

OpenIO

**Next-Gen Object Storage
and Serverless Computing**

Agenda

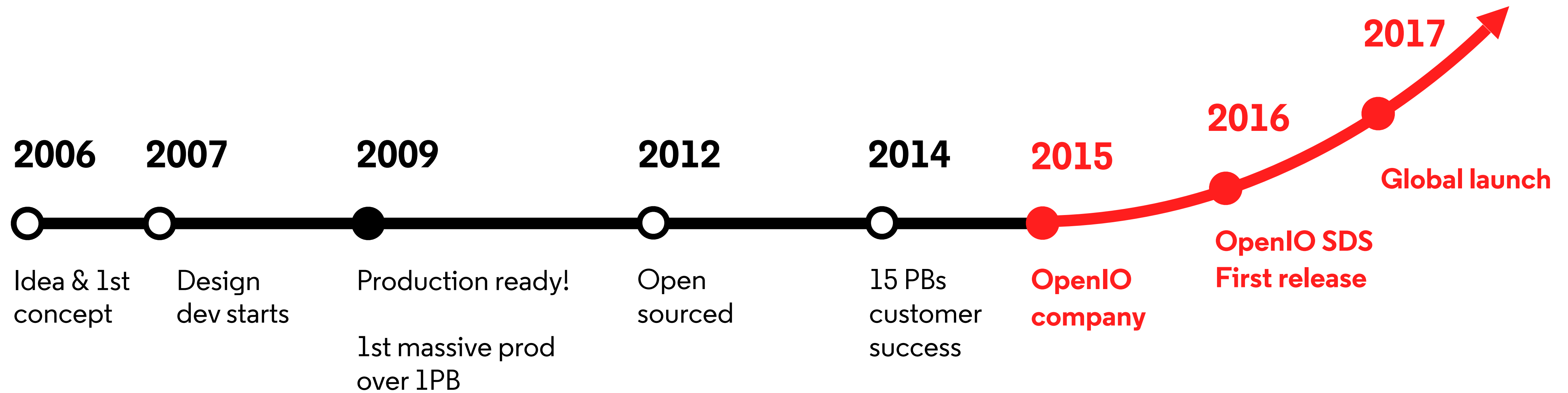
1 About OpenIO

2 SDS: Next-gen Object storage

3 Grid for Apps: Serverless Computing

About OpenIO

An experienced team, a robust and mature technology



Lille (FR) | San Francisco | Tokyo

Quickly Growing Across Geographies And Vertical Markets

35

Employees

Mostly engineers,
support and pre-sales
Growing fast

3

Continents

Hem (Lille, France), Paris,
Tokyo, San Francisco

Teams across EMEA,
Japan and, soon, US

25+

Customers

Installations ranging
from 3 nodes up to 60
Petabytes and billions
of objects

2

Solutions

OpenIO SDS, next-gen
object storage

Grid for Apps, serverless
computing framework

Customers



Large Telco Provider

Email storage



**Small objects,
high scalability**

65 Mln

mail boxes

15 PB

of storage

650

nodes

10 Bln

objects

20K

services online

DailyMotion

Media & Entertainment



**High throughput
and capacity,
fat x86 nodes**

60 PB

of OpenIO SDS

80 Mln

videos

30%

growing/year

3 Bln

views per month

Japanese ISP

High Speed Object Storage



**High number
of transactions
on SSDs**

6000

Emails per second

10+10

All-flash nodes

2-sites

Async replication

Indexing

With Grid for Apps

Teezily E-commerce Website

On-premises S3, migrated from Amazon AWS



**Private Cloud Storage
on Public Infrastructure,
Cost effective**

350TB

Very small files

10Bln

Objects

€ 400K

Money saved (1st year)

10x

Faster than S3 in most operations

Use Cases

Object Storage

- Email platforms
- Media & Entertainment
- Private Cloud Storage
- Remote Storage Consolidation
- Backup & Archiving
- Big Data

