# Data at Scale
# A Metadata Viewpoint

Jean-Thomas Acquaviva
jtacquaviva@ddn.com

CS3 2023

ddn

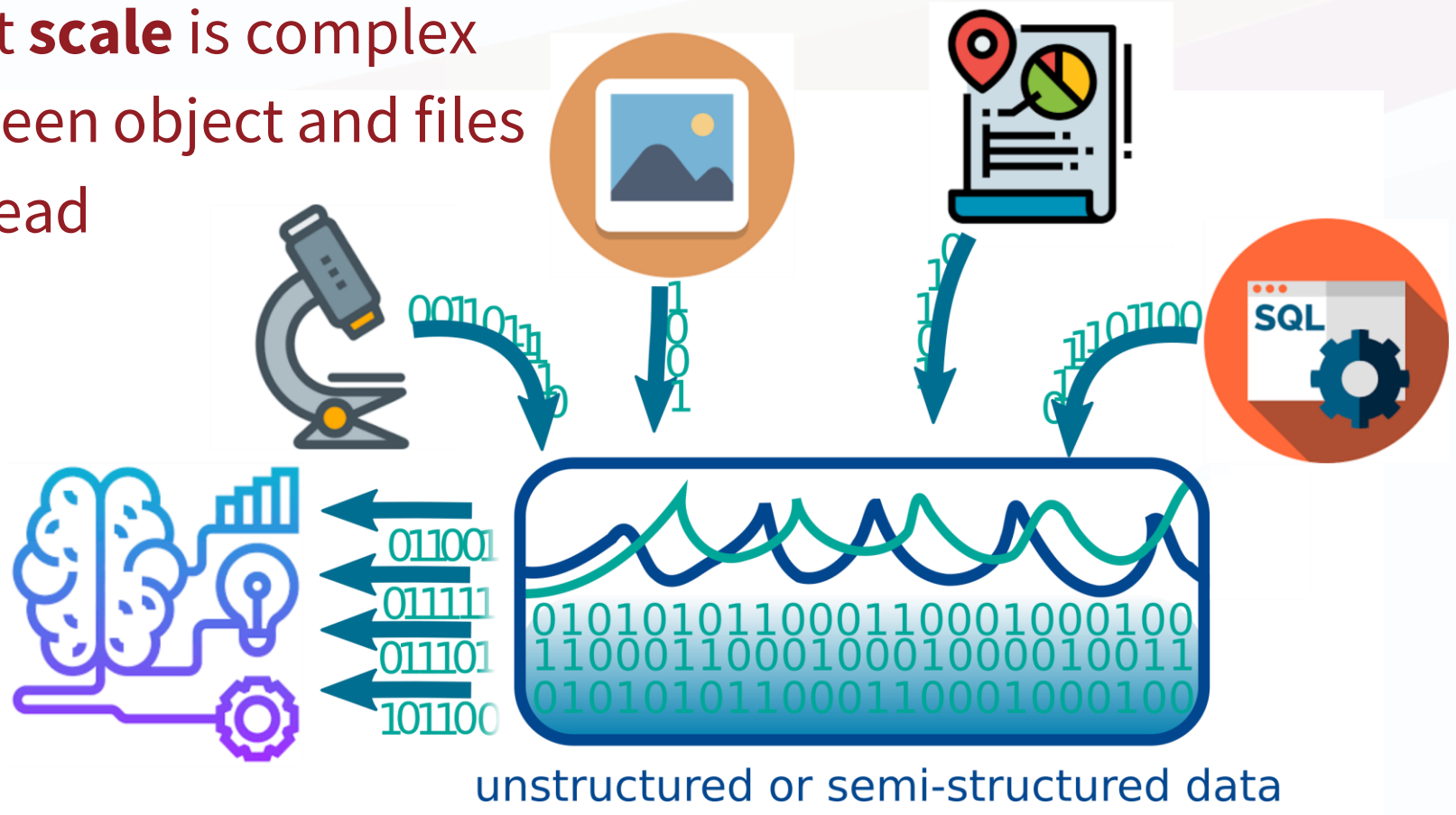
# Introduction

**A File system lives 5 years, but data live forever**

- Data life outpaces the lifetime of their substrat
  - Data migration is needed
- Data value evolves over time
    - New data tends to have higher value
  - None-reproducible data are always valuable
- Data ownership and access policy changes over time
  - Depends on organizations/policies evolution
- Data live cycle need to be manageable
  - We need data describing data to manage data

# Data Exploitation: File interface + objects

- File system at **scale** is complex
- Bridges between object and files
- schema-on-read



