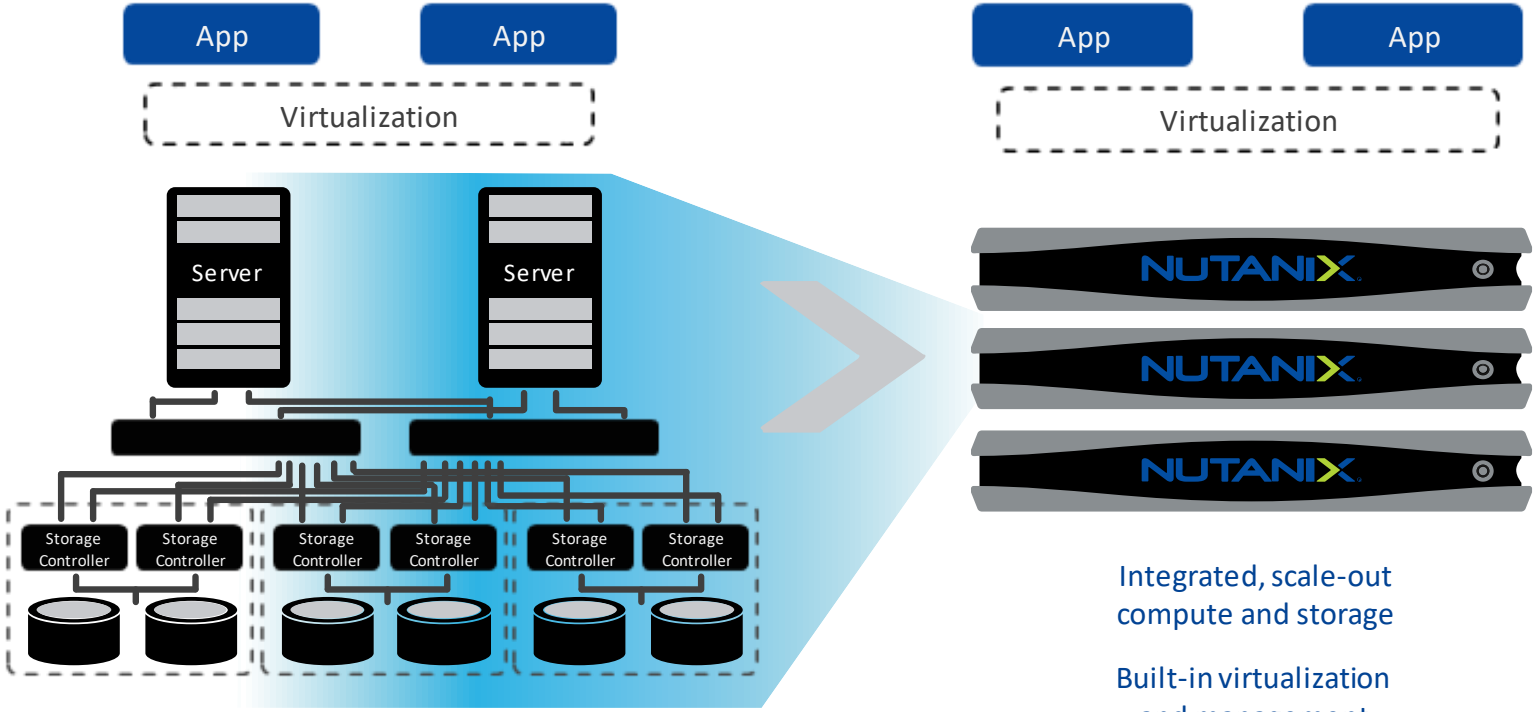
Infra CapEx Preservation

Purchase based on user needs & provide 100% utilization

VDI Pain points on legacy architecture



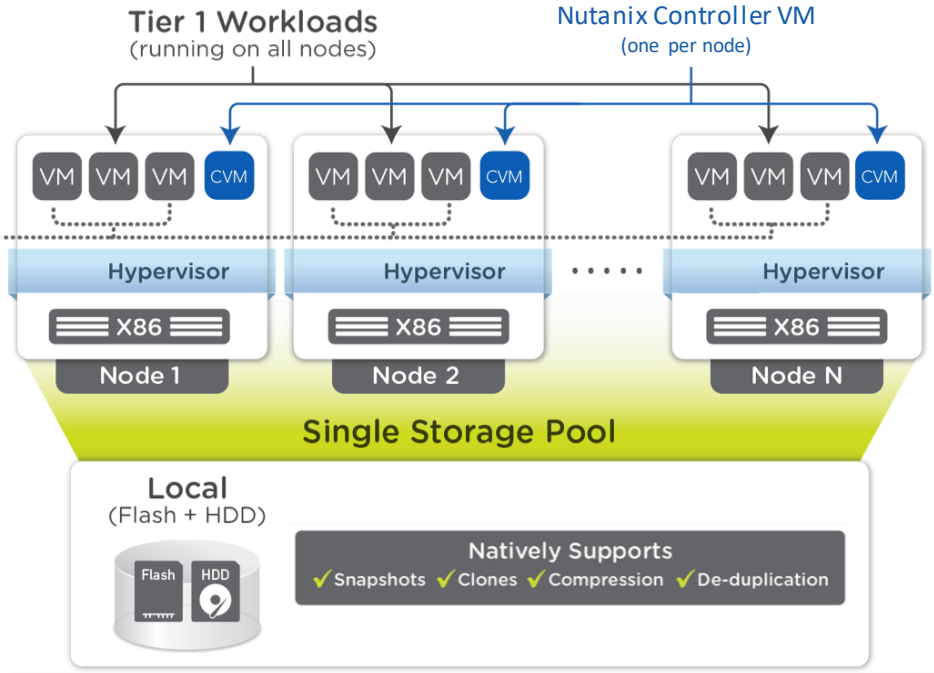
Nutanix: Web-Scale Converged Architecture