SDS: Next Generation Object Storage

Next Gen Applications

- Integrated Data Processing
- Industrial IOT
- Machine learning

Grid for Apps: Event-driven Compute Framework

SDS

Next-Generation Object Storage

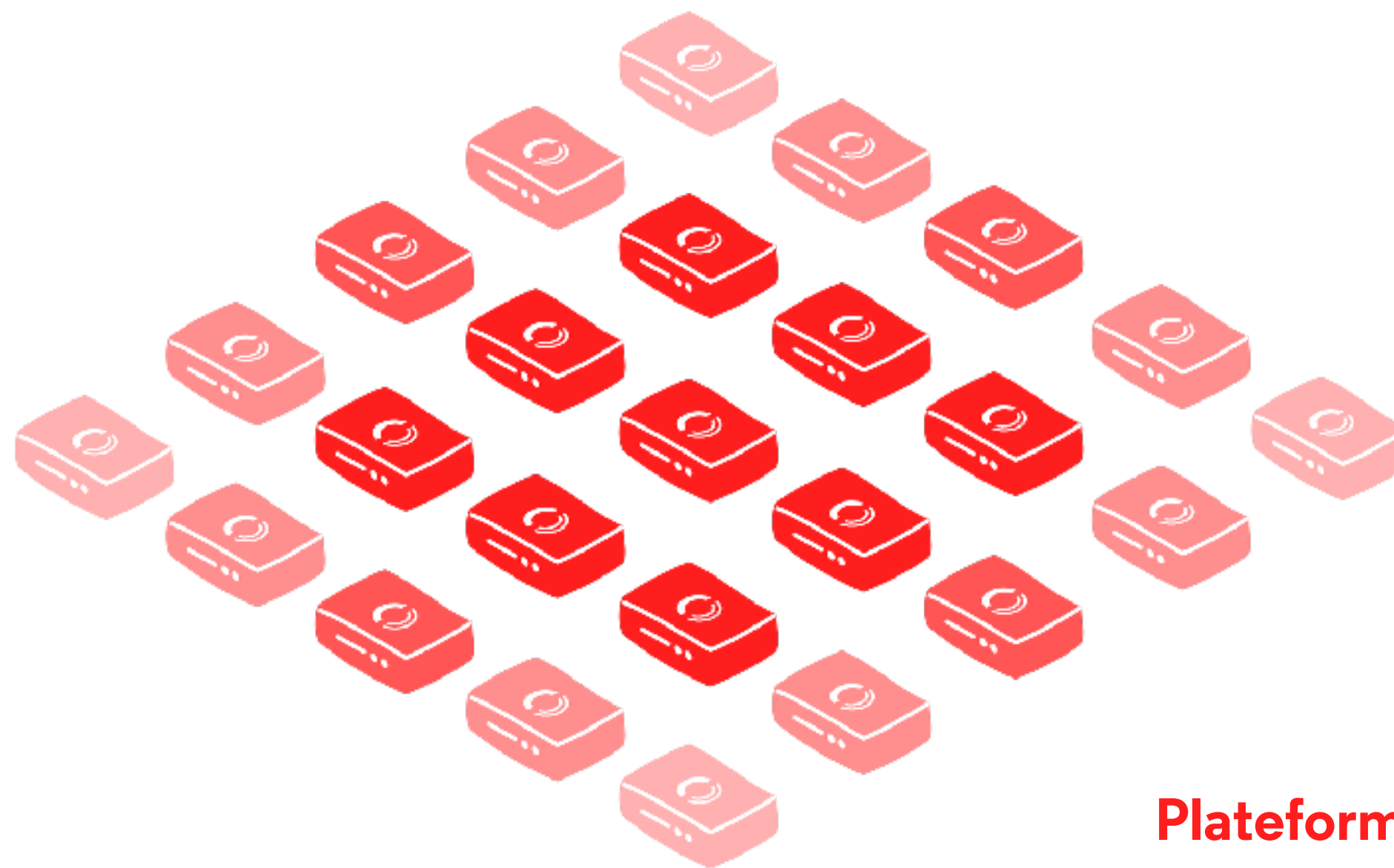
OpenIO SDS



+



=



Matériel Standard

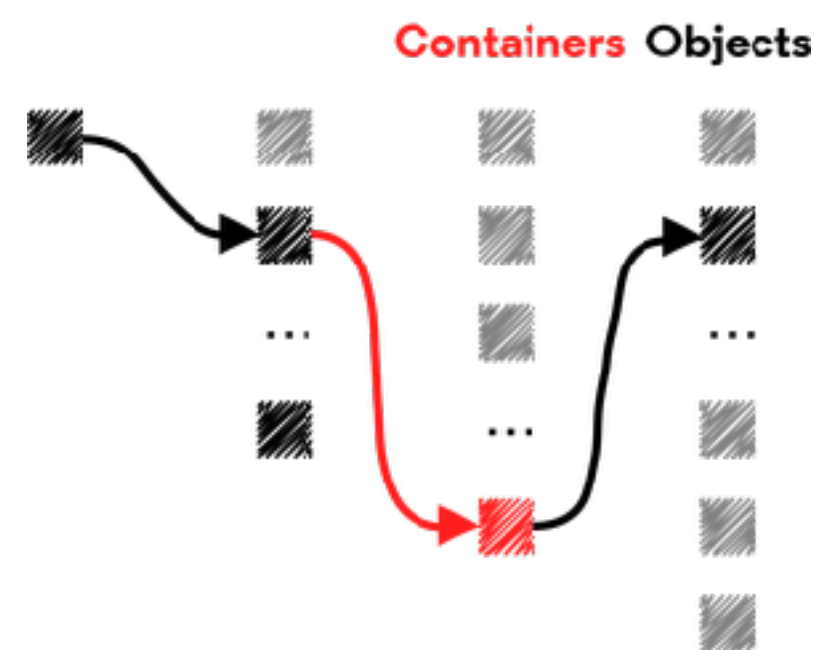
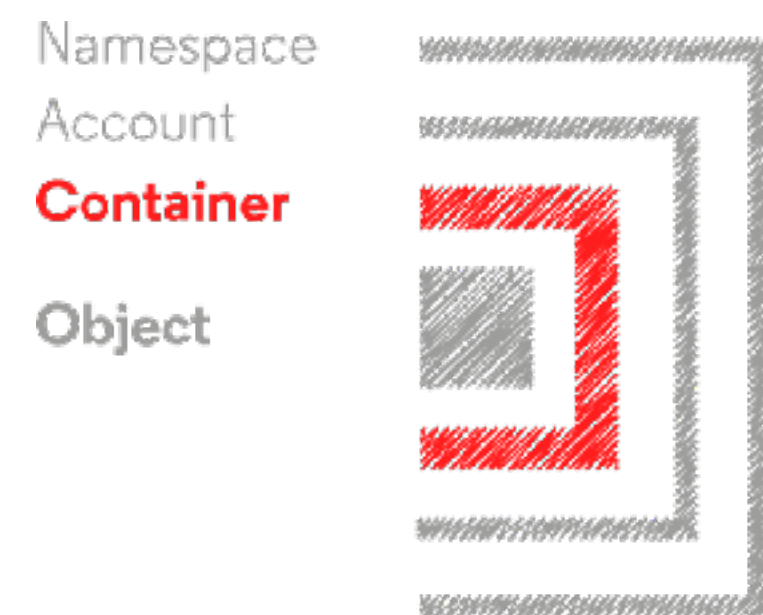
OpenIO SDS

**Plateforme de stockage
massivement scalable**

We are different...