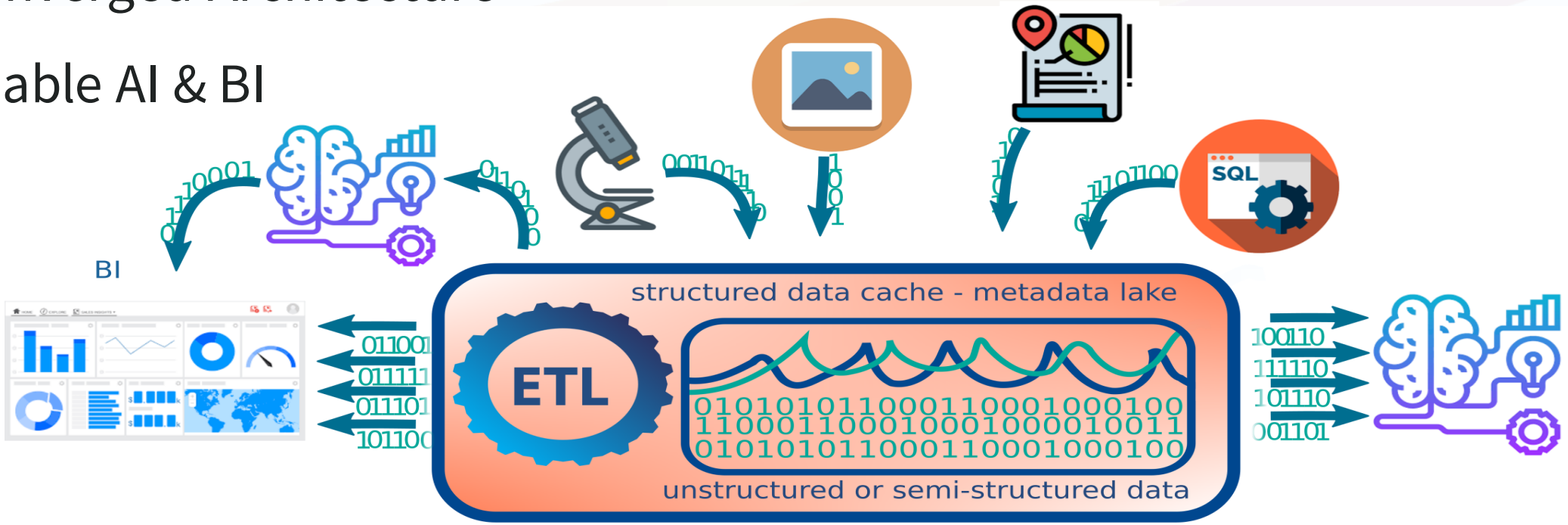
# Data at Scale

**What does it mean to operate at scale?**

- Customer site: 15 PB
  - No problem we can manage 100s of PBs
- Number of files: 1.5 Billion
  - Where are **my** experimental results?
- Lack of discipline turns Datalake to Data Swamp
  - Unstructured data warehouse tend to become unmanageable
- Structuration and self-discipline
  - No real enforcing tools for data policy
  - Relies on file path and file naming

# Data Lakehouse: structured data lake

- Converged Architecture
- Enable AI & BI



# Structuring datalake with Metadata

**Metadata allows to describe data**

- **Structural:** Defines the metadata elements that need to be collected; labels like title, author, date created, subject, purpose, etc. Defining these structural elements is typically based on a mix of organizational and system needs, along with standard schemas like Dublin Core.
- **Administrative:** Often created automatically when content is entered into the CMS, these values are used to manage the content. Administrative metadata includes things like date created or author. They can sometimes include sub-elements about rights-management or preservation.
- **Descriptive:** These values describe aspects specific to each content component, like title, subject, audience, and/or purpose.

# Scientific Community tools

## Scientific dataset


- Project RUCIO
  - https://rucio.cern.ch/
  - High energy physics but evolve toward of genericity

https://indico.cern.ch/event/1185600/contributions/5101386/attachments/2545507/4383570/MetaCat%20for%20Rucio%20Workshop%202022

- Project Phidias
  - https://www.phidias-hpc.eu/

PROJECT
PHIDIAS

**PHIDIAS is creating access services to increase the High-Performance Computing (HPC)**

and data capacities of the European Data Infrastructure in the context of the Connecting European Framework (CEF) on Open data.

