Versity
DELL Technologies

What we do

Software defined storage platform

Manage & Protect

large unstructured data collections at low cost

MASS STORAGE & LARGE ARCHIVE

Versity

Today

Why we exist?

Where are we headed?

First, where have we been?

Co-founder & CTO - Going Strong!

LSC 1986

SAM-FS

Sun 2001

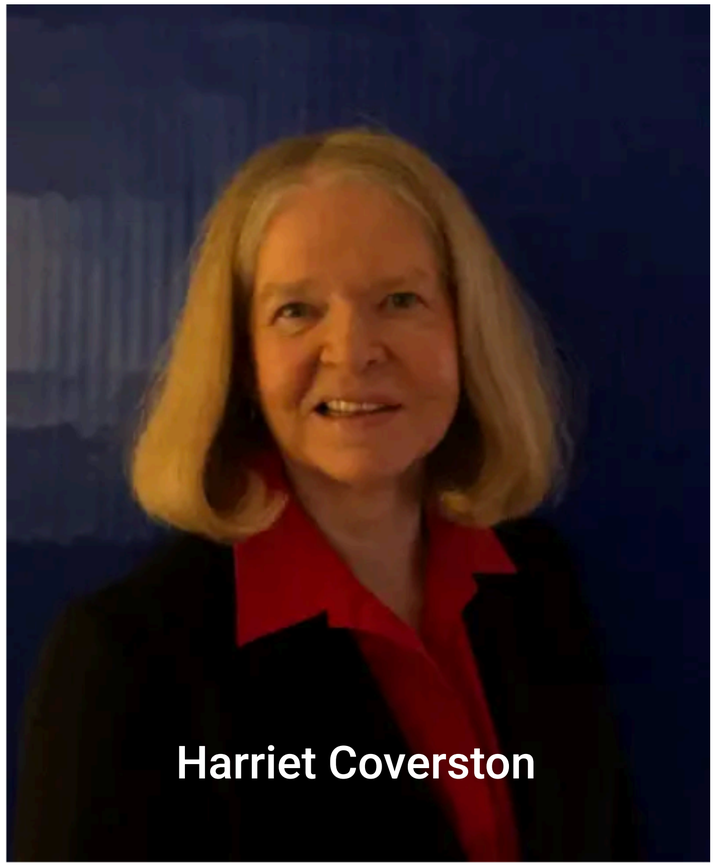
SAM-FS

Oracle 2010

OHSM (SAM-FS

Versity 2011 - now

VSM (SAM-FS) & **ScoutAM**

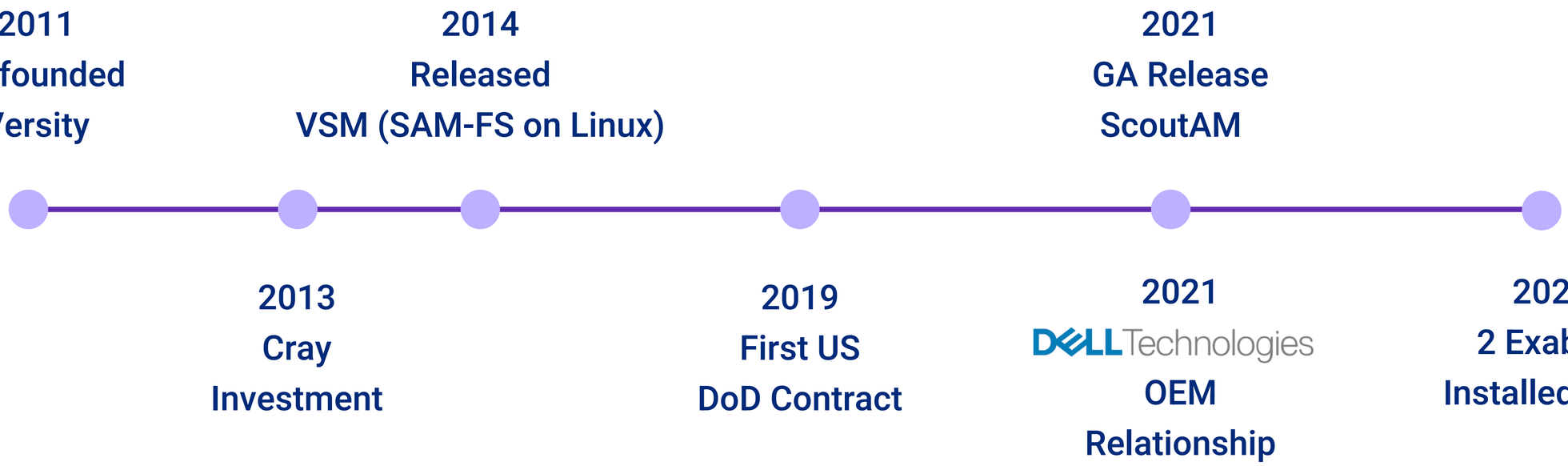


One of the longest continually employed programmers in the world!