Directory with indirections



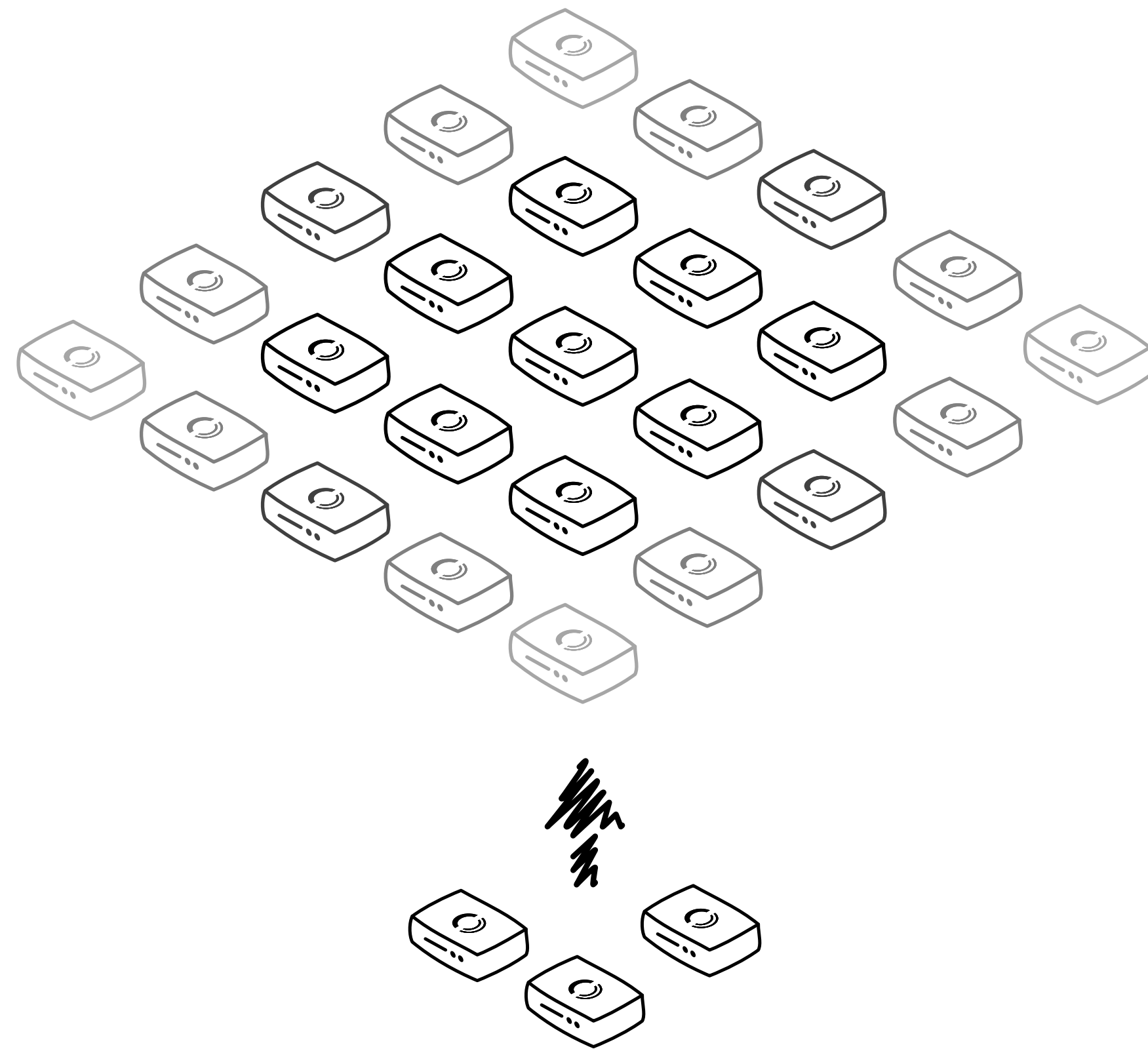
Track containers not objects

- Container and objects are stored in a persistent 3-level distributed directory
- High performance consistency, with always 3 hops to get to the data
- The directory has the form of a hash table, mapping containers' UUIDs to their services

`grid://namespace/account/container/object`

Safe, predictable and consistent at any scale

Scale out storage



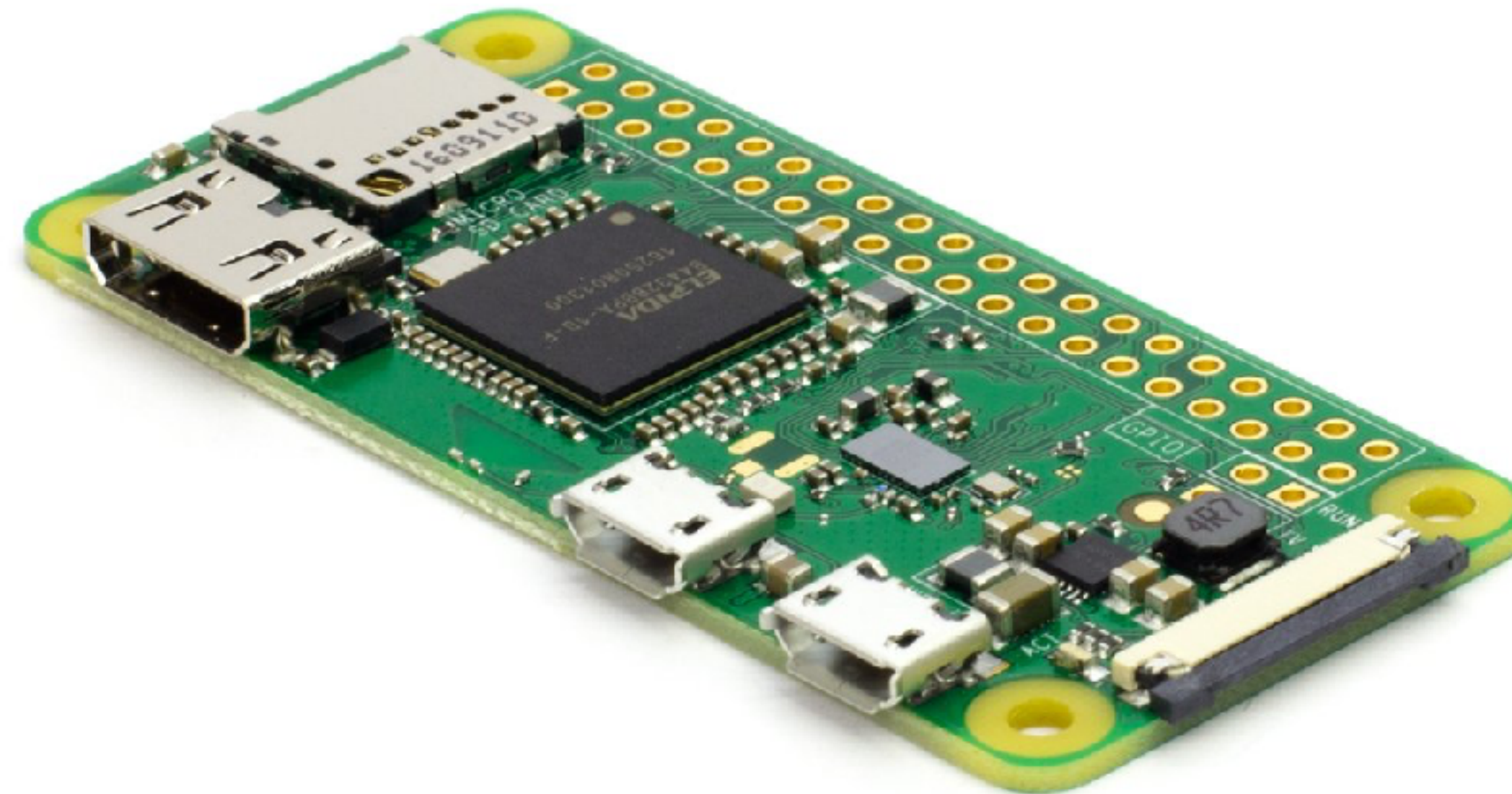
Hyperscalable storage

- **Scale-out by nature with shared-nothing model to aggregate storage capacity from independent x86 or ARM servers**
- **Limitless storage based on open source object storage technology**
- **Store thousands of PBs of data and billions of objects**
- **From 3 nodes to thousands**