Integrated, scale-out compute and storage

Built-in virtualization and management

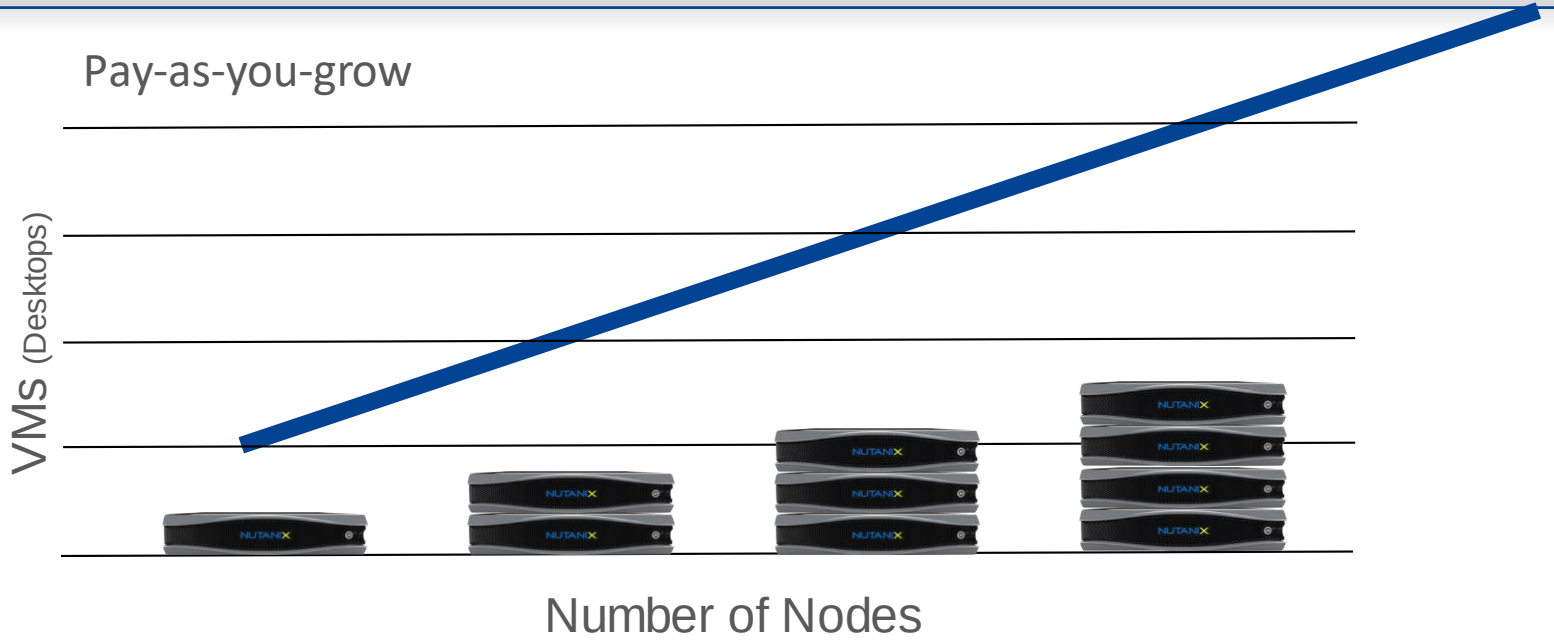
Built for Virtualization



- ✓ High performance
- ✓ Auto optimized

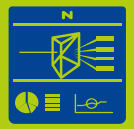
- Software-defined approach with Controller VM per node
- Pooled storage resources across the platform and scale as needed

The Result is Linearly Scaling



- Scale incrementally one node at a time
- Protect infrastructure investment by eliminating forklift upgrades
- Scale storage capacity & performance linearly

Nutanix Platform



Prism

Infrastructure / VM Management

Operational Insights

Capacity Planning

Acropolis

App Mobility Fabric

Distributed Storage Fabric

AHV (built-in Virtualization)

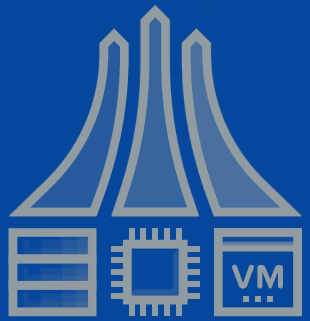
Comprehensive management solution radically simplifying datacenter operations

Turnkey infrastructure platform that converges compute, storage and virtualization to run any application, at any scale

NUTANIX

Nutanix AHV

The hypervisor built for the Enterprise Cloud



AHV: Foundation of the Enterprise Cloud OS

KVM + + =

Lean Virtualization

Enterprise Feature Set

Amazing Management

Cloud Operating System

NUTANIX

NX Appliance

Lenovo IBM DELL EMC

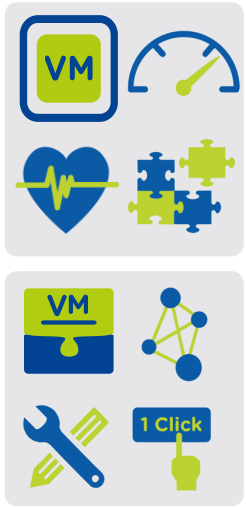
OEM

Hewlett Packard Enterprise CISCO KLAS TELECOM

3rd Party

NUTANIX

AHV Powers HCI



Security

- Data Encryption
- Certifications

Self-Healing

- Hot Spot Remediation
- Auto STIG Compliance

Management

- Prism Central
- API and CLI

Performance

- AHV Turbo
- IO Optimization

Automation

- Distributed Scheduler
- Calm Blueprints

Biz. Continuity

- HA
- Synch Rep (2019)

Auditability

- Logging
- Remote Syslog

Ecosystem

- Over 100 ISV solutions
- Backup, SAP HANA

Analytics

- Prism Pro
- Performance

Extendable

- Memory
- vCPU

VDI/ GPU Acceleration

- GPU, vGPU
- PVS, MCS

AHV: Citrix App and Desktop Virtualization

Starting with Citrix XD/XA 7.9 … AHV is integrated with MCS

Nutanix AHV is

CITRIX ready™

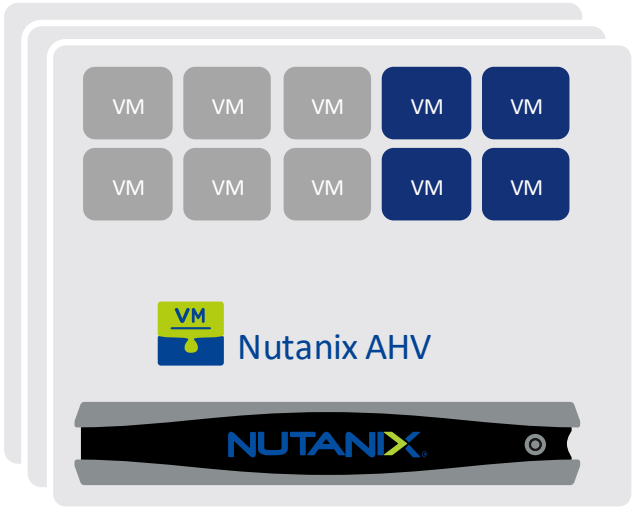
Citrix Infrastructure

CITRIX

- XenApp & XenDesktop
- NetScaler VPX
- ShareFile
- XenApp, XenDesktop support
- Cloud Connect InstantOn
- PVS and MCS plug-ins
- NVIDIA vGPU support
- High perf. + data locality