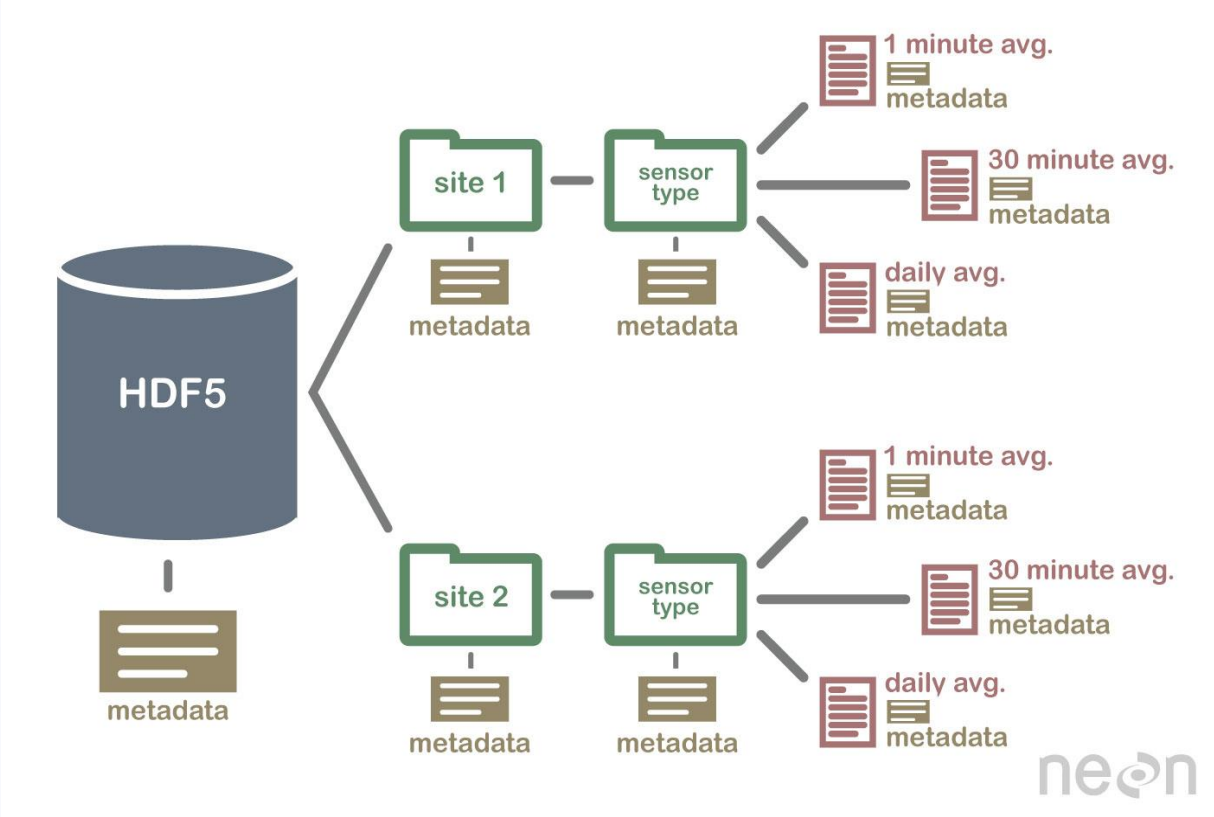
DISCOVER

# I/O Library Community

## Extended attributes

Self-descriptive data format

HDF5, NETDCDF




Courtesy of www.neonscience.org

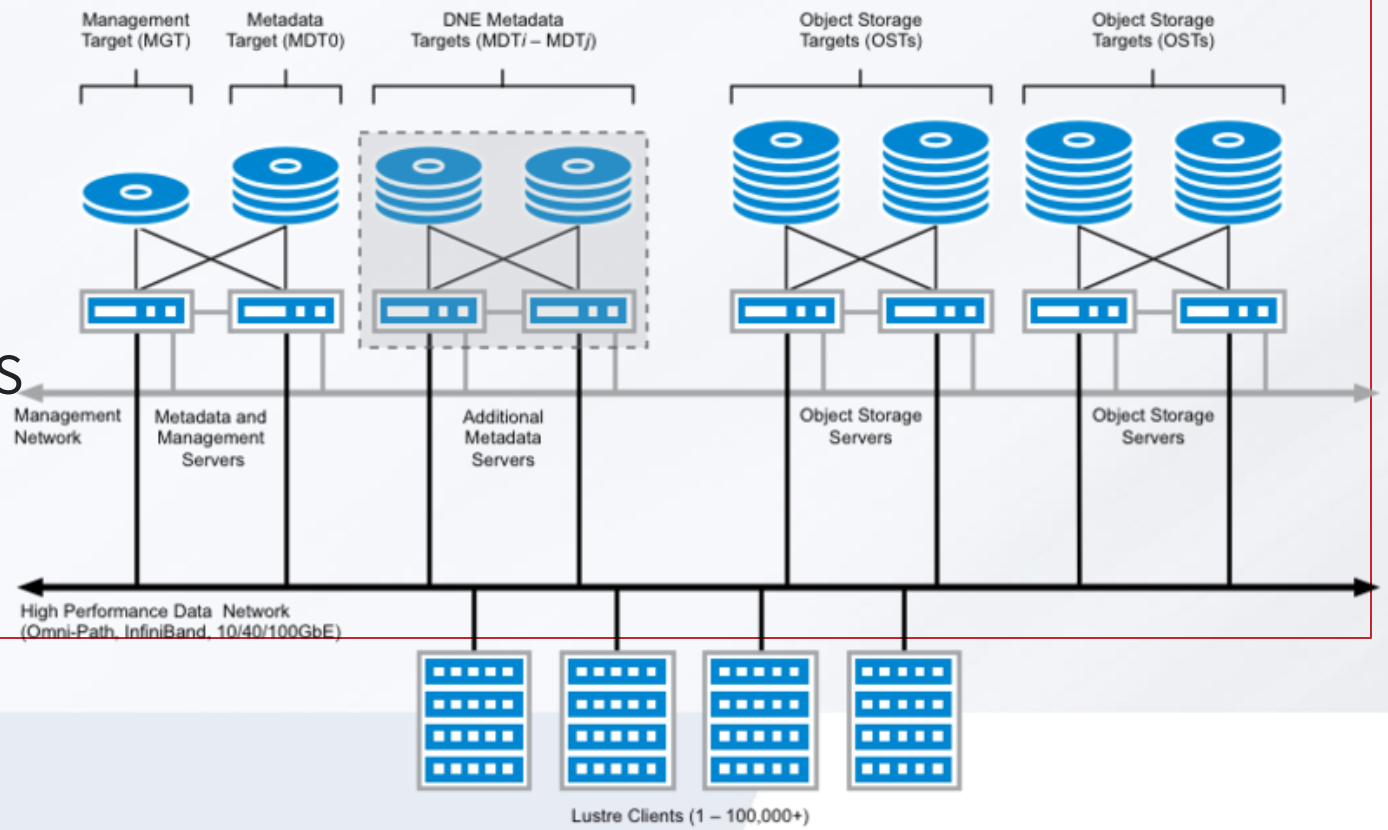
# File System Community

- Building block to implement a metadata schema
- File System extended attributes

```
%setfattr -n=instrument -v=LOFAR my_result.dat
```