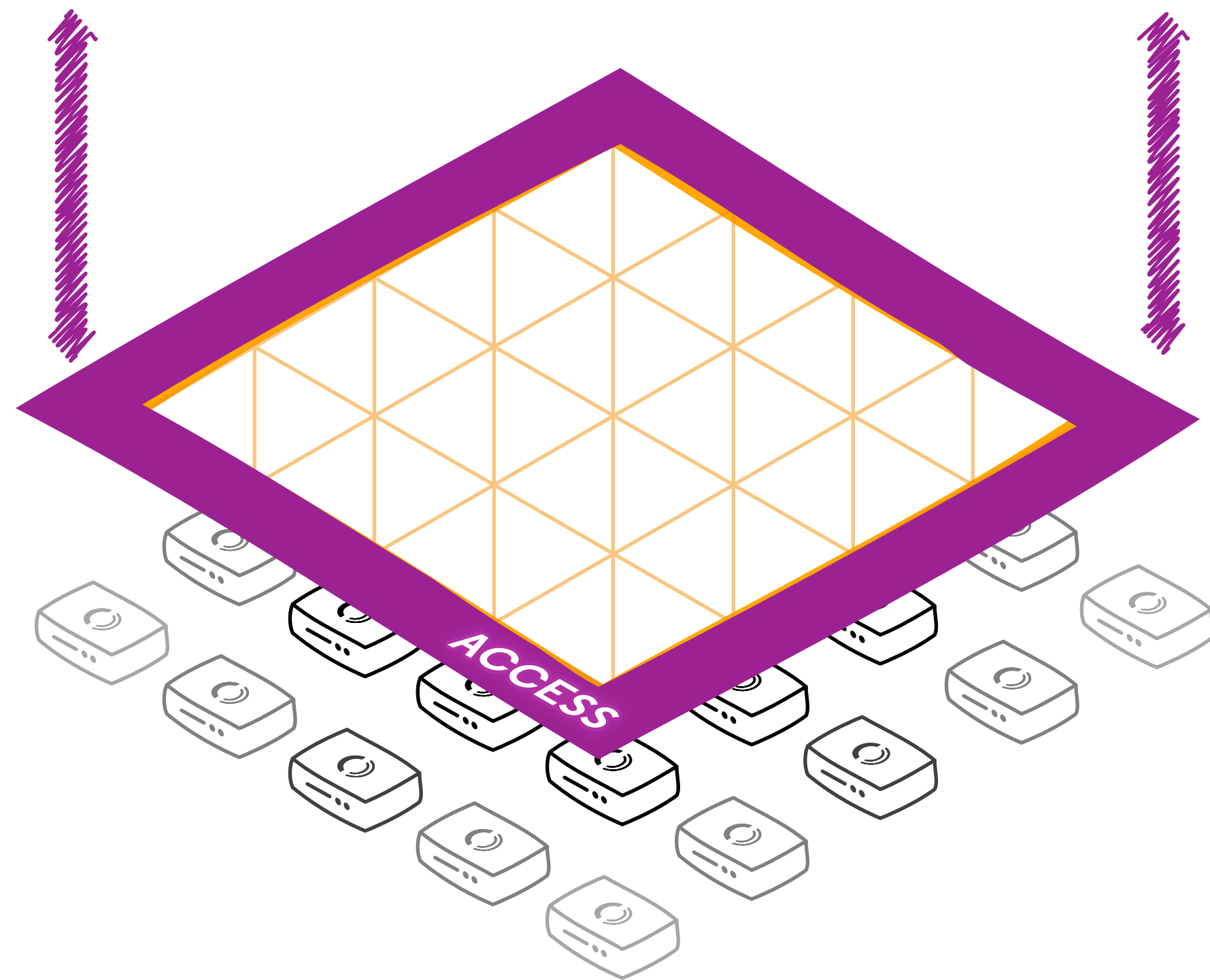
Hardware agnostic, scalable and resilient

OpenIO SDS Can Run in a \$5 Computer!

400MB RAM, 1 ARM CPU core resource footprint. Highly optimized C code



Broad access layer



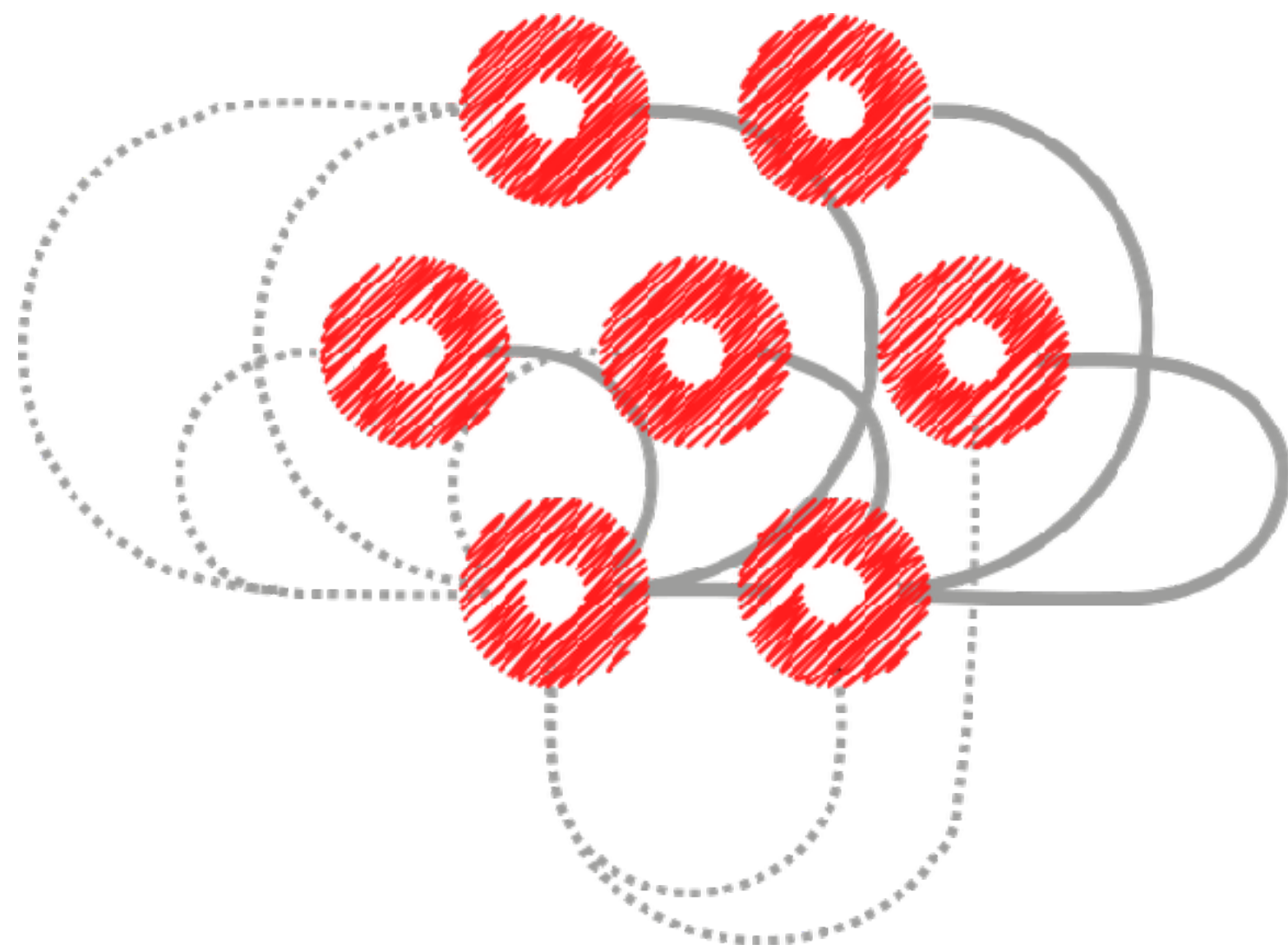
Standard APIs and file protocols

- **Standard Object APIs to leverage natively the platform: OpenIO REST/HTTP, Amazon S3 and OpenStack Swift**
- **Industry File-Sharing Protocols: NFS, SMB and FTP**

