

Unstructured Data Solutions

Marco Debnar
UDS Channel Lead (DACH)

Agenda

- PowerScale Architecture, Features & Functions
- Ransomware Protection
- ECS / ObjectScale
- Use Cases

Challenges handling unstructured data

All while IT budgets remain flat

GROWTH

3x

Expected to triple from
2019 to 2024²

QUANTITY

80%

of the world's data is
unstructured¹

LOCALITY

75%

Outside
data center

¹Source: Gartner Market Guide for File Analysis Software, March 2018

²Source: Gartner ("From 2019-2024, enterprises will triple their unstructured data stored as file or object storage.")

³Source: Gartner, Inc. "What Edge Computing Means for Infrastructure and Operations Leaders" by Rob van der Meulen, October 28, 2018

How do you tackle these challenges?

GROWTH



Data growth
without complexity

QUANTITY



Massive scale with a
single platform

LOCALITY



Diverse workloads
from edge to cloud

Uncovering the Data Landscape



Enterprise databases



ERP Systems



VMware environments

Structured Data (20%)

IT Workloads



Backup & Archive



Home directories



Video surveillance



File shares

Industry Workloads



Energy



EDA



Life sciences



Media & entertainment



Financial



Manufacturing

Emerging Workloads



Artificial intelligence



Data analytics



Assisted driving



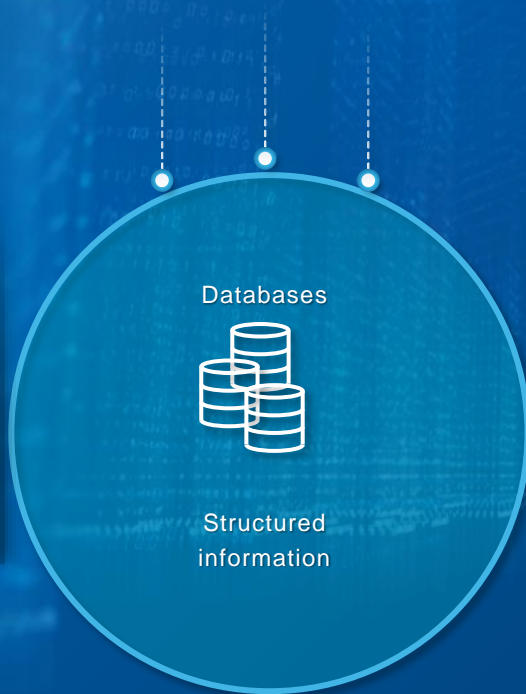
Internet of things

Unstructured Data (80%)

Whats is: Structured vs Unstructured Data

Block
Specific location on disks/memory

Tracks
Sectors



File



Specific folder in fixed logical order

File Path/Name
Date and ACL

Hierarchical:
G:\folder1\Dog.jpg

Object



Flexible container size

Unique ID
Data and Metadata

Buddy
GoldenRetriever
Likes to dig holes

DellTechnologies.com/PowerScale

Powering up your data lake with PowerScale



Performance
at scale



Capacity
at scale



Simplicity
at scale

Dell Technologies

Unstructured Data Solutions

Always in the forefront

#1

WW IDC Market share¹

60%

Fortune 500 customers

15,000+

Customers as of 2021

252

Nodes deployed in a single cluster

30EB+

Raw capacity shipped

1EB+

(raw) deployed by a single customer

¹IDC WW Quarterly Enterprise Storage Systems Tracker, 2021 Q2, September 9, 2021 – Vendor Revenue

²Gartner, Inc. "Magic Quadrant for Distributed File Systems and Object Storage" by Julia Palmer, Jerry Rozeman, Chandra Mukhyala, Jeff Vogel, October 1, 2021.

Gartner does not endorse any vendor, product or service depicted in its research publications, and does not advise technology users to select only those vendors with the highest ratings or other designation. Gartner research publications consist of the opinions of Gartner's research organization and should not be construed as statements of fact. Gartner disclaims all warranties, express or implied, with respect to this research, including any warranties of merchantability or fitness for a particular purpose.

This graphic was published by Gartner, Inc. as part of a larger research document and should be evaluated in the context of the entire document. The Gartner document is available upon request from [Dell Technologies](https://www.dell.com/technologies).

Gartner and Magic Quadrant are registered trademarks of Gartner, Inc. and/or its affiliates in the U.S. and internationally and is used herein with permission. All rights reserved.

Figure 1: Magic Quadrant for Distributed File Systems and Object Storage

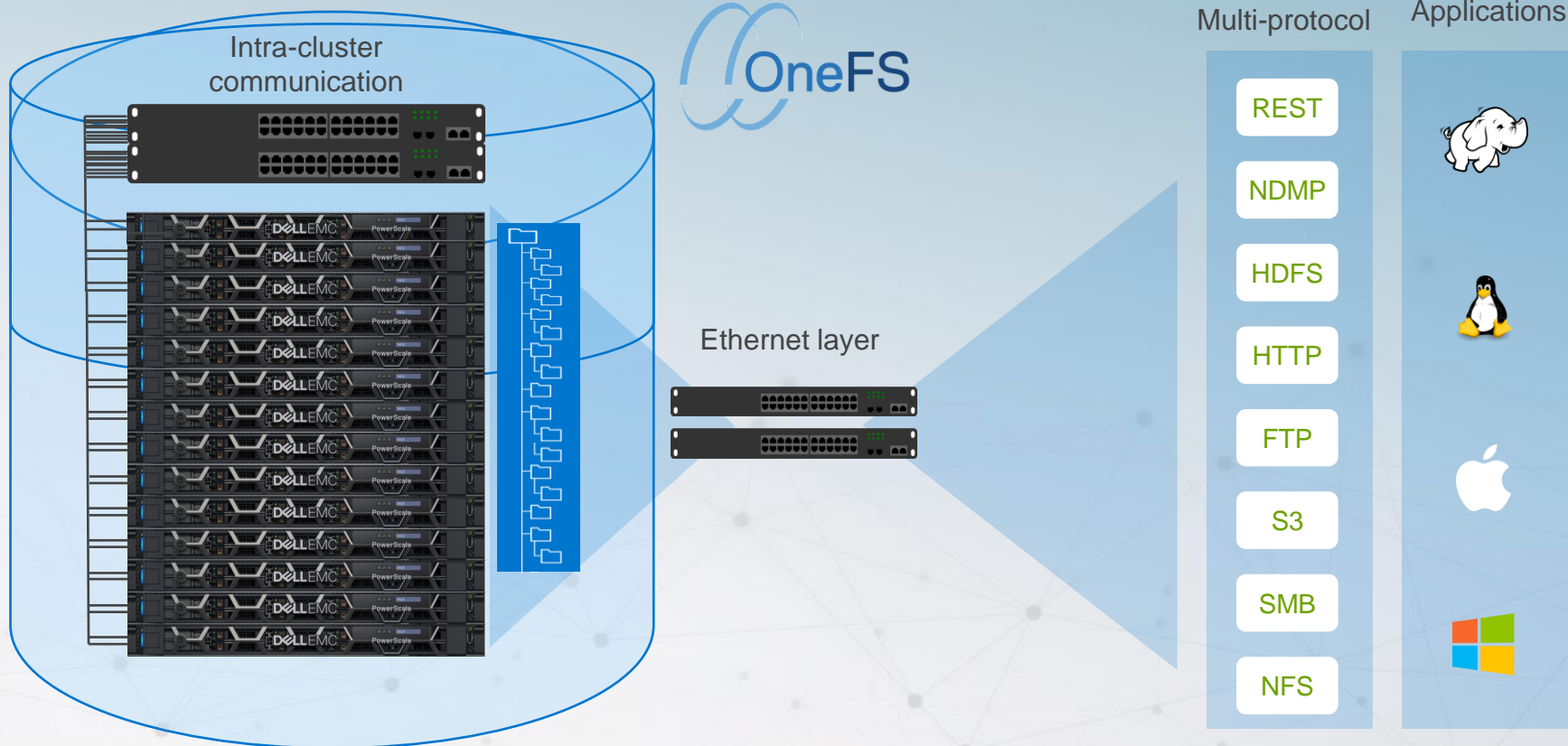


Source: Gartner (October 2021)

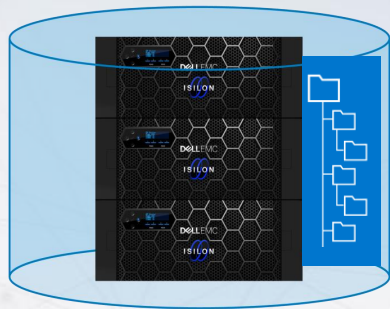
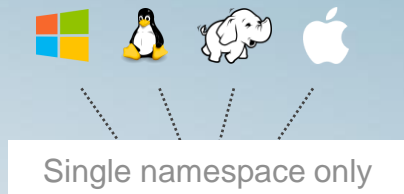
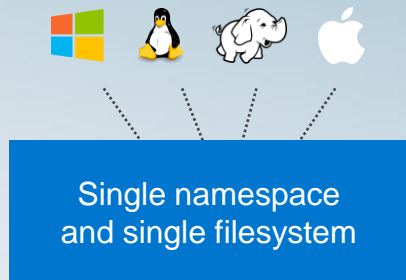
Dell Technologies recognized by Gartner® as a **Leader** in distributed file systems and object storage for **6th consecutive year²**

PowerScale Architecture

PowerScale scale-out architecture



Single namespace across a single filesystem



Traditional

No volumes to manage

All nodes managed as one

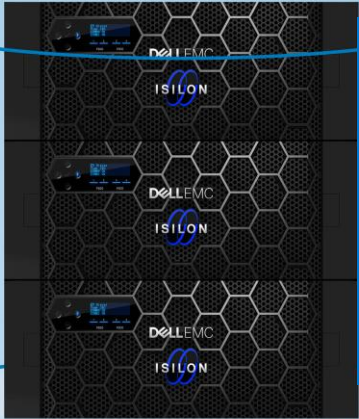
Files automatically spread evenly

Single namespace across all nodes

100PB+ on a single share or export

Single namespace across a single filesystem

Single namespace
and single filesystem



\\acme.com\share

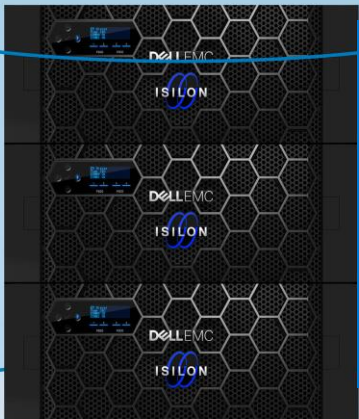


Start Simple

With up to 60+ PB on a single
share or export

Single namespace across a single filesystem

Single namespace
and single filesystem



\\navils.acme.com\finance
(10.1.2.3)