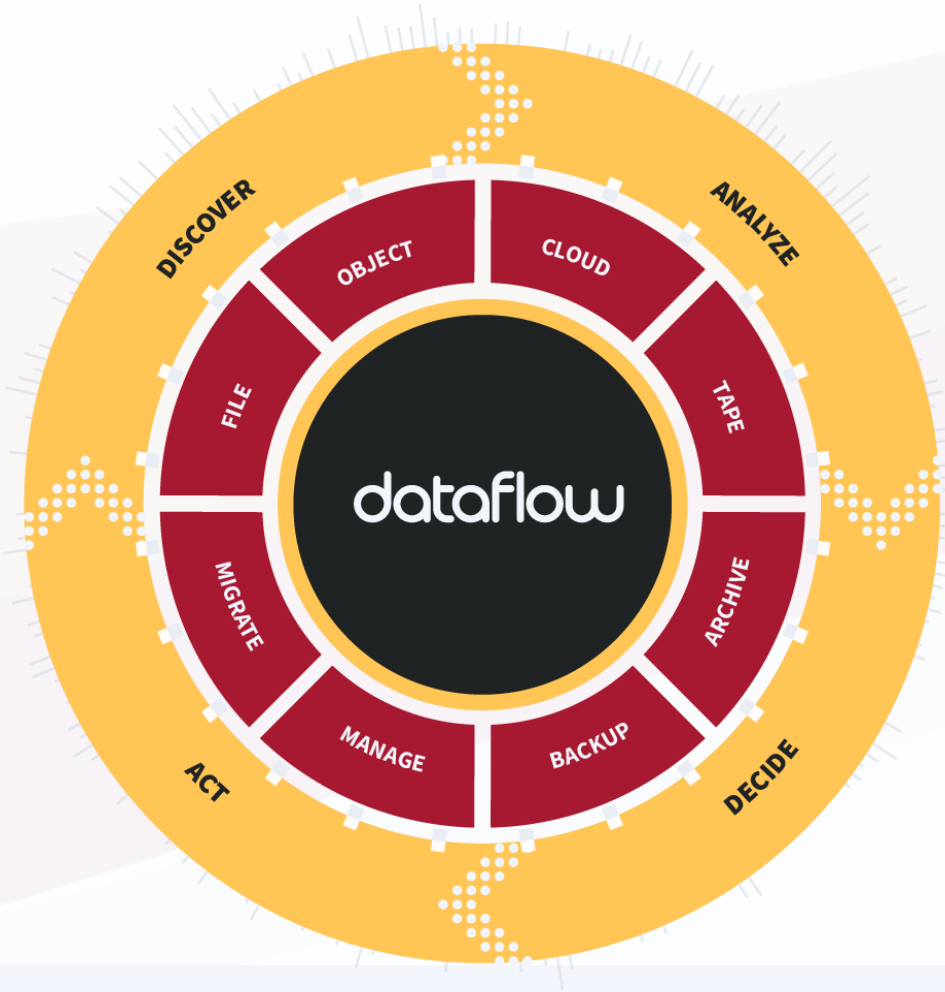
Lustre architecture

- Metadata servers
- Lustre change log
  - Track all changes occurring the FS
- Fast scan capabilities
- Snapshot



# DataFlow Analytics

# DataFlow: from datamover to data lifecycle management




- **Continuous analysis** of the entire storage infrastructure
- **Visualize, search and organize** all data through a centralized and intuitive analytics platform.
- **Automate** data manage tasks and workflows
- **Optimize data placement** for best efficiency
- Make **informed decisions** with historical trending
- Audit and enforce data governance policies

