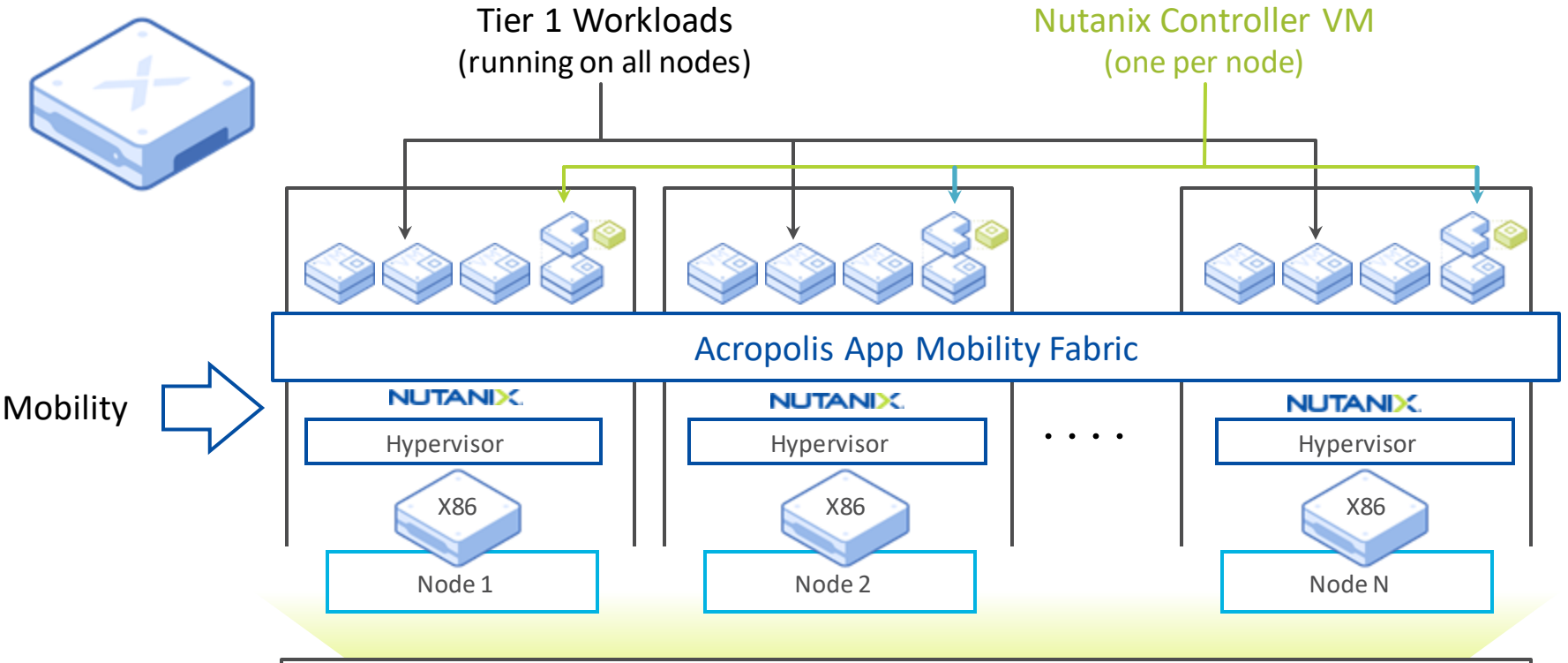
NUTANIX

Architecture Advantages

Roman Kalchenko
Solution Architect
r.kalchenko@elcoregroup.com

Nutanix Web-Scale Architecture

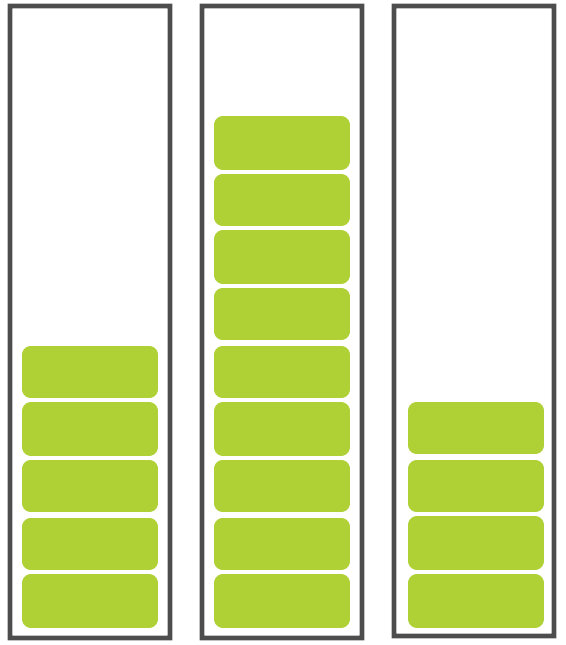
Distributed & scalable cluster, where all resources are actively used and can benefits to all workloads