\\navils.acme.com\mktg
(10.1.2.3)



\\bombs.acme.com\share
(10.2.3.4)



\\missiles.acme.com\share
(10.4.5.6)

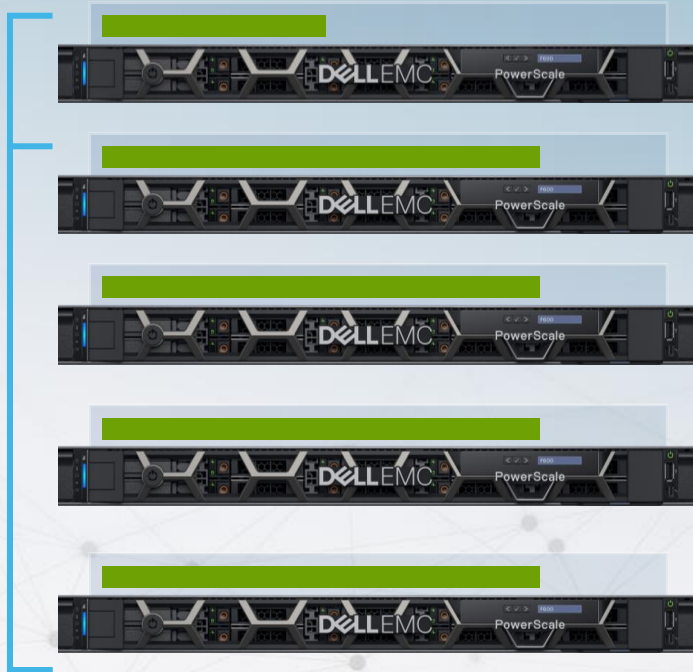


Grow complexity only as required

Support multiple:

- Departments
- Networks
- Authentication providers
- Security policies

Capacity expansion and balancing



**60 sec
Upgrade**

Automated data balancing across nodes eliminates “hot spots”

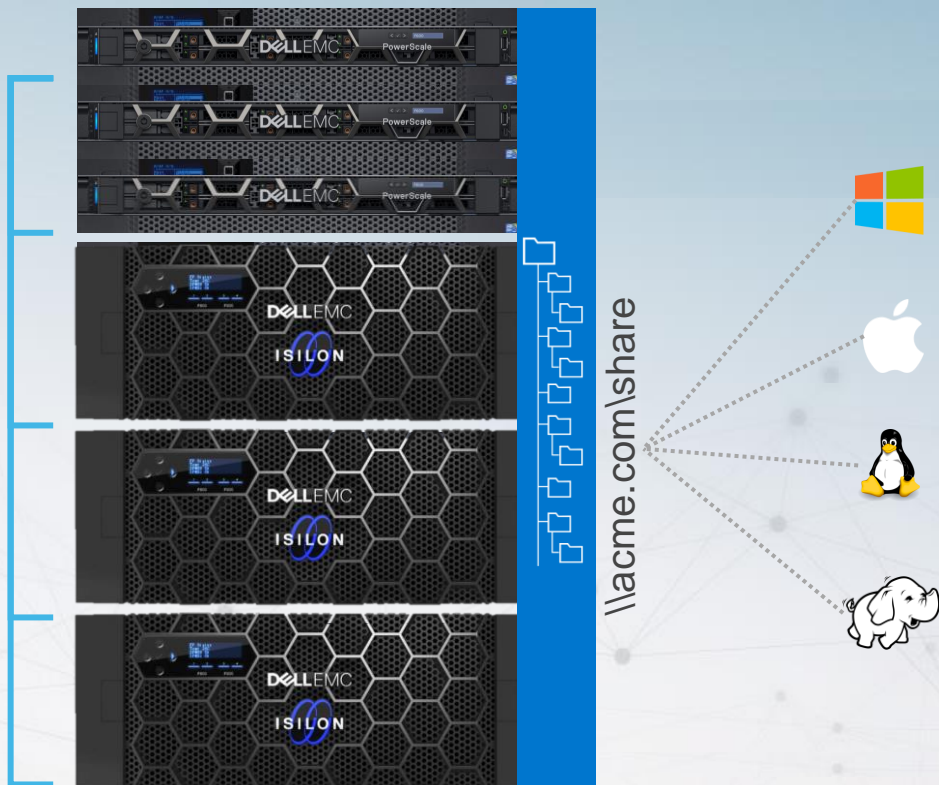
AutoBalance automatically moves content to new storage nodes