Data can be accessed by modern and legacy applications

Dynamic load balancing

Conscience technology



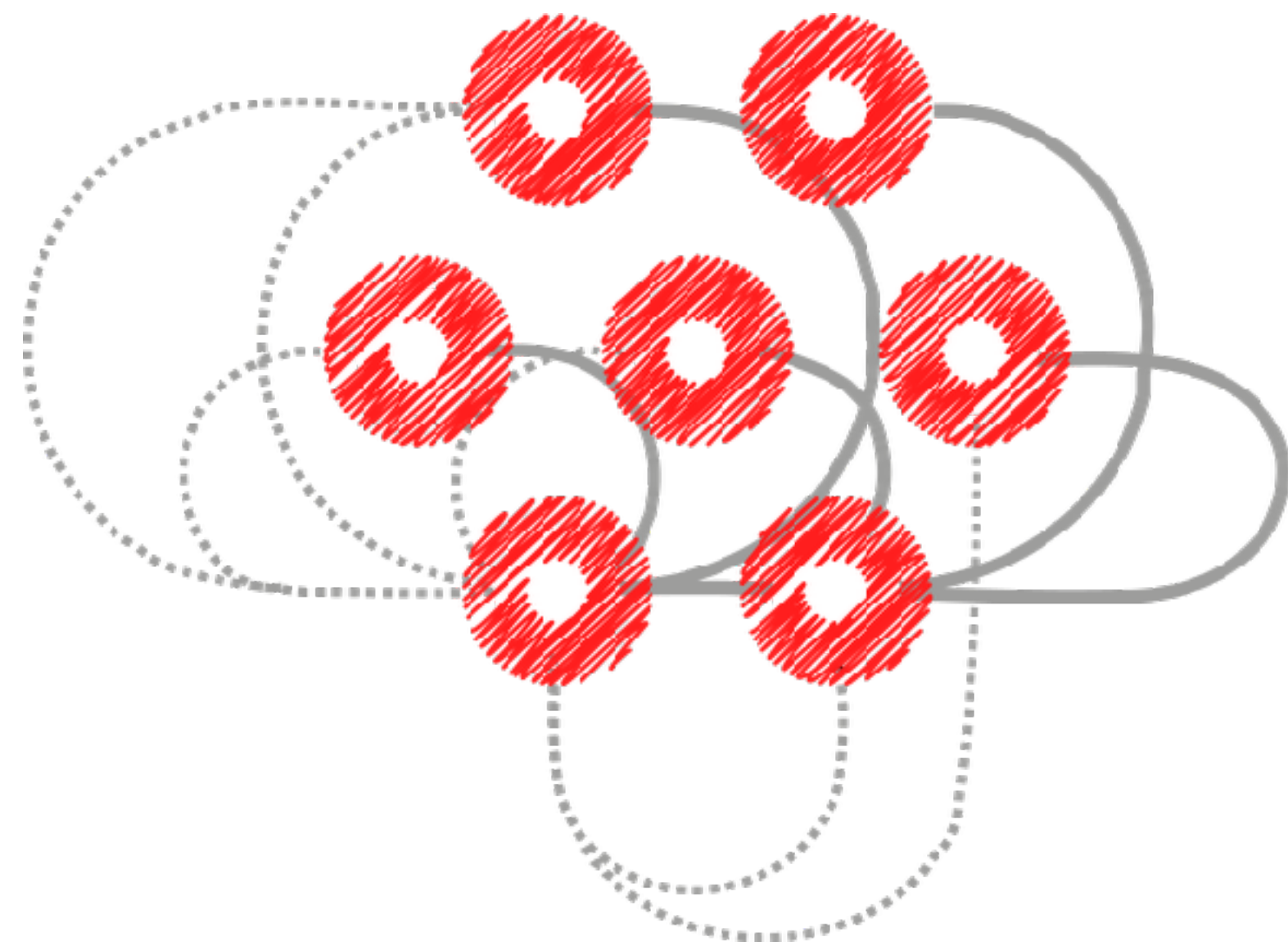
Real time load balancing for optimal data placement

- Collects systems metrics from the services of each node
- Computes a quality score for each service
- Distributes scores to every nodes and clients
- On the fly best match making for each request

