

NUG – NEC Users Group 2024

\Orchestrating a brighter world

NEC

Mass data and long term archive with Versity

Andreas Göttlicher

June 2024

NEC Deutschland GmbH

HPC Department

NUG – NEC Users Group 2024

- *Intro*
- *Market of large archives*
- *Versity and what they do*
- *POC*
- *NEC & Versity*
- *Recent success at NIWA*

NEC Deutschland GmbH

HPC Department

Munich-Bavaria

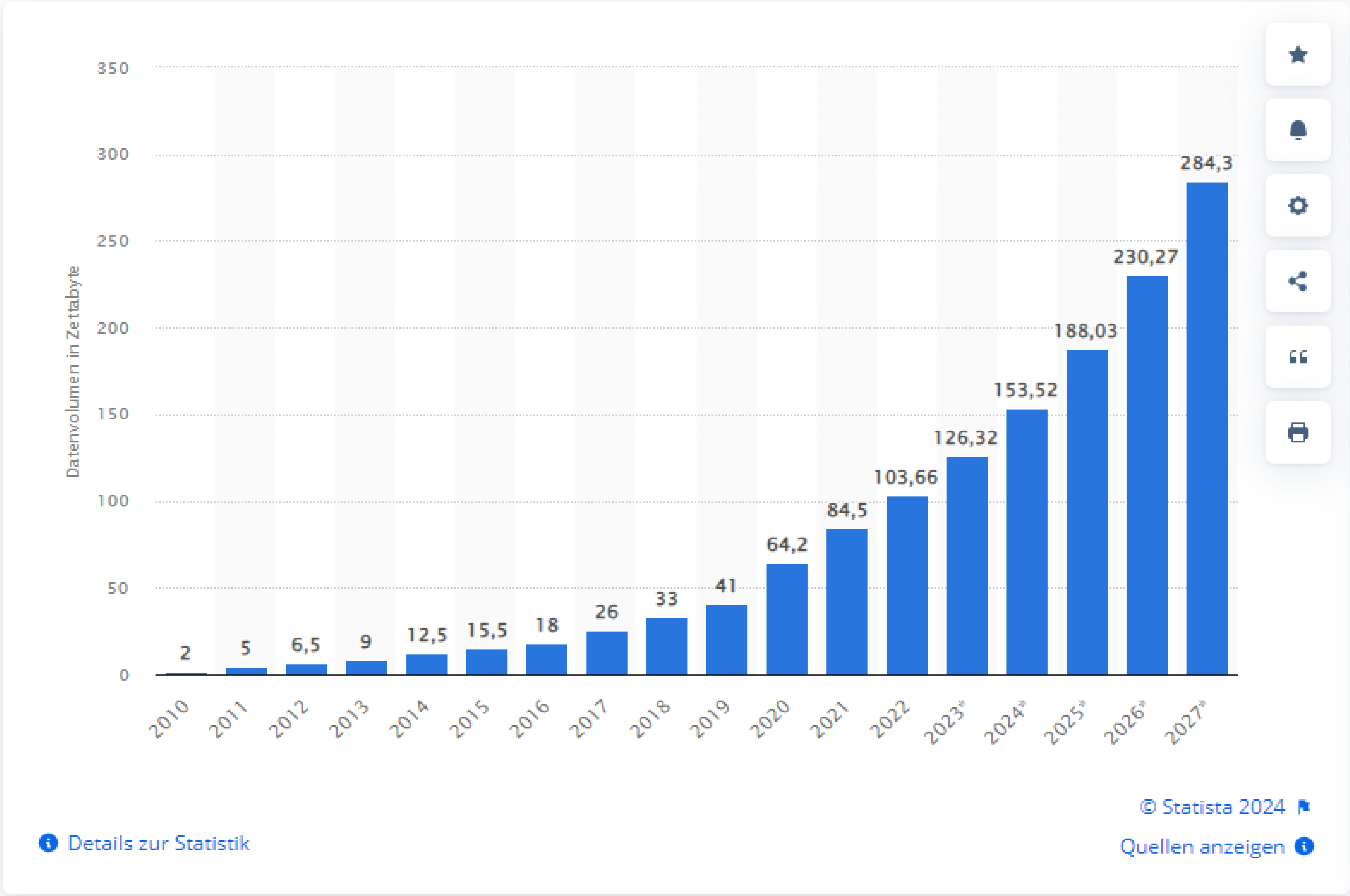
Andreas Göttlicher

- Senior Sales Manager for 22 years
- Meteo / Climate / Oceanography



Data production

(in Zettabyte)





From 80 Byte...