No administrator intervention required

Utilization increases as cluster grows

Performance scales linearly with capacity

Configuration persistence across generations



Cluster configuration lives independent of hardware configuration

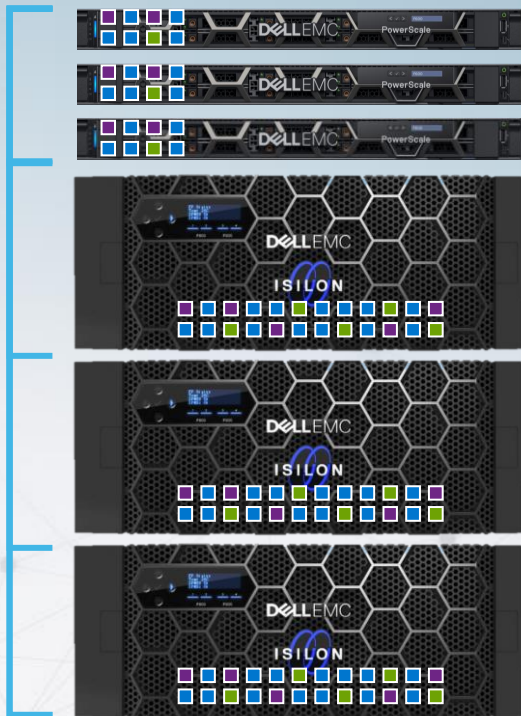
Capacity expands automatically

Shares and exports are unchanged when hardware is added or removed

No application changes required as hardware changes

Data can be spread across old and new nodes

Eliminate data migrations



Add new generation node to existing cluster

Non-disruptive access continues

Policy-driven job moves data to new nodes

Seamless and transparent to clients

No impact to users

Eliminate data migrations

Add new generation node to existing cluster

Non-disruptive access continues

Policy-driven job moves data to new nodes

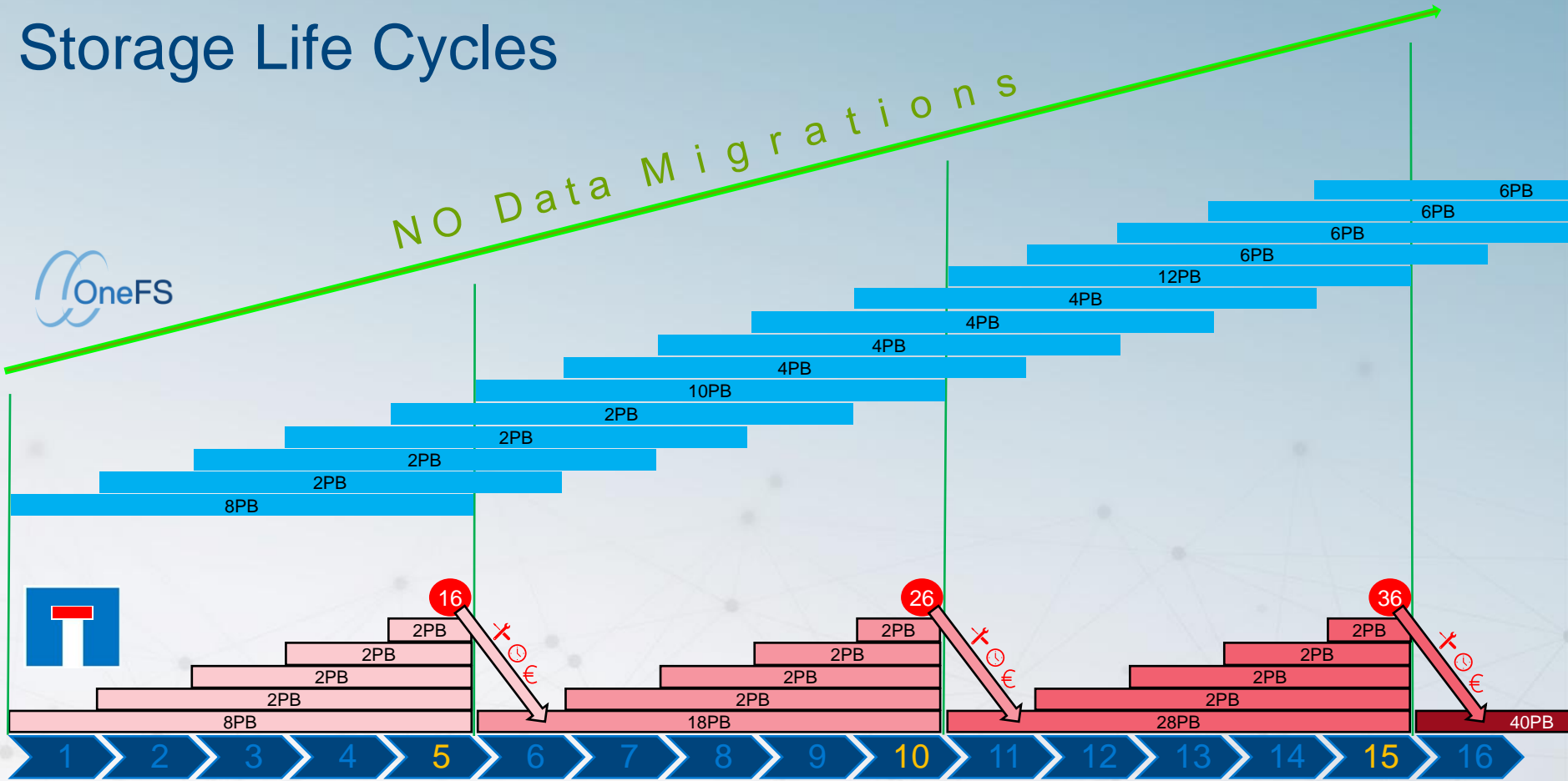
Seamless and transparent to clients

No impact to users

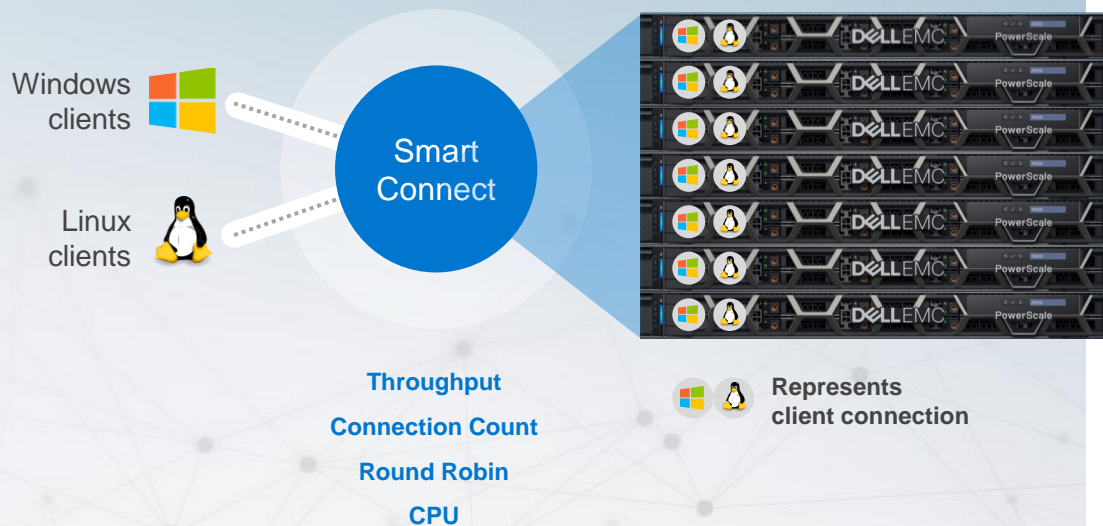


Storage Life Cycles

NO Data Migrations



Connection and CPU balancing



Client connects evenly distributed across cluster

In flight reads and writes handed off

Multiple connection balancing policies

User and application access are uninterrupted

Ensures CPU is evenly utilized across cluster

Policy based data protection

Protection with no downtime

N+1 → **N+4**



Flexible disk and node protection with lower impact

Faster rebuild

On the fly at directory and file

Multiple levels of protection

No impact to users

Policy based data protection

Protection with no downtime

N+1 → **N+4**



Flexible disk and node protection with lower impact

Faster rebuild

On the fly at directory and file

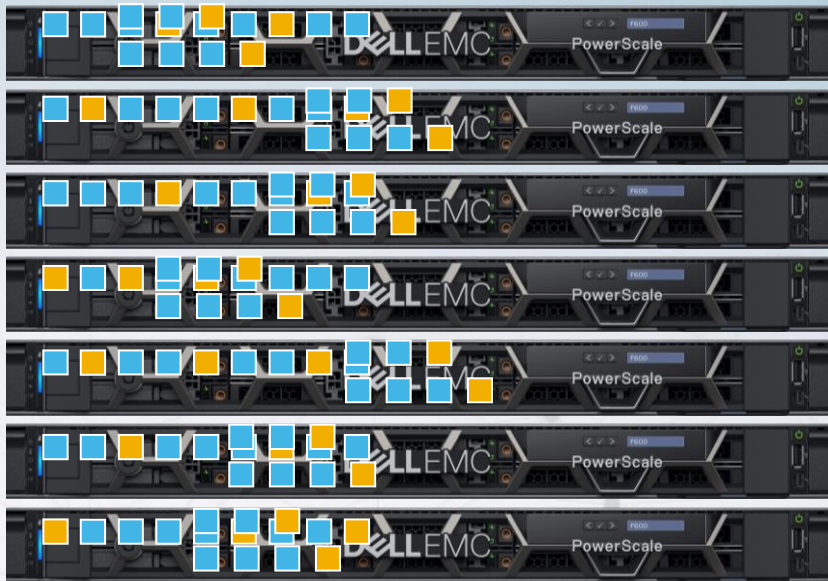
Multiple levels of protection

No impact to users

Policy based data protection

Protection with no downtime

N+1 → **N+4**



Flexible disk and node protection with lower impact

Faster rebuild

On the fly at directory and file

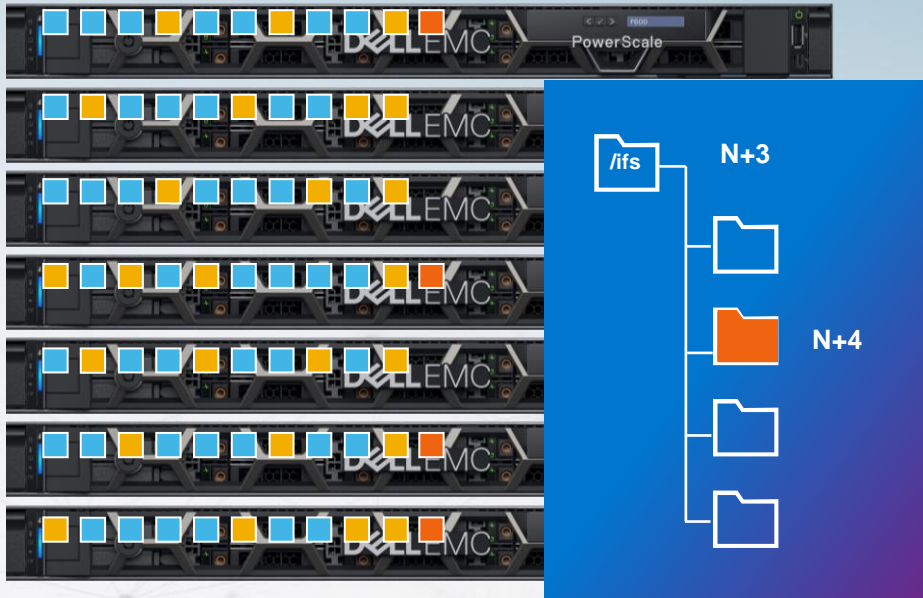
Multiple levels of protection

No impact to users

Policy based data protection

Protection with no downtime

N+3



ADVANTAGES

Change data protection on the fly

Directory and file level granularity
change just a single directory to a higher level

Policy based data protection

Protection with no downtime

N+3

Protected @ N+3



No Need for Hot Spare Drives

ADVANTAGES

Change data protection on the fly

Directory and file level granularity
change just a single directory to a higher level

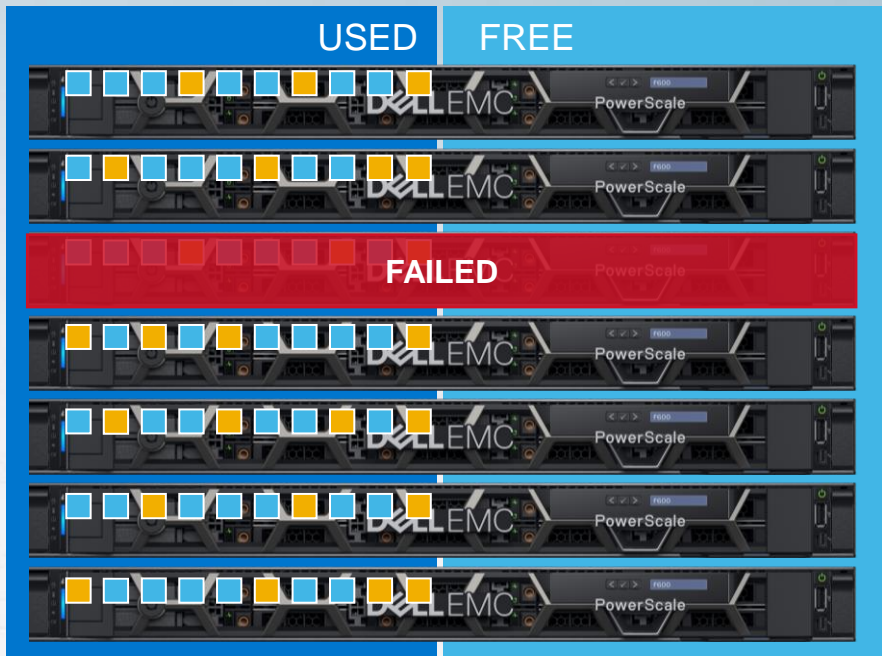
How rebuilds work

Policy based data protection

Protection with no downtime

N+3

Protected @ N+2



Node Fails, Below Protection

ADVANTAGES

Change data protection on the fly

Directory and file level granularity
change just a single directory to a higher level

How rebuilds work

Policy based data protection

Protection with no downtime

N+3

Protected @ N+3



All of the nodes rebuild the failed units into free space

ADVANTAGES

Change data protection on the fly

Directory and file level granularity
change just a single directory to a higher level

How rebuilds work

Policy based data protection

Protection with no downtime

N+3

Protected @ N+3



Failed node is replaced

ADVANTAGES

Change data protection on the fly

Directory and file level granularity
change just a single directory to a higher level

How rebuilds work

Policy based data protection

Protection with no downtime

N+3

Protected @ N+3



Data is redistributed and balanced automatically

ADVANTAGES

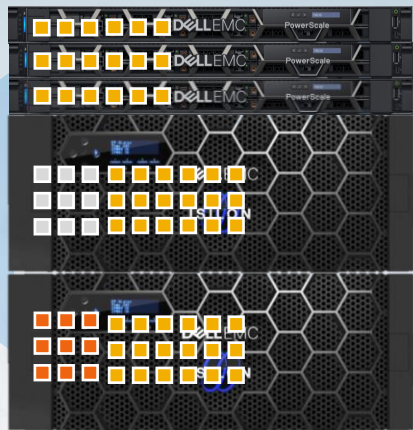
Change data protection on the fly

Directory and file level granularity
change just a single directory to a higher level

How rebuilds work

Multiple pools, one filesystem

Optimize Data Placement to Lower Blended TCO



POLICY

1 week
move to hybrid tier

1 month
move to archive tier

1 year
move to cloud



Drive lowest blended TCO
without compromising
performance

Transparently leverage
multiple pools of storage

SMARTPOOLS

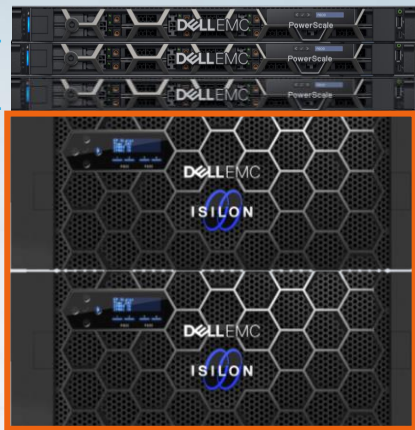
Policy based automated
tiering within one filesystem