Real time load balancing for optimal data placement

The score is computed with a configurable formula

Conscience technology

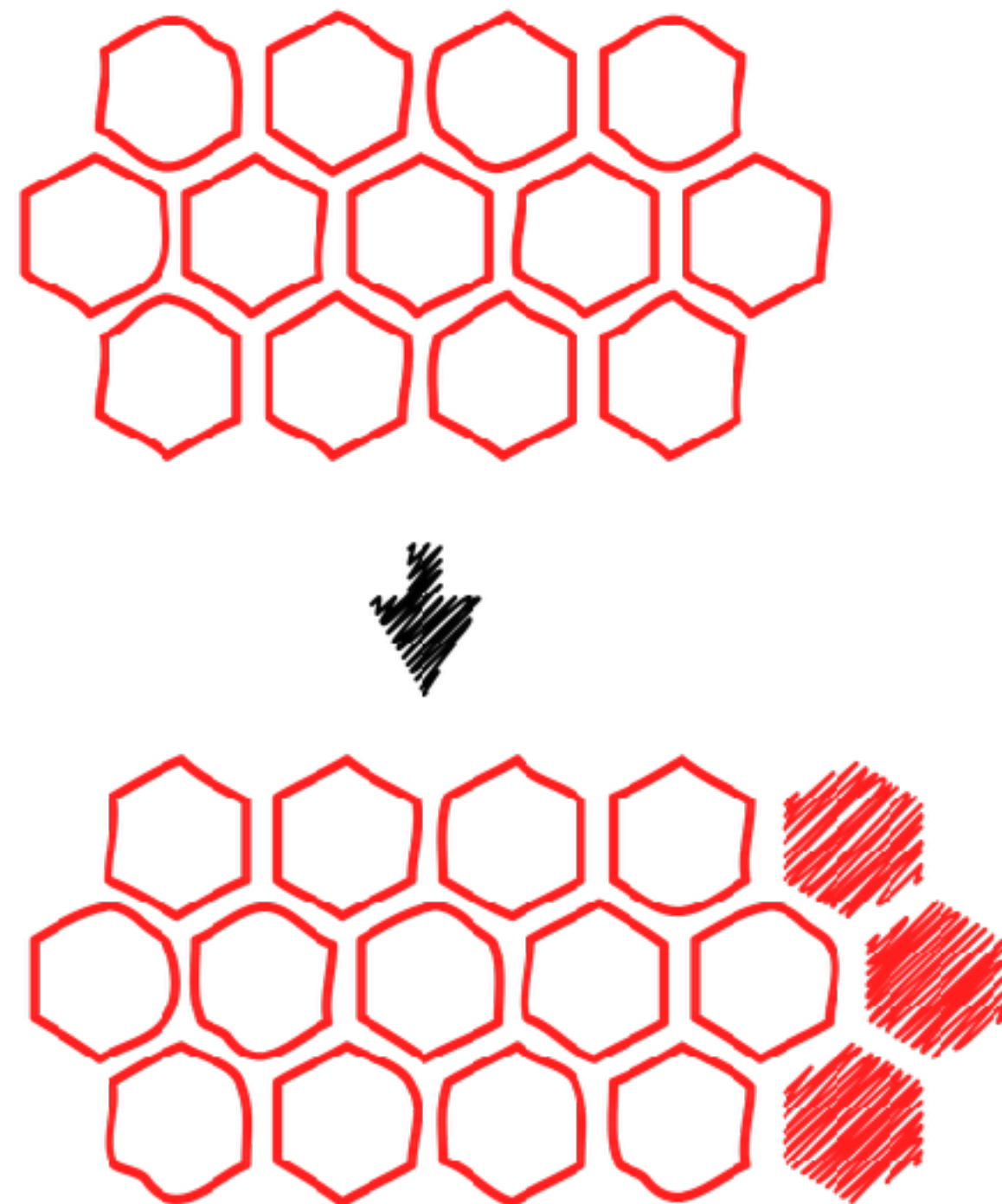


What is it good for?

- **No hot spots**
- **Scales without rebalancing data**
- **Heterogeneous hardware supported by software**
- **Storage tiering and QoS implemented at the core**

Simplified operations and capacity planning

Grid of nodes

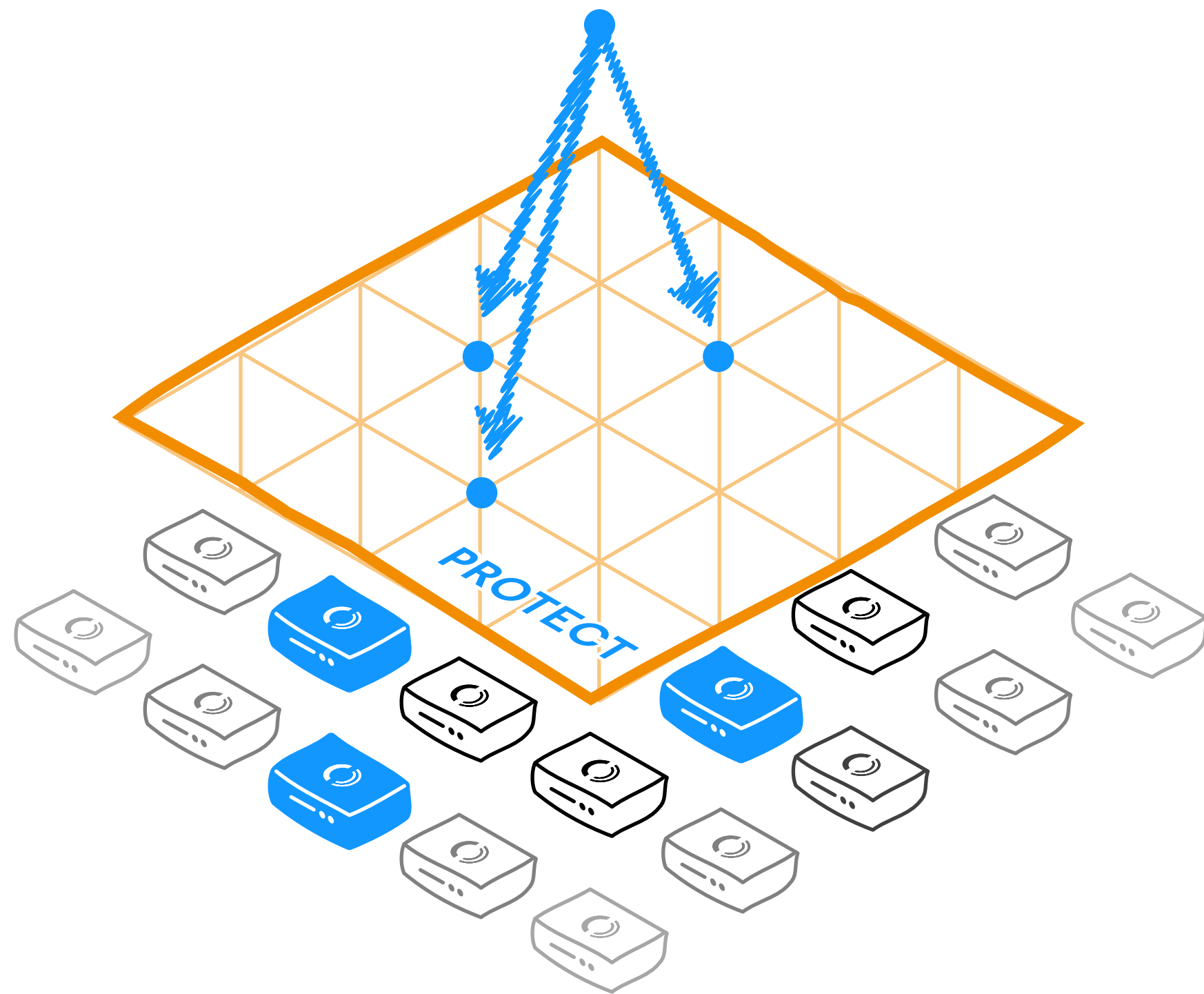


Never need to rebalance

- No consistent hashing algorithm: no recalculation of the key space
- New nodes and resources are automatically discovered and immediately available
- Nodes can be heterogenous

