Versity

Bruce Gilpin - CEO

Versity ScoutAM

Efficient Mass Storage for Cloud-Scale Data Collections

World leading archival storage platform

Versity

DELLTechnologies

ScoutAM Archive with Intelligence

ScoutAM

**Maximum scalability
with minimal complexity**

Efficient management and
preservation of exponentially
increasing collections
of unstructured data

Orchestrates extreme
scale data storage workloads

Easy to use, simple to
configure and quick to deploy

ScoutFS

**High-performance
metadata management**

Open-source Linux filesystem
specifically created to
support mass storage
and archiving applications

Enables radically faster
search results in extremely
large file and object populations

Versity Gateway

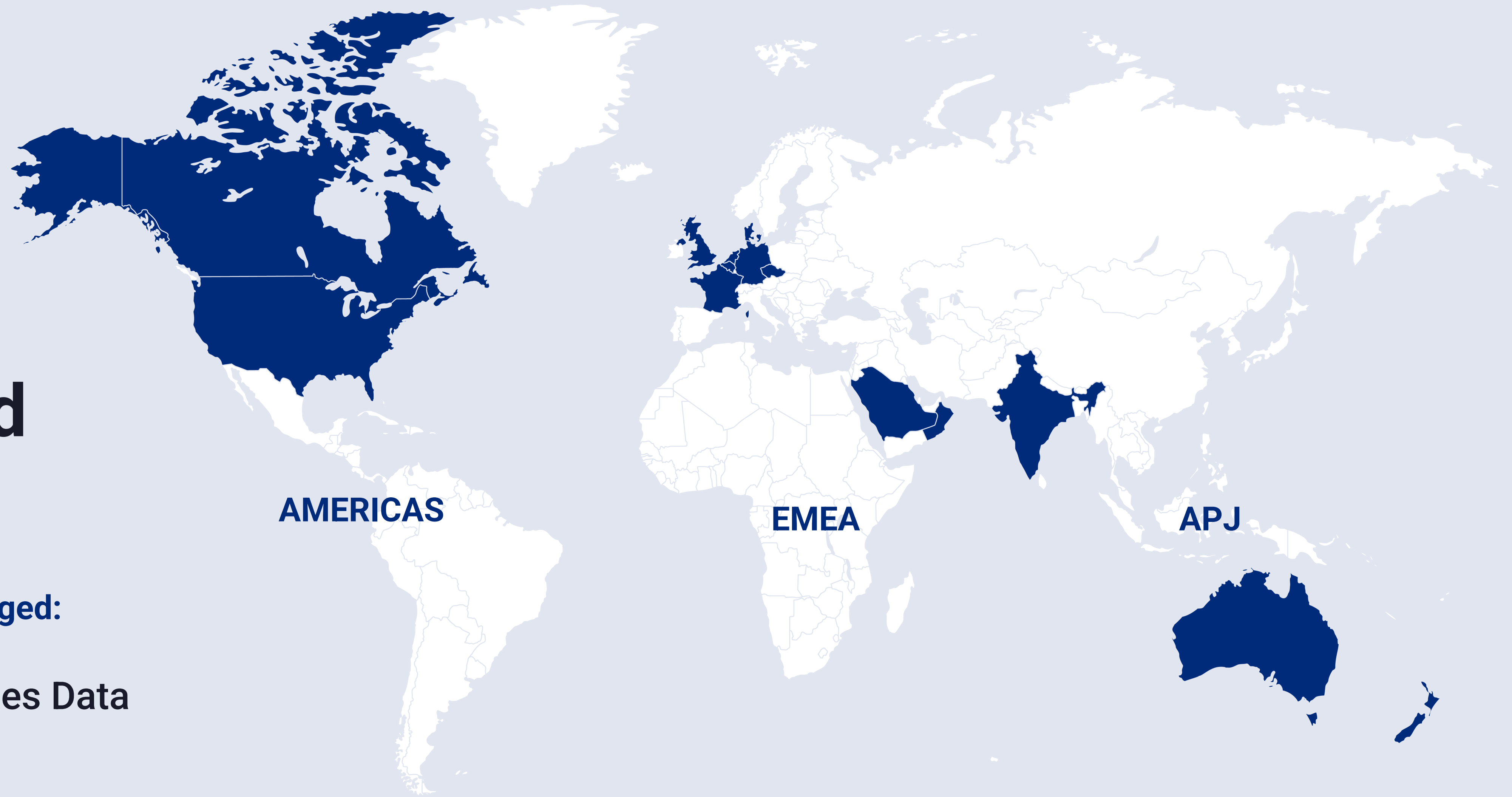**High-performance
object-to-file translation tool**