CLOUDPOOLS

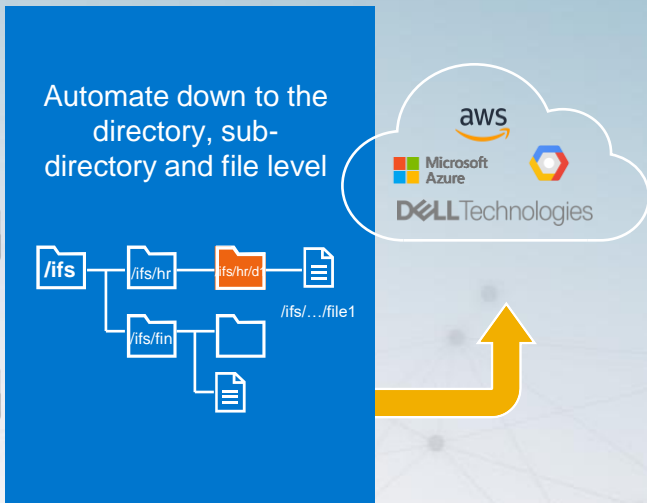
Extends filesystem to the
cloud or on-prem object store

Multiple pools, one filesystem

Optimize Data Placement to Lower Blended TCO



Archive Pool



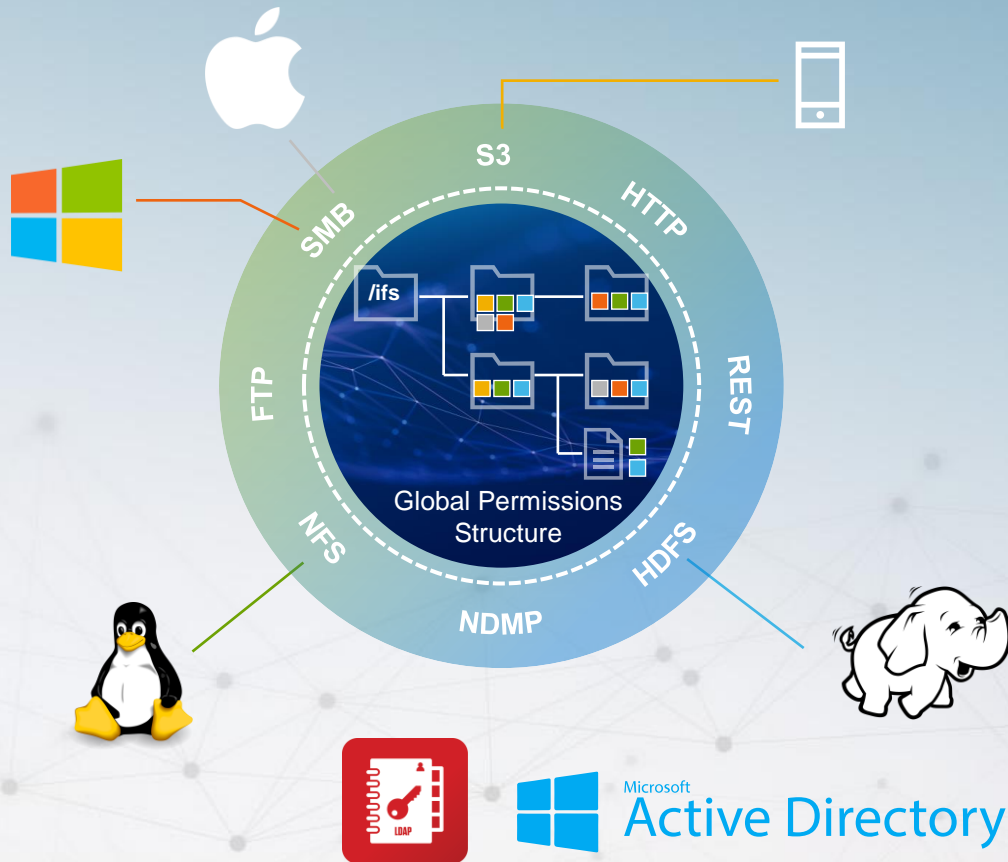
SMARTPOOLS

Policy based automated tiering within one filesystem

CLOUDPOOLS

Extends filesystem to the cloud or on-prem object store

Secure multi-protocol access



Global permissions structure shared across ALL users with ALL protocols

Unified Authentication and Identity Management across all protocols

Full control of protocol stack and modular architecture to add new protocols easily

All protocols get all data services

Single or multiple authentication providers

Object access with S3



Unified access

File and object applications share same dataset

Write data over NFS and perform analytics over S3, write backups over S3 and read to tape, and more



Rich data services

Inline efficiency, snapshots, and more

Objects written over S3 take advantage of most of the data services that PowerScale OneFS offers



All NVMe performance

Ideal for high performance object access

Analytics and other applications requiring high performance access can take advantage of NVMe flash

SHARE DATA BETWEEN TRADITIONAL AND CLOUD NATIVE APPLICATIONS

S3 FEATURES

- Create, Delete and List Buckets
- Create, Delete and List Objects
- Put and Get Objects
- Put and Get Bucket ACLs
- Get Bucket and Object Metadata (Head)
- Multipart Uploads

READ PERFORMANCE²

- F600: Up to 7K TPS per node
- F810: Up to 13K TPS per node

CREATE PERFORMANCE²

- F600: Up to 1200 TPS per node
- F810: Up to 1800 TPS per node

¹ Accessing data that is tiered to secondary cloud / object with CloudPools is not possible

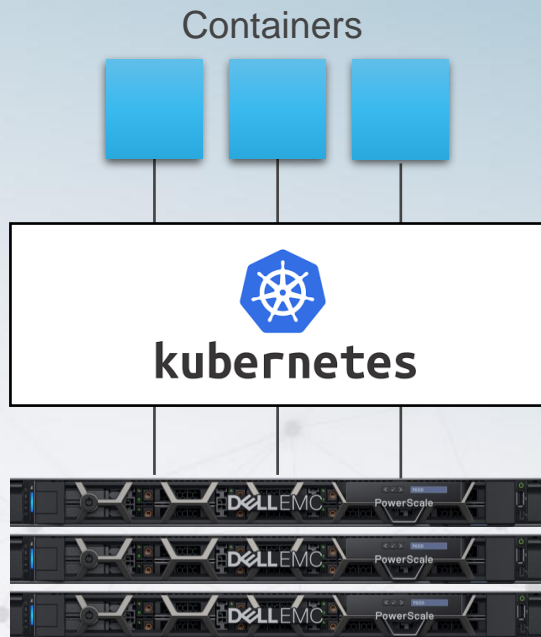
² Performance numbers are for 10KB object size

Multi Tenancy, Security & Compliance



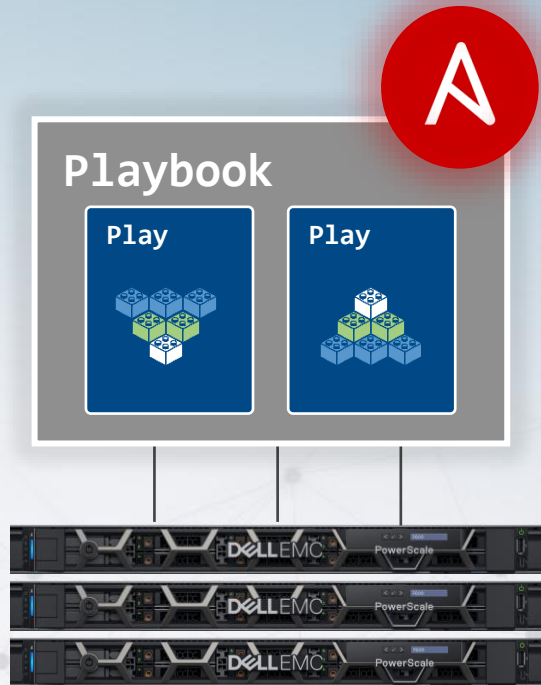
- Rollbased Access Control (RBAC)
- Secure Access Zones
- WORM (KPMG certified)
- File System Auditing
- Data-at-Rest-Encryption (DARE) with Self-Encrypting-Drives (SED)

Persistent Storage for Containers



- Kubernetes Container Storage Interface (CSI) driver for Isilon NFS that conforms to CSI spec
- Modern applications running in containers can take advantage of Isilon's enterprise scale-out storage capabilities
- Provides persistent storage with dynamic provisioning / decommissioning in stateless containers
- Stateful security with AD authentication / authorization

Automation with Ansible playbooks



- Automate IT operations and modern DevOps workflows with Ansible – one of the most popular open source tools
- Supports creation / automation of shares, zones, snapshots, users, smart quotas and more
- No specialized programming skills needed – Ansible modules do the heavy lifting of coding Ansible interactions with Isilon
- Use Ansible modules as building blocks to execute end to end workflows reliably and at scale