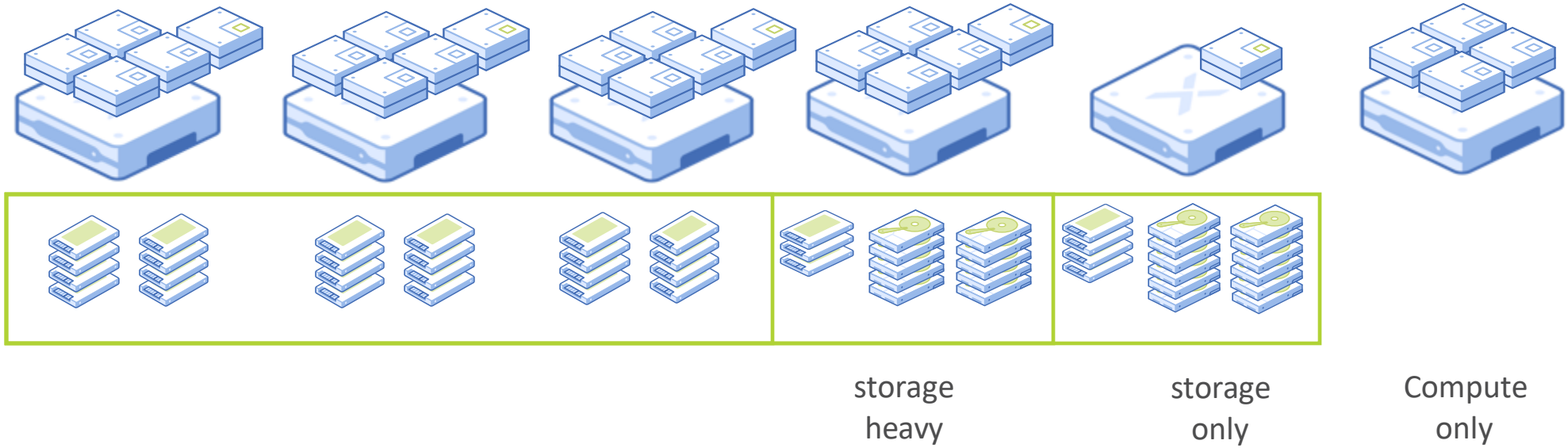


✓ Performance ✓ Resiliency ✓ Scalable

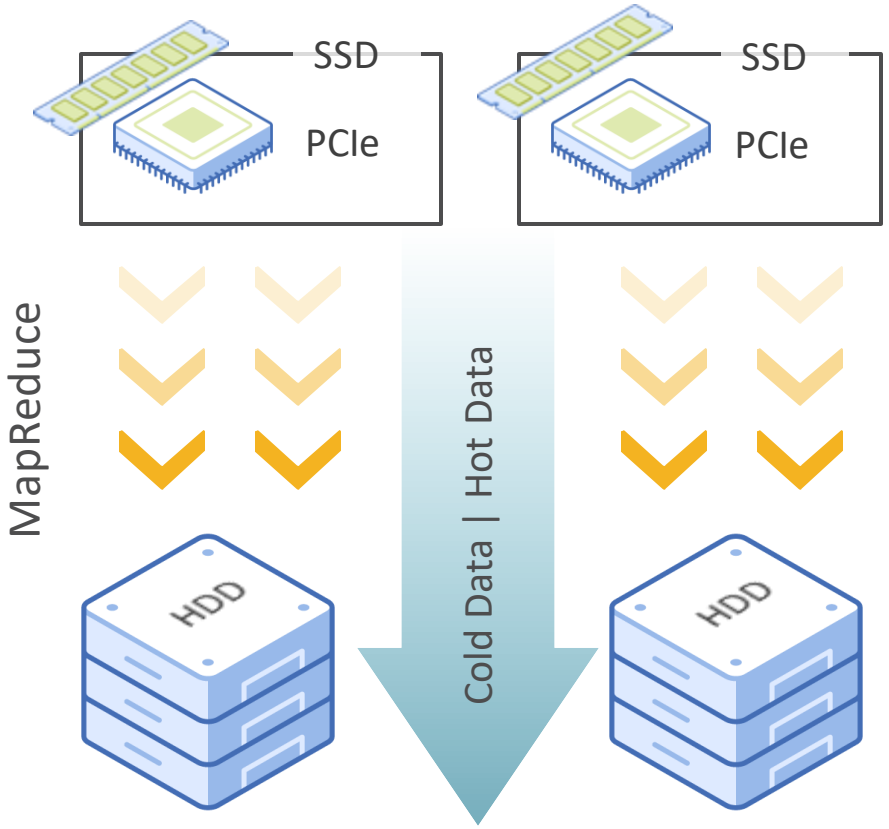
Flexibility

- Linearly scale performance and capacity
- Separate compute and storage expansion
- Combine the best node types
- Data is automatically rebalanced