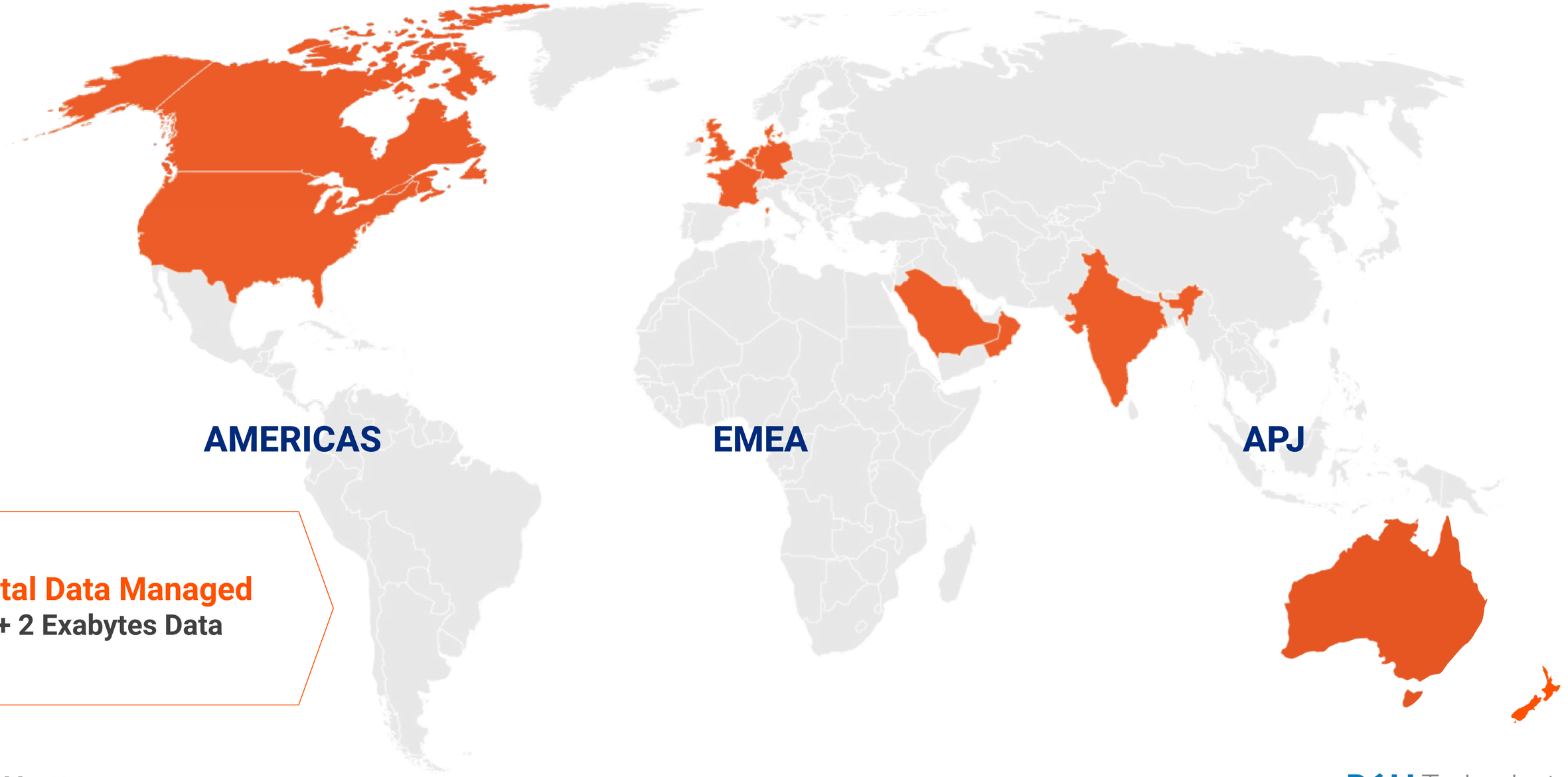
Livermore 1967 - Present = 56 years

Versity

Versity History - 12 Years Old Already!