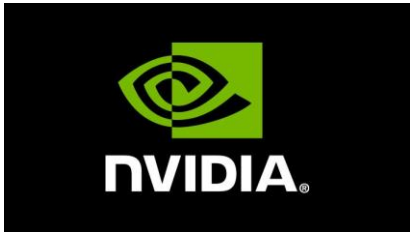
NUTANIX

End user desktops



Nutanix-NVIDIA Strategic Partnership

- Nvidia: Industry leader in Visual Computing Technologies and GPU accelerators
- AHV
    - First commercial kernel-based virtual machine (KVM) to support vGPU
    - Fully supports NVIDIA virtual GPU technology (GRID)
        - Quadro Virtual Data Center Workstation (vDWS)
        - nVIDIA GRID Virtual PC
        - GRID Virtual Applications



AHV: Crystal Clear Graphics

GPU

vGPU

vGPU HA

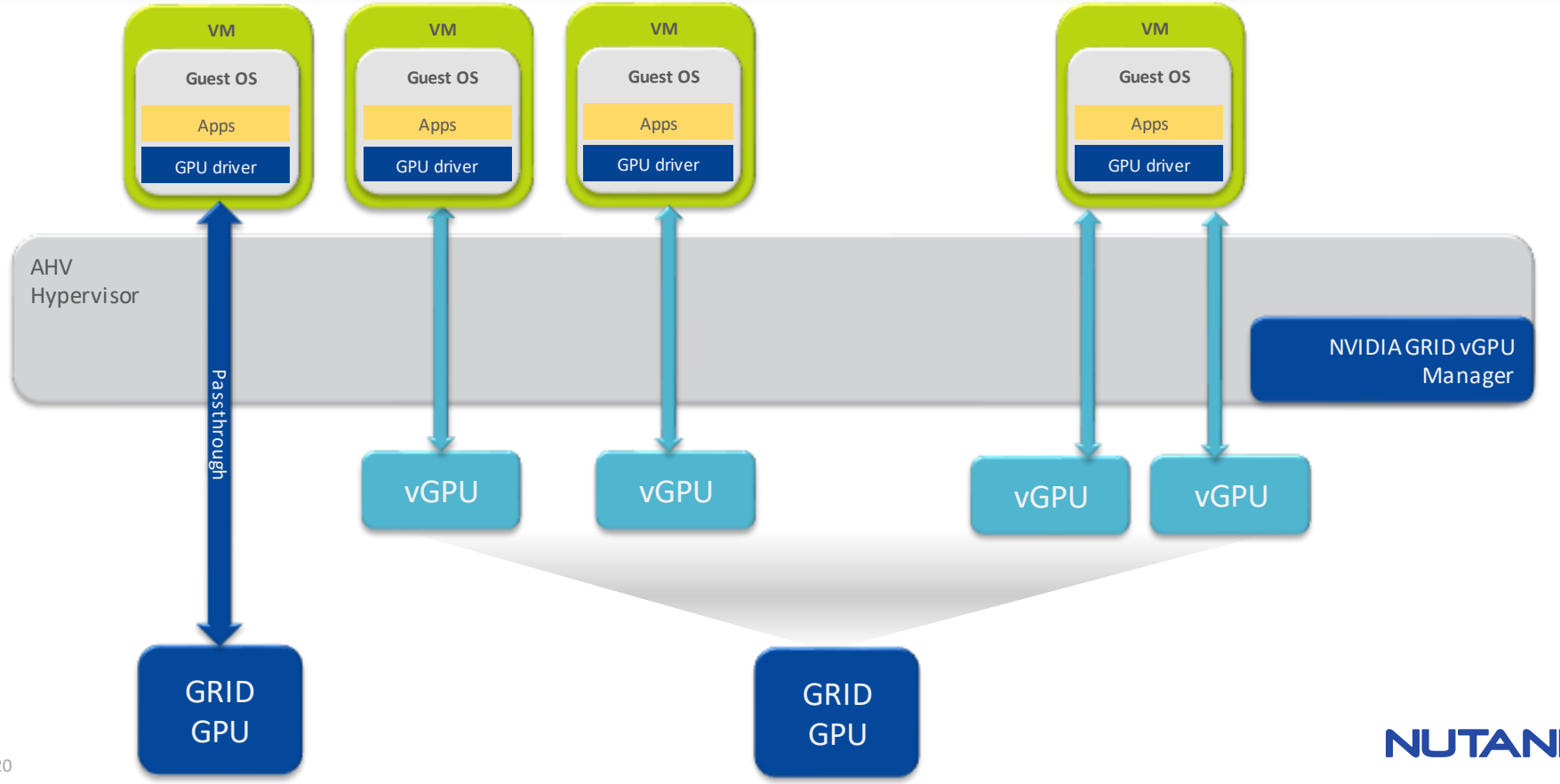
Multi-vGPU VM

(Coming Soon)