The multi-dimensional filesystem

SMALL CLUSTER

to meet specific application needs

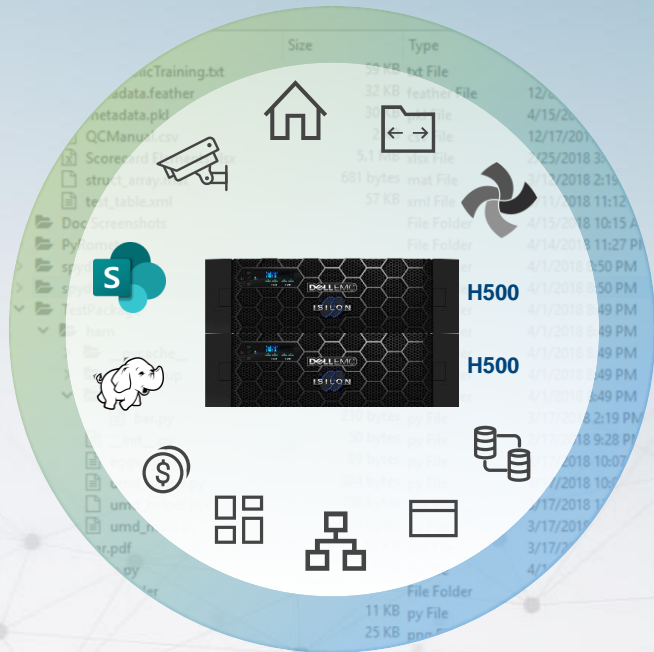


OneFS File System

As use cases expand,
so do capacity requirements

SUPPORT
SCALE
SIMPLICITY

The multi-dimensional filesystem

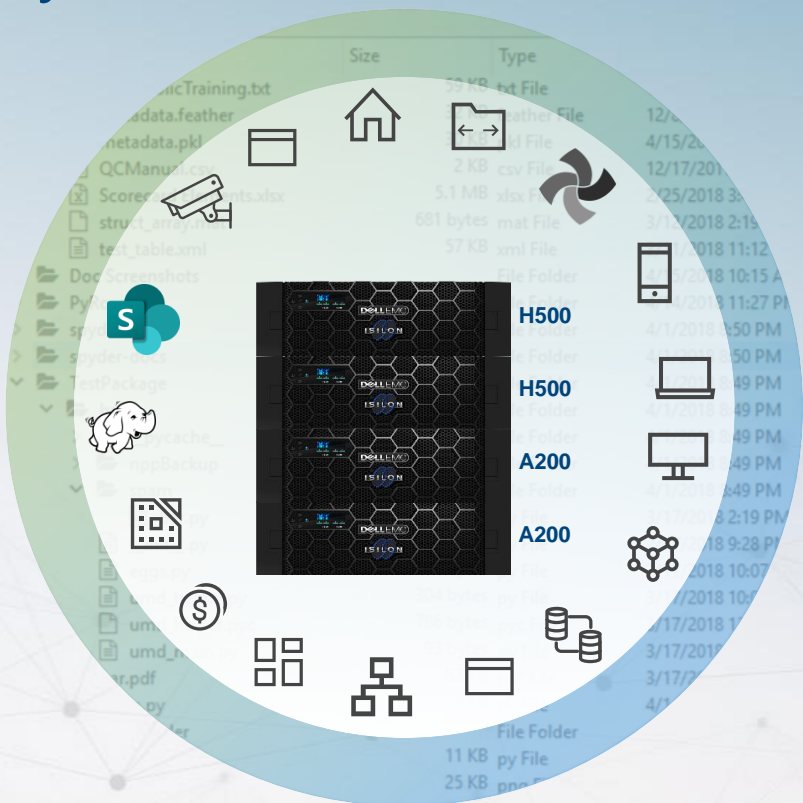


OneFS File System

As use cases expand,
so do capacity requirements

SUPPORT
SCALE
SIMPLICITY

The multi-dimensional filesystem



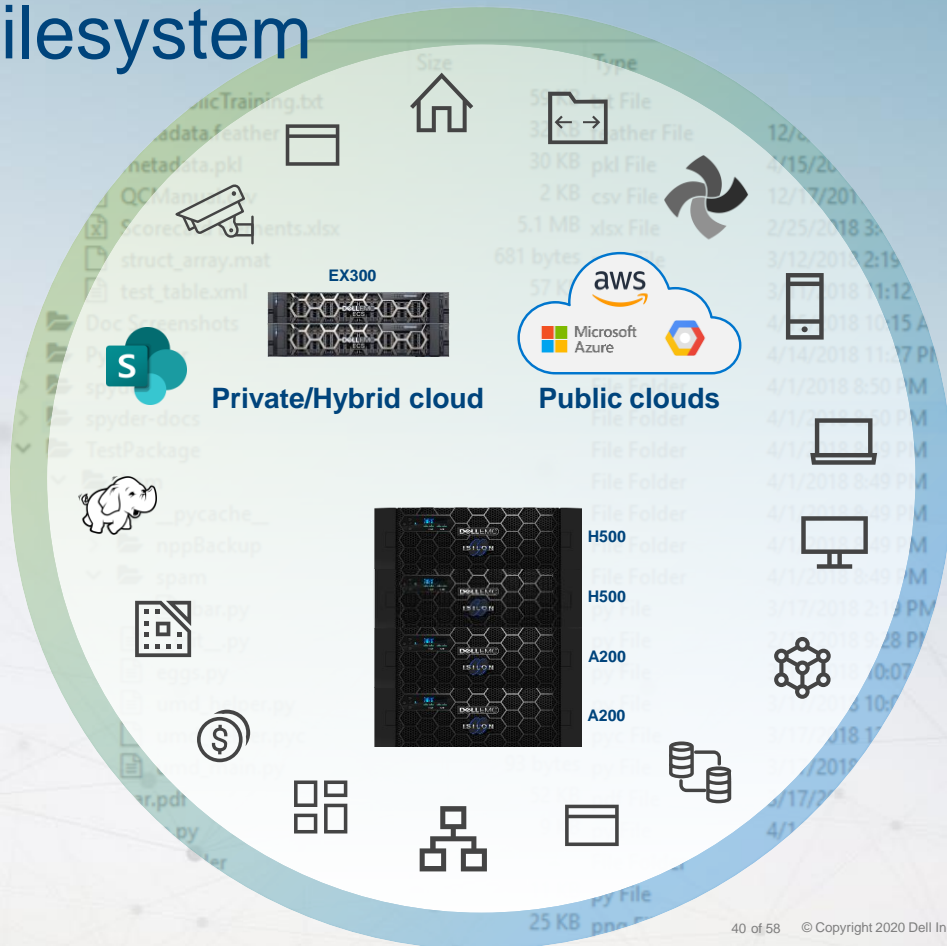
File system is automatically expanded

No file migration

Simple policy places files and based on file age, type, directory or other metadata property

SUPPORT
SCALE
SIMPLICITY

The multi-dimensional filesystem



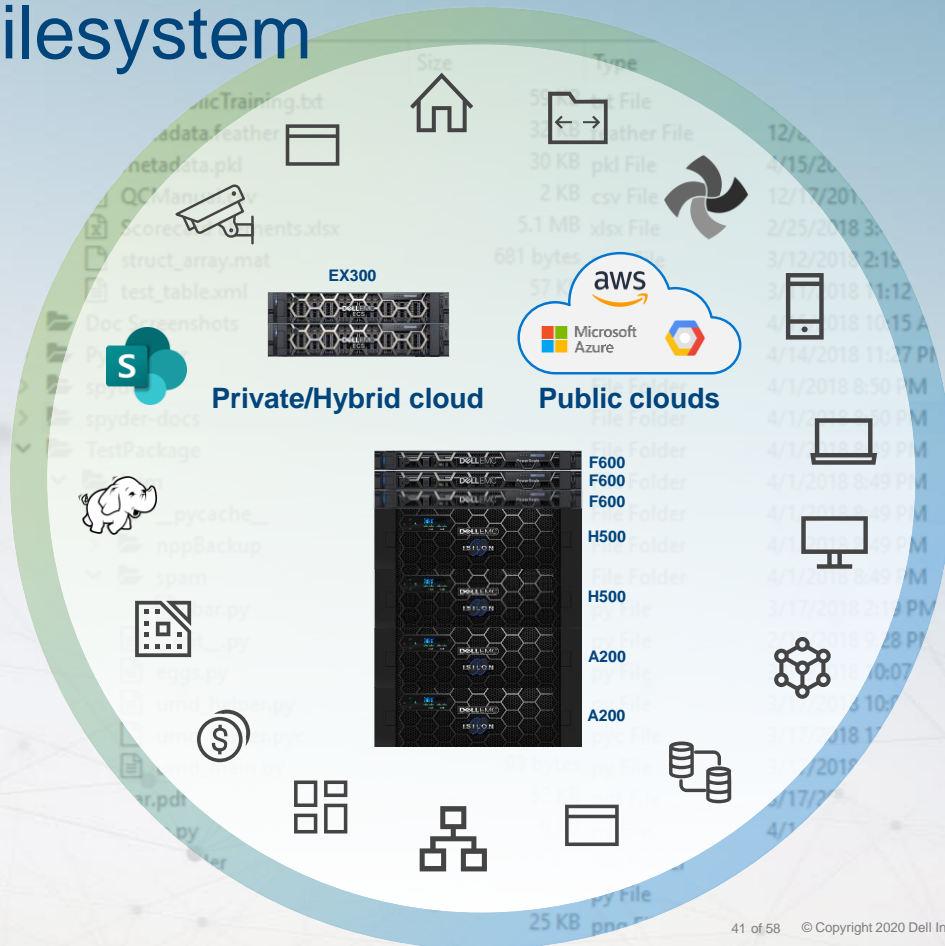
CloudPools

Take advantage of even better capacity economics with Public, Private or Hybrid Cloud **(or all three!)**

Retain self healing, resilient pointers to the object location

Policies can be rewritten as required to relocate data from archive tiers (in-cluster or in-cloud) seamlessly