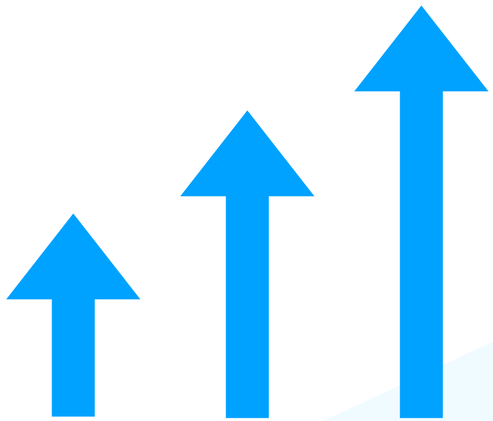


Versity Installed Base



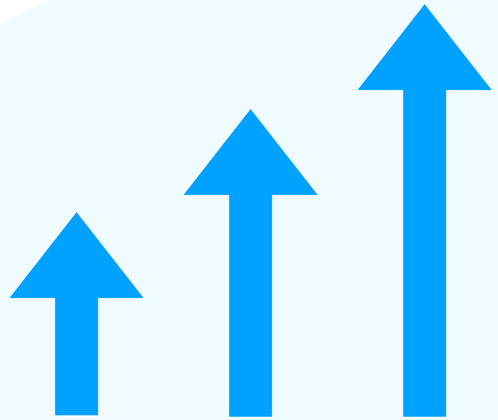
Total Data Managed
+ 2 Exabytes Data

Compute & Archive



FLOPS

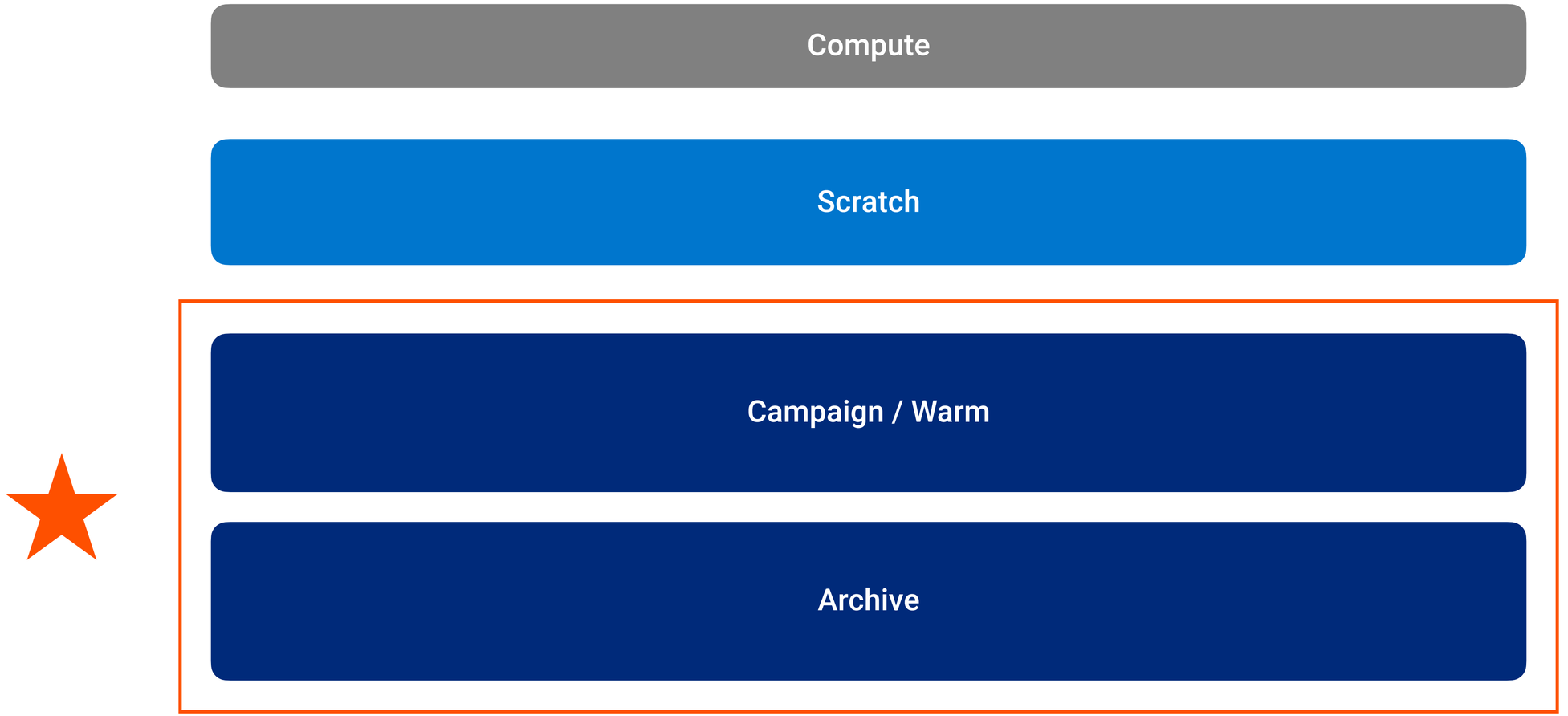
Innovation: HIGH



Petabytes

Innovation: LOW

Where Versity Plays - HPC



Today

OUR MISSION

Develop new & innovative archiving technologies

Versity

Why has technology fallen behind?

Structural issues

Entrenched players

High switching costs

No disruptive force

Problem = Low innovation

Solution

Versity ScoutAM

VSM/SAM-FS Hit 3 Roadblocks (after 35 years)

Namespace Scalability

- ✗ 100's of Millions
- ✓ 10's of Billions

- ✗ Namespace scans
- ✓ Advanced Query Interface

- ✗ Full dump / restore
- ✓ Incremental dump / restore

- ✗ Local namespace
- ✓ Namespace replication

Data Parallelism

- ✗ Central MDS with clients
- ✓ Scale Out Cluster

- ✗ Data Only
- ✓ Data & Metadata

- ✗ Legacy tape handling code
- ✓ All new tape handling code

- ✗ Hard limits to data in flight
- ✓ Unlimited data in flight

Overall Value

- ✗ Difficult to install
- ✓ Easy to install

- ✗ Difficult to run & manage
- ✓ Easy to run & manage - GUI

- ✗ Disjointed architecture
- ✓ Coherent architecture

- ✗ Monolithic
- ✓ Modular

ScoutAM - All New Technology

- Clean full implementation - written in Go GOLANG Google
- Not one line of code from VSM/Sun Oracle
- New commercial archiver runtime application
- New open source shared block filesystem (ScoutFS)

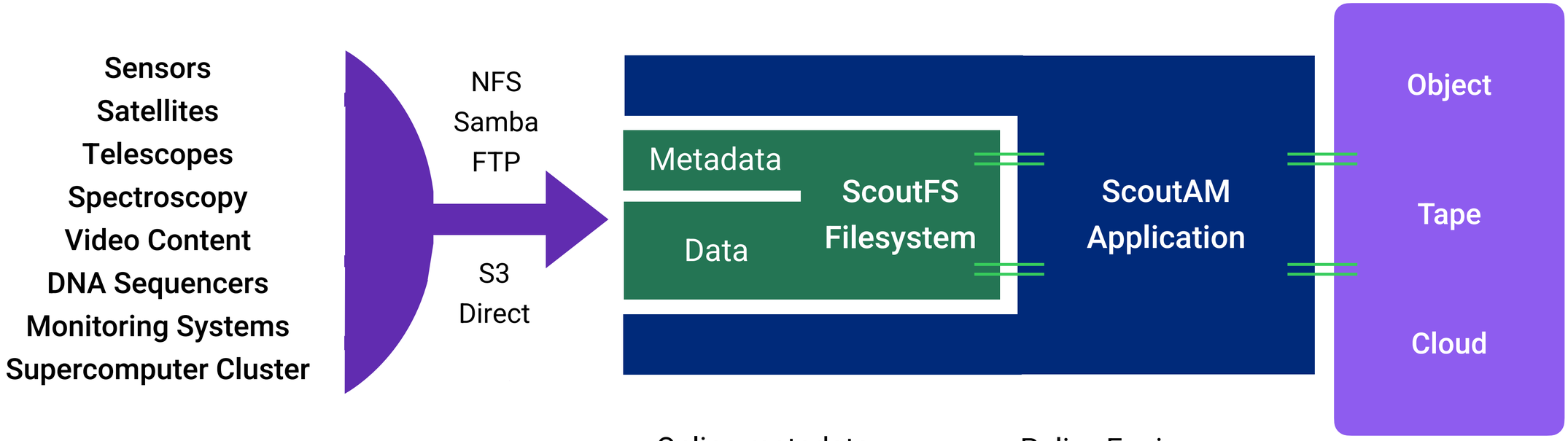
Versity Confidential

Benefits of new

- Scalable
- Modular
- Easy for customers to install and configure
- Easy for Versity to change and evolve
- Only mass storage platform with a modern code base

