Overview of the Storage Landscape: Trends, Evolution and Key Developments

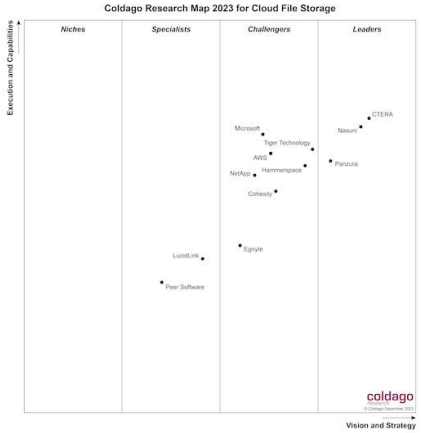
Philippe NICOLAS
pnicolas@coldago.com
Paris - May 28, 2024

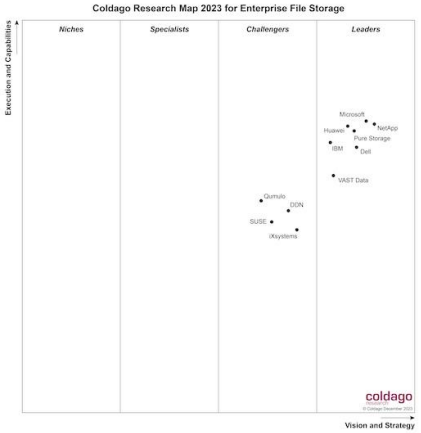
coldago research

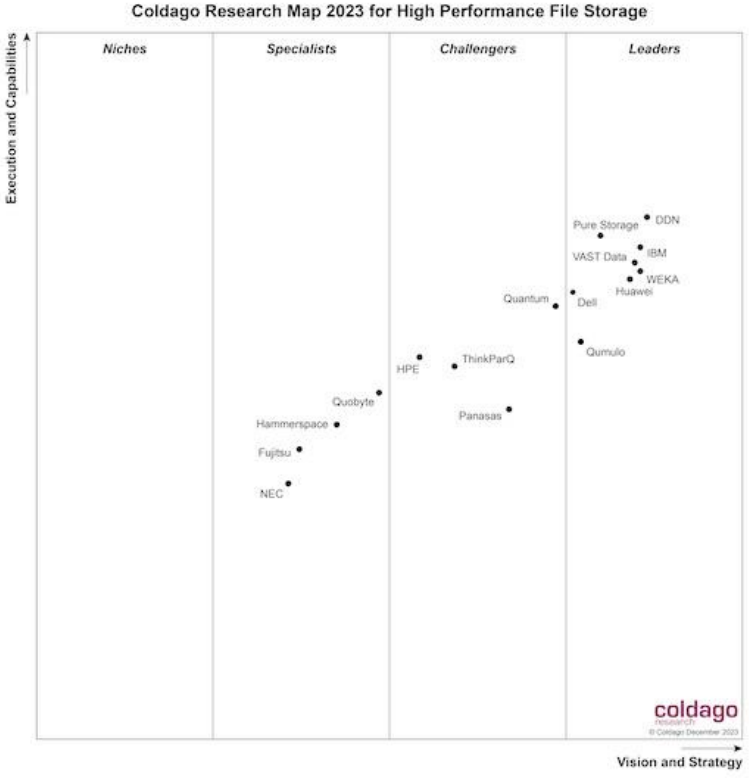
A few words about me and my activities

- Me: In the data storage industry since the beginning of 90's
  - Working at Compaq, Silicon Graphics, Veritas Software, Symantec, Brocade Communications & Scality plus SNIA
- Plus
  - Coldago Research: 2009 2018 Map Unicorn Gems End-Users Survey coldago.com
  - The IT Press Tour: 2009 56 300 500 itpresstour.com
  - StorageNewsletter: 1988 2007 2010 75000 storagenewsletter.com
  - The French Storage Podcast: 2020 Covid 200 tfsp.fr

Coldago Map 2023 for File Storage







A hot and dynamic market

coldago research

Primary & Secondary Storage

Cloud Storage

SaaS/PaaS/IaaS

Security and Storage

Data Protection

Software, Software, Software

HDD, SSD, Flash, NVMe

Access Methods

Data Lake/Lakehouse

Energy/ESG

New Use Cases & Workloads

Container/Kubernetes

Open source

M&A

AI & Stockage

Composability, Disagg...

Startup

Data (Storage) Platform

coldago research

- Notion of DataStore
- Convergence, Unification, Consolidation, Composability, Disaggregation…
- Multi purposes, Multi use cases
- “Sum of silos” with some horizontal behavior
- “Storage Data Lake”
- Virtualization - Confirmation of the SDS wave (Block, File and Object)
- Convergence between Primary and Secondary Storage
- HDD, SSD, Flash, NVMe… RDMA, CXL
- Multiple Access Methods (files, S3) + block mode but also HDFS
- U3 - “Unified, Universal & Ubiquitous Storage”

Media

- From a simple world with HDD and Tape to a complex rich one with Flash in all flavors, HDD, tape, Cloud
- Tape is dead? Yes for backup, Tape is a (deep) archive media
- HDD shift to secondary as well, everything else is Flash
- Tape lost their capacity advantage not with HDD but also with SSD (Solidigm, Huawei…)
- NVMe and its network companion
- Next world? Flash/SSD + tape

Architecture

coldago research

- Scalability and Availability plus Durability => New ideas
- Shared-nothing and Shared-everything (NVMeoF with Disagg.)
- Fast networks
- Modularity with containerization of services (+ Kubernetes)
- Storage services for new application models (microservices)
- Any to Any - Decoupling layers (access vs. capacity/storage)
- Asymmetric model to scale (Metadata and Data segregation)
- SFSI and cacheless at scale
- GPU impact on system design (LPU…)

Object Storage

coldago research

- Origin
- S3 Storage
- Shift in use cases and adoption (primary and secondary storage)
- 3 flavors for 3 domains: High Performance, Capacity oriented and Edge
- Vertical vs. Horizontal model (OID vs. Vol. Mgmt./FS)
- VTL 2.0 with S3-to-Tape
- Key/Value Store popularity
- Erasure Coding

File Storage

coldago research

- No real new new disk file system, neither network or distributed flavors
- NFS and SMB still very popular
- Performance options: RDMA, nconnect, multipathing, multichannel…
- Parallelism (proprietary vs. standard, open source…)
- Popularity of Parallel FS thanks to AI beyond HPC
- File + S3 to access same content
- Gateways, global access (GFS, NFV)

AI & Storage

coldago research

- “AI Storage” - Storage for AI
  - HPC heritage, Parallel domination
  - Driven by latency, IOPS and BW constraints
  - GPU, NPU & GPUDirect Storage
- “AI in Storage” - AI in or for Storage
  - “Vertical” ML, LLM, NLP…
  - New way to manage data, information… learning and generation
  - Data protection, pattern recognition, content indexing, autonomous data surveillance, predictive analysis, decision phase…
  - AIOps

And Also

coldago research

- Data Management
- Data Protection & Security (Ransomware)
- ESG
- Open Source
- Container & Kubernetes
- and don’t forget what Europe loves (regulations, norm, standards…)

Q&A

coldago
research