Nutanix Intelligent Data Tiering



Automatic Performance Optimization

- ✓ Leverage multiple tiers of storage
- ✓ Continuously monitors data access patterns
- ✓ Optimally places data for best performance
- ✓ No user intervention required
- ✓ VM Flash Mode – Pin VM/VMDK To Flash Tier

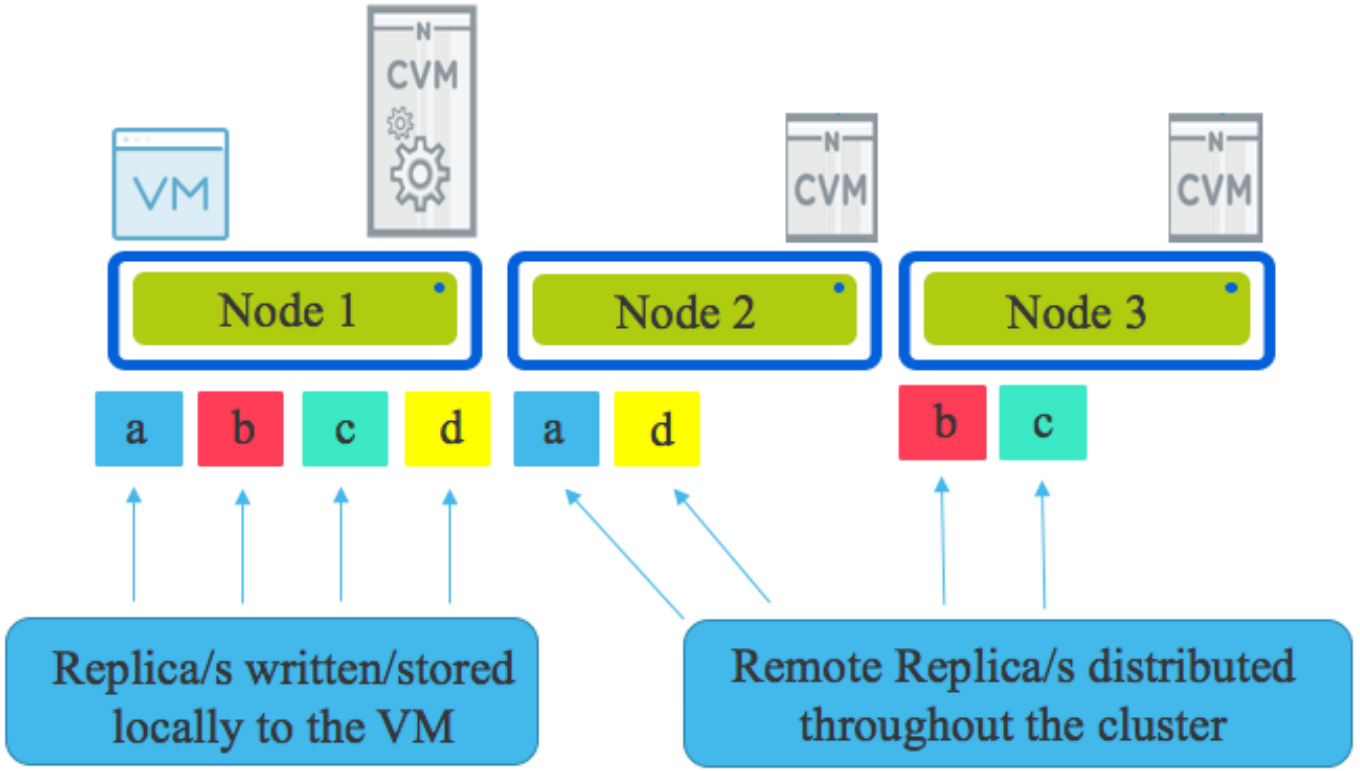
Hot Data SSD

- Random data
- Persistent tier
- Maximum performance

Cold Data HDD

- Sequential data
- Highest capacity
- Most economical

Distributed Storage Fabric



New Flash Technologies Driving Performance

NVMe

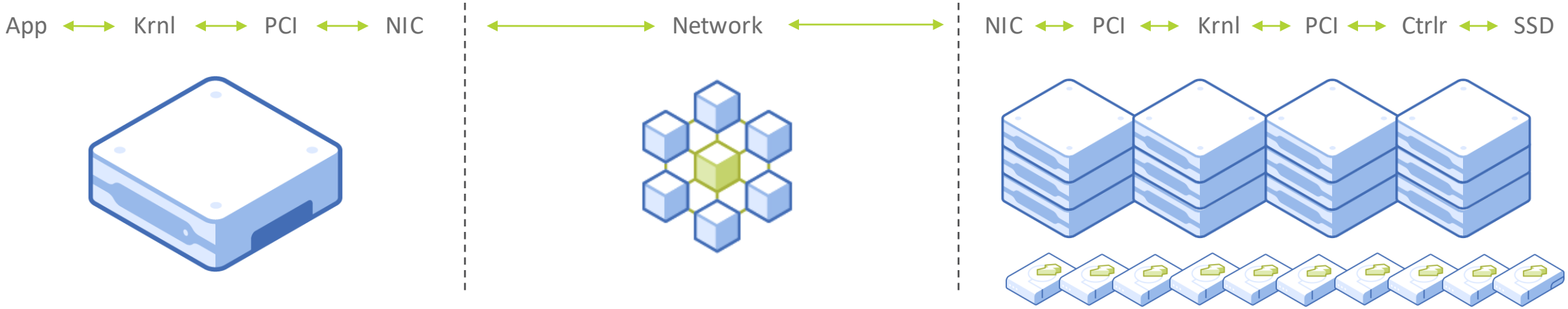
- Replaces SATA and SAS with a PCIe-based standard
- Fabric topologies replace networks

3D Xpoint (Optane)

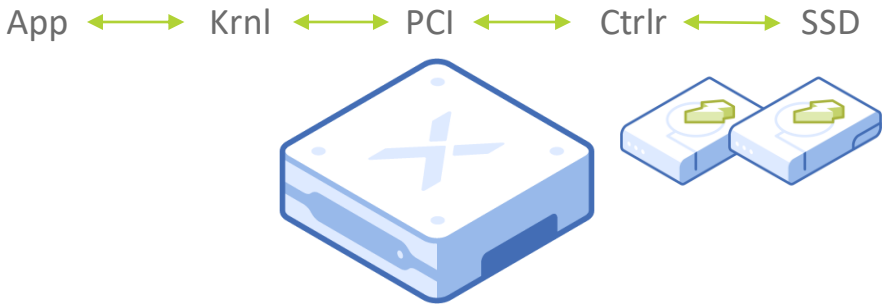
- New nonvolatile storage memory from Intel and Micron
- Much greater performance than today's flash (1,000X)
- Ultra Low Latency (<10us)
- Network latency becomes very important

All-Flash SAN: Long I/O Data Path

SAN I/O Data Path



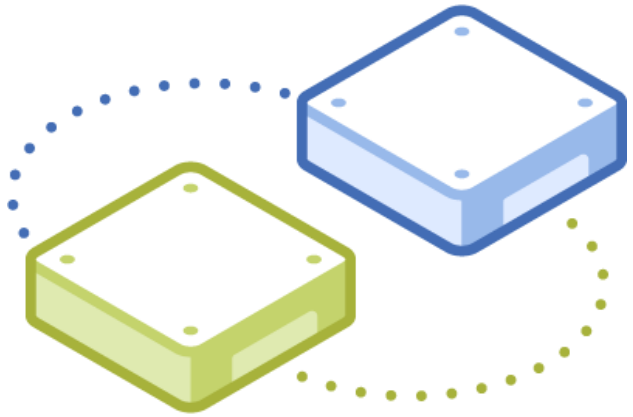
HCI Has a Shorter Data Path … But Not All HCI Are Created Equal