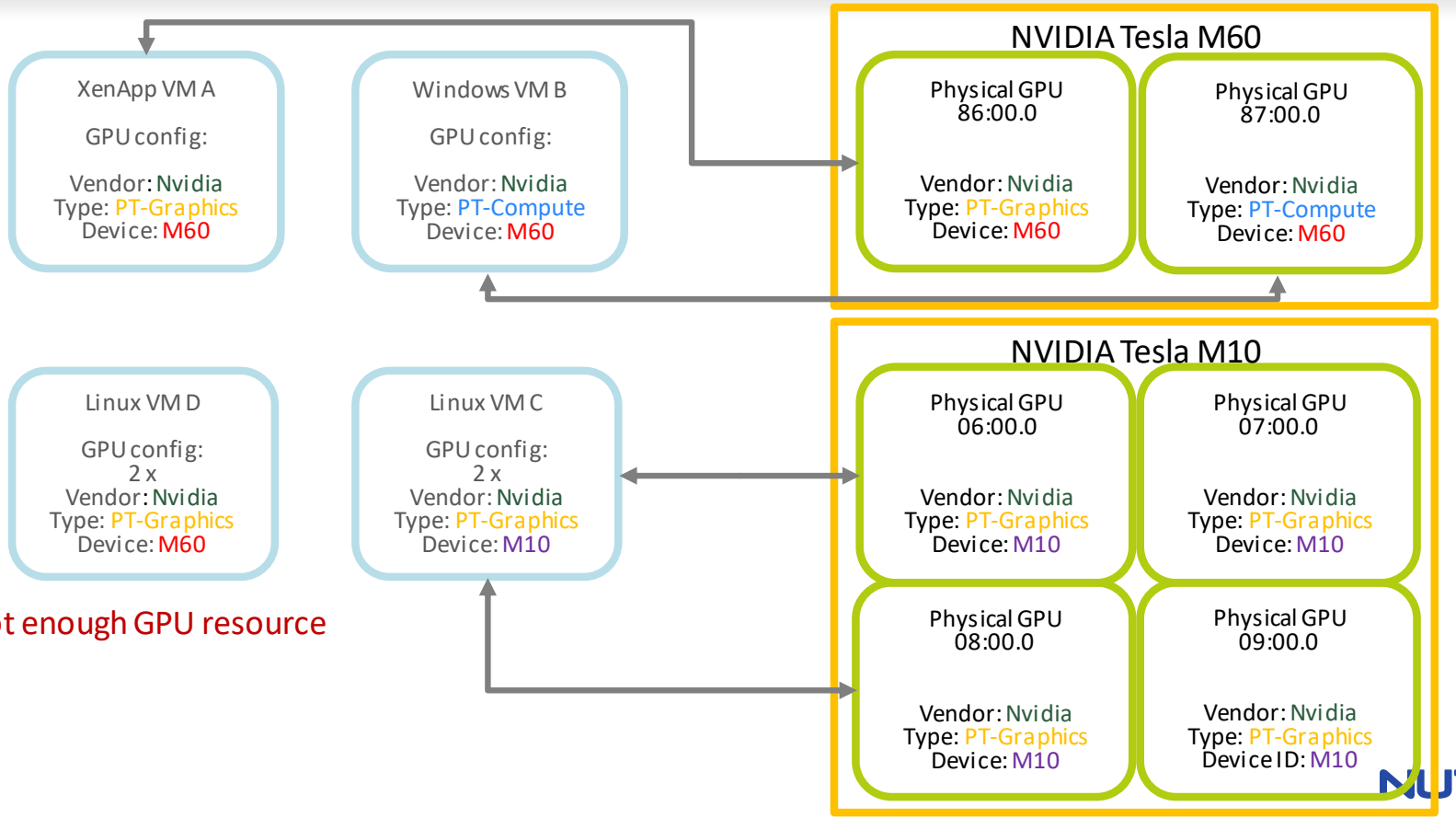
vGPU Live Migration

(Coming Soon)