Seamless cluster expansion without performance impact

No compromise data resiliency

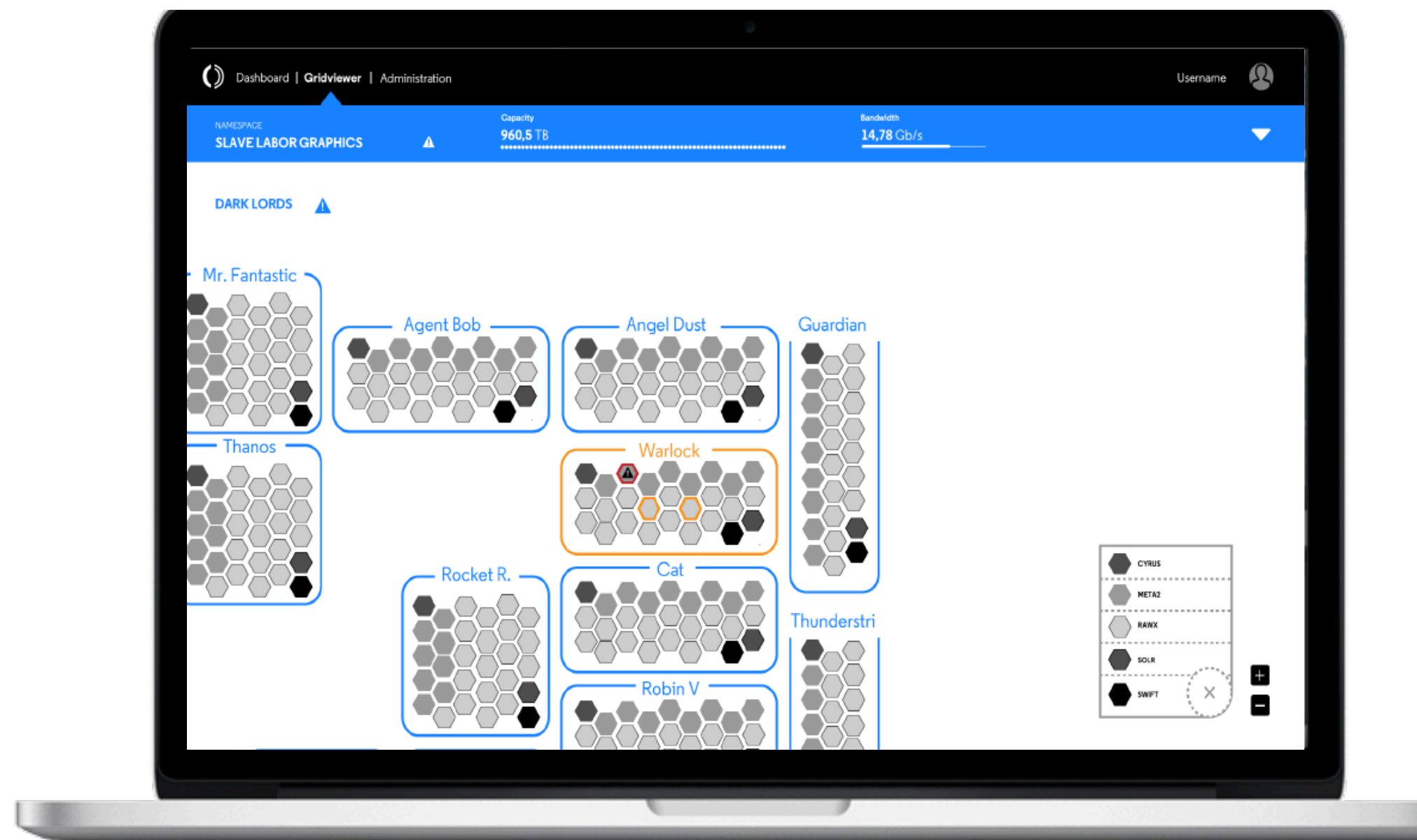


Multiple protection schemes

- N data copies
- Erasure coding based on Reed-Solomon
- Dynamic data protection policies
- Various topologies from 1 Data Center to multiple or stretched cluster across geos
- Synchronous and/or asynchronous replication
- Storage tiering
- Data encryption

Efficient data protection for any workload

Ease of use



Full operational control

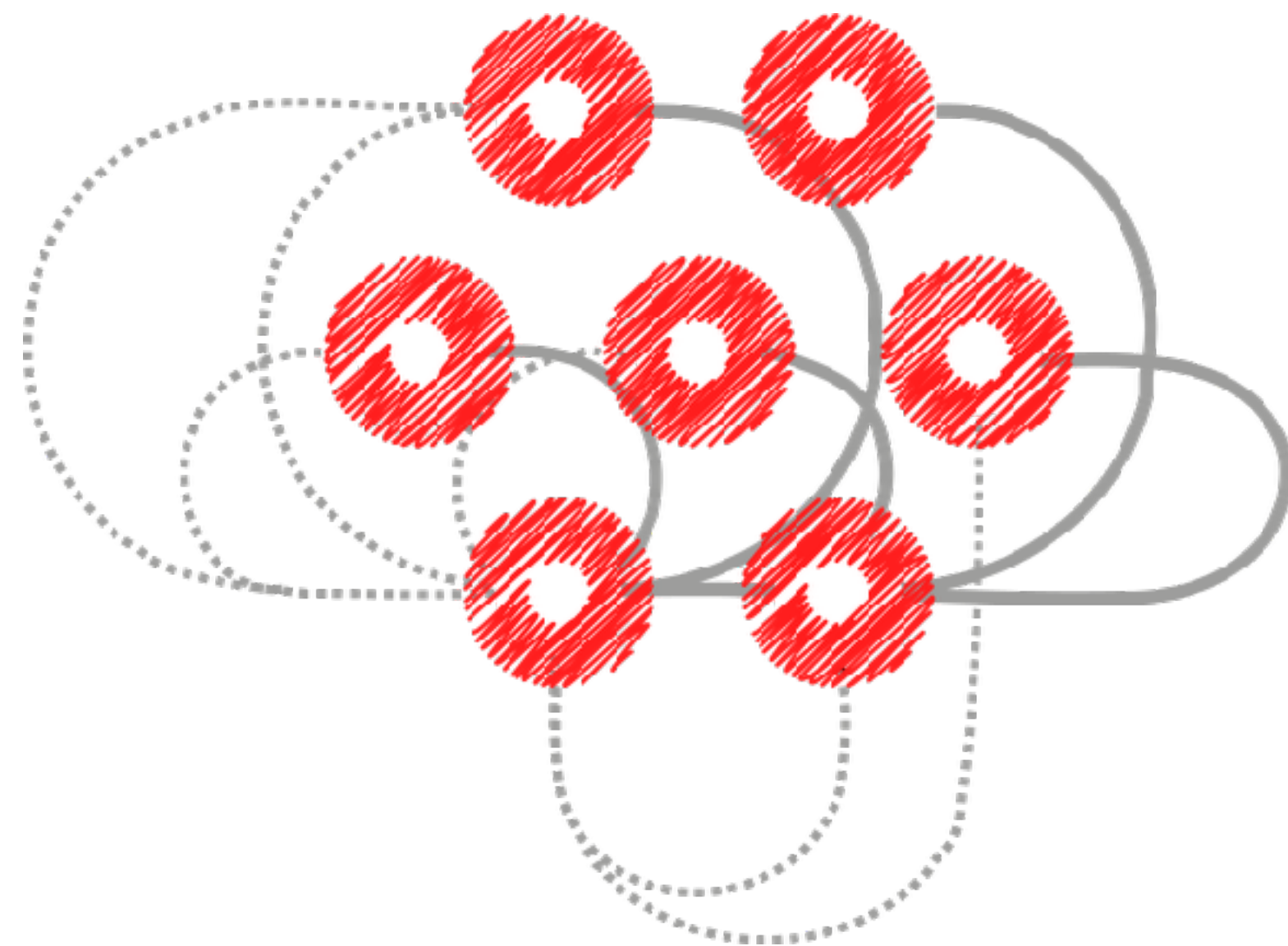
- Consistent and simple Command Line Interface
- Web User Interface for management, monitoring and reporting
- Chargeback API for billing
- Ansible and Puppet scripts for massive node deployments