Nutanix Has Data Locality by Design

- ✓ Keep primary data copy on the same node as VM
- ✓ 2nd/3rd (RF2/RF3) copies are distributed throughout the cluster
- ✓ All read operations localized on same node
- ✓ If the VM moves, all new data is written locally
- ✓ Reads of remote data trigger ILM to transparently re-localize data
- ✓ Reduces network chattiness significantly

Nutanix Modern Self-Healing Systems



Self-healing system

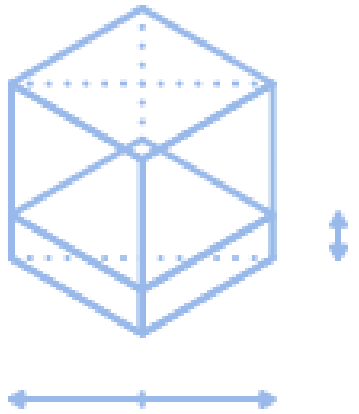
Fault isolation with distributed recovery

- SSD or HDD Failure = offline drive
- Parity rebuilt throughout cluster
- Larger the cluster, faster the recovery
  - Typically, min rather than hours/days
- Don’t need to replace drive until capacity needed = no fire drill
- SSD/HDD replacement adds capacity back to the cluster

Nutanix Requires Minimal Free Space

For Nutanix all write I/O goes to either the Extent Storage or Oplog, both of which are housed on the SSD tier. All random writes are serviced by the Oplog until it reaches 95% capacity at which point the oplog is bypassed.