Scale Out Archive Manager: ScoutAM

A modern data management platform

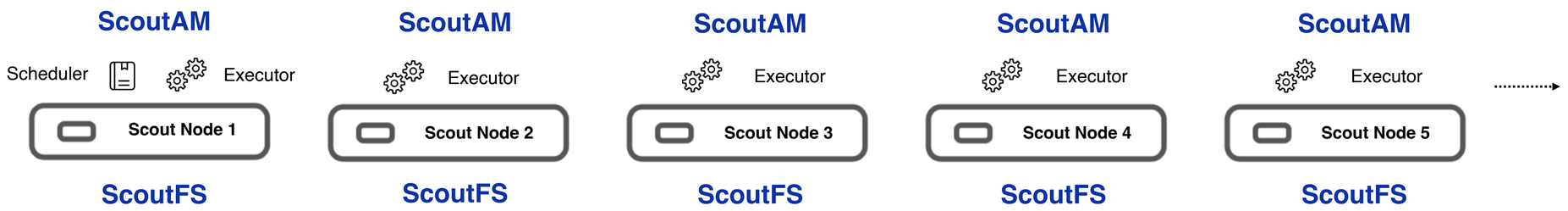


- Online metadata
- Indexed attributes
- Scalable namespace

- Policy Engine
- Scheduling
- Parallel Data Movement

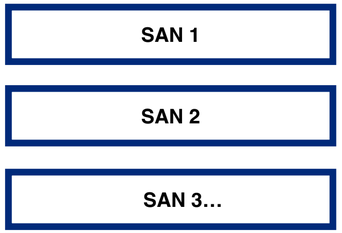
Versity

Scale Out Architecture - Scalability & HA



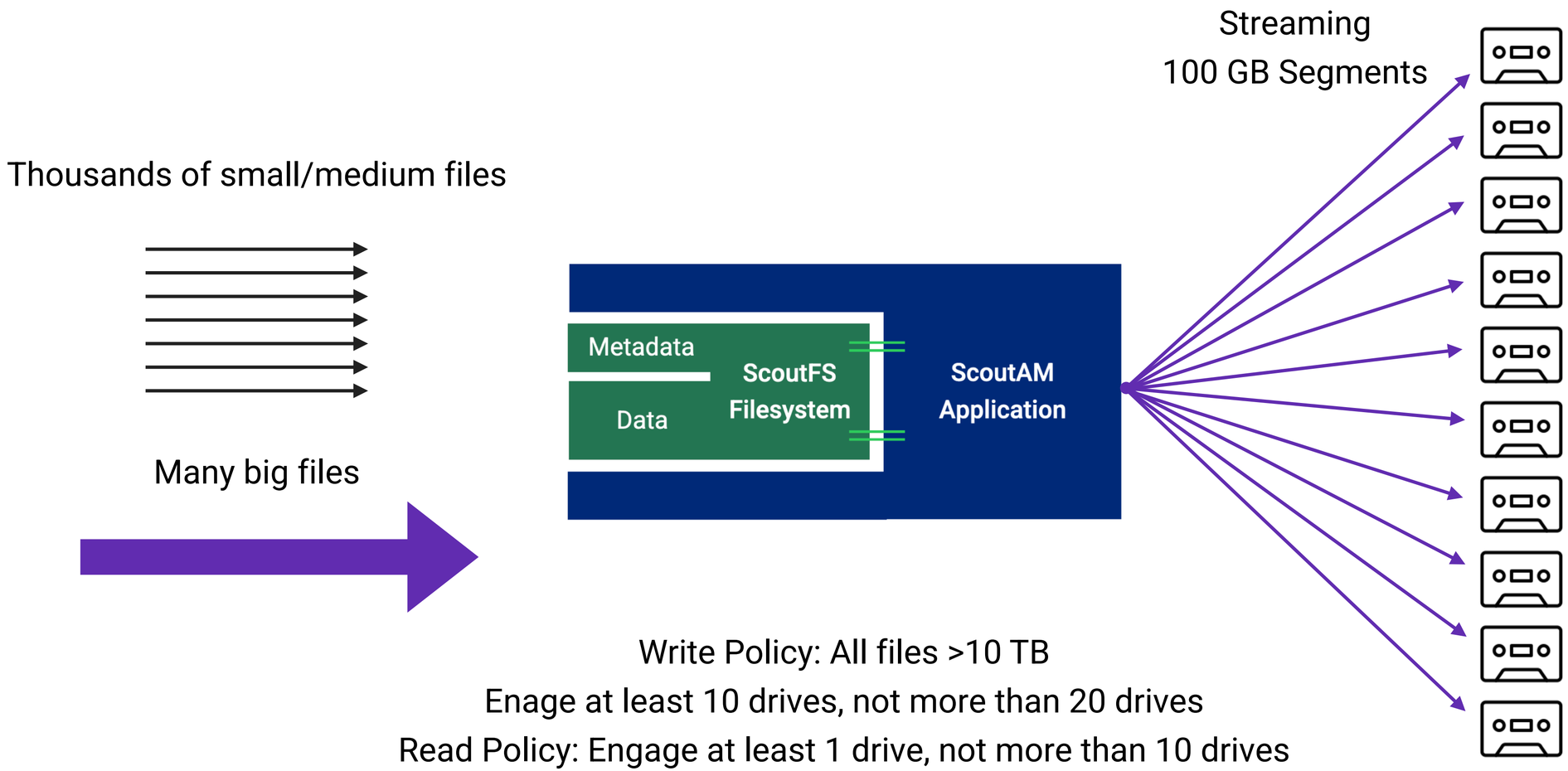
Scales by number of server nodes

Scales by number of SAN devices



+ Dual Site or Multi-Site Replication
+ Incremental Dump & Parallel Restore
+ High Availability (2 or more nodes)

Data Parallelism



Write Policy: All files >10 TB

Enage at least 10 drives, not more than 20 drives

Read Policy: Engage at least 1 drive, not more than 10 drives

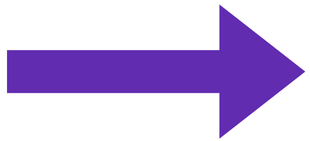
Namespace Protection & Redundancy

OLD

Monolithic Dump/Restore

One giant dump file

Long recovery window



NEW

- Incremental Dumps
- Streaming replication - DR site
- Streaming réplication - primary site
- Ongoing background dump/restore

