# Fremtidens AI-forskningsmiljø på ét fælles datalag

## Case: Aalborg Universitet - ét datalag fra lab-instrument til GPU


**Tommy Nielsen**
Sales Director
Tommy.Nielsen@advania.dk

**Jens Melhede**
Principal Architect
Jens.Melhede@advania.dk

# Agenda

30 minutter, tre områder:

- **Del 1:** Den nye flaskehals er data, ikke FLOPS - konvergensen mellem HPC- og AI-workloads som metodeskift
- **Del 2:** Et operativsystem for data: ét datalag, alle protokoller, alle tiers fra flash til aktivt S3-on-tape-arkiv
- **Del 3:** Fra arkitektur til beslutning: valgene bag Aalborg Universitets nye AI-forskningsmiljø

# Den nye flaskehals er data, ikke FLOPS

Forskningen har blandet simulering og AI sammen. Infrastrukturen har ikke fulgt med.

**Det dyreste i en AI-installation er ikke regnekraften - det er GPU'er, der venter på data**

- Før målte vi flops pr. watt. Nu er spørgsmålet, hvor meget træning vi får for strømmen - epochs ☺ pr. watt

**Simulering og AI er blevet det samme arbejde**

- Resultatet af én simulering er træningsdata til den næste model - og modellen kan både overtage de tungeste dele af regnestykket og blive brugt, mens regnestykket kører

**De to verdener stiller modsatte krav til lagringen - og det er hele forklaringen på siloerne**

- HPC skriver få, meget store filer i ét stræk. Træning læser millioner af små stumper i tilfældig rækkefølge




*Regnekraften blev billigere. Vejen hen til den blev flaskehalsen.*

# Tre systemer til de samme data

Arven fra en tid, hvor lagring kun kunne én ting ad gangen

### Tre systemer

- Parallel file system - til HPC og checkpointing
- Object storage - til AI-pipelines og deling
- Tape archive - til det, ingen rører

### Og hvem der betaler

- Tre kopier af de samme data, tre adgangsveje, tre teams
- Forskeren venter på staging, før jobbet overhovedet går i gang
- Det, der ikke kan findes, bliver ikke brugt - men bliver gemt alligevel

**Parallel file system**
kopi 1

**Object storage**
kopi 2

**Tape archive**
kopi 3

*Samme datasæt, tre steder - og ingen af systemerne ved, at de to andre findes*

DEL 2

# Et operativsystem for data

Jens Melhede

Principal Architect, Advania Danmark

PAST




VAST

# The VAST AI Operating System

## DataEngine
Scalable, Event-Driven Computing

Triggers, Functions, Containers

**SyncEngine**
Index & DataRouter

Massive-Scale

**InsightEngine**
Real-Time RAG

Embedding Creation

**AgentEngine**
Tools & Orchestration

Agents, Tools, Observability

## DataStore
Universal Storage Infrastructure – Multiprotocol Access

NFS, SMB, S3, NVMe/TCP

## DataBase
Transactional & Analytical Database

Kafka, Python, Parquet, SQL, Similarity (Vectors)