As such, performance remains high until 95% capacity. Therefore only 5% free capacity is required to ensure high performance.

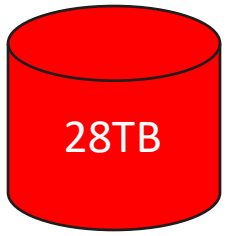


Real World Example

How much less capacity does provide competitor in a 4 node AF cluster?

- 4 Nodes of All Flash
- 6 SSDs per Node
- Competitor Configured with 2 disk groups/node
  - Each disk group requires cache drive
  - Cache drive does not contribute to capacity
- Competitor 10% Operation Reserve – maximum 80% usable space

Nutanix

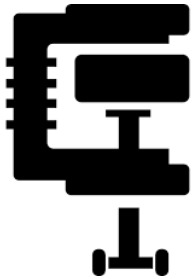


Competitor

	Nutanix	Competitor
# of Nodes	4	4
# of Drives / Node	6	6
Capacity / Drive (TB)	3.84	3.84
RAW Drive Capacity	92.16	92.16
# of Disk Groups	0	2
Cache Drives / Node	0	2
Total Cache Drives	0	8
RAW Cache Drive Capacity	0	30.72
Available RAW Capacity	92.16	61.44
Advertised Capacity (RF2/FTT=1)	46.08	30.72
% Oplog /Operations Reserve required	5	10
Usable Capacity	43.78	27.65

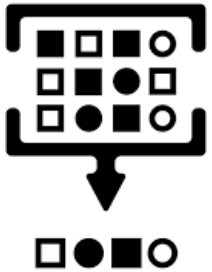
Space Efficiency mechanisms

Compression



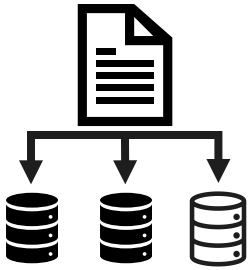
- Configured at storage container level
- Enabling/Disabling applied live to new data. Post-processed on existing data to preserve performance.
- Fast algorithm in-line for performance, better algorithm applied post-processing to increase efficiency while been impact less on performance

Deduplication



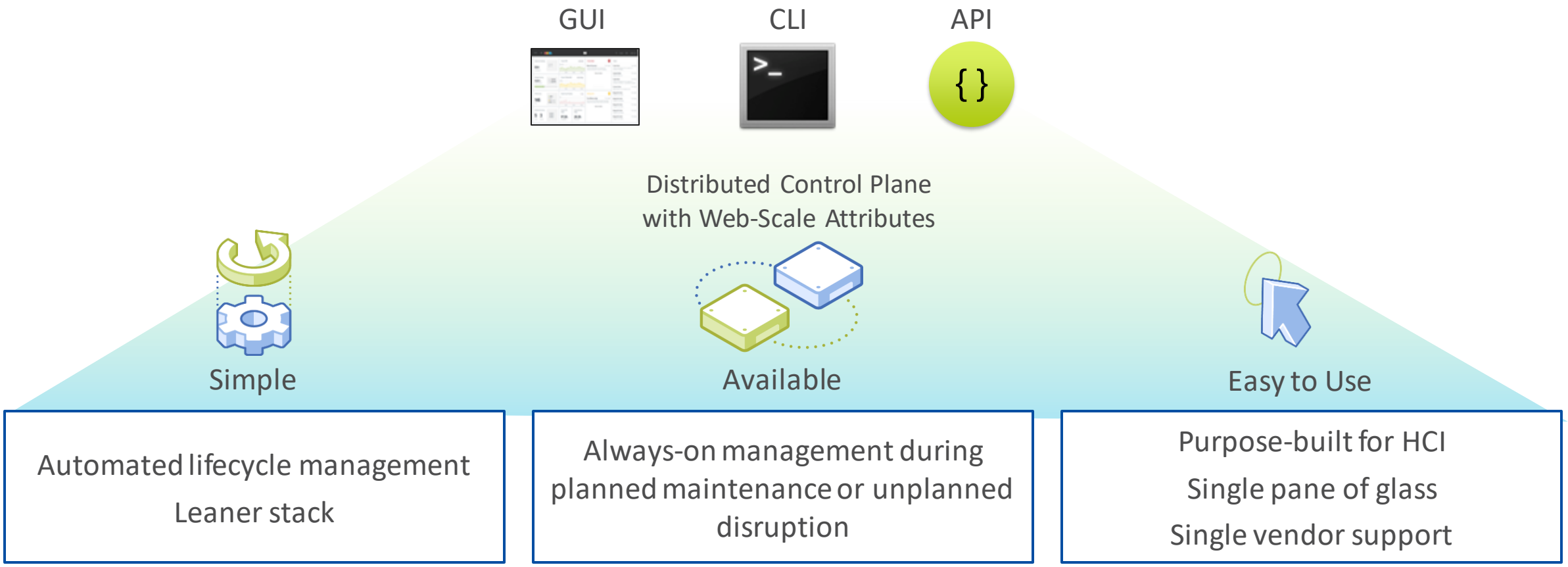
- Configured at storage container level, independent but compatible with compression
- Enabling/Disabling applied live to new data. Post-processed on existing data to preserve performance.
- Deduplication global to whole cluster => better efficiency
- Reducing bandwidth usage during Resync between host & Replication between clusters.