Brute force restore
@ 250,000 files / second
and going higher

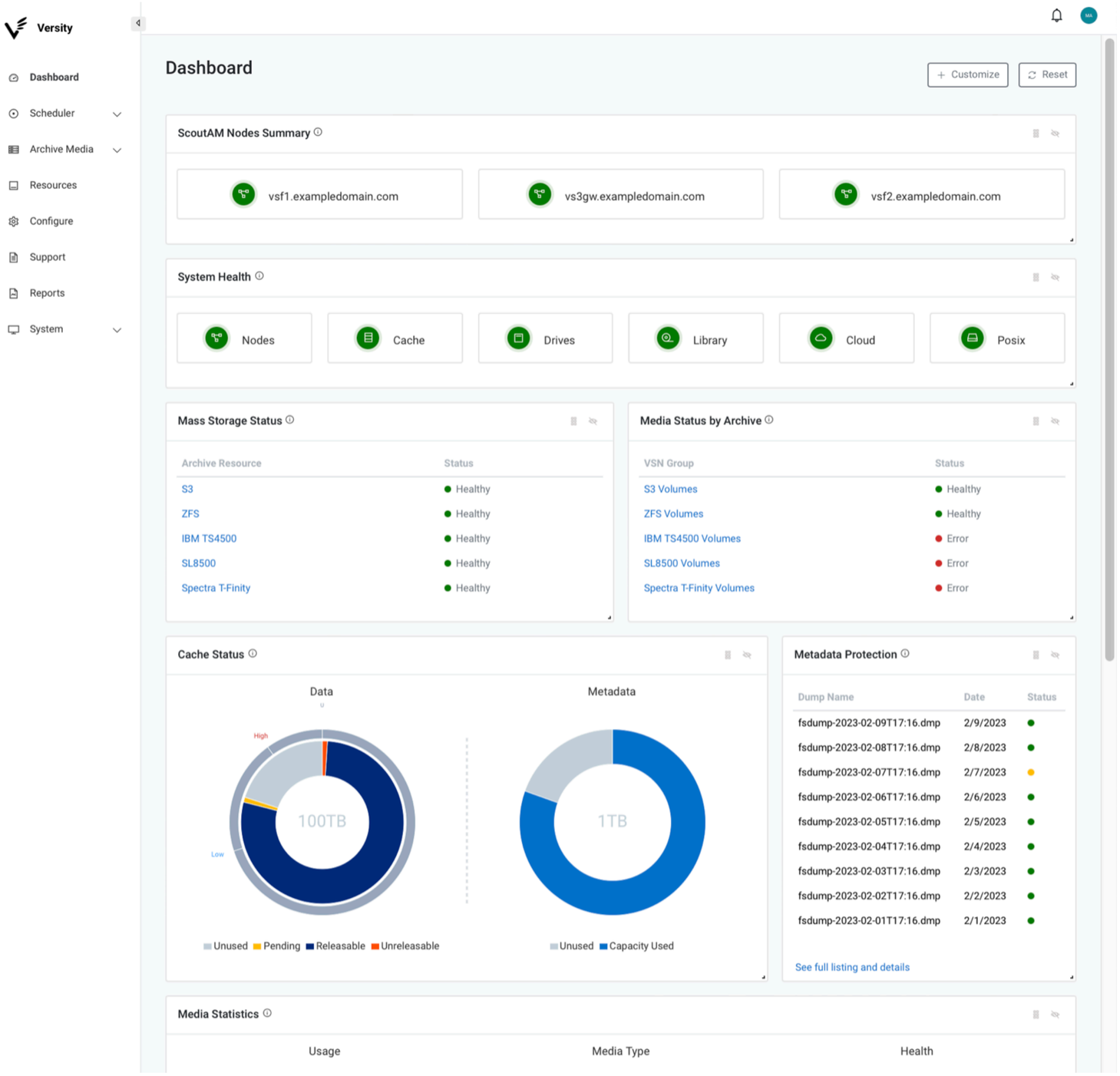
Versity

Versity Confidential

Overall Value

- Easy to install - 10 Minute RPM install
- Easy to manage - 5:1
  - Modern Graphical User Interface
  - REST API driven - full API coverage
- Modular and expandable
- Decoupled from hardware