Seamless file storage capabilities
for S3-based applications

Apache 2.0 licensed
open-source project

Scaleable with modular
backend support

Open Standards + Open-Source

Versity

DELLTechnologies

Global Installed Base

Total Data Managed:

2 Exabytes Data

90 Sites

AMERICAS

EMEA

APJ

Versity

DELLTechnologies

Versity's Global OEM Partner

DELLTechnologies

Turnkey Solution
Customer tailored and validated turnkey configurations
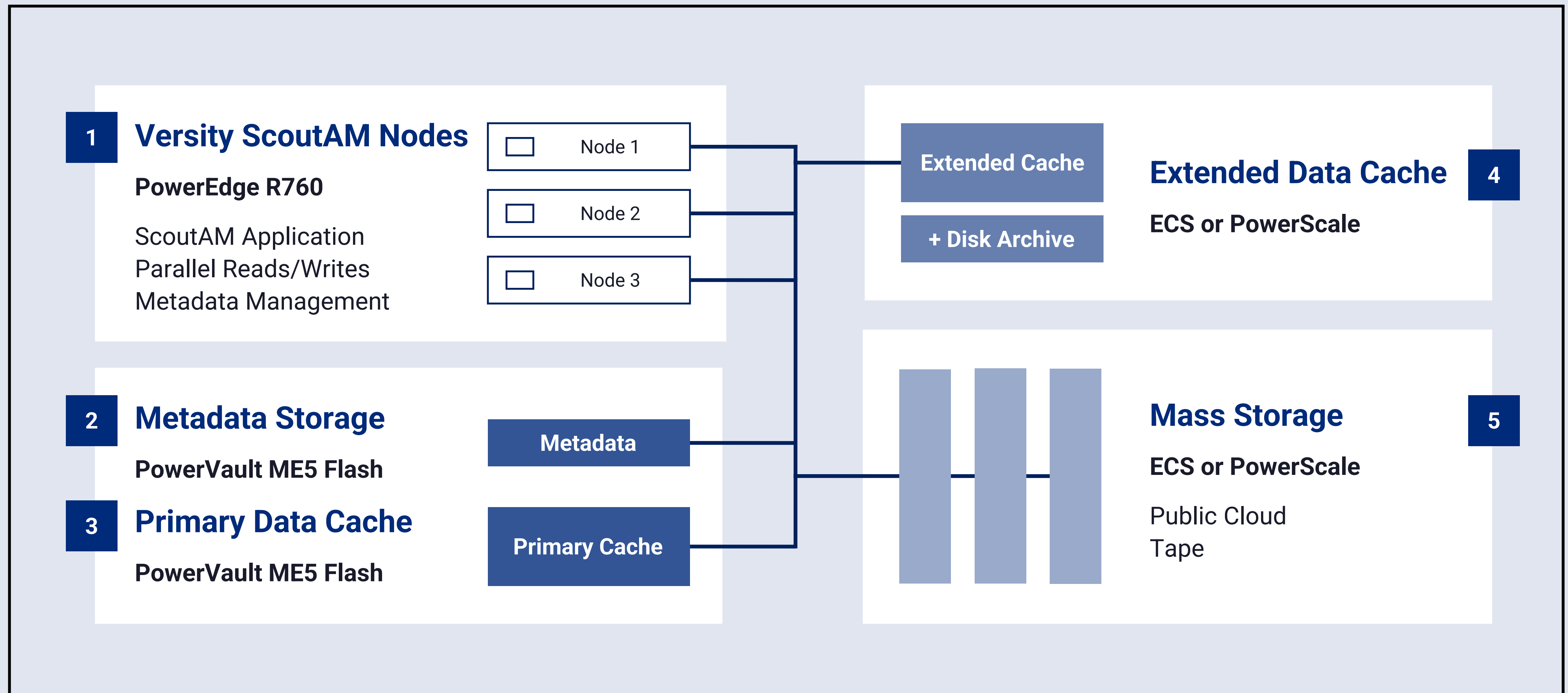
Solves Unique Needs
Meets customer technology needs in specific verticals & use cases