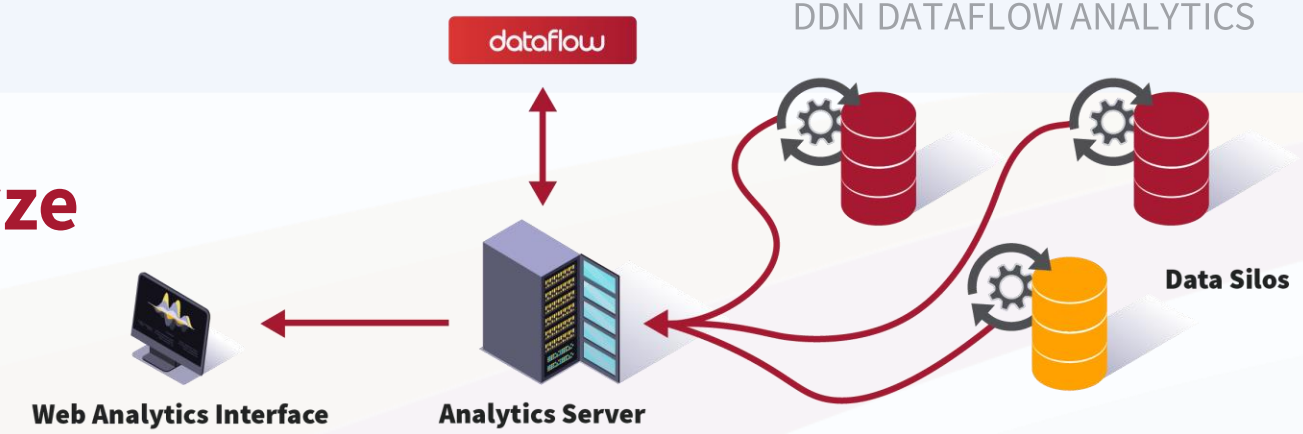
# Data Management

**Discover, Analyze, Search and Act on your Data**

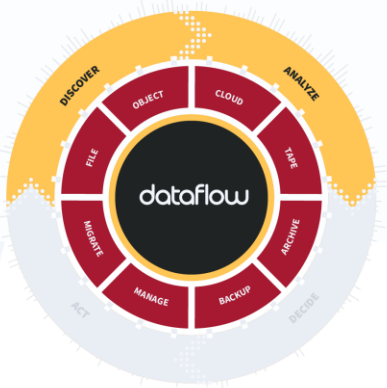
Catalog Server 1

- Build **central metadata repository** including customized tagged data
  - Structuration around a centralized data catalog
- **Automated, continuous and incremental** data crawling
- Available for **al common file  storage systems**: Lustre, NFS..
- Achieve **optimal data movement** across multiple data silos

# Analytics -Discover and Analyze

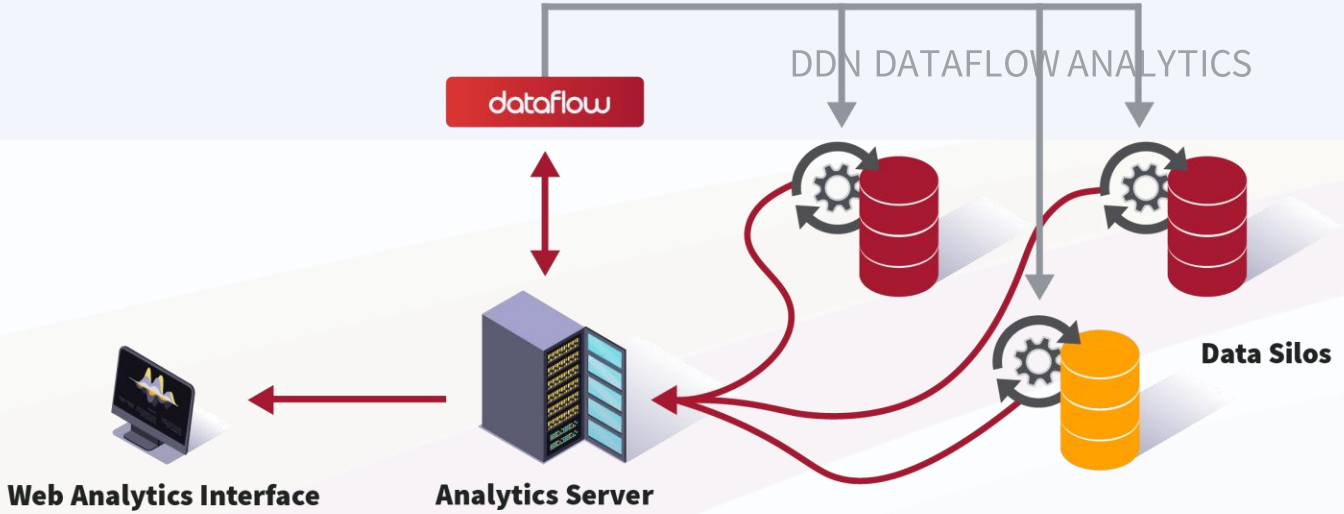



- Discover, visualize and search data based on Elasticsearch
- Discover Agent :
  - online and offline crawling modes available
- Data collected
  - **Objects**: Directories, files, symlinks, hardlinks
  - **Data**: Relative and last known full path, Btime, Mtime, Ctime , Atime, size, owner and inode
- Rest APIs
  - For Search and Analytics
  - For Administration

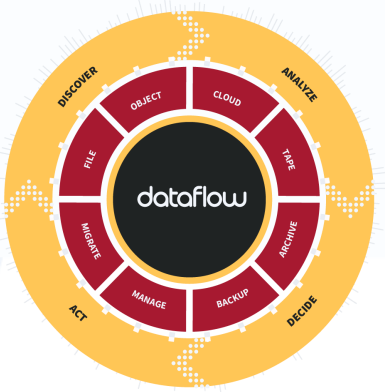


ddn dataflow

# Analytics - Decide and Act

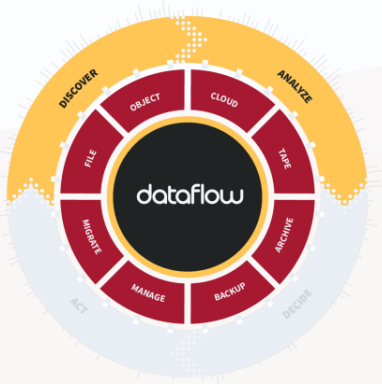



- Analytic Discovery Agents (crawler containerized)
- Data movement based on collected data
  - Task definition based on multiple metrics: *file and directory name format, file size, access time, modification time*…
- Add Cloud Storage Support
  - S3, GCP, Azure, Swift