The multi-dimensional filesystem



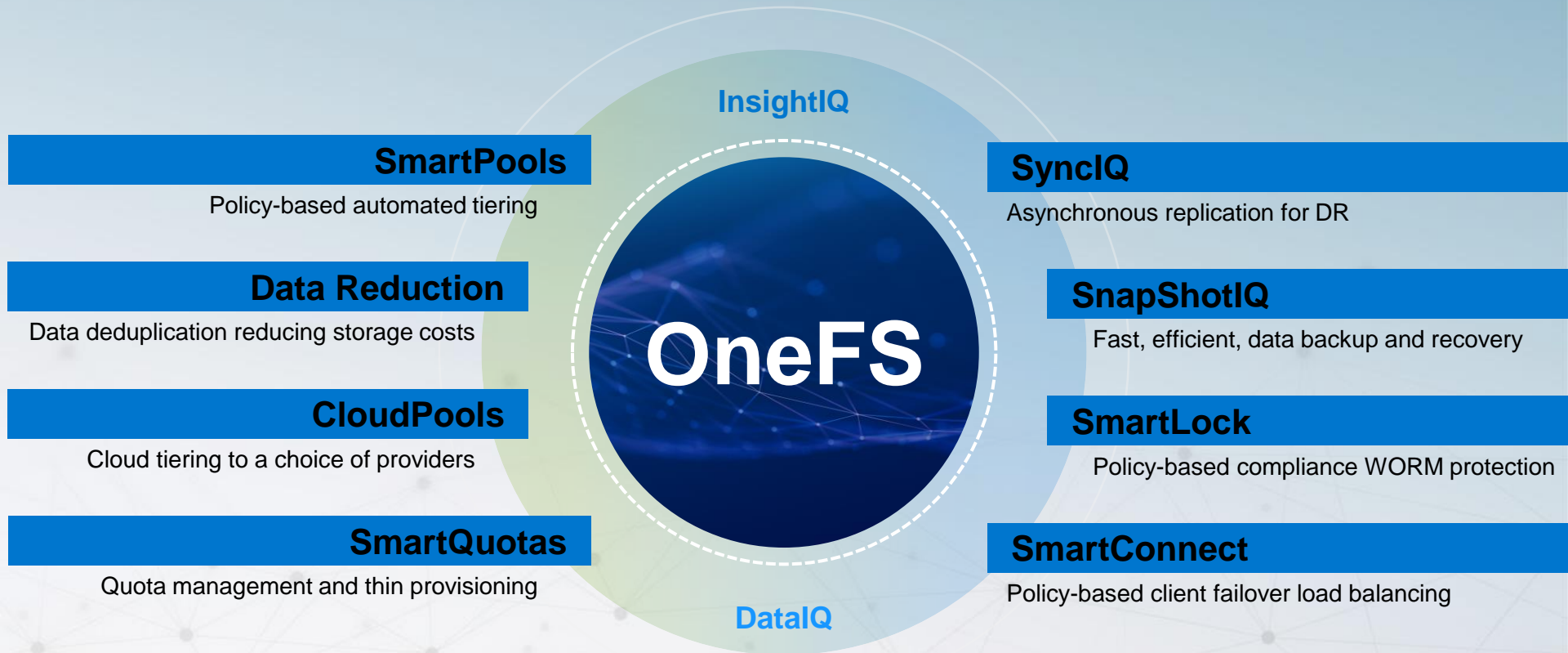
Storage platform refresh

Add new nodes to the existing cluster with ease

Leverage simple to configure file policies

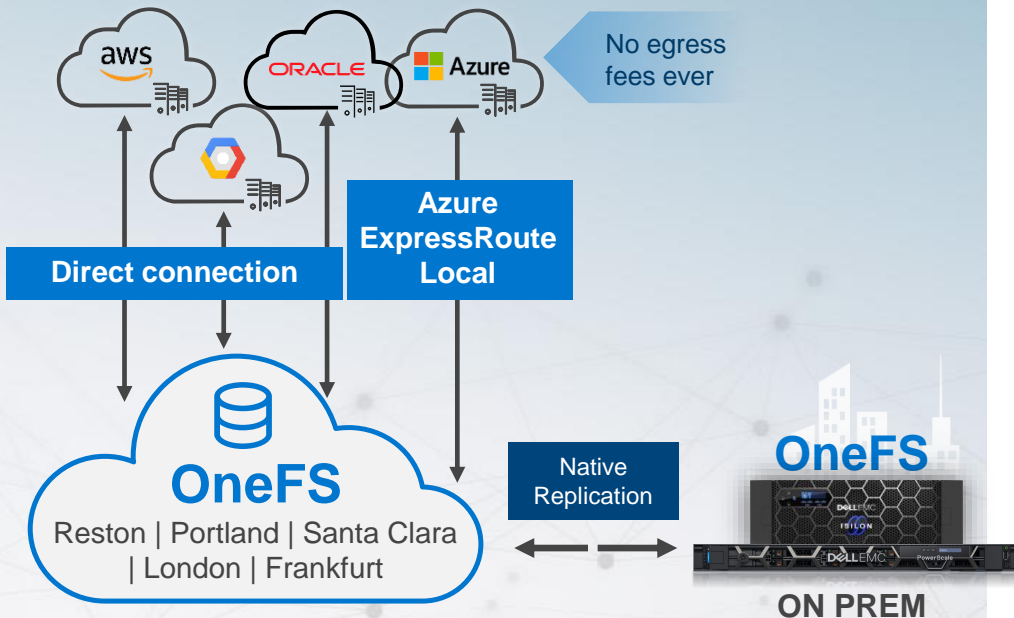
Data movement is invisible to the end users and touchless for administrators.

OneFS features: manage what matters



PowerScale for multi-cloud

Cloud storage services for multi-cloud



Native SyncIQ replication

Multiple clouds simultaneously

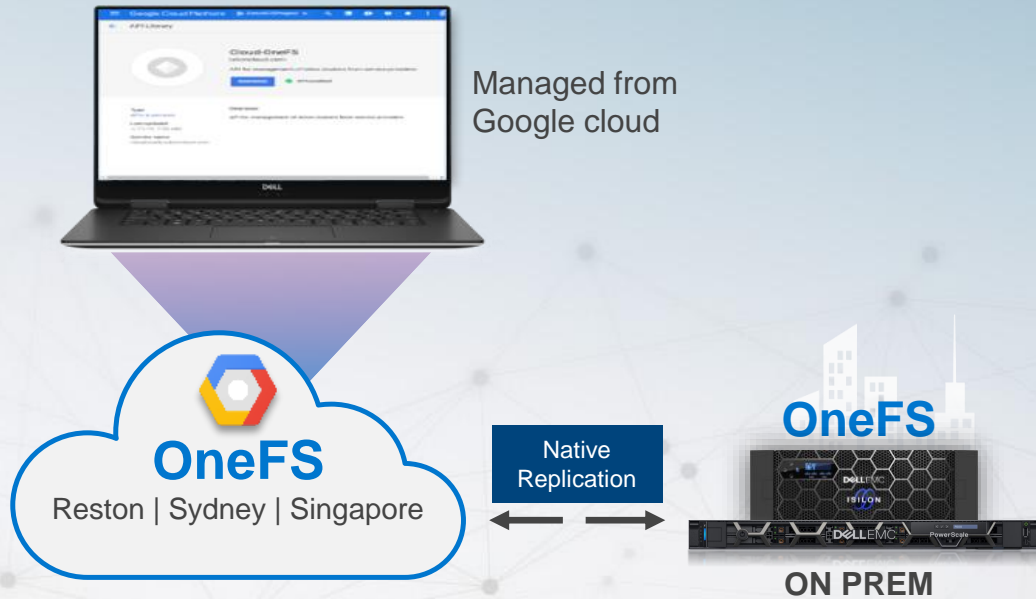
One copy of data

White-glove managed service

Advanced Azure integration

Native cloud integration

PowerScale OneFS for Google Cloud



Native SyncIQ replication

Multiple file tier options

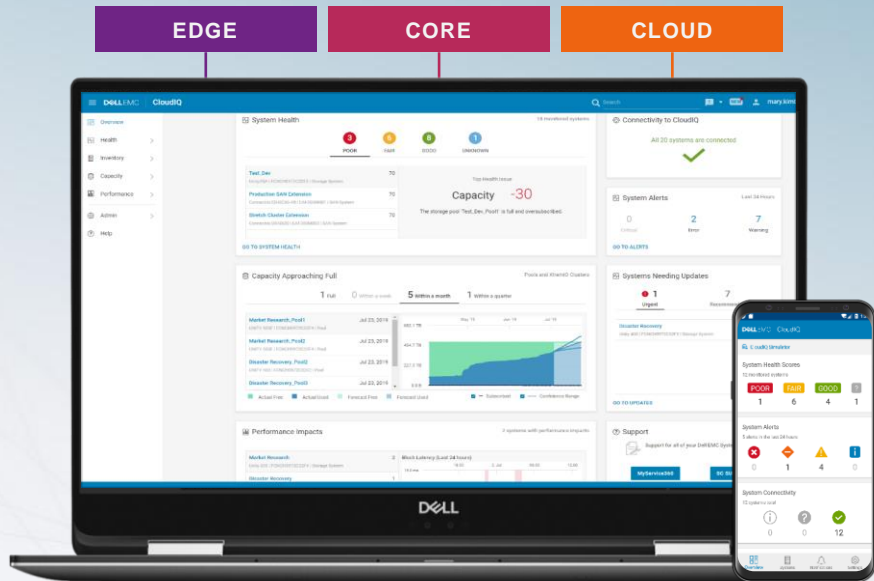
Fully integrated GCP solution

Natively billed through GCP

Integrated support model

Intelligent infrastructure insights

With CloudIQ



FASTER TIME TO INSIGHTS¹

¹Based on an April 2020 Principled Technologies Report commissioned by Dell EMC, "Dell EMC CloudIQ streamlined the user experience in five cloud-based storage preventive management tasks", compared t HPE InfoSight with an HPE Primera array vs. CloudIQ with a Dell EMC Unity array. Actual results may vary. Full report: <http://facts.pt/m8a5u3v>



Reduce risk

Leverages Machine Learning to identify and avoid issues and expedite troubleshooting



Plan ahead

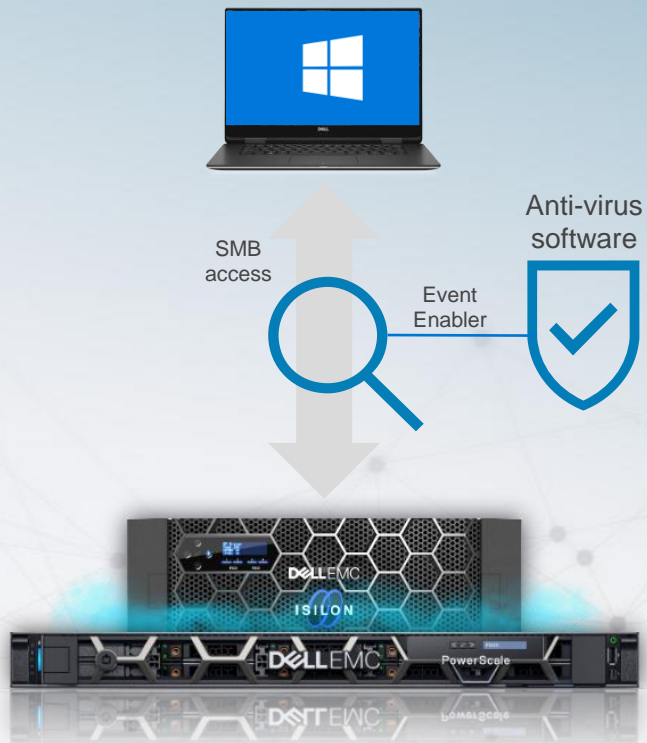
Anticipate business needs and avoid outages



Improve productivity

Improve staffing and infrastructure productivity with a single pane of glass view

Anti-Virus Protection



Key Features

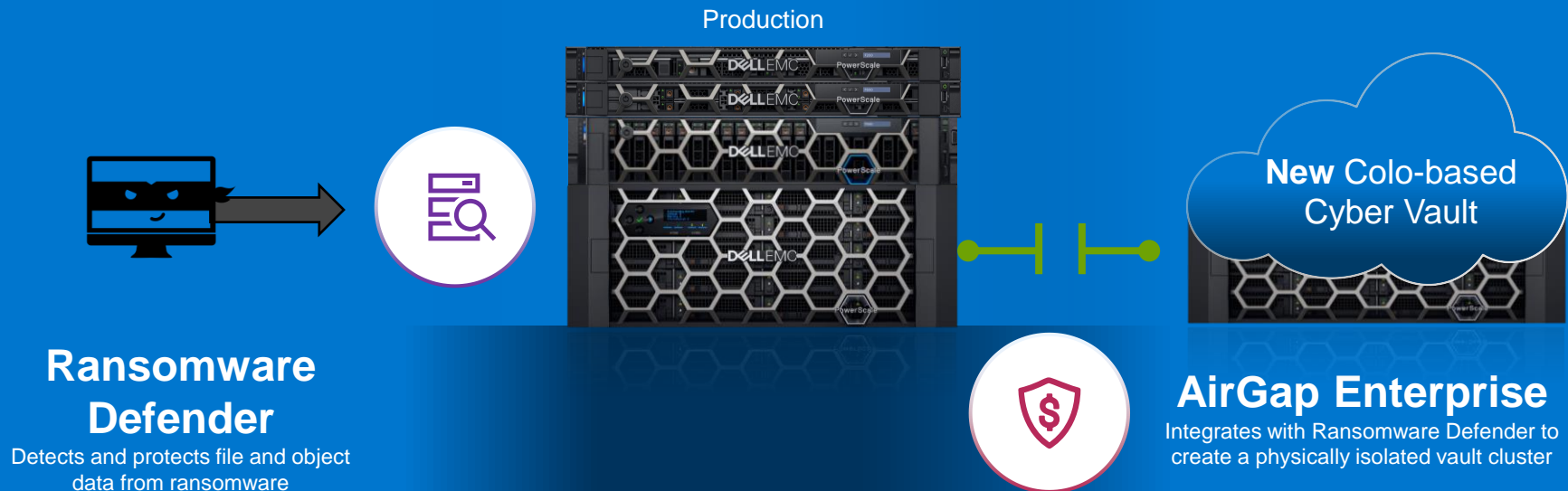
- Leverage external anti-virus software to scan files dynamically via Dell EMC's event enabler agent
- Option to scan based on schedule or access - on access scans can be at file open, file close or both
- Supports SMB protocol initially