## DataSpace
Globally-Distributed Data Computing

Hewlett Packard Enterprise · Lenovo · Supermicro · Cisco — Hardware Platforms

aws · Google Cloud · Azure · Oracle — Traditional CSPs

CoreWeave · core42 · Lambda and more… — GPU Clouds

# VAST Data's DASE Architecture

Disaggregated, Shared Everything

## Linear Scalability

100,000s of CPUs/GPUs, Exabytes of Data

## Real-Time At Scale

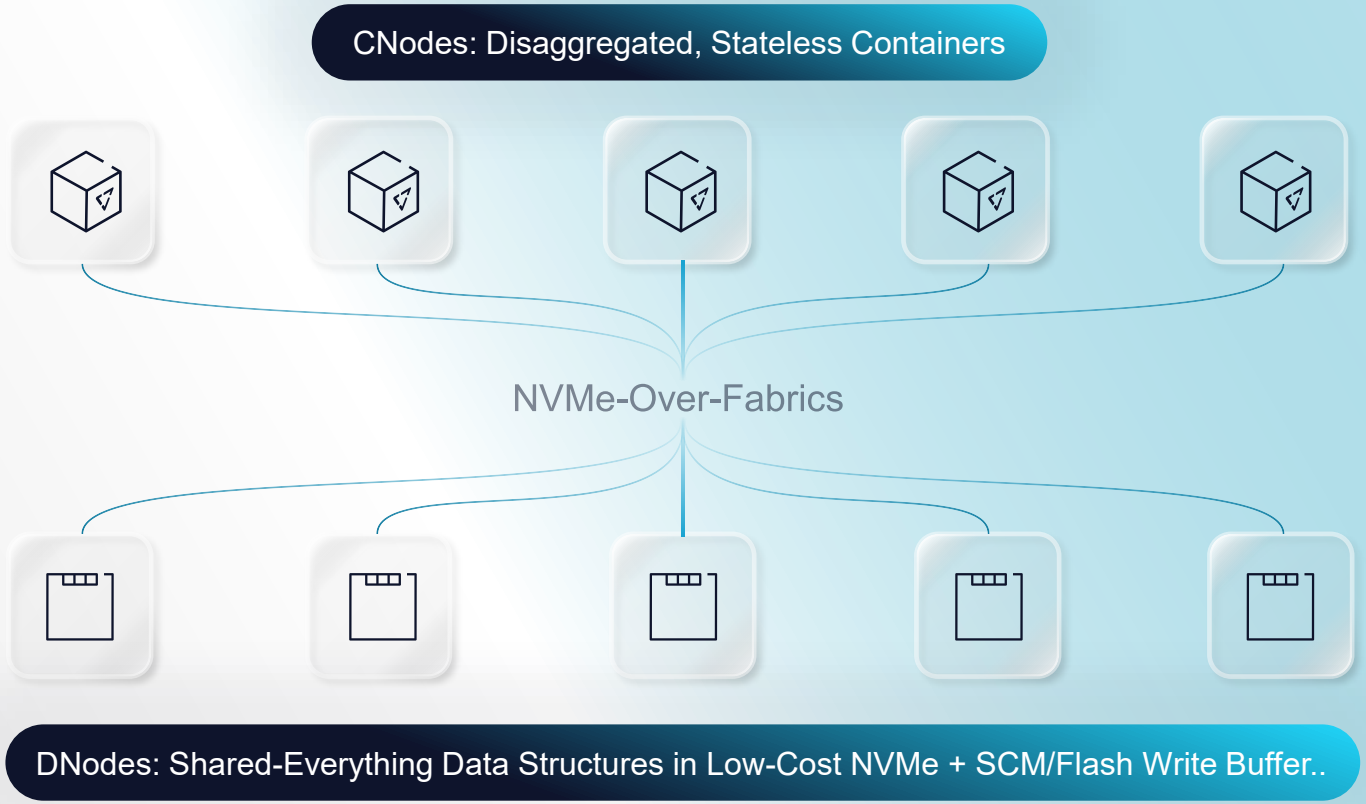
Parallel File I/O, Events, Vectors, Queries

## Resilient at Extreme Scale

99.999+% Uptime In The World's Harshest DCs

## Affordable

Accelerate CPUs, Lower SSD cost, **no DRAM cost**



# Ingen klientsoftware på noderne

Standardprotokoller i stedet for proprietære agenter - det er dér, driften bliver enkel

### Sådan når noderne data

- Standard Linux NFS-klient - ingen kernel module, ingen agent
- NFS over RDMA med multipath over alle front-end-noder
- GPUDirect Storage: direkte ind i GPU-hukommelsen
- S3 til objekter, også over RDMA
- InfiniBand og Ethernet betjenes samtidig

### Hvad det betyder i drift

- Ingen kernel module-matrix at holde i sync med nodeimages
- Et nyt cluster er en mount, ikke et udrulningsprojekt
- Front-end-noder er stateless - klienten skifter sti selv
- Mere båndbredde er flere front-end-noder, ikke en ny arkitektur

**Traditionelt parallelt filsystem (Lustre, GPFS, BeeGFS mv.)**
Klient-software og kernel module på hver eneste node, matchet til kernen