AHV: Modes of GPU usage

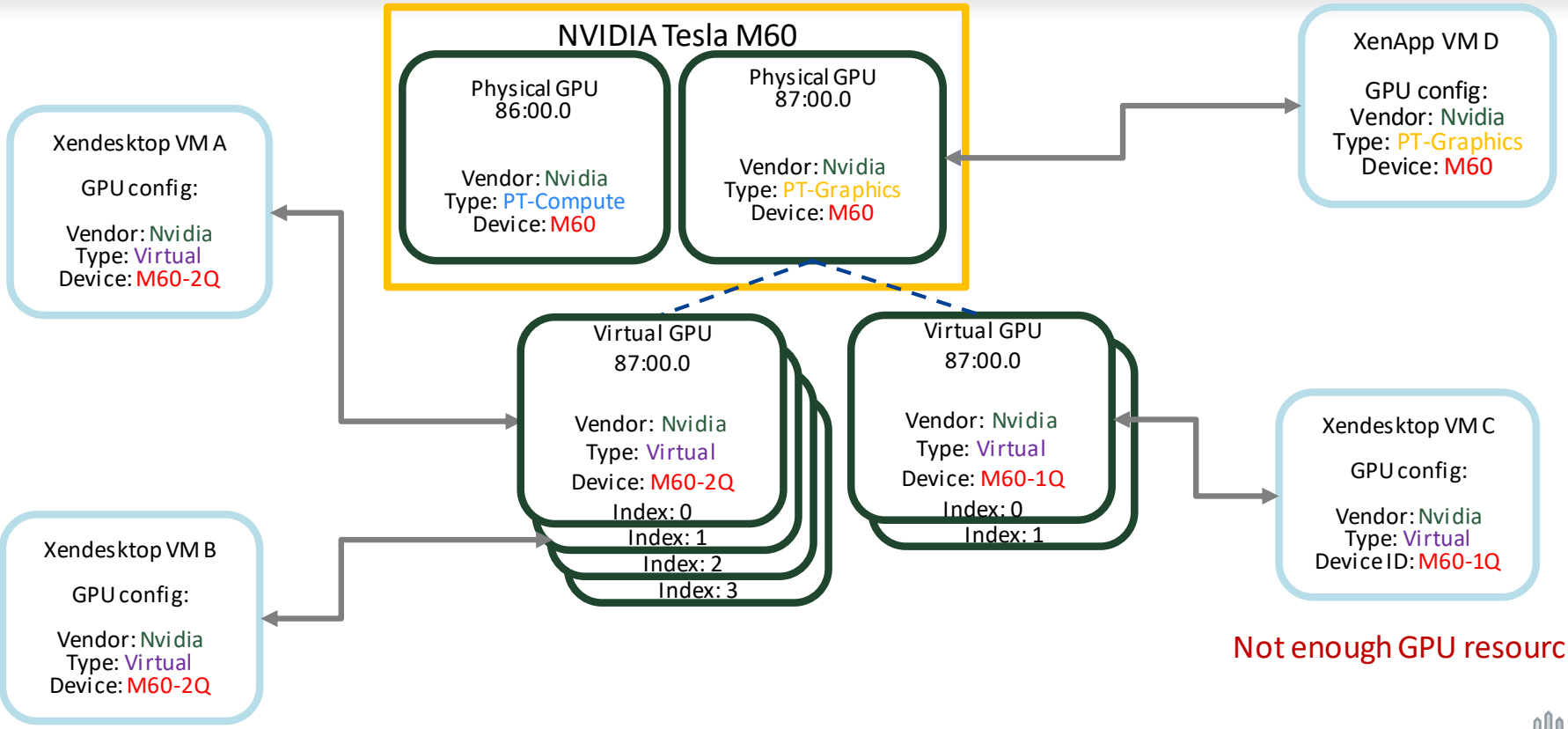


Acropolis GPU resource management concepts

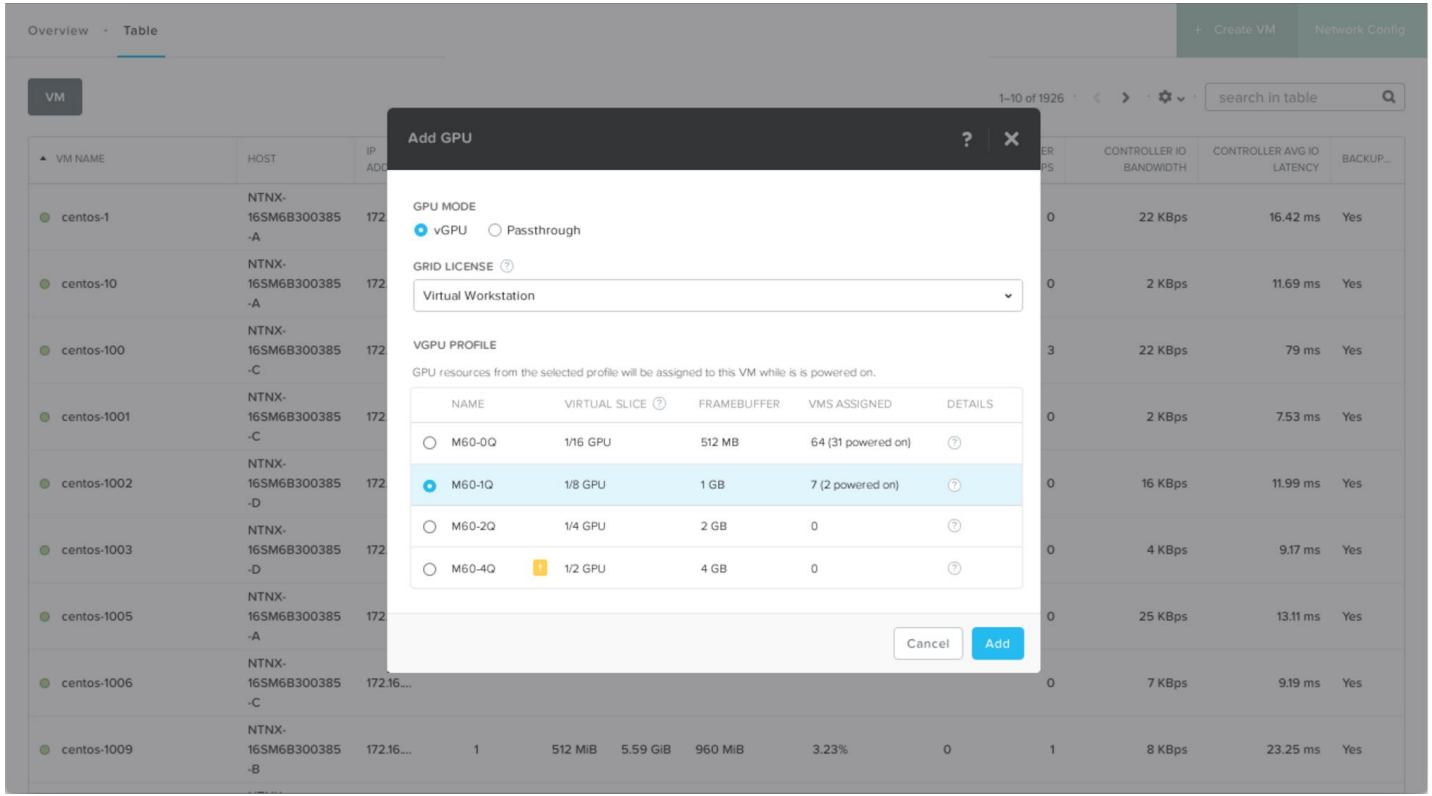


Not enough GPU resource

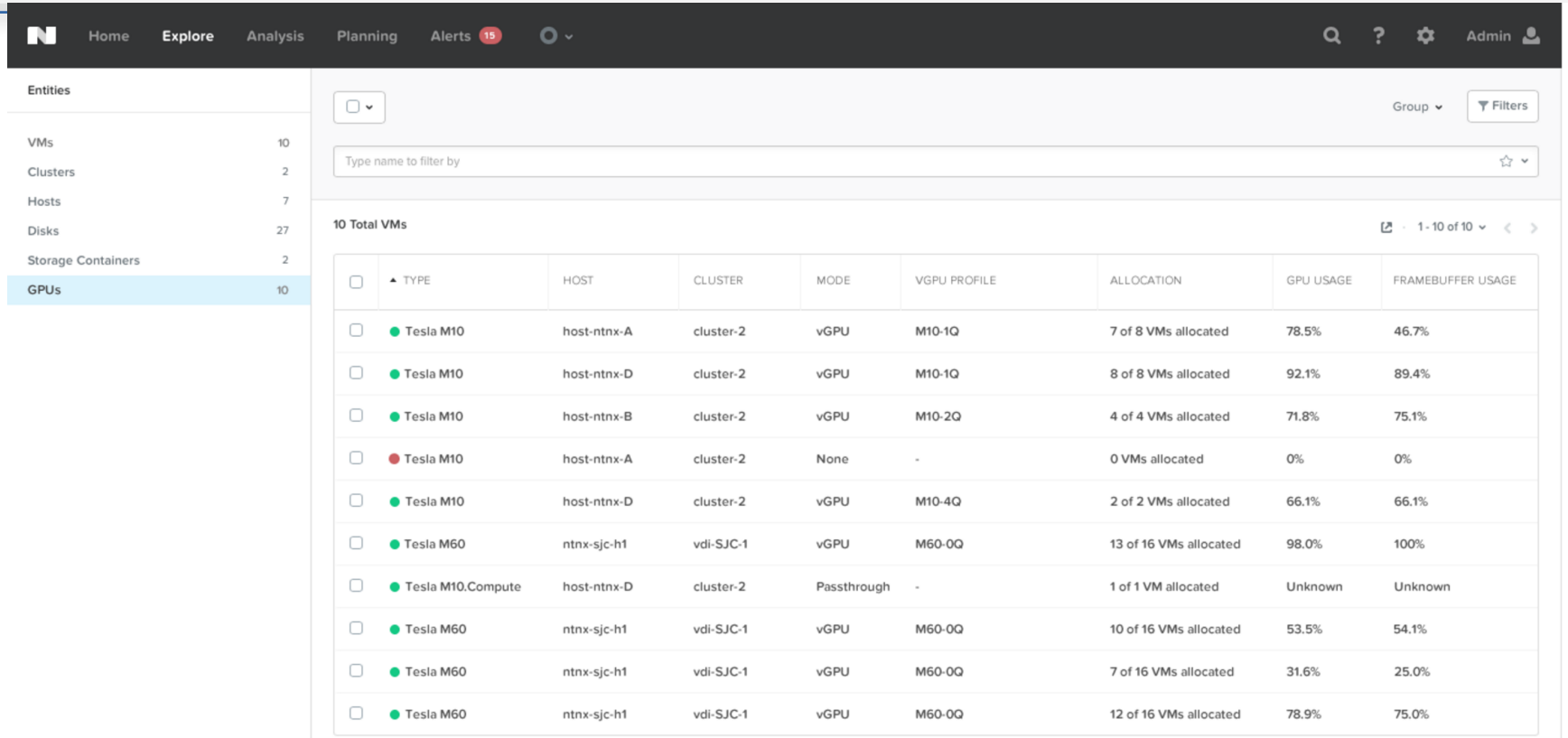
Acropolis GPU resource management concepts