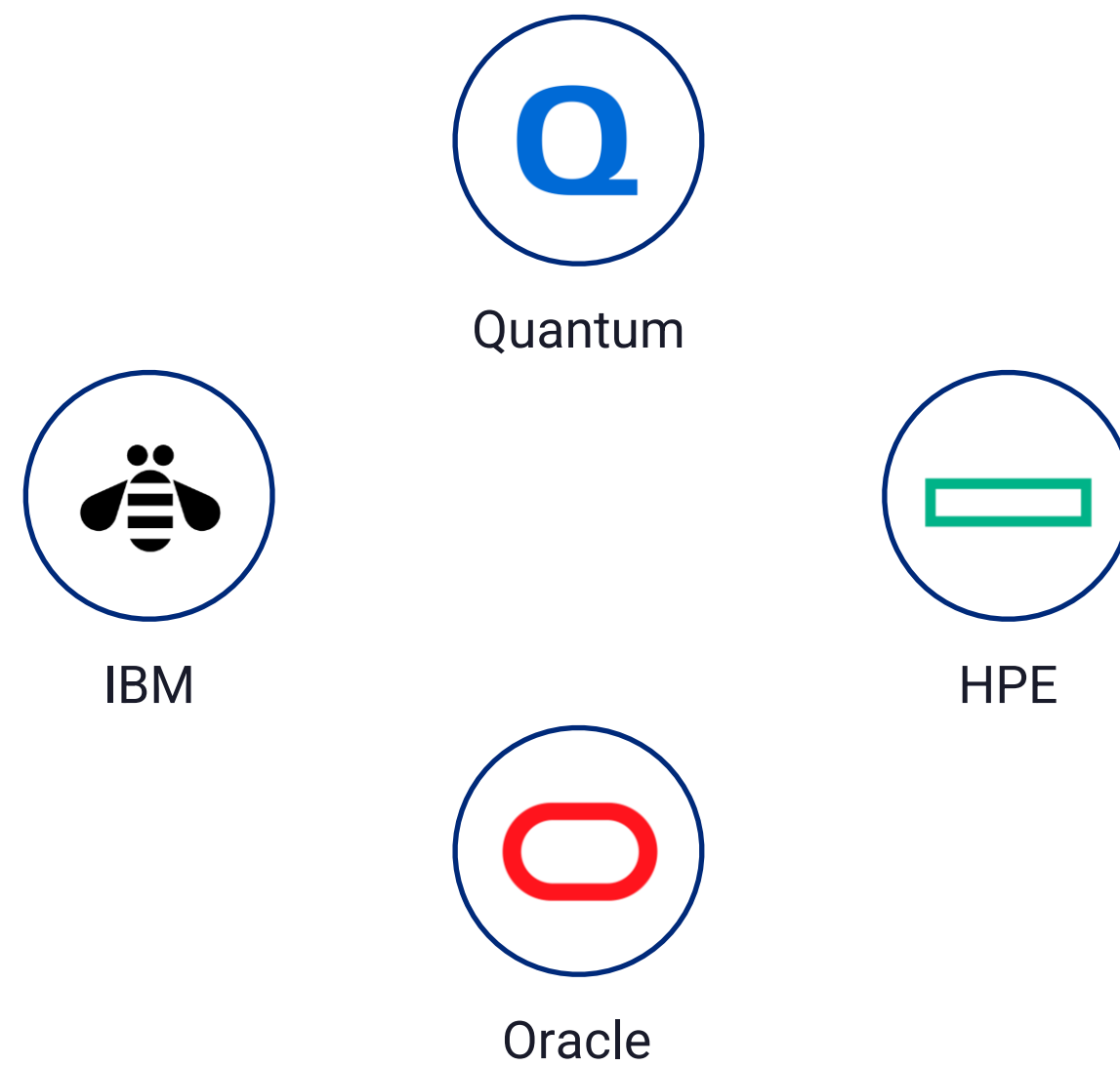


...up to 50TB

Large Archive Landscape

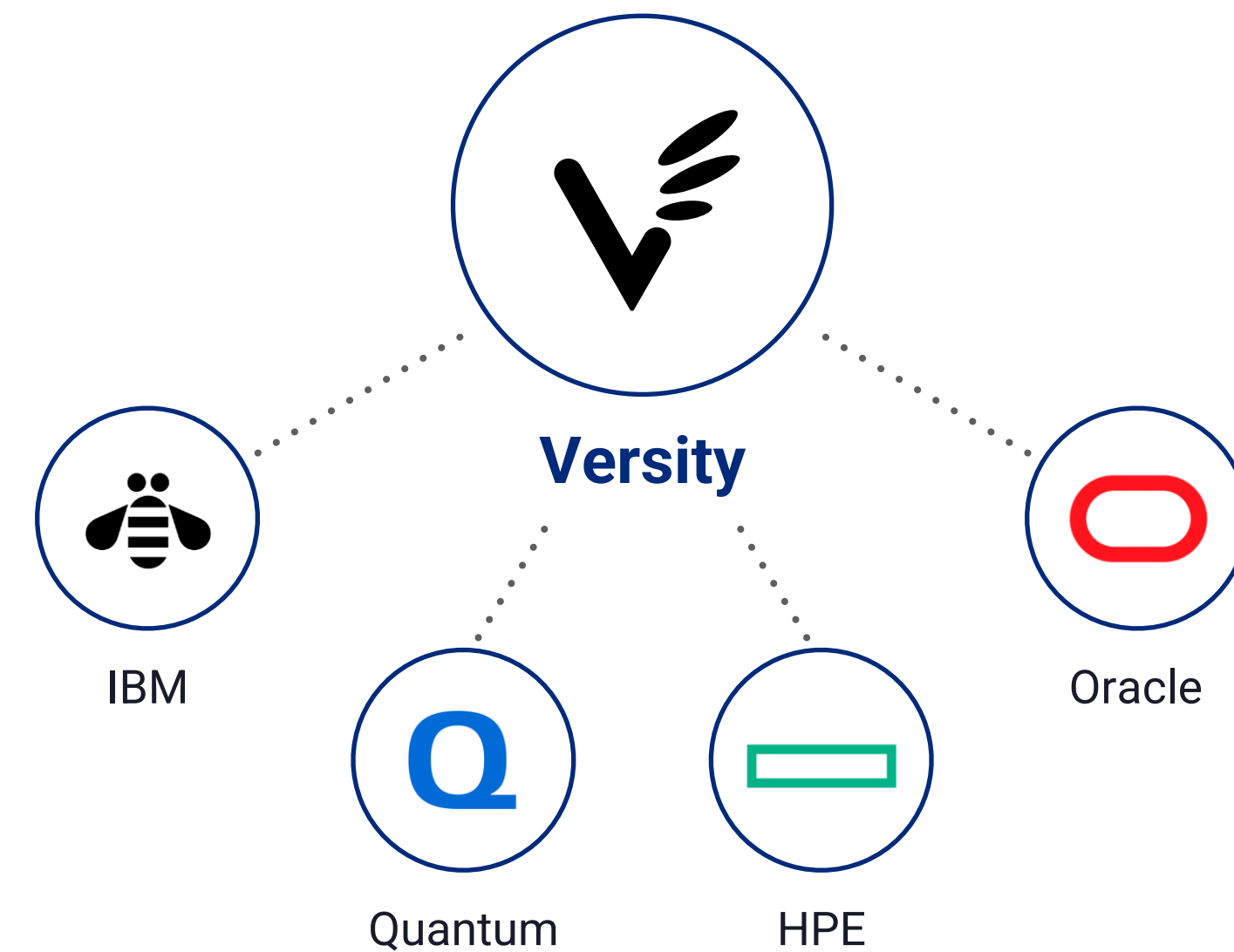