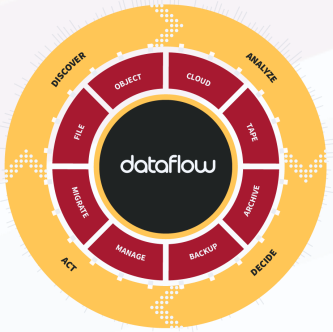
# Analytics

## Discover and Analyze




- Analytics engine to discover, visualize and search Data
- Standalone implementation
- Online and offline crawling modes available
- Support Disk, NAS and PFSs

## Decide and Act

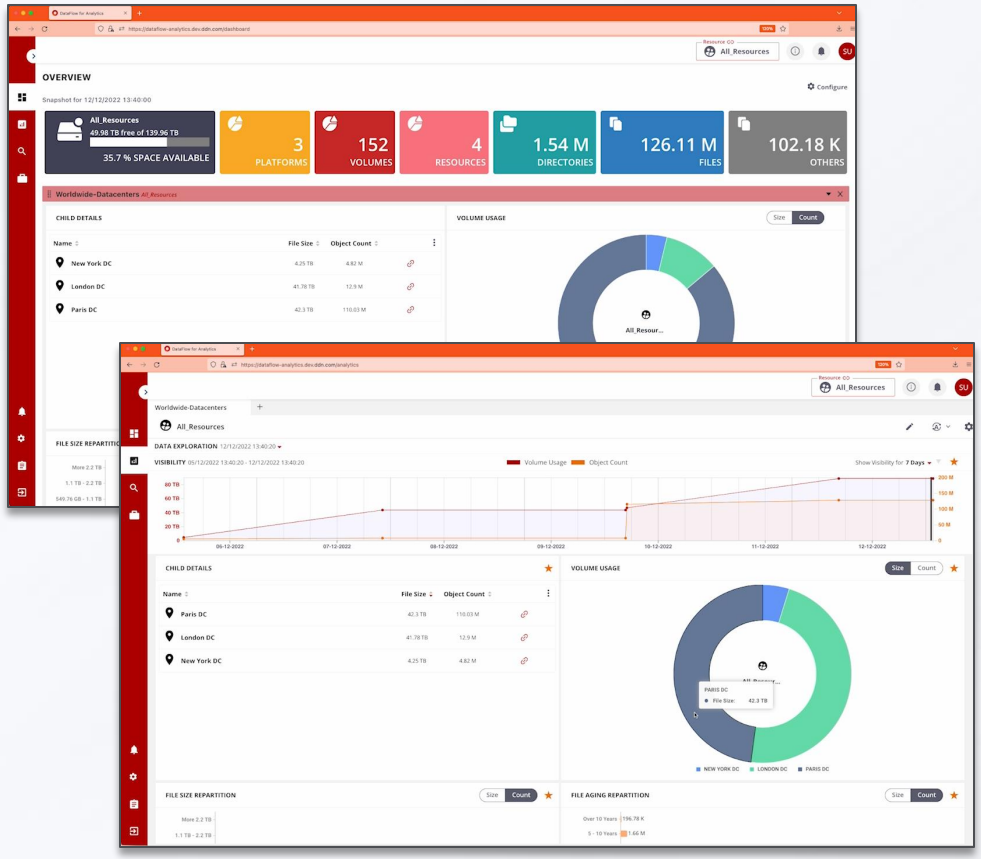



- Fully integrated into DataFlow Management Software
- Cloud Storage Support
  - S3, GCP, Azure, Swift