Erasure Coding

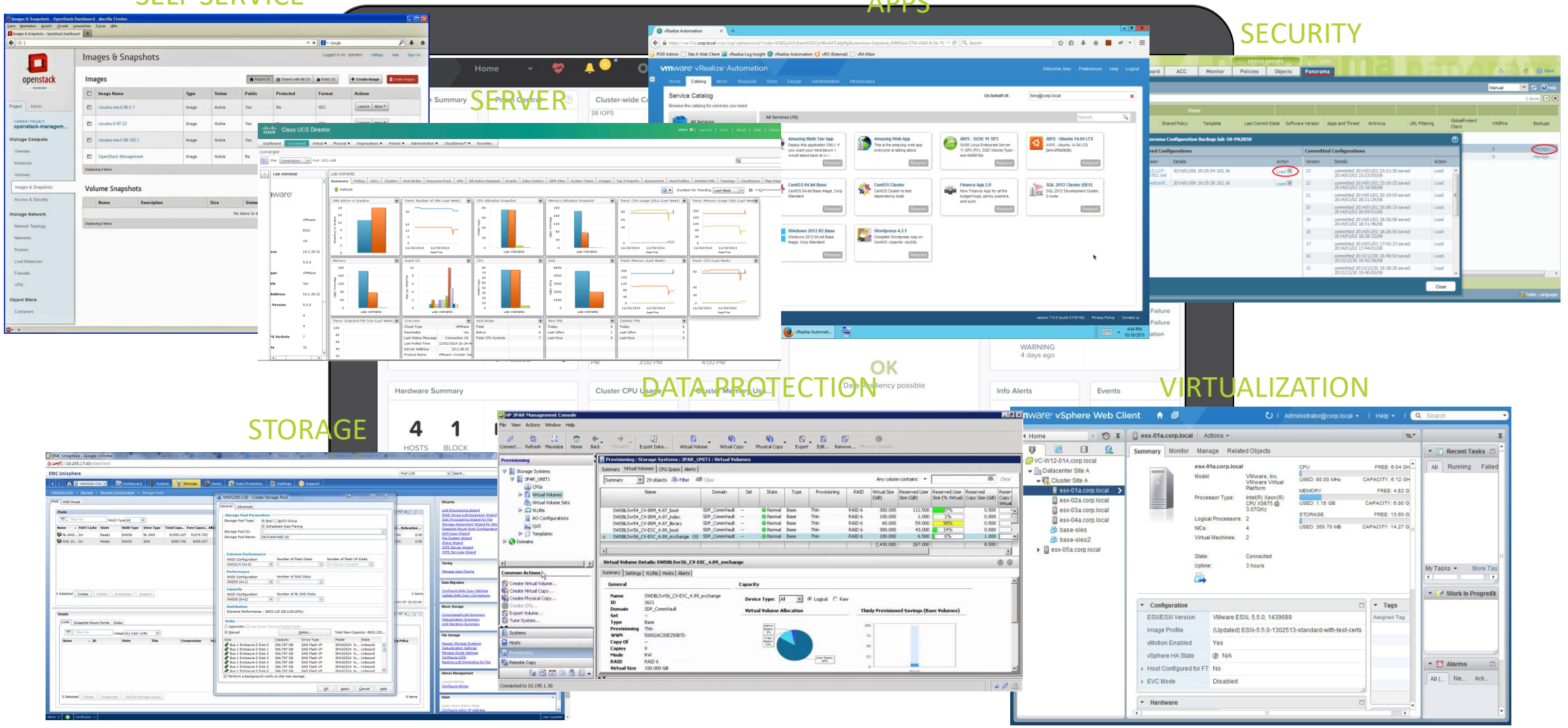


- Configured at storage container level
- Strip size adjusted to cluster size to protect against host failure without recalculating the full strip => faster reprotection
- Apply only on cold data => lower efficiency but impact less on performance