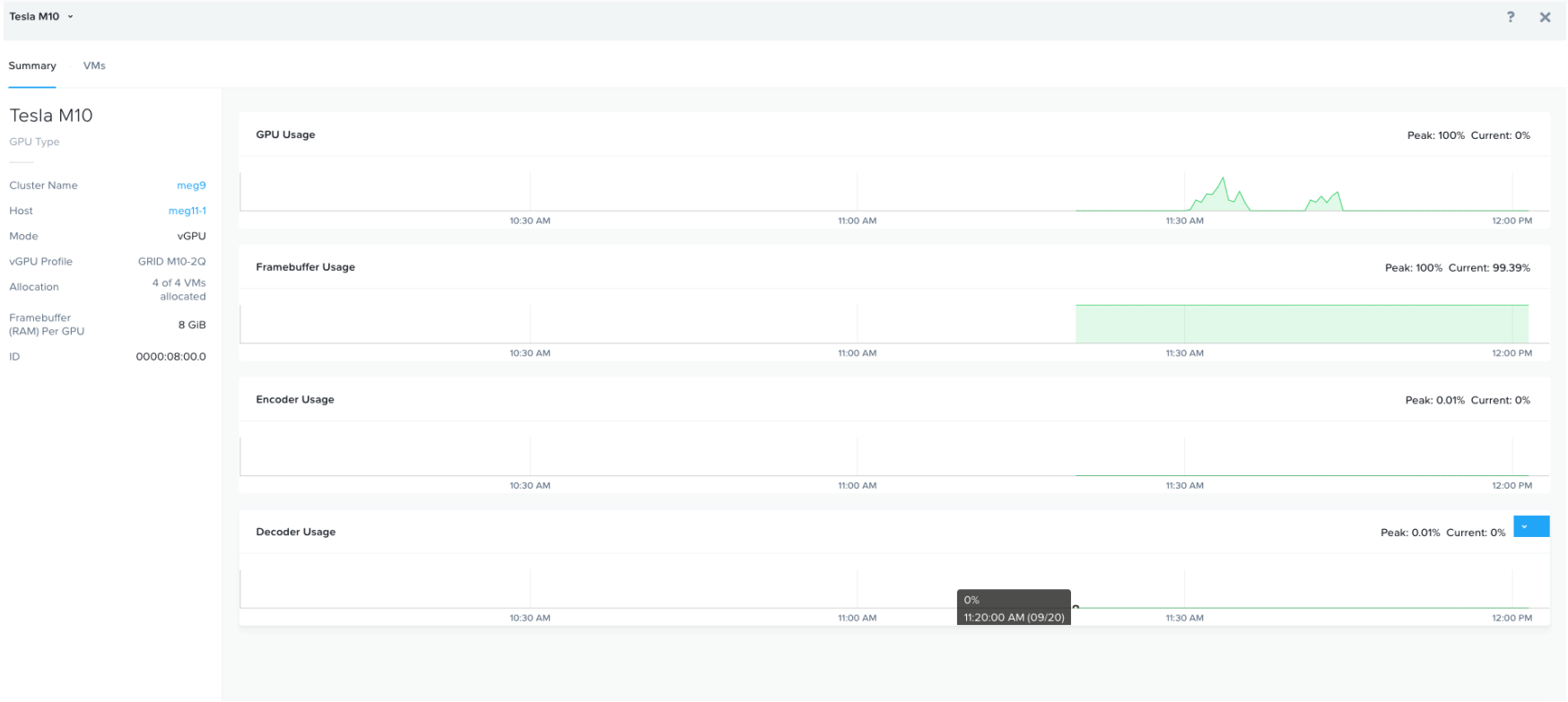
AHV VM GPU resource configuration



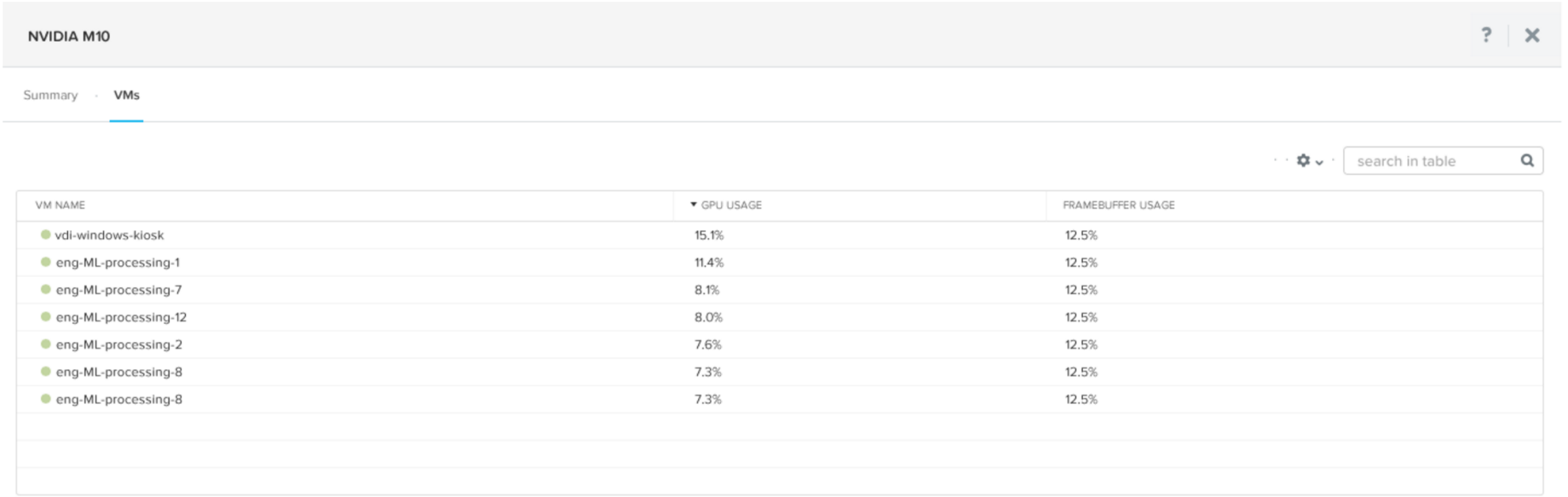
Prism UI - Physical GPU overview



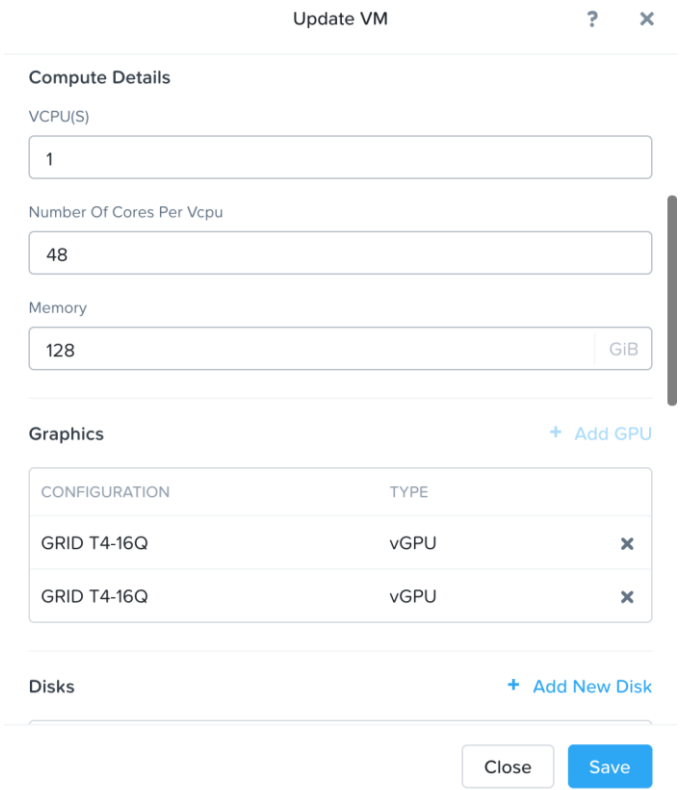
Prism UI - Physical GPU metrics



Prism UI - Virtual GPU metrics



Prism UI – Multi-vGPU per VM



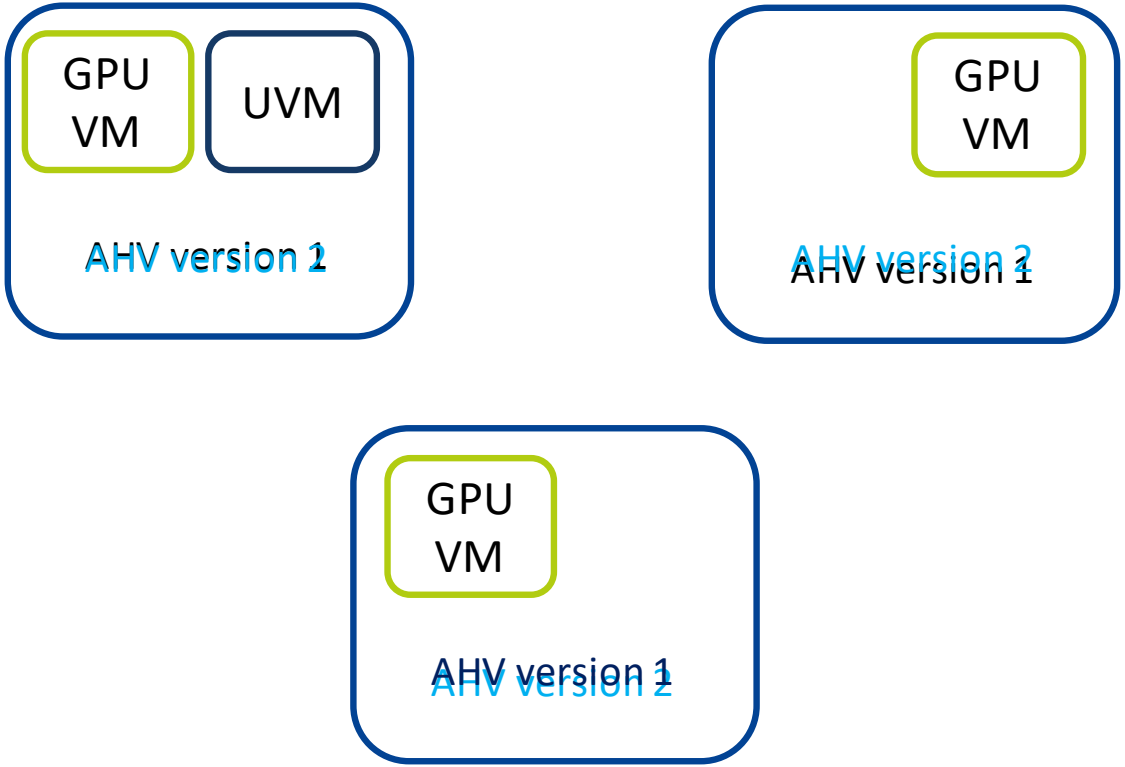
Powerful REST API for all GPU resource information

hosts

Show/Hide | List Operations | Expand Operations

GET	/hosts/gpu_profiles	Get the list of virtual GPU profiles for all physical Hosts.
GET	/hosts/host_gpus	Get the list of GPUs for all physical Host.
GET	/hosts/{uuid}/host_gpus	Get the list of GPUs for a physical Host.