Benefits

- Ensures files stored on PowerScale do not have any threats
- Offloads virus scanning to ICAP & CAVA enabled anti-virus software: Symantec, McAfee, TrendMicro, Sophos and Kaspersky
- Efficient, flexible and dynamic policy based scanning

Protect your file data from cyber attacks

Over 1000 customers* have adopted Ransomware Defender to boost the cyber resiliency of their unstructured data



Detect

AI-based system scans for behavior such as mass deletions and encryption activity

Protect

Air gap can be automatically triggered to create a physically separate cyber vault copy

Recover

If needed, recovery happens at storage speed, saving hours, days or weeks

Cyber attacks: a continuous threat to business

A **cyber attack** happens every ...

100

Seconds

Threat vectors are increasing

Both in frequency and sophistication



* *By cartoonist John Klossner – May 2016*

156 million phishing emails sent
WW every day



16 million make it through filters



8 million are opened



800,000 LINKs are click'd



All it takes is **ONE**

State of Cybercrime

The Damage & Costs Continue to Rise

Cybercrime to cost the world \$10.5 trillion annually by 2025

- If it were measured as a country, then cybercrime, which is predicted to inflict damages totaling \$6 trillion USD globally in 2021, would be the world's third-largest economy after the U.S. and China.

Cybercrime Magazine, Nov. 13, 2020



Evolution of Cyber Threat Actors

Different Motivations, Techniques, & Goals

CRIME



Theft & extortion for financial gain

INSIDER



Trusted insiders steal or extort for personal, financial, & ideological reasons. Increasingly targeted because of privileged access to systems

ESPIONAGE



Corporate or Nation-state actors steal valuable data

HACKTIVISM



Advance political or social causes

TERRORISM



Sabotage & destruction to instill fear

WARFARE



Nation-state actors with destructive cyber weapons (Not Petya)

What can happen to data?

Depends on the motive of the attack

Gain access
to critical data



Encrypt and
demand ransom



Permanently
delete



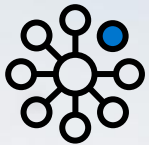
Sell data on
the dark net



Trade secrets
corporate espionage



Characteristics of a Cyber Recovery Solution



Isolation

Physical & logical separation of data



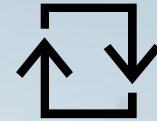
Immutability

Preserve original integrity of data



Intelligence

Machine learning based threat detection, alerting and reporting



Recovery

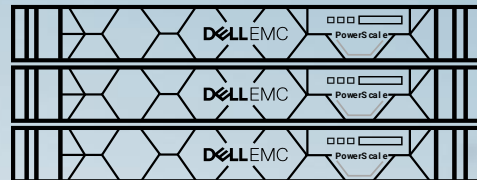
Fast recovery for minimal operational impact

Two solutions based on workloads

PowerProtect Cyber Recovery for Backup Data



Ransomware Defender for Unstructured Data



Offering the same solution components

Already adopted by 1000+ Customers



Isolation

Physical & logical separation of data



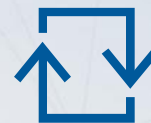
Immutability

Preserve original integrity of data



Intelligence

Machine learning based threat detection, alerting and reporting



Recovery

Fast recovery for minimal operational impact

Ransomware Defender for File Data



Detect & Respond

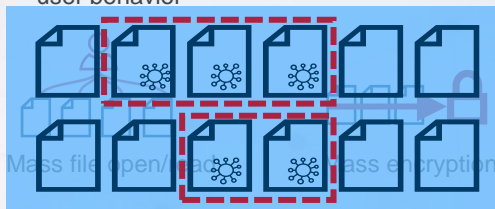
Stop an attack when it happens



Anomalies in user behavior



Mass deletion

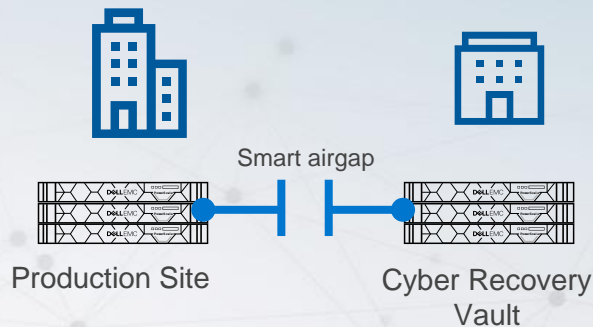


Mass file operations



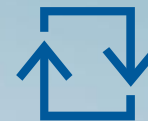
Protect your vault

With a Smart Airgap



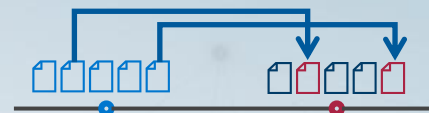
Production Site

Cyber Recovery Vault



Recover

With minimal business impact



Granular Data recovery
From PROD, DR or CR



Operational recovery

Powered by **Superna Ransomware Defender**

DELLTechnologies

DR versus Cyber Recovery



Disaster recovery site

Cyber recovery

 Protection against
 Network connection
 Location
 Admin privilege
 Recovery purpose

Natural disasters, localized outages

Cyber attacks

Connected for continuous replication

Isolated with intermittent replication.

Remote

Local or remote

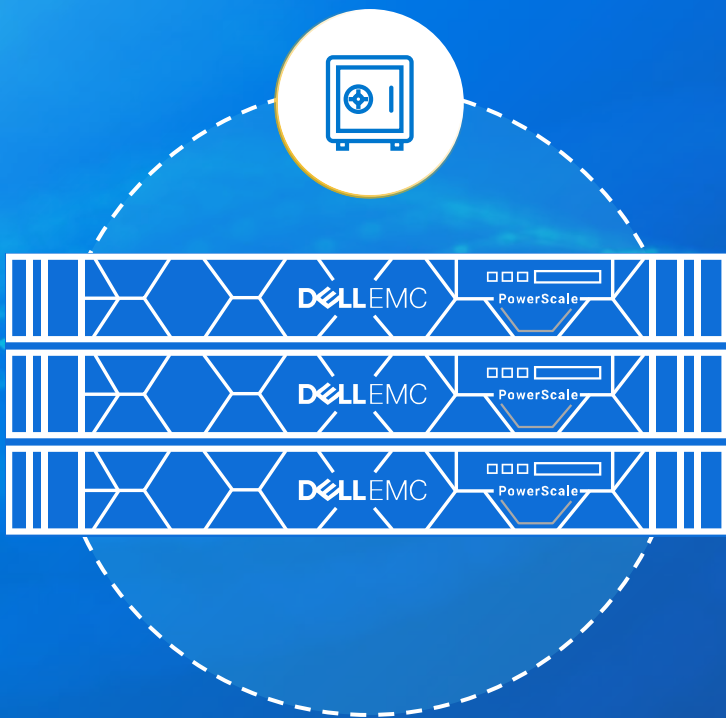
IT admins

Restricted to CISO

To support most of the business operations

To support only the most critical operations

Key features



Recover 1 PB of data in a few hours

No other vault storage, on-prem or cloud, comes even close to PowerScale's data recovery speed



Immutability with worm lock

Data immutability makes sure attackers cannot alter or delete data



AI powered threat detection

Monitoring production data and alerting of suspicious activity puts IT a step ahead of attackers



Scalable to multiple clusters

Single pane of glass for threat detection and data isolation to protect multiple PowerScale clusters

Hardware

New server-based PowerScale all-flash family

New

PowerScale F900



2U

Capacity Density

48 cores | 736GB memory | 46TB to 369TB NVMe per node

The image shows a 2U server rack with a honeycomb-patterned front panel. The Dell EMC logo and PowerScale branding are visible. An orange tag with a blue string is attached to the right side of the server.

PowerScale F600



1U

Performance Density

16 cores | 128GB to 384GB memory | 15TB to 123TB NVMe per node

The image shows a 1U server rack with a honeycomb-patterned front panel. The Dell EMC logo and PowerScale branding are visible. A purple tag with a blue string is attached to the right side of the server.

Q4 CY21

Coming: 48 cores 736GB memory option for performance density!

PowerScale F200



1U

Entry & Edge

10 cores | 48GB or 96GB memory | 4TB to 31TB SAS flash per node

The image shows a 1U server rack with a honeycomb-patterned front panel. The Dell EMC logo and PowerScale branding are visible. A red tag with a blue string is attached to the right side of the server.