Nutanix Delivers Natively Integrated Management



Introducing Prism: The Answer to Your Frustrations



Infrastructure Management and IT Operations into one UI

PRISM

API / GUI

Financial Resource Insights
- VM Cost Metering
- Chargeback
- Budgeting

Operations Automation
- Codeless Task Automation
- Action Gallery

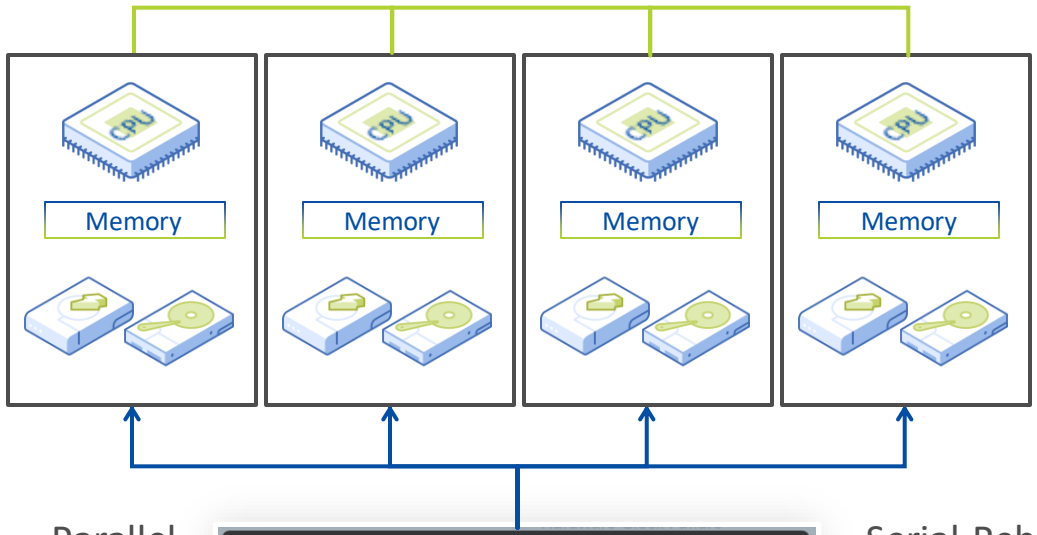
Application Insights
- Application Discovery
- Application & Non-AOS VM Monitoring

Predictive Optimization & Remediation
- Capacity forecast
- Capacity Optimization
- Just-in-time planning

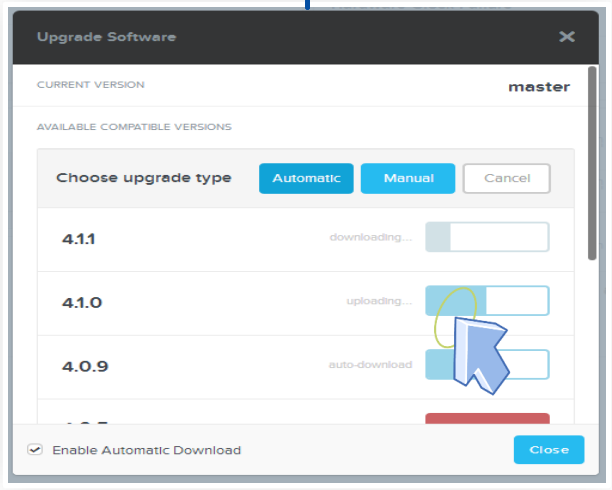
Personalized Insights
- Comprehensive search
- Customizable Dashboards & Reports
- Anomaly Detection

Multi-Hypervisor Full-Stack Management
- Cluster Management
- VM Management
- Storage & Network Management
- Self-Service (RBAC)
- Centralized Upgrades

Nutanix Delivers One-click Upgrades



Parallel Upgrades



Serial Reboots

What is it

- Automatically upgrade Nutanix software, Hypervisors and firmware non-disruptively with no manual intervention

Benefits

- Nodes upgraded in parallel
- Automatic sequencing of reboots
- Done in minutes with zero touch
- No downtime while upgrade happens
  - (CVM Autopathing)
- As easy as upgrading iOS