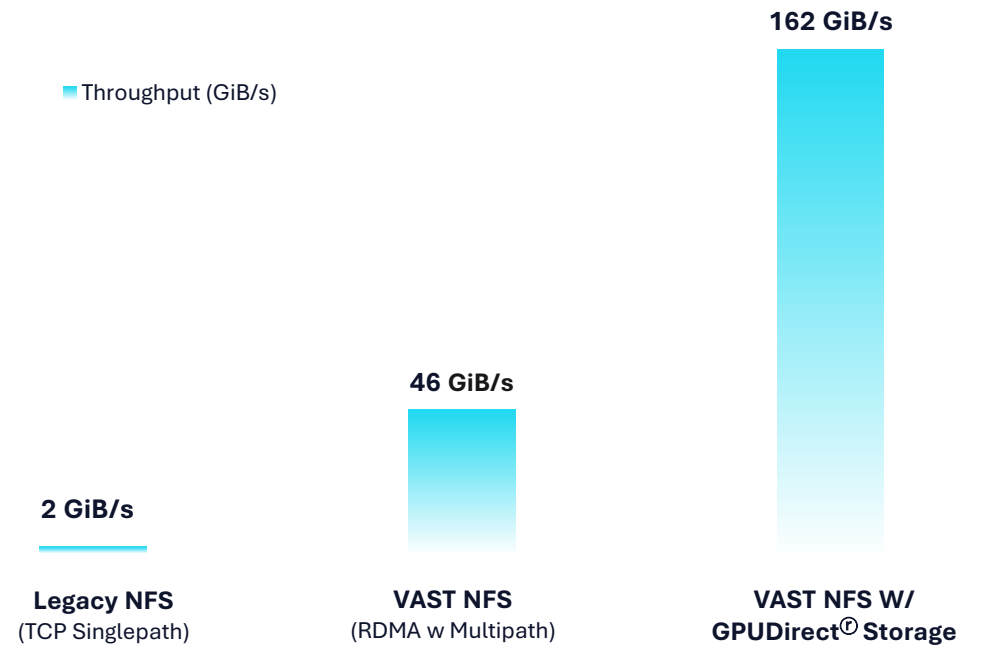
**Standardprotokoller**
Den NFS-klient, der i forvejen er en del af Linux

*Der skal ikke installeres noget på compute-noderne*

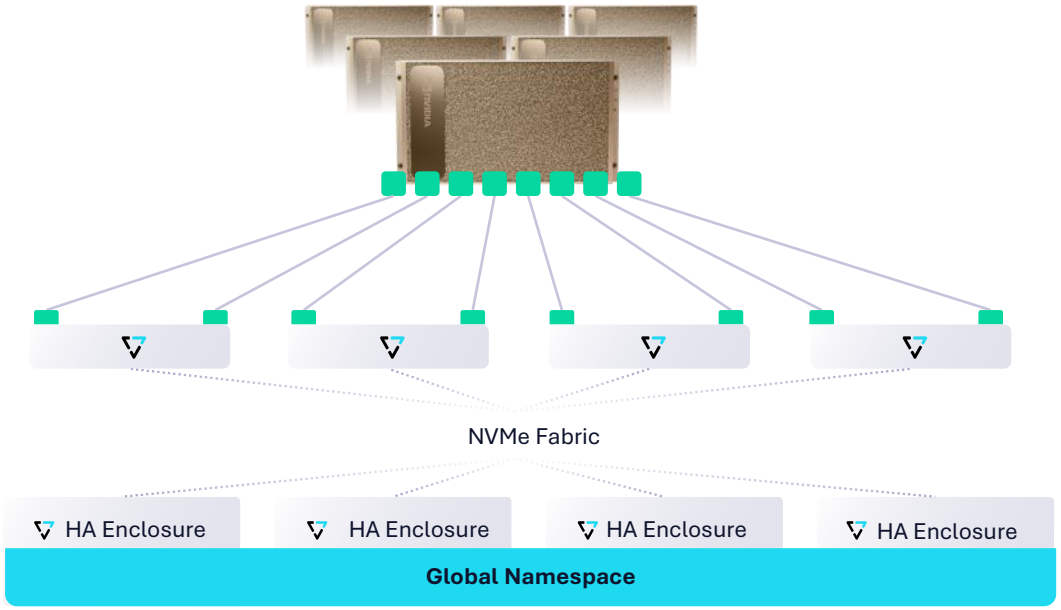
# AI Performance + NAS Simplicity

## GPU Optimizations

### 1x DGX A100 / Single Mountpoint Performance Testing




*https://vastdata.com/whitepaper/#VAST-nfs-parallel-file-system-speed,-nas-simplicity*