PowerEdge server-based with choice of cores, memory and capacities to get right **performance-capacity** mix

Up to **70% performance** improvement with OneFS 9.2

Nov CY21: New high-performance F600 option with bigger CPU and more memory

- Performance of the **F900** ... at the form factor and capacity of the **F600**
- Ideal for **EDA** and other performance dense workloads

CY22: NVMe 15TB and 31TB QLC for **density** and even **lower \$/TB**

New PowerScale hybrid and archive appliances

New

ARCHIVE

PowerScale A300 | A3000



6 cores | 96GB memory
30TB to 320TB per rack unit

HYBRID

PowerScale H700 | H7000



16 cores | 192 or 384GB memory
30TB to 320TB per rack unit

More Value

- Purpose-built appliances for best **capacity density** for archive/hybrid workloads (up to 320TB per RU)
- Improved **performance** and **availability** with more cores, larger DRAM and bigger SSD cache
- Support for **100Gb**, **25Gb** and Infiniband interface, **FIPS encryption** and **node compatibility** with Gen 6*

Less Cost

- **Inline data reduction** across the full PowerScale family -- all-flash, hybrid and archive nodes
- Combine all-flash, hybrid and archive under a single filesystem **with data reduction** for best \$/TB
- More flexibility with **post-process dedupe** --- now included at **\$0** cost on all PowerScale nodes

*Includes H500, H5600, H400, A200 and A2000

Dell Technologies

Isilon Family Re-branded to PowerScale :



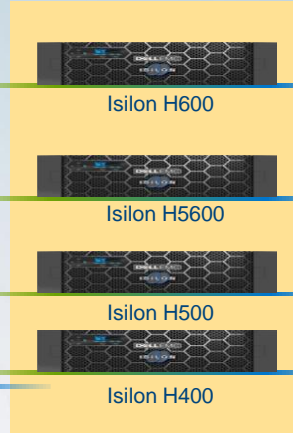
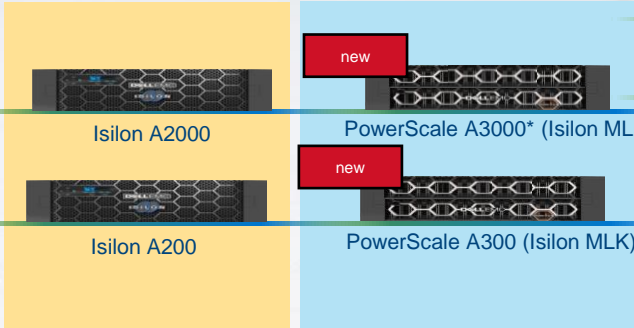
All-Flash

High-level transition plans

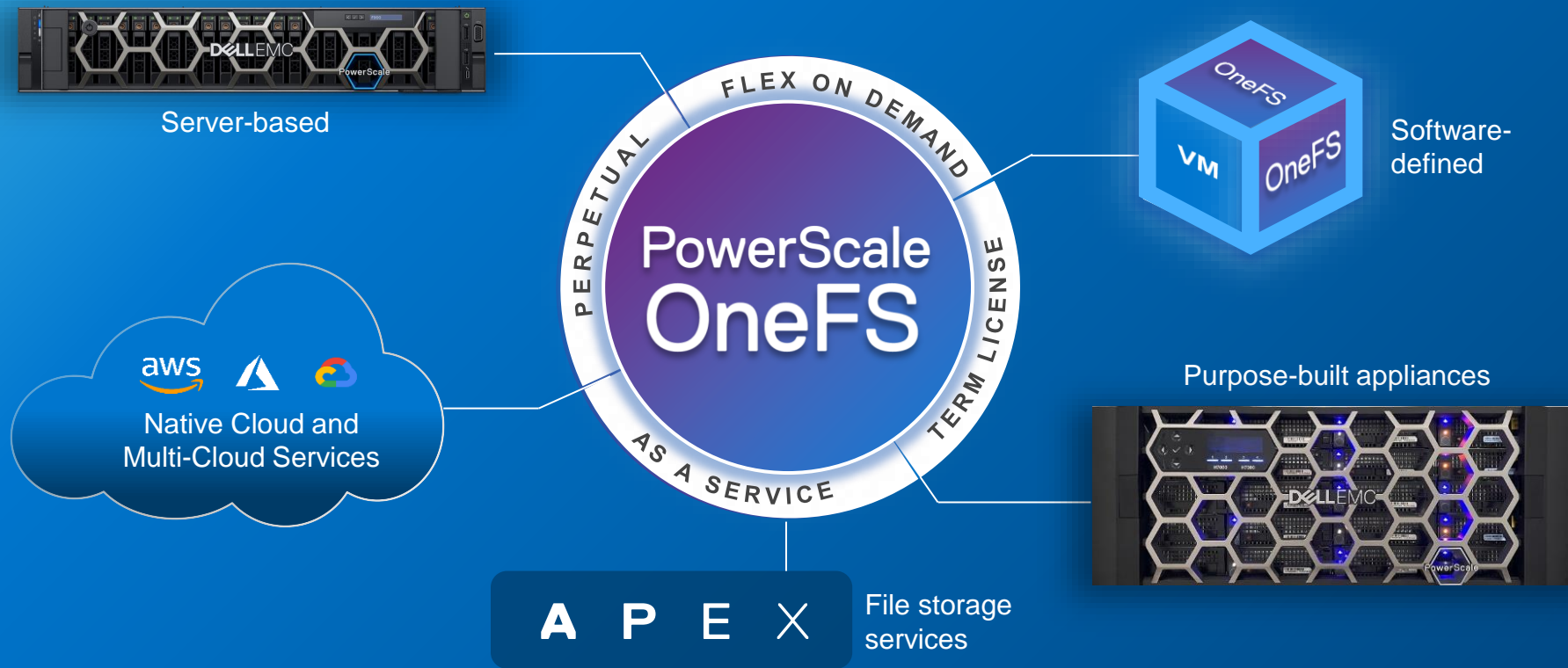
Isilon Gen 6 (Infinity Chassis)	Isilon Gen 6 MLK – rebranding to PowerScale; (Infinity Chassis)	PowerScale (PowerEdge Chassis Based)
A200	A300	
A2000	A3000	
H400	A300 and H700	
H500	H700	
H600		F200, F600 and F900
H5600	H7000	

Hybrid

Archive



PowerScale: *one* filesystem, *many* ways to consume

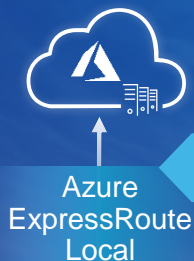


**Simulation and non-production use only for now.
Production availability in 1H 2022.*

PowerScale public cloud services

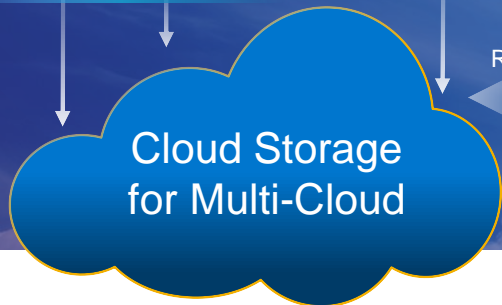
Enterprise features, performance and scale as a service in the cloud

Multi-cloud service



No egress charges

Native Replication



RESTON | PORTLAND | SANTA CLARA | LONDON | FRANKFURT

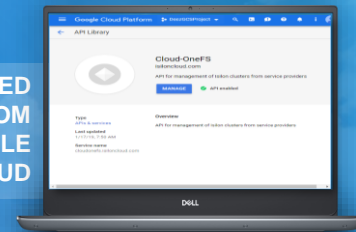


PRIVATE OR COLOCATED DATA CENTER

Native Replication

Native cloud service

MANAGED FROM GOOGLE CLOUD



RESTON | IOWA | SYDNEY | SINGAPORE

UDS Use Case (summary)

Enterprise Workloads



Home Directories & File Shares



Video Surveillance



Modern (Cloud) Applications



Archive & Backup Target

Media & Entertainment

Content Creation



Broadcast



Content Delivery



Healthcare & Life Sciences



Radiology & Cardiology PACS

Digital Pathology, Oncology



OBGYN, Sonograms



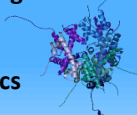
Research Data Repository



Genome Sequencing



Proteomics



Automotive

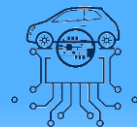
CAD



CAE



ADAS (Advanced driver assistance systems)



Finance



EDA



AI & Analytics

Pivotal



cloudera

splunk



Public

State & Local Government



Higher Education



Police



Certification

Certification Path

Please navigate to the following URL:

<https://education.dell.com/content/emc/en-us/home/certification-overview/find-exam.html?k=Storage>

TRACK	LEVEL	TRAINING	FUNCTION	EXAM DESCRIPTION	EXAM #	PRACTICE TEST
DCA-ISM	Associate	Information Storage and Management	Technology / Associate	Information Storage and Management v4 Exam »	DEA-1TT4	Launch »

TRACK	LEVEL	TRAINING	FUNCTION	EXAM DESCRIPTION	EXAM #	PRACTICE TEST
DCS-TA	Specialist	PowerScale (Isilon)	Design	Specialist - Technology Architect, PowerScale (Isilon) Solutions Exam »	E20-555	Launch »
DCE	Expert	PowerScale (Isilon)	Design	Expert - PowerScale (Isilon) Solutions Exam »	DEE-1421	Launch »

TRACK	LEVEL	TRAINING	FUNCTION	EXAM DESCRIPTION	EXAM #	PRACTICE TEST
DCE	Expert	PowerScale (Isilon)	Deploy	Expert - PowerScale (Isilon) Solutions Exam »	DEE-1421	Launch »
DCS-IE	Specialist	PowerScale (Isilon)	Deploy	Specialist - Implementation Engineer, PowerScale (Isilon Solutions) Exam »	DES-1423	Launch »

TRACK	LEVEL	TRAINING	FUNCTION	EXAM DESCRIPTION	EXAM #	PRACTICE TEST
DCS-IE	Specialist	Elastic Cloud Storage ECS	Deploy	Specialist Implementation Engineer Elastic Cloud Storage (ECS) Exam »	DES-1B21	Launch »

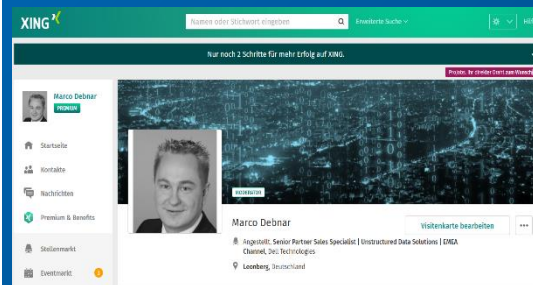
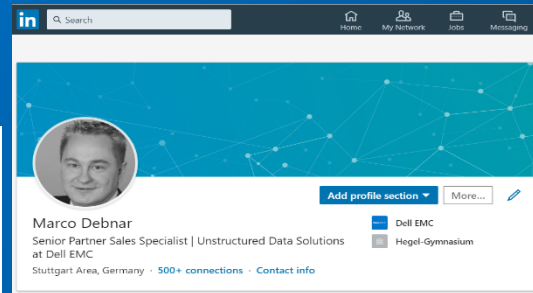
Marco Debnar

UDS Channel Lead (DACH)
Unstructured Data Solutions
Dell Technologies | Channel EMEA

mobile: +49 160 90 93 0447
email: marco.debnar@dell.com

Dell GmbH
Neue Ramtelstraße 4, 71229 Leonberg, Germany

DellTechnologies.com



DELLTechnologies