Versity Confidential



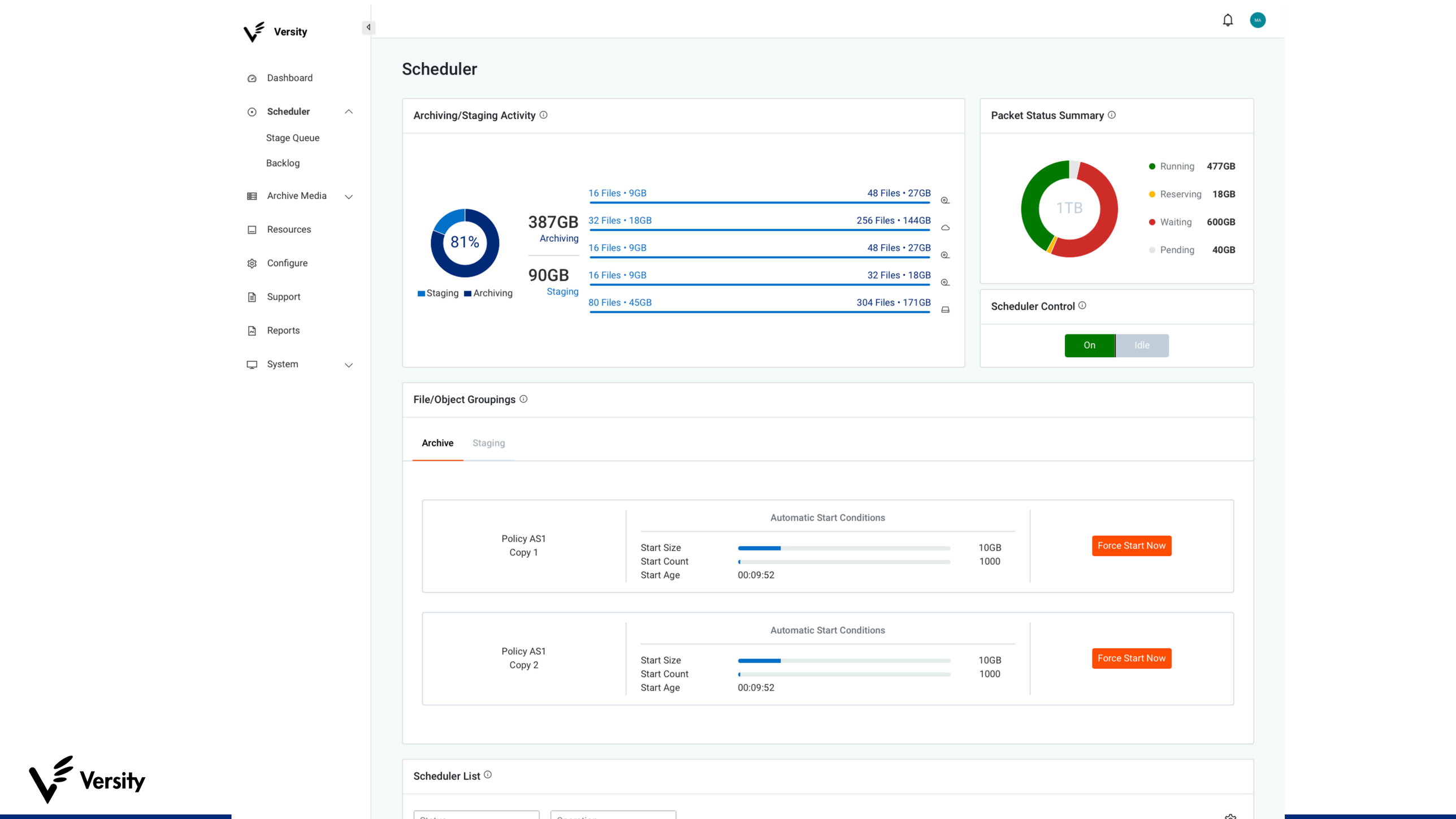
Interface

Modern REST API Driven

Easy to use

Easy to install

Versity


Versity
Dashboard
Scheduler
Stage Queue
Backlog
Archive Media
Resources
Configure
Support
Reports
System
Scheduler
Archiving/Staging Activity
81%
Staging
Archiving
387GB
Archiving
90GB
Staging
16 Files • 9GB
48 Files • 27GB
32 Files • 18GB
256 Files • 144GB
16 Files • 9GB
48 Files • 27GB
16 Files • 9GB
32 Files • 18GB
80 Files • 45GB
304 Files • 171GB
Packet Status Summary
1TB
Running 477GB
Reserving 18GB
Waiting 600GB
Pending 40GB
Scheduler Control
On
Idle
File/Object Groupings
Archive
Staging
Policy AS1
Copy 1
Automatic Start Conditions
Start Size 10GB
Start Count 1000
Start Age 00:09:52
Force Start Now
Policy AS1
Copy 2
Automatic Start Conditions
Start Size 10GB
Start Count 1000
Start Age 00:09:52
Force Start Now
Scheduler List

Versity

- Dashboard
- Scheduler
  - Stage Queue
  - Backlog
- Archive Media
- Resources
- Configure
- Support
- Reports
- System

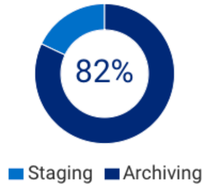
MA

Scheduler / Backlog

Backlog

Backlog Summary

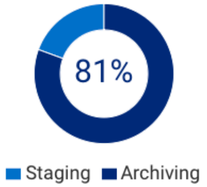
Total Space



547GB
Archiving

120GB
Staging

Number of Files



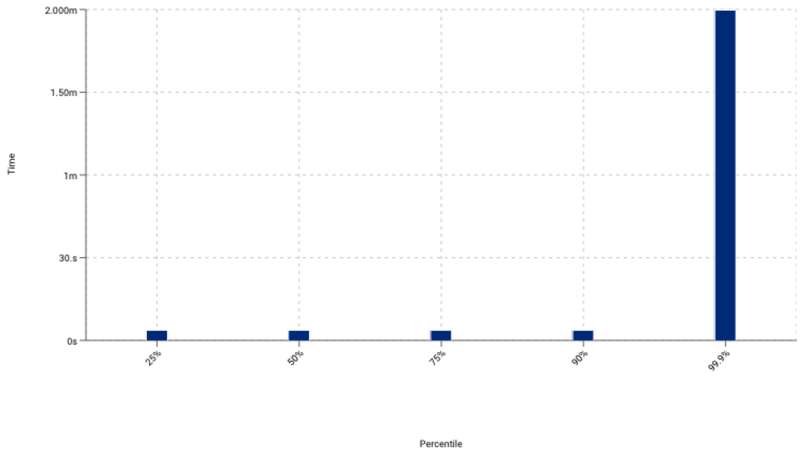
1392
Archiving

335
Staging

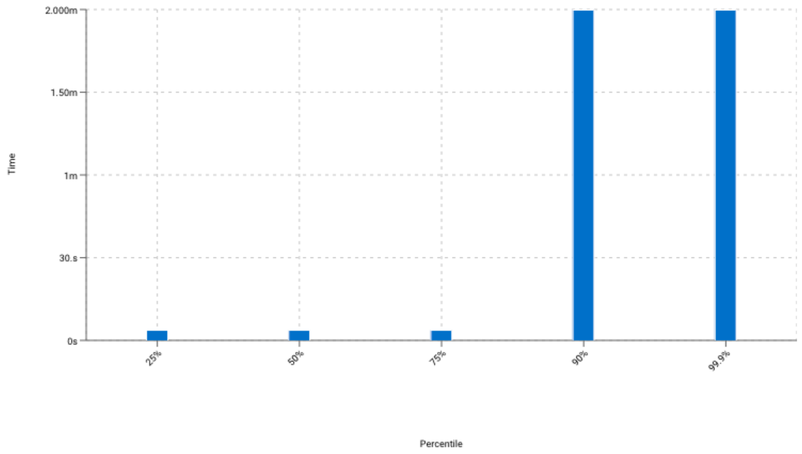
Longest Queued

Rank	Time Queued	Data Size	File Count	Operation
1	2m	10GB	16	Archiving
2	2m	10GB	16	Archiving
3	2m	10GB	16	Archiving
4	2m	10GB	16	Archiving
5	2m	15MB	5	Staging

Time on backlog by percentile: Archiving



Time on backlog by percentile: Staging



Versity

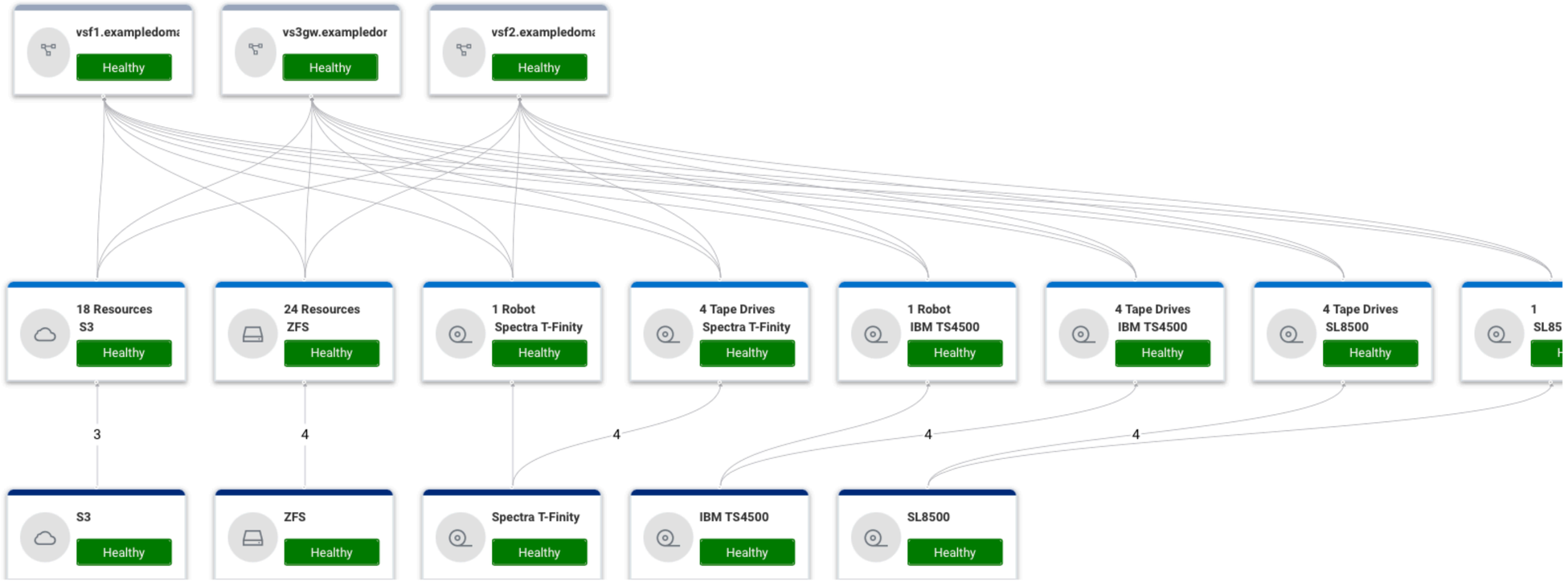
- Dashboard
- Scheduler
  - Stage Queue
  - Backlog
- Archive Media
  - Media Pools
- Resources
- Configure
- Support
- Reports
- System
  - System Visualization
  - Notifications