GET	/hosts/{uuid}/host_gpus/{gpu_uuid}	Get a particular GPU for a physical Host.
GET	/hosts/{uuid}/host_gpus/{gpu_uuid}/stats/	Get stats for a particular GPU for a physical Host.

Acropolis 1-click operations and GPU resources



Innovating in Three Fronts for VDI/GPU accelerated env

Architecture

Scale out, software-defined architecture

Technology

Powerful data fabric and control fabric capabilities

Infra CapEx Preservation

Purchase based on user needs & provide 100% utilization

Nutanix Data & Control Fabric Solutions

Software-defined Snapshots & Clones

Offloads virtualization tier → higher ops performance

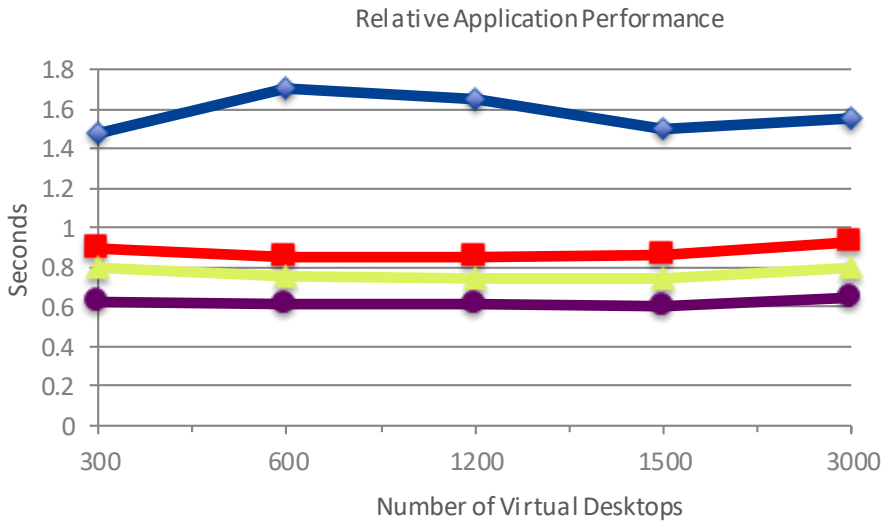
Array-based quick-clones for efficient provisioning