Past

Highly Fragmented Market - No Center



Present

Consolidating Around Clear Leader





Bruce Gilpin
CEO and Co-Founder

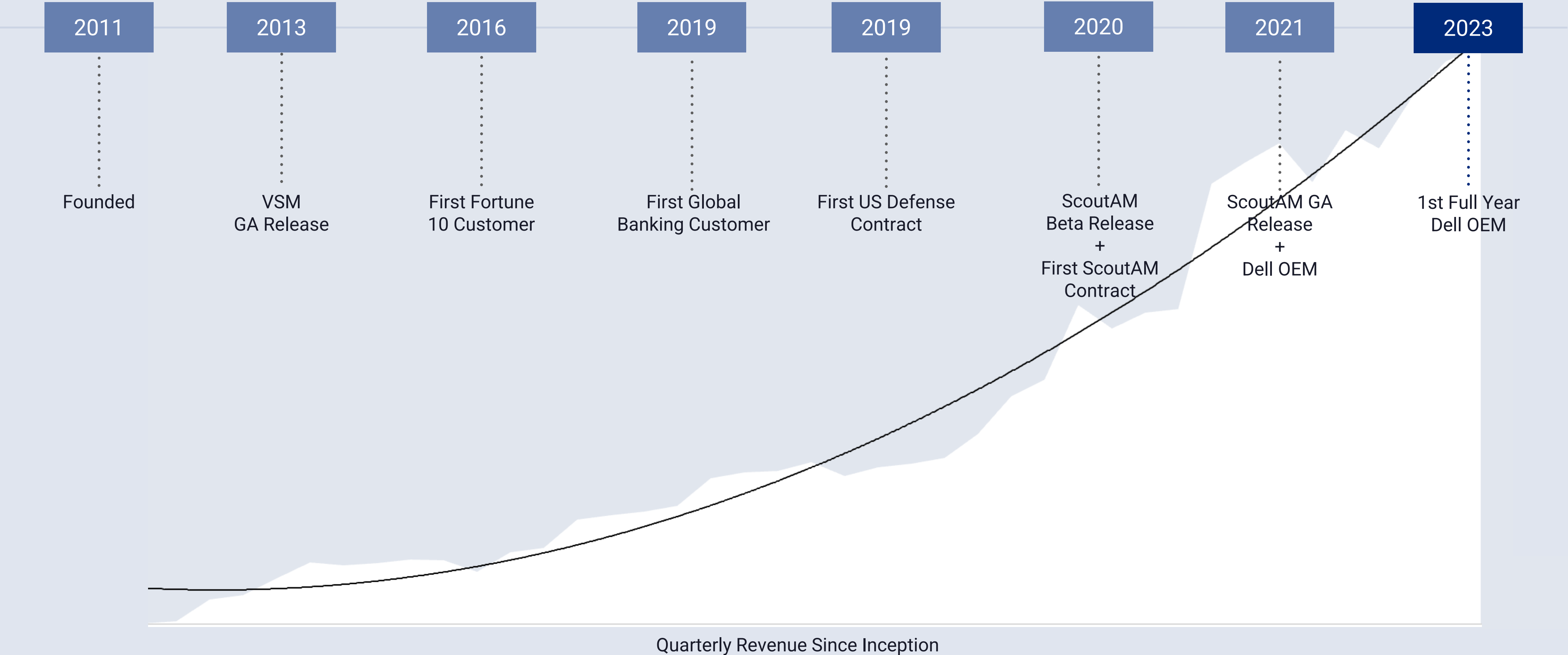


Harriet Coverston
CTO and Co-Founder

[Harriet Coverston - Wikipedia](#)