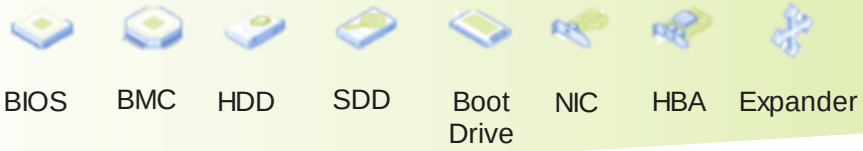
LCM: Non-Disruptive Upgrades

Easy to Use

Removes Complexity

Scalable

Software & Firmware Full Stack Upgrades

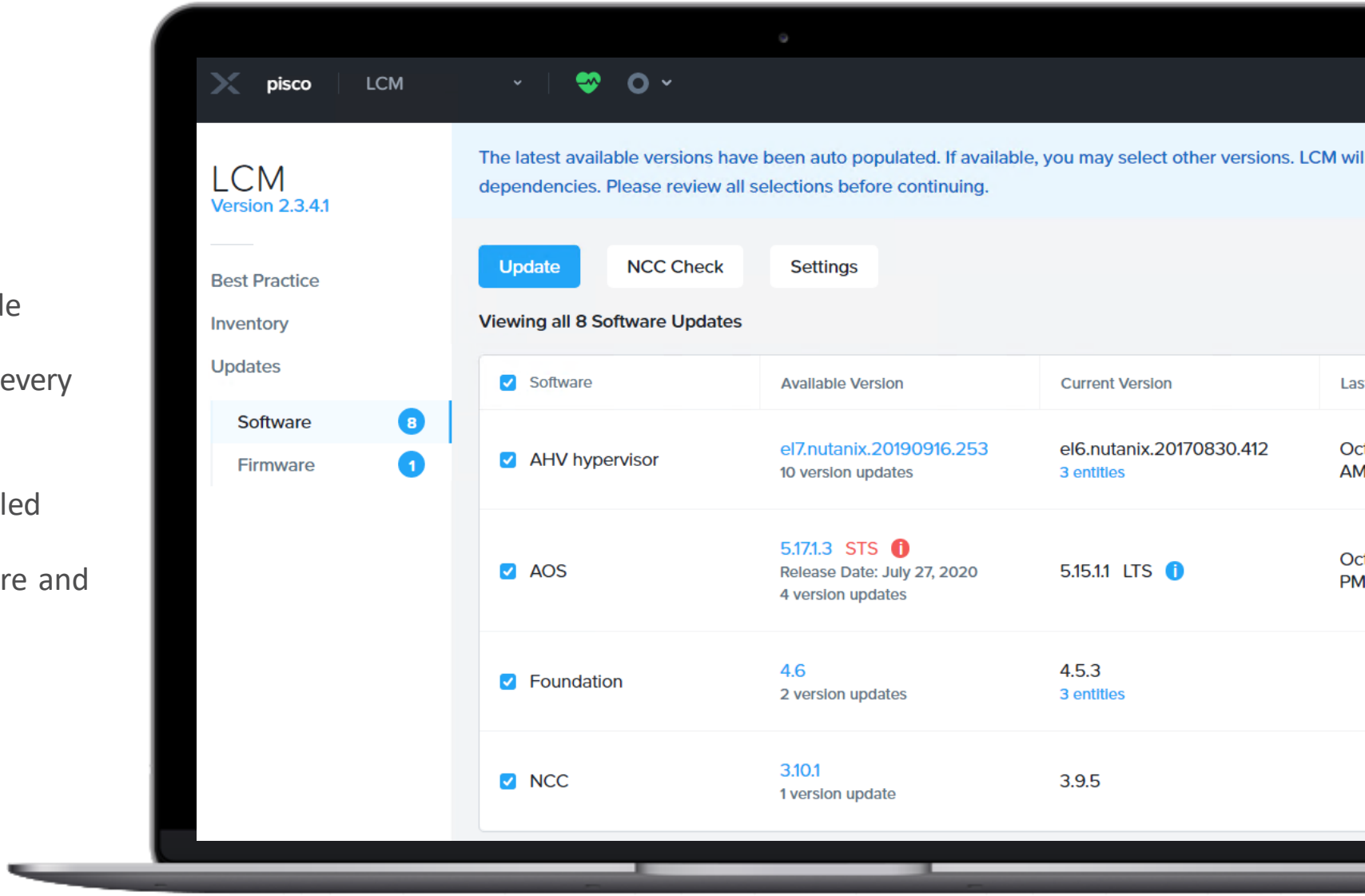


Automatic Dependency Management

Unified Upgrade Process

LCM: worry-free upgrade for the whole stack

- From hardware firmware and software components to full-stack solution upgrade
- Can updates all at once or granularly for every components
- All dependencies are automatically handled
- Single-source of truth for Nutanix software and system upgrades



Elcore

PEOPLE
TECHNOLOGIES
WE INTEGRATE