Native VM-centric snapshots

Nutanix Shadow Clones

Distributed caching of vDisks and VM data read by multiple CVMs

~50% reduction in boot time



Consistent response time while incrementally scaling blocks

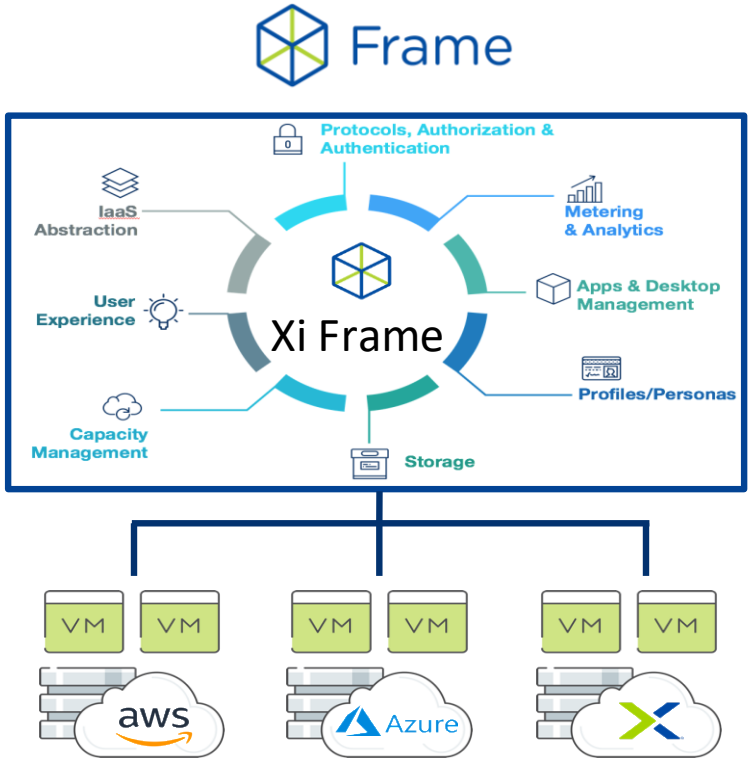
Nutanix Frame: Desktop-as-a-service

Nutanix Frame

Browser based Desktop-as-a-service platform

Multi-cloud Infrastructure (orchestrates workloads in public clouds or on-prem (AHV)

Exceptional End-user & Admin UX



NUTANIX

Innovating in Three Fronts for VDI/GPU accelerated env

Architecture

Scale out, software-defined architecture

Technology

Powerful data fabric and control fabric capabilities

Infra CapEx Preservation

Purchase based on user needs & provide 100% utilization

Need for GPU powered & oversubscribed infra

- Increased need for GPUs
  - AI Composability & Virtualization
  - Data Simulation
- 20x Faster > CPUs
  - Fast Complex Calculations & Simulations
  - Massively parallel, multi-core GPUs = 1000s of Compute Cores
- Infra CapEx
  - GPUs more expensive than CPUs
  - Underutilization of GPU resources (< 15%)

AHV: Graphics by Day / Compute by Night

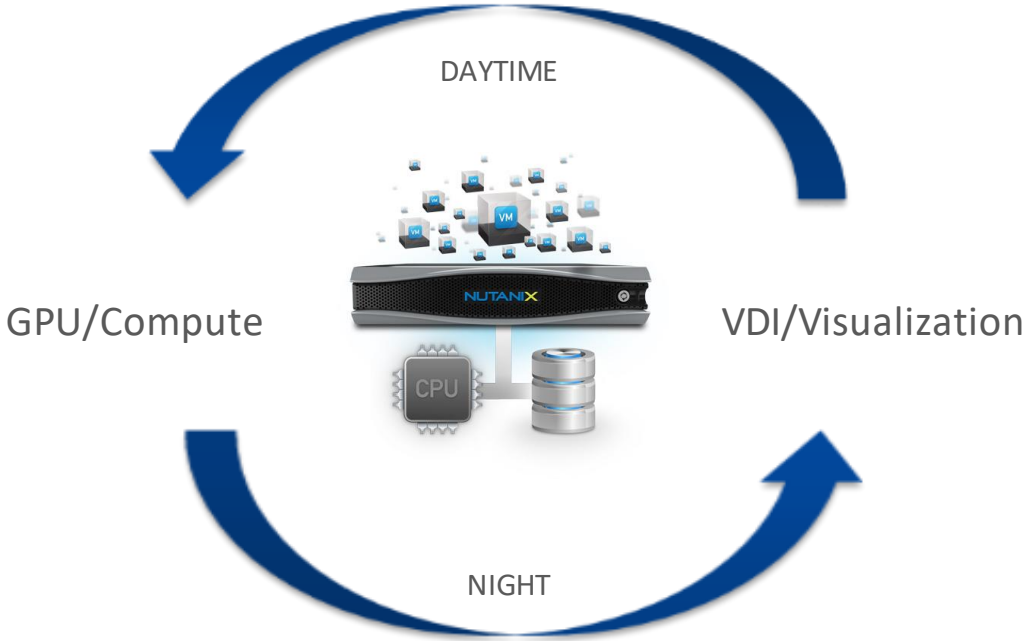
100% GPU CapEx Utilization!

Display by Day

- Graphics Mode
- VDI
- Design & Visualization

Compute by Night

- CUDA Mode
- Simulation, Processing
- Analysis
- Number Crunching



Follow the Sun !

Future developments

- Prism App for Nvidia Dashboard
- vGPU live migration *(Coming soon)*

- Support for Nvidia’s cutting-edge Quadro RTX technology

Conclusion: AHV drives GPU infrastructure elasticity

- Fully Supports Nvidia GRID technology
- Intuitive GPU management workflow
- Built-in scale out technology
  - Storage scale out and elasticity
  - Data locality
  - GPU config simplicity (metrics, usage, alerts)
  - Storage scale out and elasticity
- Infrastructure elasticity
  - 100% GPU utilization
  - CUDA/Graphics mode support
  - CapEx & TCO savings
- Workload support
  - HPC
  - AI/ML (Inferencing, TensorFlow)
  - High End Design, Simulation

NUTANIX™

DEMO – Multi-vGPU per VM


NUTANIX
DEMO – Graphics by Day &
Compute by Night


NUTANIX
Thank you!
Questions?

Your Enterprise Cloud Platform