Streamlined Purchase Experience
Easy and flexible purchase experience for Dell Technologies customers

DELLTechnologies

Dell's Modular Reference Architecture

1 **Versity ScoutAM Nodes**

PowerEdge R760

ScoutAM Application
Parallel Reads/Writes
Metadata Management

Node 1

Node 2

Node 3

Extended Cache

+ Disk Archive

Extended Data Cache **4**

ECS or PowerScale

2 **Metadata Storage**

PowerVault ME5 Flash

Metadata

3 **Primary Data Cache**

PowerVault ME5 Flash

Primary Cache

Mass Storage **5**

ECS or PowerScale

Public Cloud
Tape

Versity

Versity Confidential

DELLTechnologies

What is Unique?

- Independent company

- Open-source file system

- Open source file format

- Free and perpetual read-back

- Ultimate disaster recovery

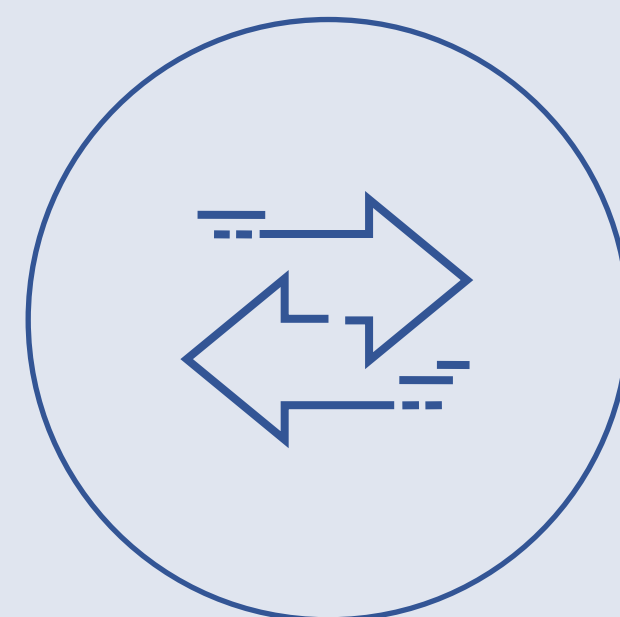- HPSS, TSM, DMF & OHSM format compatibility

Versity S3 Gateway

A free open-source S3 object storage translation tool

S3 ⟶ POSIX

Enables S3 workloads to utilize POSIX
filesystems and tape via ScoutAM

S3 ⟶ Object

Enables S3 workloads to utilize non
S3
object storage systems like Azure

Created By

Versity   Los Alamos NATIONAL LABORATORY   PAWSEY

Versity   DELL Technologies

Why is this interesting ?