System / System Visualization

System Visualization

System Visualization

Export as PDF



ScoutAM

Better product needed better distribution

Dell OEM Relationship

Official Dell OEM Engineered Solution since mid 2021

- Dell Reference Architecture
- Solution is tested and certified on Dell hardware
- Dedicated Dell Lab @ Versity
- Support

Turnkey Solutions

Customer designed and validated turnkey configurations

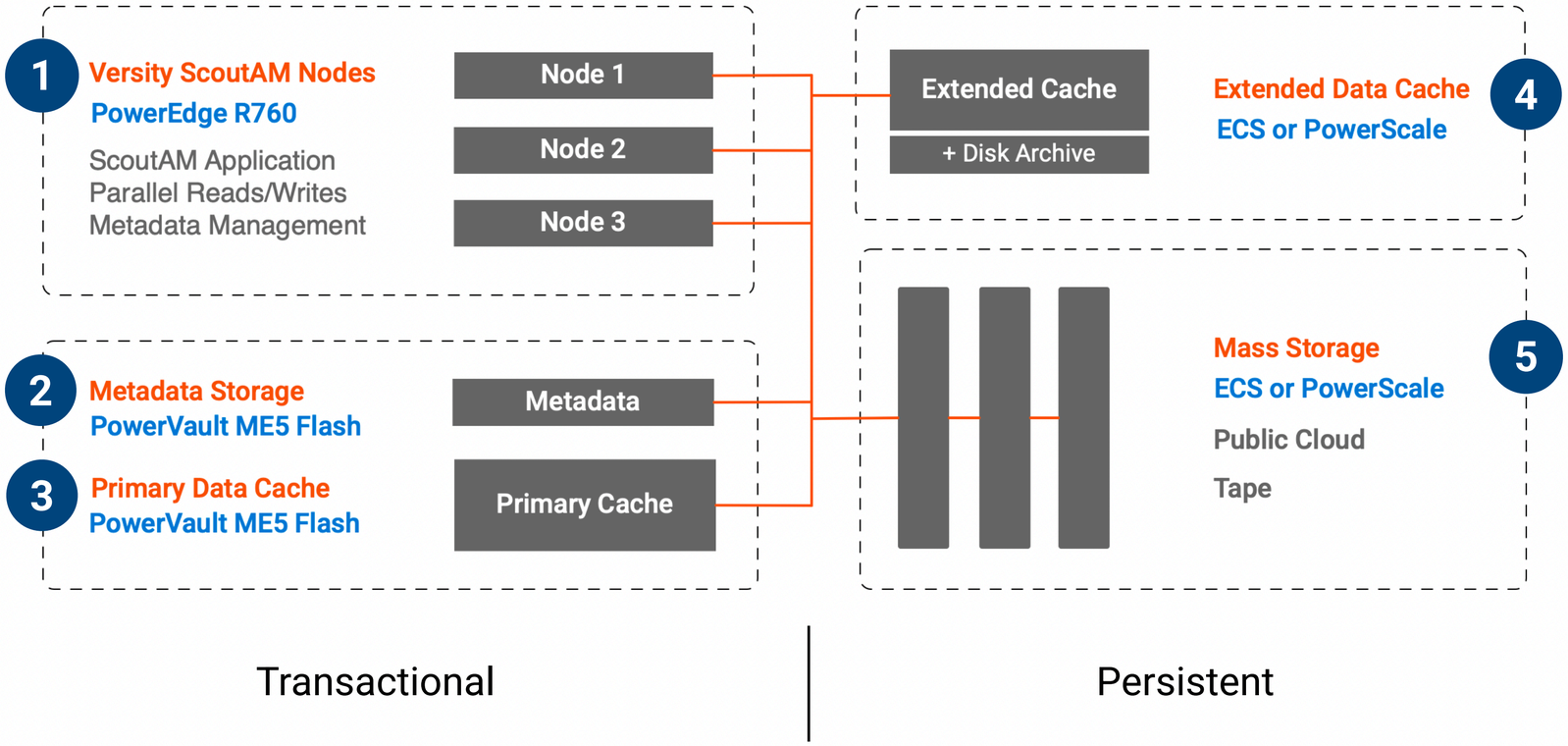
Solves Unique Needs

Meets additional customer technology needs in specific verticals or use cases.

Streamlined Purchase Experience

An easy and flexible purchase experience for Dell Technologies customers

Dell + Versity OEM Reference Architecture



SPECIAL ANNOUNCEMENT

From June 13, 2023

Who ?

- Versity - Project Sponsor / Owner
- Collaborator: Los Alamos National Laboratory
- Collaborator: Pawsey Supercomputing Research Centre

Versity is announcing a new open source project

Versity Gateway

Friendly Licensing & Reliable Sponsor

Versity

Why ?