# Analytics – Visualize and Optimize

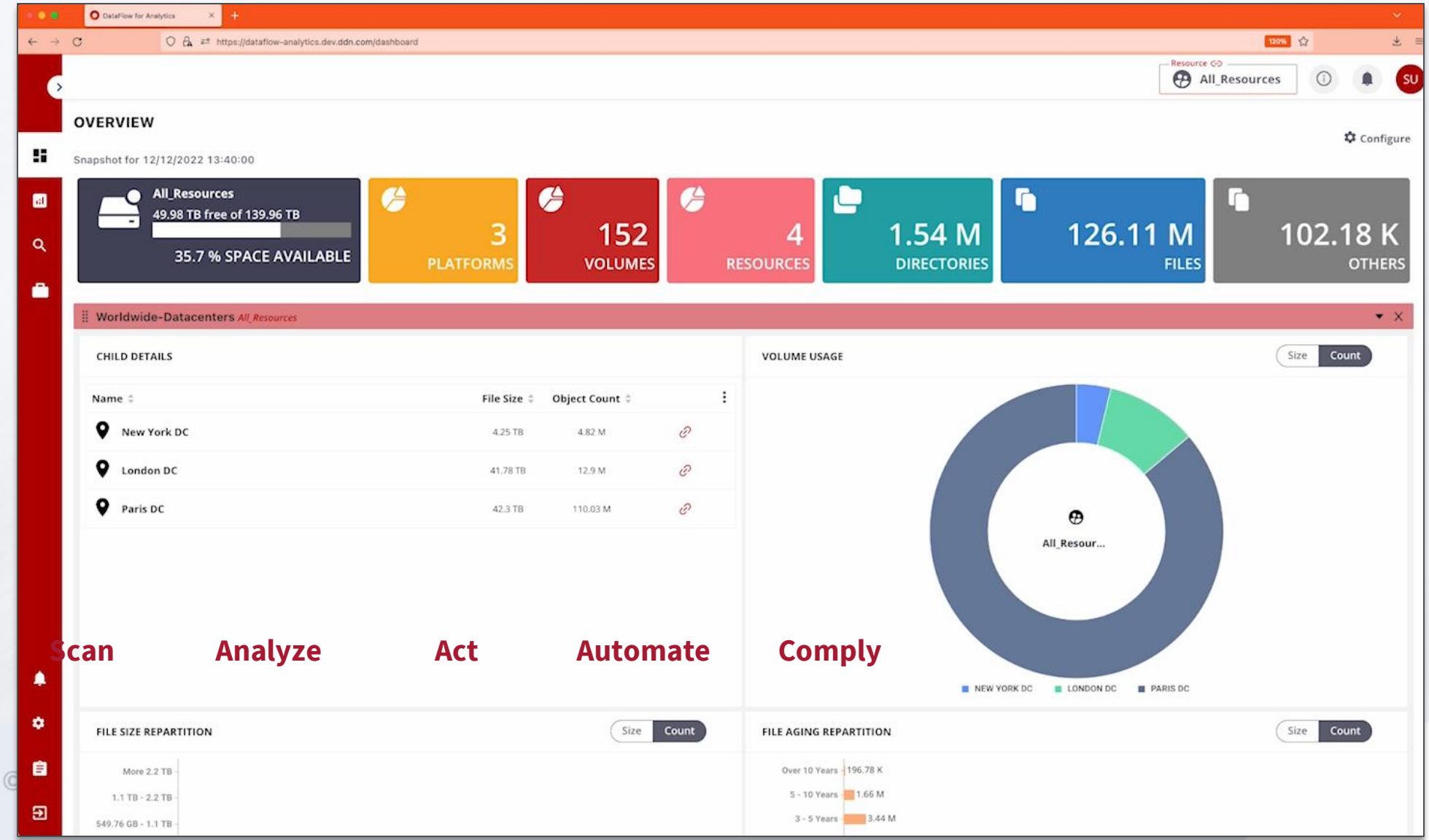
## Single Pane of Glass Aggregates all your Data Silos

- Consolidated view of Data Distribution across Storage Platforms
- Time-based navigation to highlight trends and forecast capacity growth
- Go beyond built-in reports with customizable, contextual search engine