**NFS Simplicity + Performance** ✓ RDMA ✓ Multi-Path ✓ GPUDirect™ Storage

INTRODUCING VAST

# Similarity-Based Data Reduction

## The World's Most Powerful Global Data Reduction Engine

2.9:1 Average Data Reduction on Largely Pre-Compressed and Unstructured Data



INTRODUCING VAST

# Similarity-Based Data Reduction

## The World's Most Powerful Global Data Reduction Engine

2.9:1 Average Data Reduction on Largely Pre-Compressed and Unstructured Data



# Hvad det betyder i rackspace og kroner

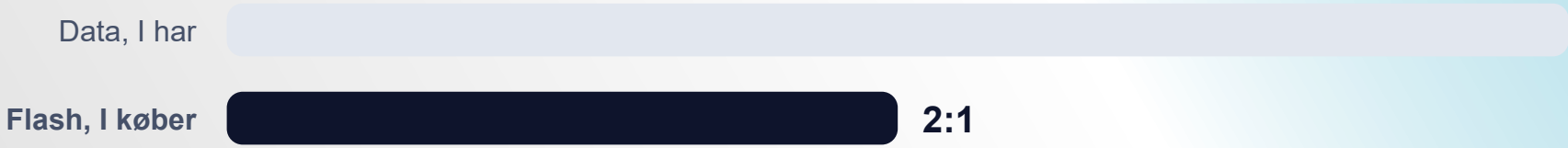
To mekanismer, der ændrer regnestykket

**Similarity data reduction**

- Global og fine grained på én gang
- Virker også på data, der allerede er komprimeret
- Reduktionen gælder hele clusteret og alle datatyper

**DataBase uden DRAM-kravet**

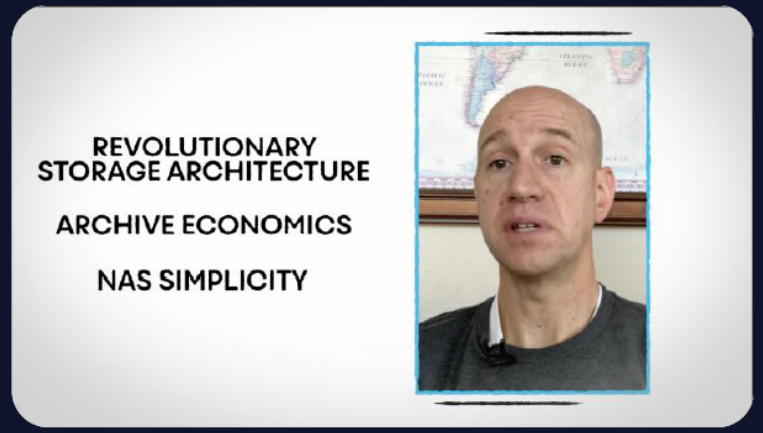
- Metadata og indeks ligger i Storage Class Memory, ikke i dyr DRAM i hver storage C-node
- Den strukturerede del (VastDB) kan skalere uden partitionering og sharding
- Færre noder og mindre flash betyder færre racks, mindre strøm og mindre køling