1. Makes it easy for end users to support S3 workloads in a mixed storage environment
2. Makes it easy for OEM's to support S3 from NAS and other file based solutions
3. Makes it easy to scale up total ingest bandwidth for high performance applications
4. Makes it easy for S3 applications to work on Azure
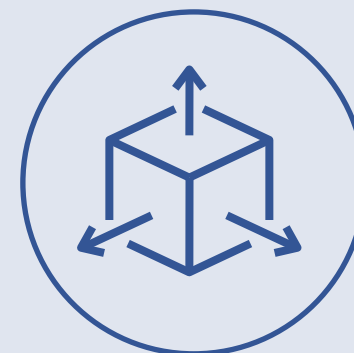5. Clean and friendly open-source license: Apache 2

Very Fast

Written from scratch, Go / GoFiber
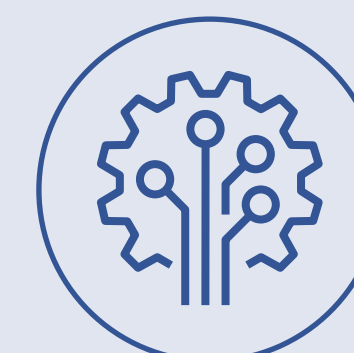Fast processing and response times

Easy Scale Up

Multiple stateless gateway servers per physical server + multiple physical servers. Just run more instances & load balance across them

Modular Back End Storage

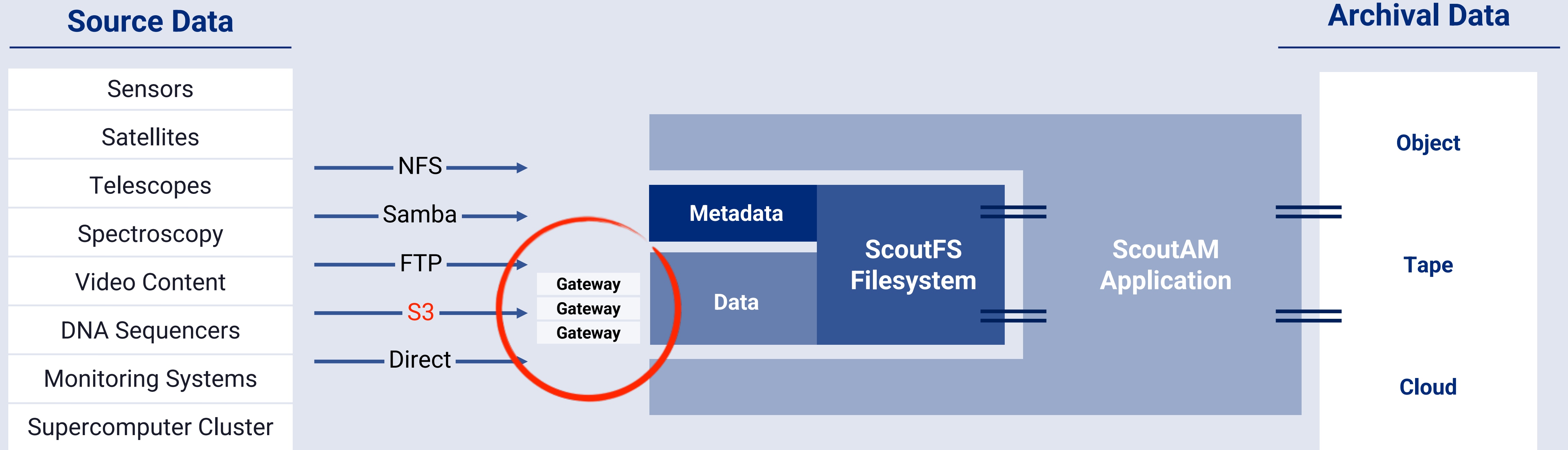Generic POSIX and ScoutAM backends supported, easy to add more

Commercial Quality

Maintainers and key contributors work at Versity. Support contracts available with 7x24x365 global coverage from Versity

Versity

DELL Technologies

Versity Gateway Example - ScoutAM

Source Data

- Sensors
- Satellites
- Telescopes
- Spectroscopy
- Video Content
- DNA Sequencers
- Monitoring Systems
- Supercomputer Cluster