NEC Deutschland GmbH
HPC Department

History & Growth





What We Do at Versity:

**Manage large data collections
at the lowest possible cost**

- A Modern Software-Defined Storage Platform
- For Mass Storage & Large Archive Systems

Efficient Mass Storage for Cloud-Scale Archival Data Collections

Our solution provides:

- On-premises glacier infrastructure for ultra-efficient unstructured data preservation
- Secondary storage for HPC and data lake platforms with efficient archival storage functionality
- Highly differentiated replacement for first generation systems including IBM HPSS, IBM Spectrum Archive (TSM), HPE DMF, Oracle OHSM and Quantum StorNext Storage Manager



Versity Customers

- Have extremely large data collections that are too big for backup
- Require long-term preservation
- Actively read & write large data sets
- Need cost efficiency, versatility, control, and performance



Versity Global Installed Base

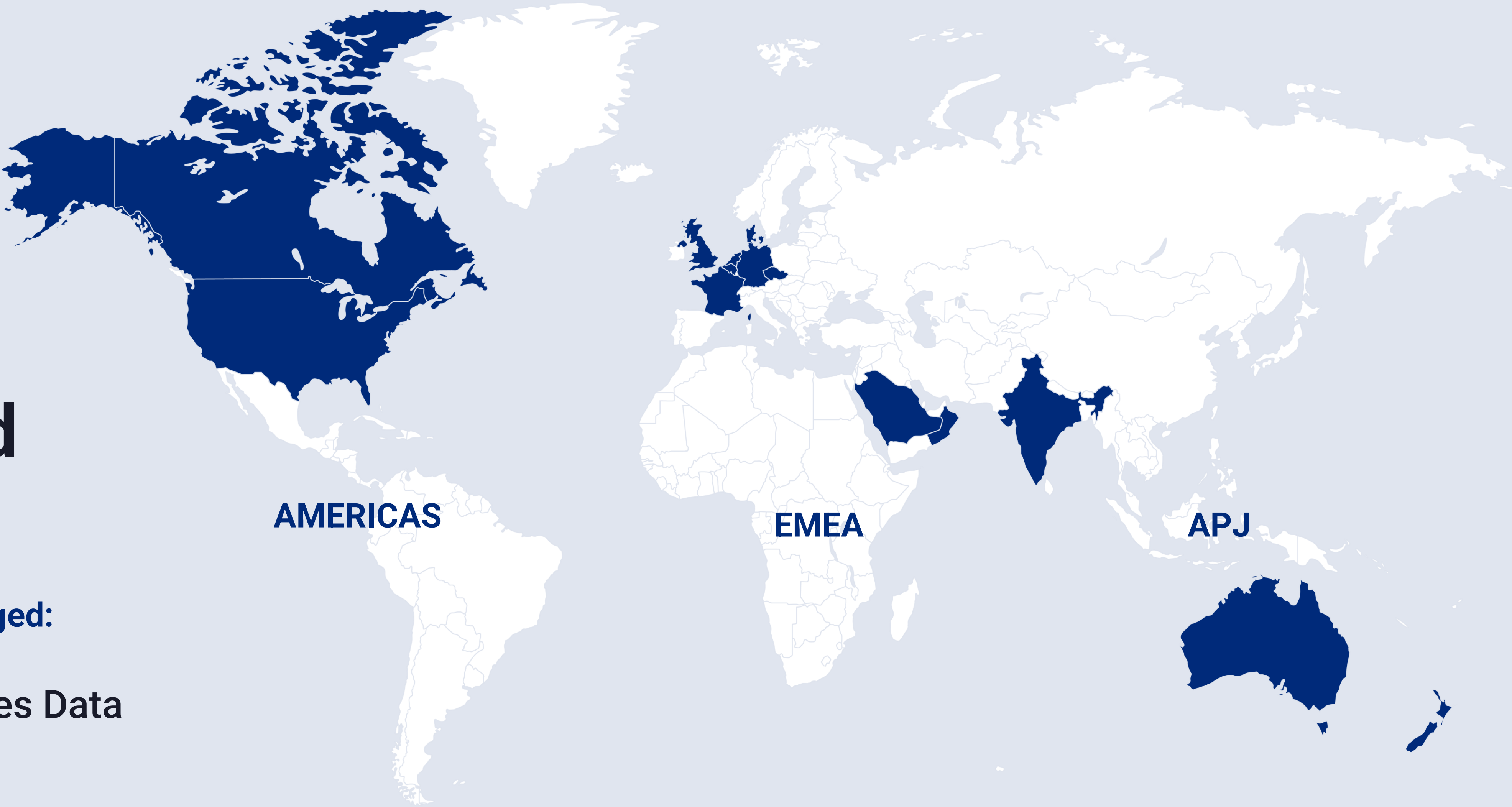
Total Data Managed:



2 Exabytes Data



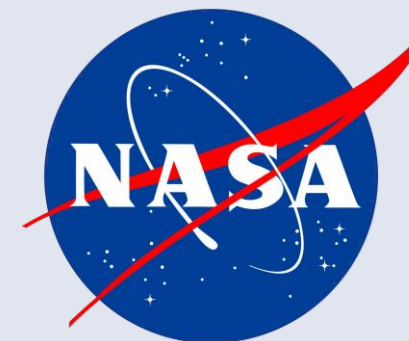
95 Sites



Versity Customers



Versity HPC Customers



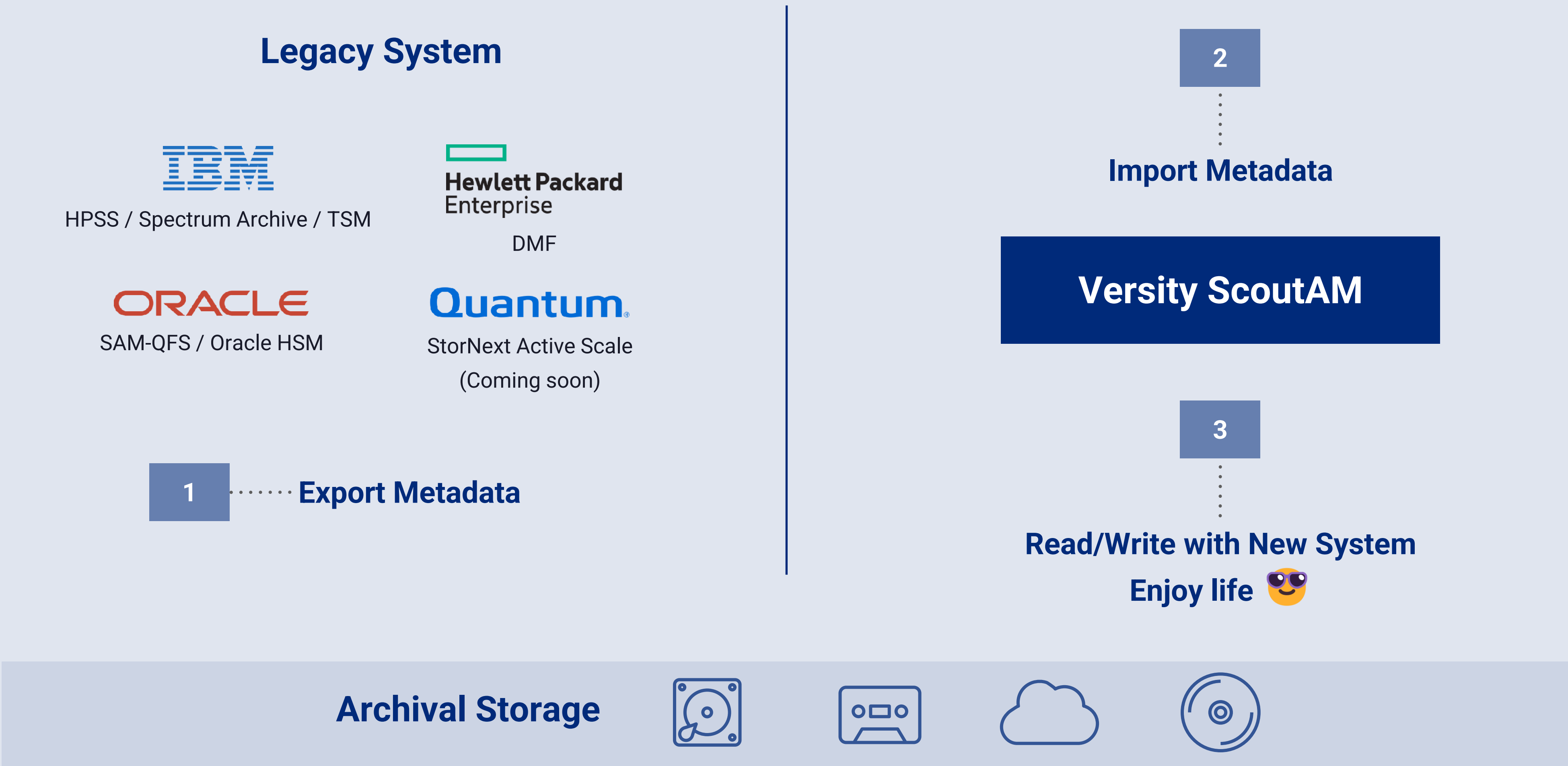
Why Versity ?

The **ONLY** company offering:

- Open-source file system & S3 Gateway
- Open source archival file format
- Free and perpetual read license
- Ultimate disaster recovery