- MinIO deprecated their widely adopted S3 gateway
- No credible replacement emerged
- Community needs a viable tool
- Community needs new features and active development

Versity Gateway

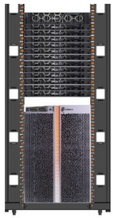
- A high-performance inline S3 to file translation tool
- Brand new and developed by Versity from scratch
- Open Source - Apache 2 licensed
- Scalable stateless architecture - scale out
- High performance - written in Go and on gofiber server framework
- Modular back end support - generic POSIX + Optimized ScoutFS ++

PUBLIC USE CASES

NCSA Solution Architecture

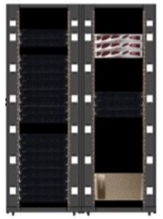
Delta HPC Cluster

Radiant

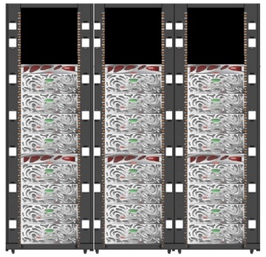


HOLL-I

Dedicated Industry Partner System #4



Hal Cluster - AI and ML workloads



Global Filesystem
Lustre

globus

Cluster Export Nodes
NFS/SMB/S3/Globus

Versity ScoutAM Solution

ScoutAM Nodes
5x Dell R730

ScoutAM Intelligent Cache
Data 2.6PB

60 x RAID 6 - 6TB HDD (8+2) Luns
Metadata 4.5TB
3 x mirrored Luns - 1.5 TB SSDs

34 x JAG7 Tape Drives | 2 x RIM Connections | 60 PB

Versity

NCSA Solution Architecture

Scientific Data

Compute

Global Lustre Filesystem (Taiga)

Cluster Export Nodes
NFS/SMB/S3/Globus

globus

Versity ScoutAM Solution

ScoutAM
Nodes
5x Dell R730

ScoutAM
Intelligent Cache
Data 2.6PB

60 x RAID 6 - 6TB HDD (8+2) Luns
Metadata 4.5TB
3 x mirrored Luns - 1.5 TB SSDs

Mass Storage

Versity

Pawsey reduces technical debt to accelerate innovation

Industry:
Scientific Research

Challenges:
Legacy technology with limited roadmap
Data accessibility required costly upgrade

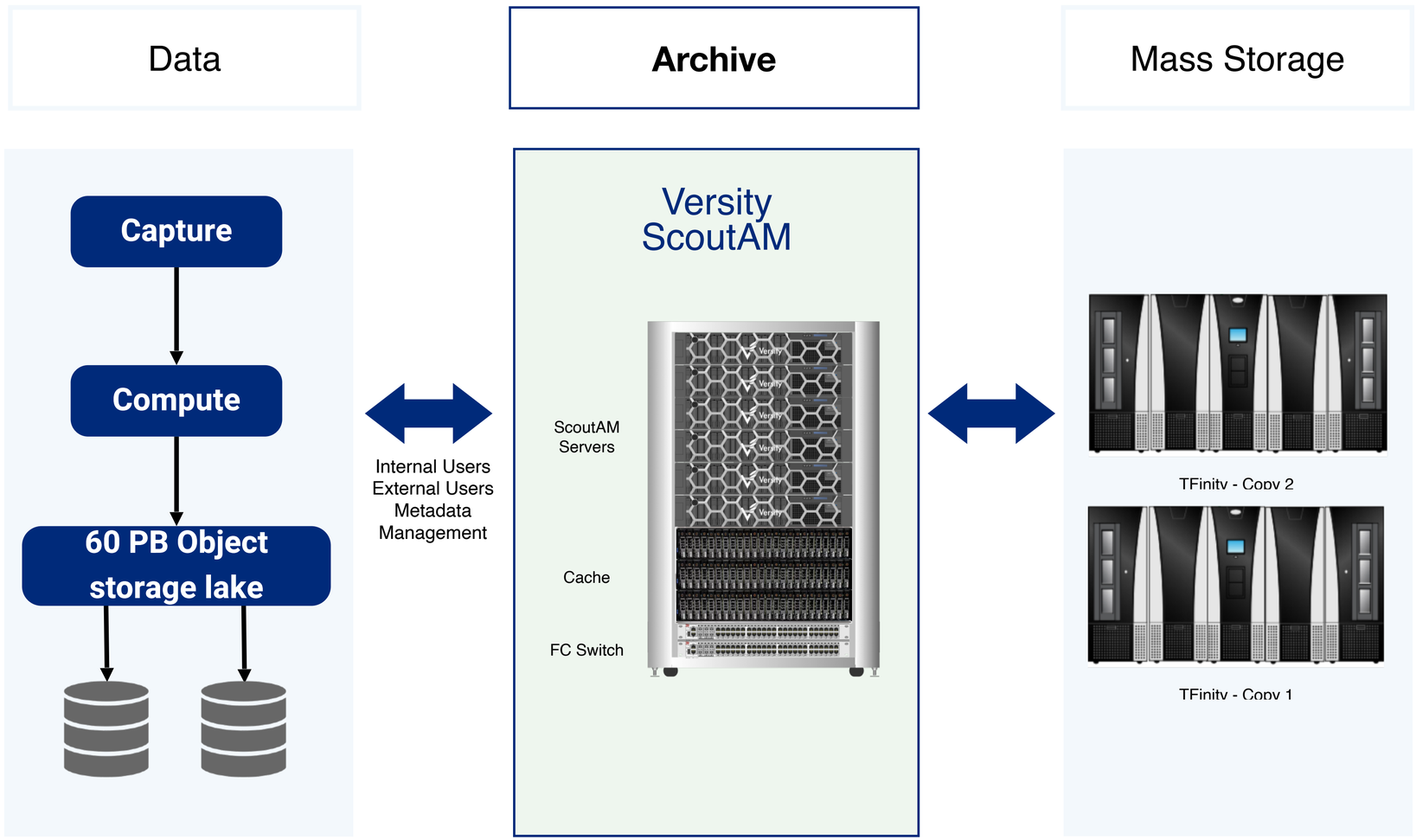
Versity Delivered 150 PB, 34 Tape Drive High Performance Tape Management System

- **High-performance tape management solution: managed 150 PB data in one week with zero data migration**
- **Improved TCO: eliminated the need for expensive upgrade**
- **Ability to access data and plug it into multiple workflows: internal/external researchers, metadata system**

"Versity gave us an easy off-ramp from older proprietary systems and allowed us to tap into and get the most value out of our data. It has been a game changer for our teams."

Mark Gray, Head of Scientific Platforms

Pawsey Solution Architecture



Versity

Thank You

versity@dell.com

sales@versity.com

@versitysoftware