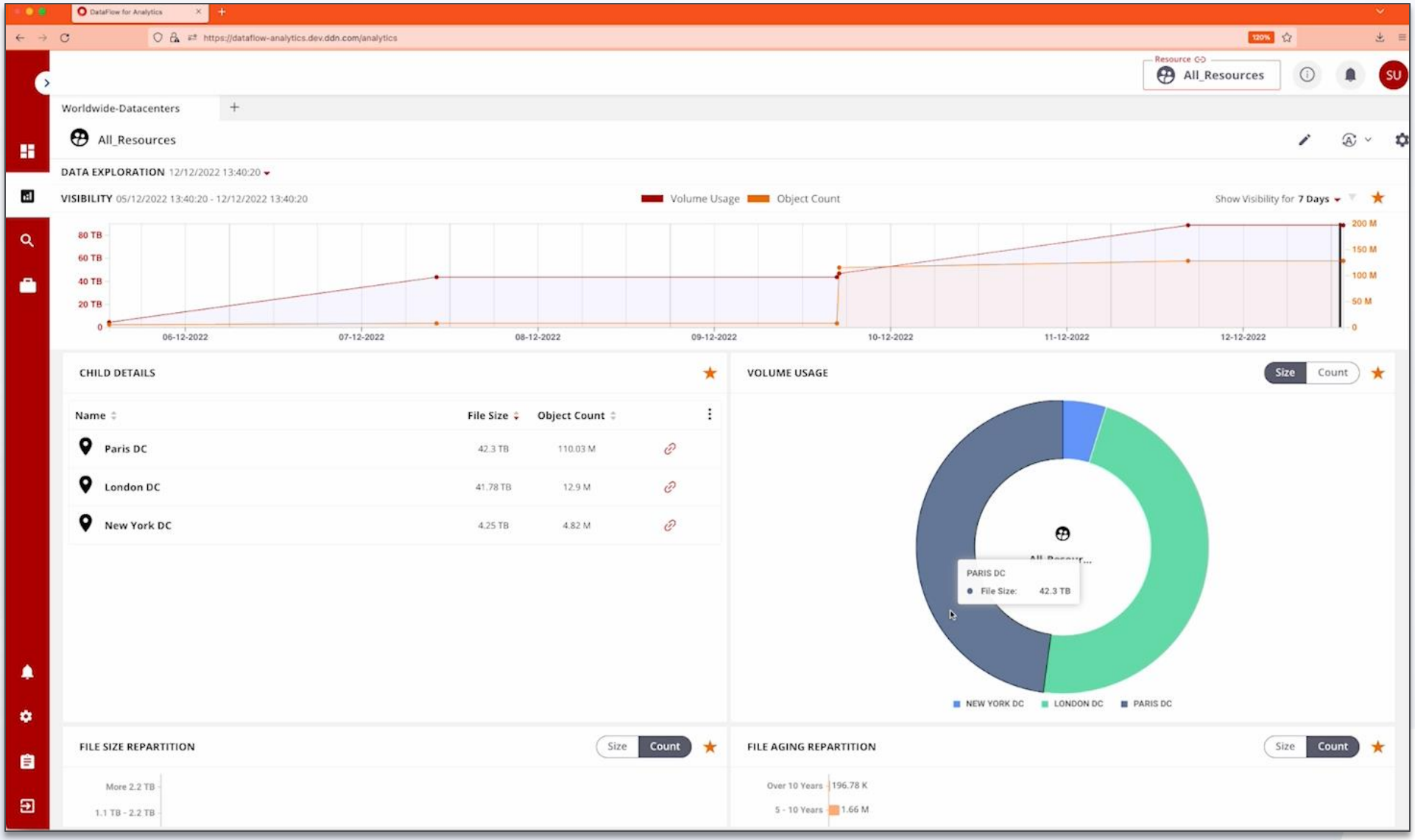

**Scan** **Analyze** **Act** **Automate** **Comply**

# Analytics – Data Analysis view




Scan Analyze Act Automate Comply

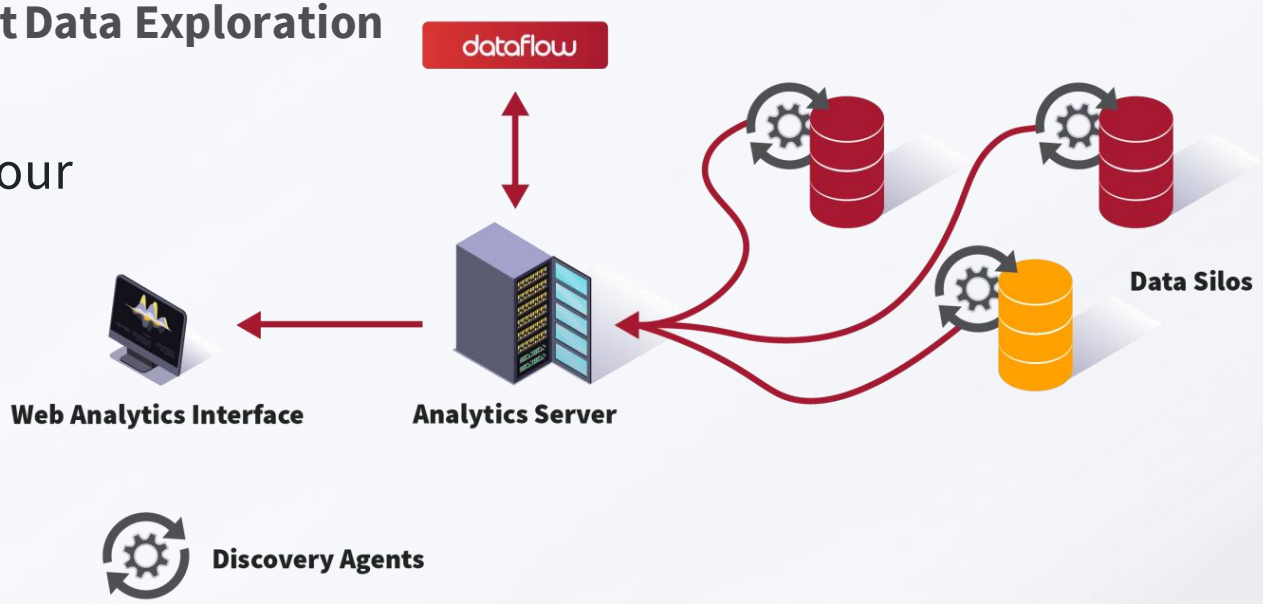
# Analytics – Data Analysis view



# Metadata Analytics – Architecture

## Scan Local or Remote Filesystems in parallel for Fast Data Exploration

- Analytics engine to discover, visualize and search your data based on Elasticsearch
- Filesystem scan agent runs locally or remotely, facilitating offline collection from remote storage
- Snapshot of filesystem metadata is stored in local database, providing an optimised metadata query engine for fast filesystem analytics
- Custom Reports can be downloaded in CSV or queried via the integrated Rest API



# Toward more comprehensive solutions

## Extension of large file system for a better metadata management

- Analytics and decision based on user-defined metadata
- collection from remote storage: remote data catalog
- Snapshot of filesystem metadata is stored in local database, providing an optimised metadata query engine for fast filesystem analytics
- Integration with community developed tools (e.g. RUCIO)


ddn