- IBM HPSS, IBM Spectrum Archive (TSM), HPE DMF & and Oracle OHSM format compatibility



Converting to Versity with Zero Migration



Zero Data Migration

Versity Confidential

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

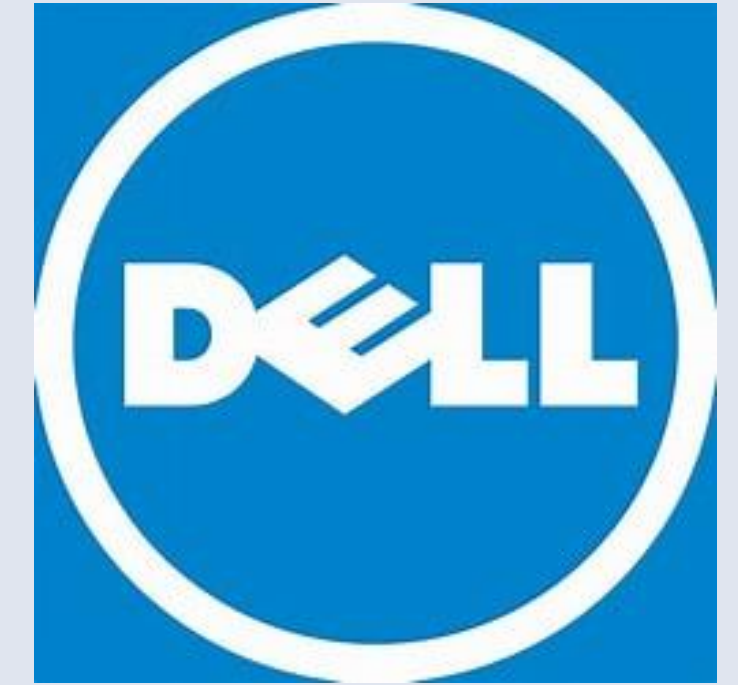
**High-performance
object-to-file translation tool**