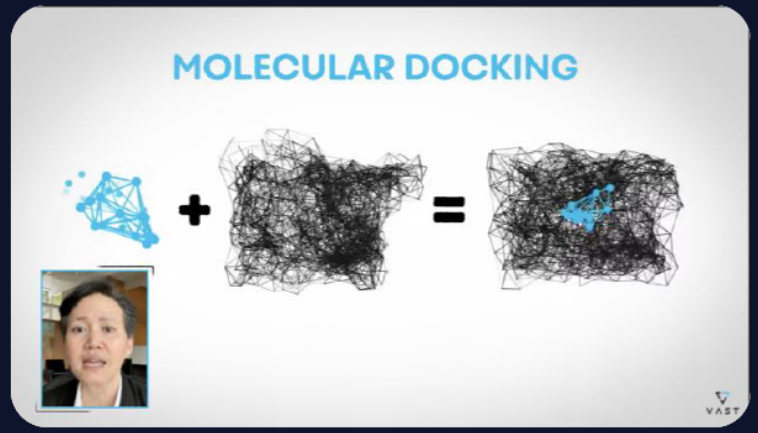



*TACC måler 2:1, LLNL 4:1 på fildata, og VAST angiver 2,9:1 i gennemsnit - også på data, der allerede er komprimeret*

# Lawrence Livermore National Laboratory (LLNL.gov)

## Using VAST as a Hub for Multiple Supercomputers




- Used at LLNL for Drug Discovery
- Billions of small files
- CORONA Cluster
- AI + Modeling + Simulation

**50.000 GPUs**

**LLNL Drug Discovery**

- Sharing Data
- Ease-of-use with industry standard protocols
- All-Flash high performance
- Industry-leading Data Reduction (4:1 at LLNL)

"VAST Data is very interesting to us," Goldstone says, saying that the company's Universal Storage has a lot of promise. "One thing that's very interesting is from the client side, it looks like NFS, meaning it uses the standard NFS client that's part of Linux and is part of the OS stack that runs on all our clusters, versus something like Lustre or GPFS, where you have a client that's provided with the file system, but it is a piece of software that has to be integrated into our software track. On the back end it's nothing like NFS"

"We expect that it will work like a parallel file system… we have seen is that it just does work, it hasn't melted down and that in itself is an important criteria. A lot of times you bring in a new technology that's not true. When we first brought in Lustre that wasn't true"

# TACC: hele rejsen fra HPC til AI på samme platform

Texas Advanced Computing Center - flere supercomputere, ét datalag

**Konsolideringen**

- Omkring 1.900 Stampede3-noder og 856 Vista-noder mounter platformen over standard NFS
- Evalueringen forud for Stampede3 blev kørt mod VAST fra over 4.000 Frontera-noder - det hidtidige parallelle filsystems metadata-server klarede kun omkring 350 noder
- Horizon, TACCs nye 2026 leadership-class system, kører file, S3 og VastDB på én platform

**Driftsbilledet, med deres egne ord**

- Opgraderinger af operativsystemet uden nedetid
- Tre til fire SSD-fejl i alt over flere år i produktion
- Uger hvor filsystemet ikke kræver opmærksomhed - mod cirka to alvorlige lagringsalarmer om ugen tidligere
- Omkring 90 procent af systemafbrydelserne blev tidligere sporet til filsystemet

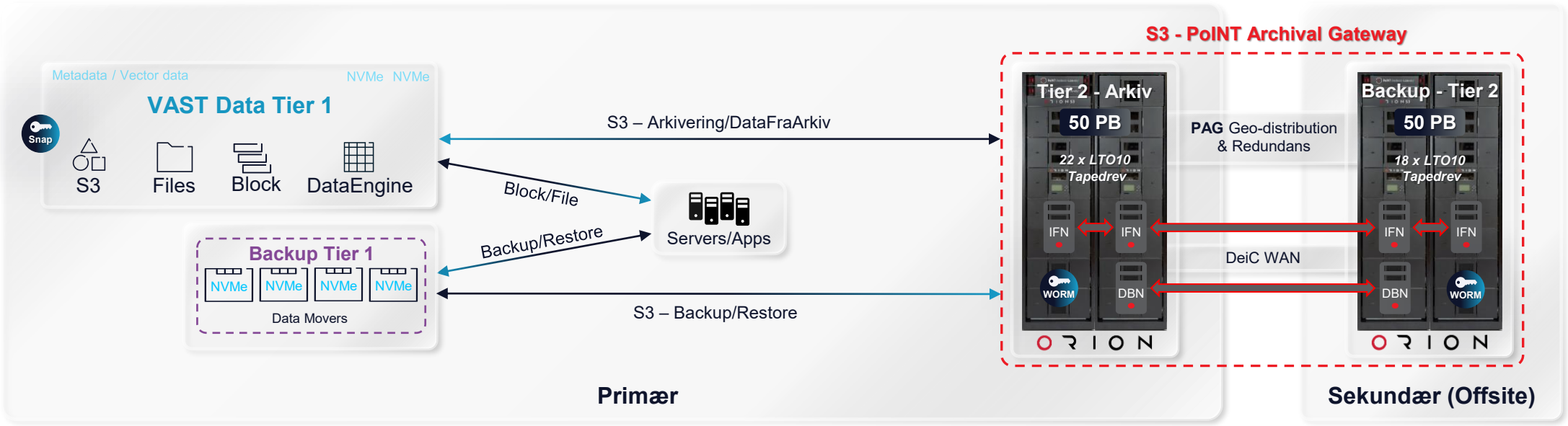
**VAST Is Redefining What Good Looks Like In HPC**