NFS
Samba
FTP
S3
Direct

Gateway
Gateway
Gateway

Metadata

Data

ScoutFS Filesystem

ScoutAM Application

Archival Data

Object

Tape

Cloud

- Parallel ingest (multi-part uploads) across n servers
- Far superior to NFS
- Runs on same server nodes as our application
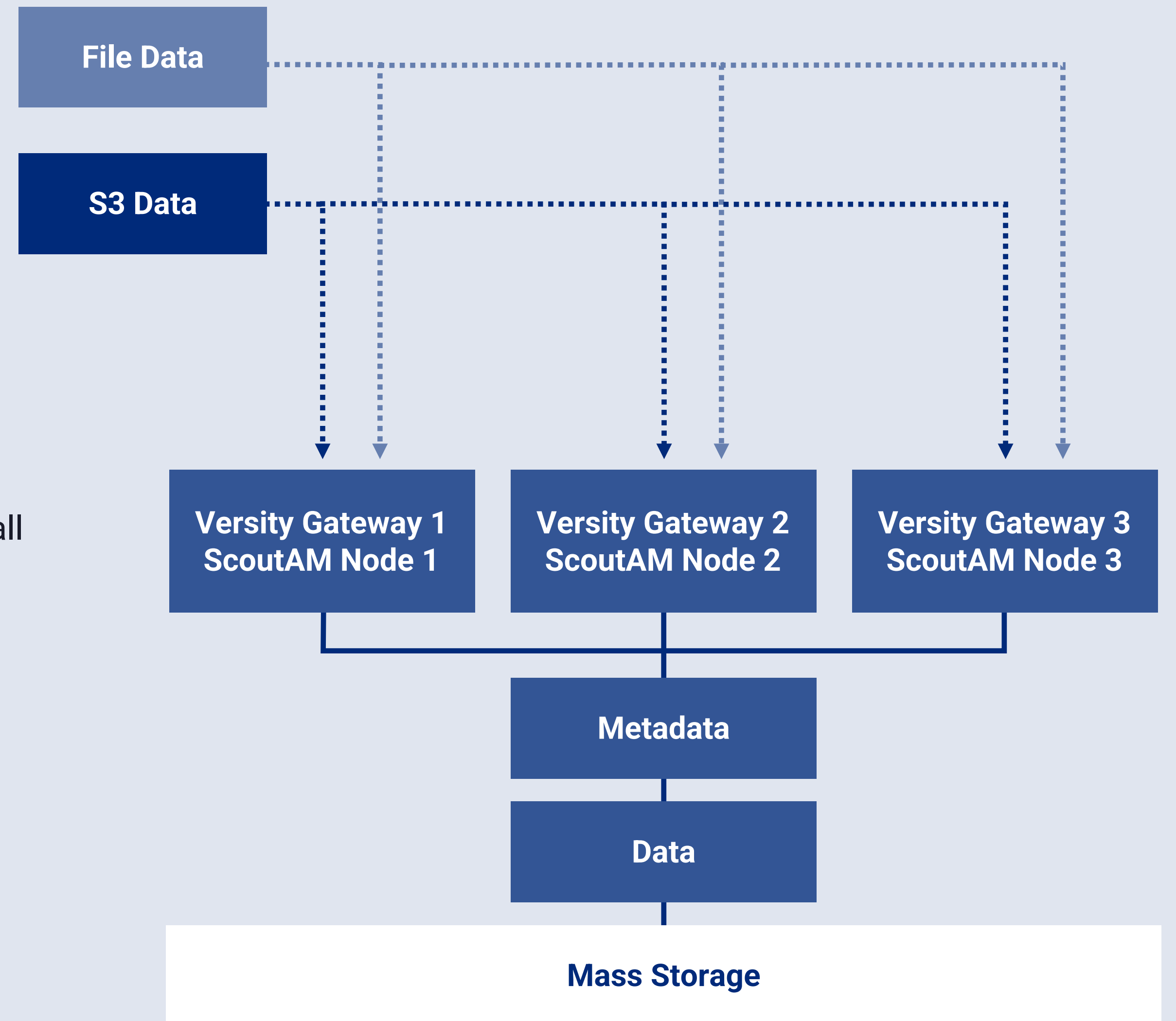- Dell R750's can handle a lot of throughput per server inexpensively