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 ScoutAM – Proof of Concept (POC)

Andreas Göttlicher, NEC

System overview

Dell PowerEdge R730



- 2x Intel Xeon Gold 5418Y (2.0 GHz, 24C/48T)
- 256 GB ECC-RAM
- 2x 480G SATA SSD for OS-Installation
- PERC HW-RAID Controller
 - 12x 3.84 TB SAS-SSD
- 1x Qlogic FC HBA (Access to IBM TS4500)
- 1x Dual-port 25G NIC (Access to customer network)

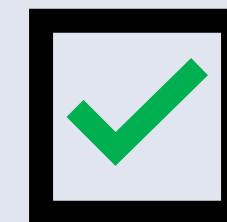
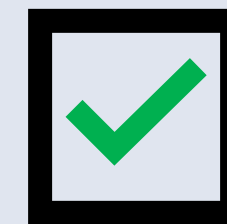
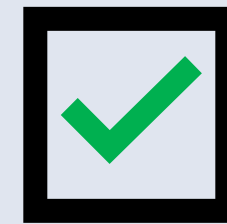
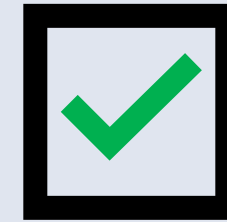
Storage configuration

- 1x RAID1 (1+1 drive) for Metadata
- 2x RAID5 (4 +1 drive) for Data

System was sized for functionality tests, not for performance tests!

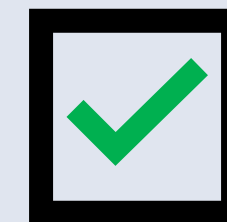
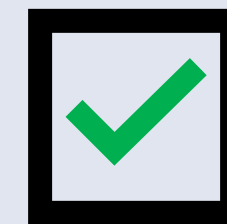
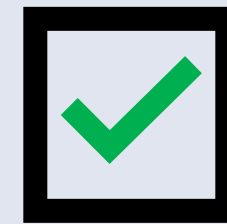
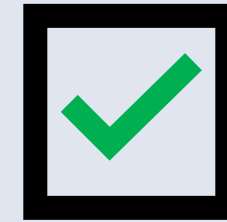
Functionality tests

- Import of an HPSS written tape
 - Read metadata from HPSS labelled tape and stage data from native tape
- Migrating/Staging
 - Read/write data from/to archive storage pool, verify data gets stored on tape
- Forever archive
 - Write three identical copies to different storage targets
- Backup
 - Test cluster was backed up to storage pool daily
- File Server
 - File Server software was compiled, installed and functionality verified on ScoutAM host



Functionality tests - continued

- Data management component
- Data management application was installed to ScoutAM host and functionality tested
- Tape recycling
- HPSS tape contents where read into ScoutAM and tape was relabelled for future use
 -
 - Read of a ScoutAM written TAR archive from tape
 - Native read from ScoutAM written tape in TAR format
 - Access via S3 interface
 - ScoutAM Gateway was tested for read/write functionality
- Overview ScoutAM management (CLI and GUI)
 - Functionality of CLI and GUI components was demonstrated



Summary

- ❖ The POC installation showed that all archive related workflows are successfully serviced by ScoutAM.
 - ❖ The provided system ran stable and did not require much manual efforts during test operations.
 - ❖ A migration from a HPSS system needs careful planning (but not scaring).
 - ❖ The openness of ScoutAM using only OSS formats is regarded very positively.
 - ❖ Interested in more details or an onsite POC? Please send me a message.





Thank You

andreas.goettlicher@emea.nec.com

sales@versity.com



@versitysoftware