Dan Stanzione – Executive Director

*"VAST is smooth and graceful for what we used to call bad I/O patterns, which we now just call I/O patterns."*

TACC

# Det moderne aktive S3 arkiv med VAST og Tape




**Registreres, når data lander**
Ingen crawl, ingen genindeksering

**Metadata og vektorer bliver på flash**
Kun bytes flytter til tape

**Én forespørgsel**
Fakta og indhold på samme tid

**Recall er en bestilling**
Ikke en afbrydelse midt i et job

Arkivet holder op med at være en blindgyde!

DEL 3

# AAU: Fra arkitektur til beslutning

Tommy Nielsen

Sales Director, AI Foundation,
Advania Danmark

# Udgangspunktet på Aalborg Universitet

Et forskningsmiljø, der skulle kunne flere ting på én gang:

- Overtage adskellige petabyte fra en ældre CEPH-baseret platform
- Stigende efterspørgsel på AI, side om side med klassisk HPC
- Forskning inden for bæredygtighed, sundhed og ingeniørvidenskab - med vidt forskellige krav til data
- Et ønske om at fremtidssikre HPC-infrastrukturen frem for at lappe på den

# Kravene - og hvad de betød i arkitekturen

Fire krav, fire arkitekturvalg

**Kravene:**

- Én platform til både HPC- og AI-workloads
- Governance til følsomme data
- En GPU-klar data vej
- Vækste nemt – uden at det er et nyt projekt hver gang

**Svarene:**

- Ét single global namespace, alle protokoller på de samme data
- Multi-tenancy, adgangsstyring og fuld sporbarhed på dataflowet
- NFS over RDMA og GPUDirect - ingen staging før træning
- Kapacitet og performance skalerer uafhængigt af hinanden

# Beslutningen - og det, der viste sig at være svært

Et nyt on-prem AI-forskningsmiljø på VAST AI OS, leveret af Advania på Cisco-hardware

- Adskellige petabyte migreres væk fra CEPH platformen – mens forskningen kører
- Metadatakvalitet er det, der afgør, om data kan findes bagefter
- Governance på følsomme data skal aktiveres, før AI-arbejdet begynder
- Én platform der forener data, compute og AI services

# Er jeres infrastruktur AI-klar?

Et par små tests / et par spørgsmål – I kan tage med hjem:

- Kan jeres forskere her i 2026 søge på tværs af alt data – som jeres institution har produceret siden 2015, uden at bede nogen om at restore noget?
- På mandag:
  - Tæl jeres kopier. (Der er penge at spare – for de fylder – og det koster ☺)
  - Find ud af, hvor stor en del af jeres data, der ikke er læst i 12 måneder. (Så skal det flyttes til et aktivt S3-to-tape arkiv)
  - Undersøg, om jeres arkiv er søgbart. (Et aktivt S3 arkiv er løsningen – der kan være meget værdi at hente i jeres arkivdata)
- Vil I høre, hvad AAU også gør?
  - Jacob-Steen Madsen, Vicedirektør AAU IT Services, i Sal A klokken 13.30: Mistral på AAU: Hvordan bygger man en AI-service til forskning?