Versity

DELLTechnologies

Versity Gateway - Supported Optimizations

Optimized Writes

- Multi-part upload segments are written to the underlying storage
- Segments are then combined into a single file with just a single system call
- **Eliminates one full read/write cycle**

S3 Glacier Emulation Mode

- Object storage class can be changed to Glacier (offline)
- Data is moved to tape
- Users can see the the object is offline

File Data

S3 Data

Versity Gateway 1
ScoutAM Node 1

Versity Gateway 2
ScoutAM Node 2

Versity Gateway 3
ScoutAM Node 3

Metadata

Data

Mass Storage

Versity

DELLTechnologies

Versity Gateway Workflow
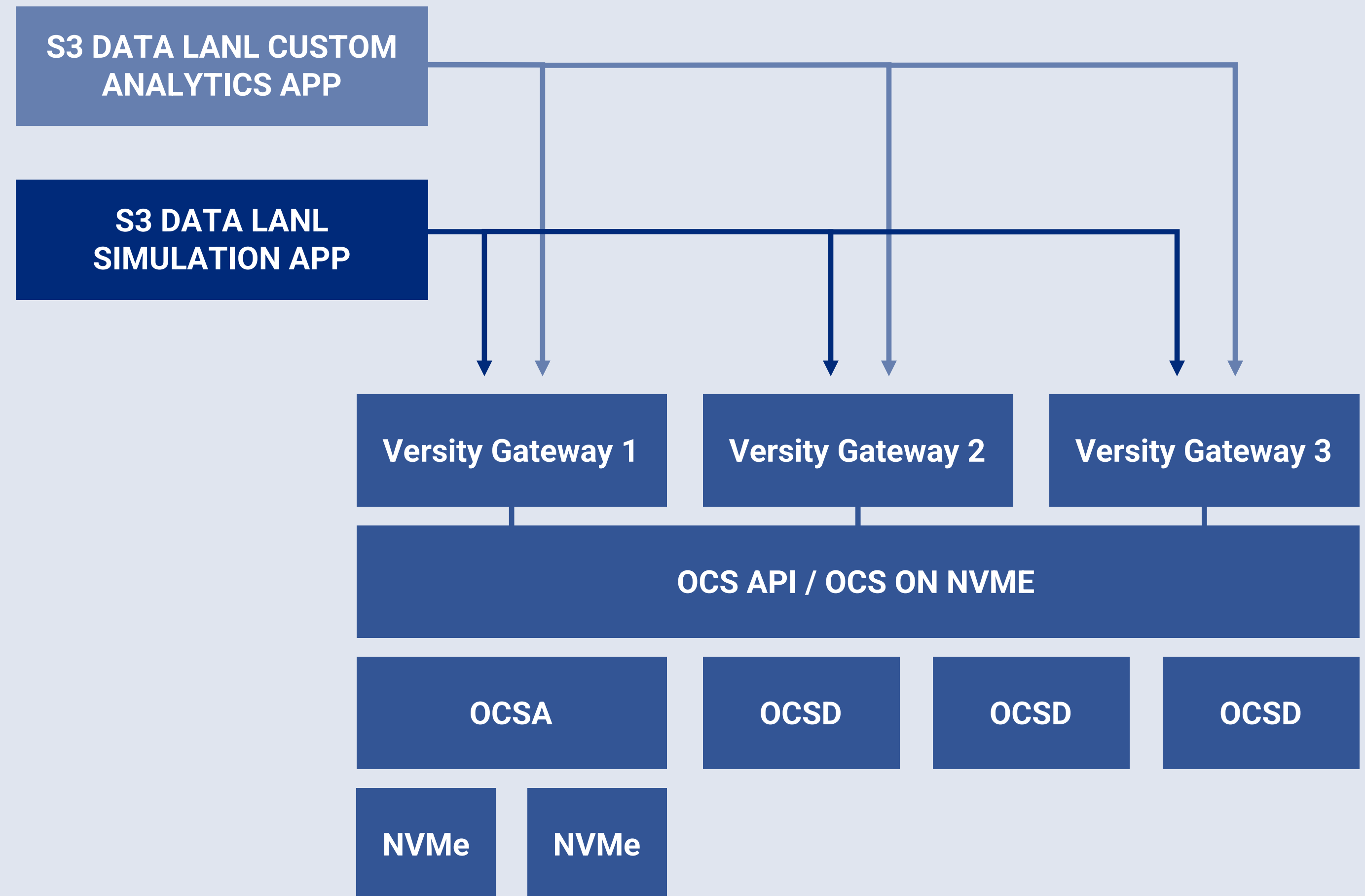
Los Alamos
NATIONAL LABORATORY

Versity S3 Gateway Enables:

- Direct communication with computational storage devices, particularly NVMe devices

- Focus on enabling content queries and select object content APIs to be processed directly on NVMe devices

Anticipated Outcomes:

- Expected advancement in handling and processing large-scale object data

- Potential for a single host to efficiently manage data volumes up to petabytes

- Improved efficiency in data management and retrieval processes, reducing I/O bottlenecks

- Optimize data access times

S3 DATA LANL CUSTOM ANALYTICS APP

S3 DATA LANL SIMULATION APP

Versity Gateway 1	Versity Gateway 2	Versity Gateway 3

OCS API / OCS ON NVME

OCSA	OCSD	OCSD	OCSD

NVMe	NVMe

Versity

DELLTechnologies

Identity and Access Management Integrations

- **Internal IAM:** Stores accounts in a specified directory, accessible across all gateway hosts for consistency

- **LDAP IAM:** Stores accounts in an LDAP directory service, with specific configuration options for access

- Authentication: access key/secret key

- Authorization: user/admin

- User attributes: UID, GID, and Project ID

Versity

DELLTechnologies

Gateway Roadmap

Current Functionality

Multi-Tenant / IAM

Glacier Emulation Mode

Audit Logging

Select Object Content

Containerized Versions

Bucket Notifications

S3 Proxy

Current Focus

Test Suite Using Common Apps (aws cli, s3cmd, etc)

Performance Optimizations

ScoutAM Integration

RPM Packaging

ObjectLock

Azure Back End

Upcoming

StatsD Metrics

Additional IAM Service Integrations

AD, Redis, Etcd

Future

Additional Backend Support

Frontend Versioning Support

Object ACLs

Roadmap subject to change

Versity

DELLTechnologies

Teaser

Major Innovation Coming Here!

ScoutAM

**Maximum scalability
with minimal complexity**

Efficient management and
preservation of exponentially
increasing collections
of unstructured data

Orchestrates extreme
scale data storage workloads

Easy to use, simple to
configure and quick to deploy

ScoutFS

**High-performance
metadata management**

Open-source Linux filesystem
specifically created to
support mass storage
and archiving applications

Enables radically faster
search results in extremely
large file and object populations

Versity Gateway

**High-performance
object-to-file translation tool**

Seamless file storage capabilities
for S3-based applications

Apache 2.0 licensed
open-source project

Scaleable with modular
backend support

Open Standards + Open-Source

Versity

DELLTechnologies

Thank You

sales@versity.com   versity@dell.com

@versitysoftware

Versity

DELLTechnologies