Simplified deployment and management

Grid for Apps

Serverless Computing Framework

Advanced scheduling

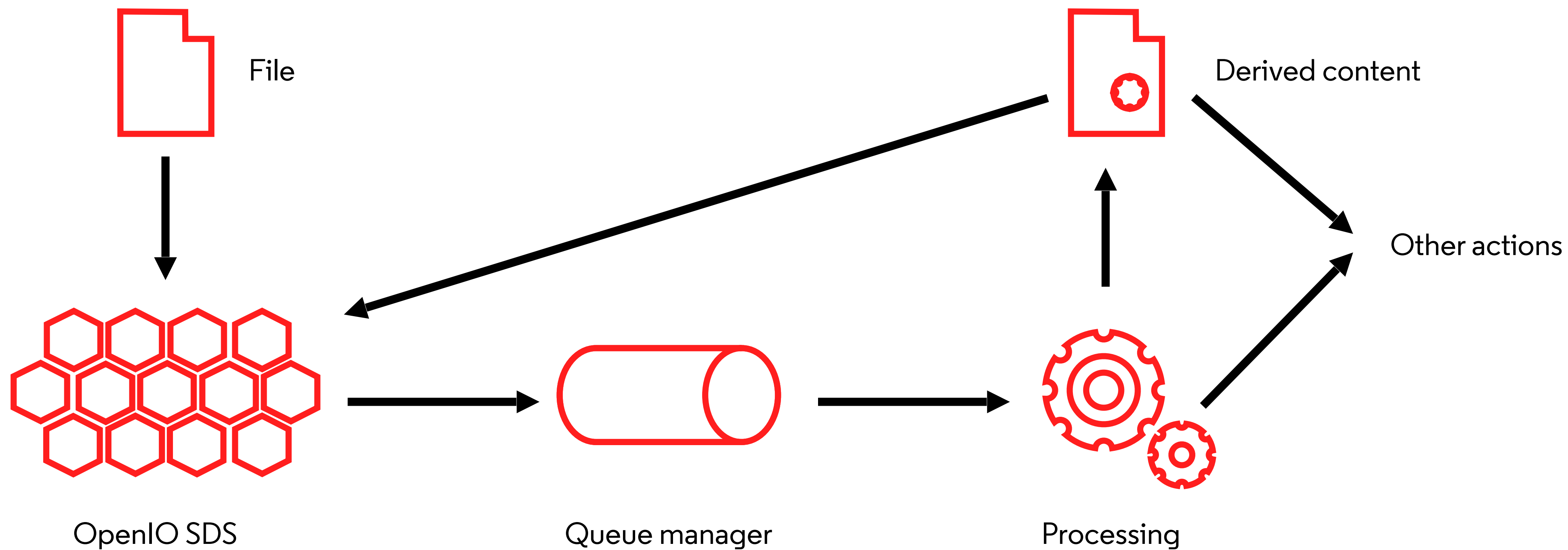


Conscience based

- **Tasks, Functions and Jobs are allocated on most available nodes**
- **Nodes with specific characteristics (i.e. GPU) can be tagged and selected for specific workloads**
- **All the resources are continuously monitored and jobs/task can be re-allocated if a node/process fails**
- **Container isolation provides resource allocation and security**

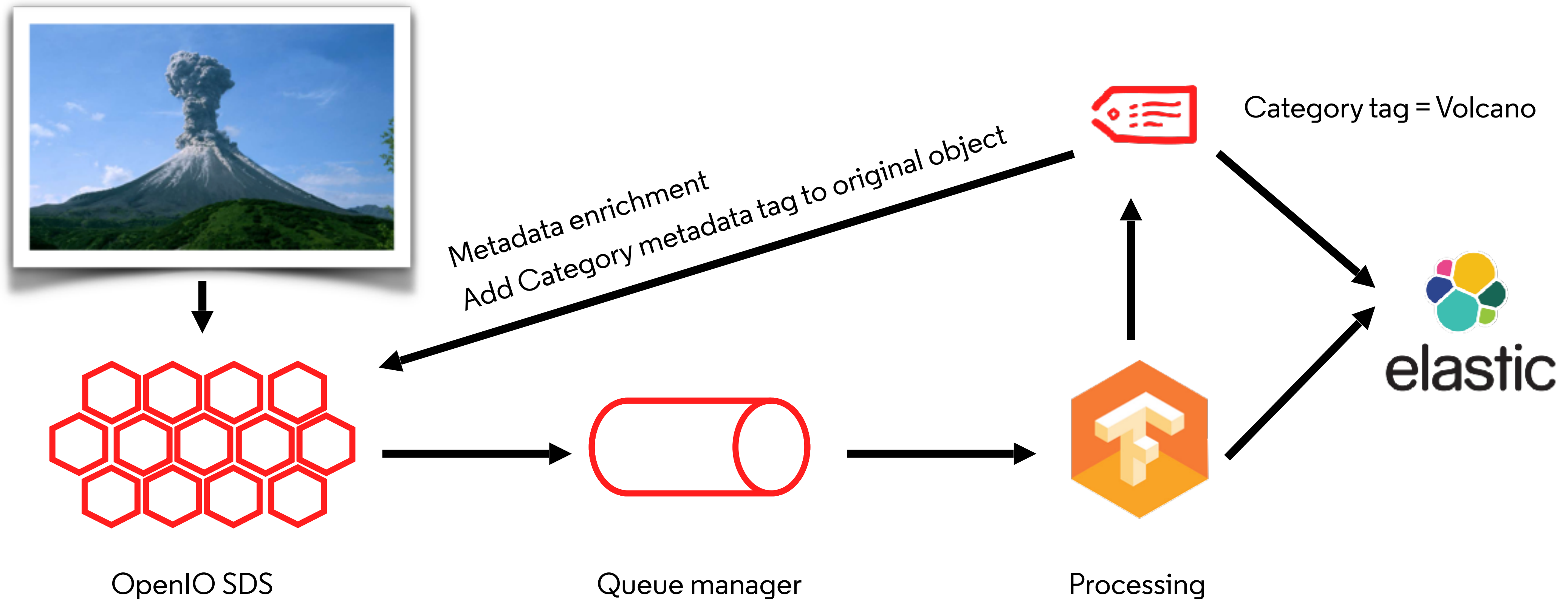
A complete solutions, for every need

Grid for Apps in practice



Grid for Apps in practice

Everything happens into SDS (no external resources)



ANF

Resources

Enjoy OpenIO

Links

- <http://docs.openio.io>
- <http://slack.openio.io>
- guillaume@openio.io

Projects

- Institut du Cerveau et de la Moelle Epinière
- Datawan

ICM

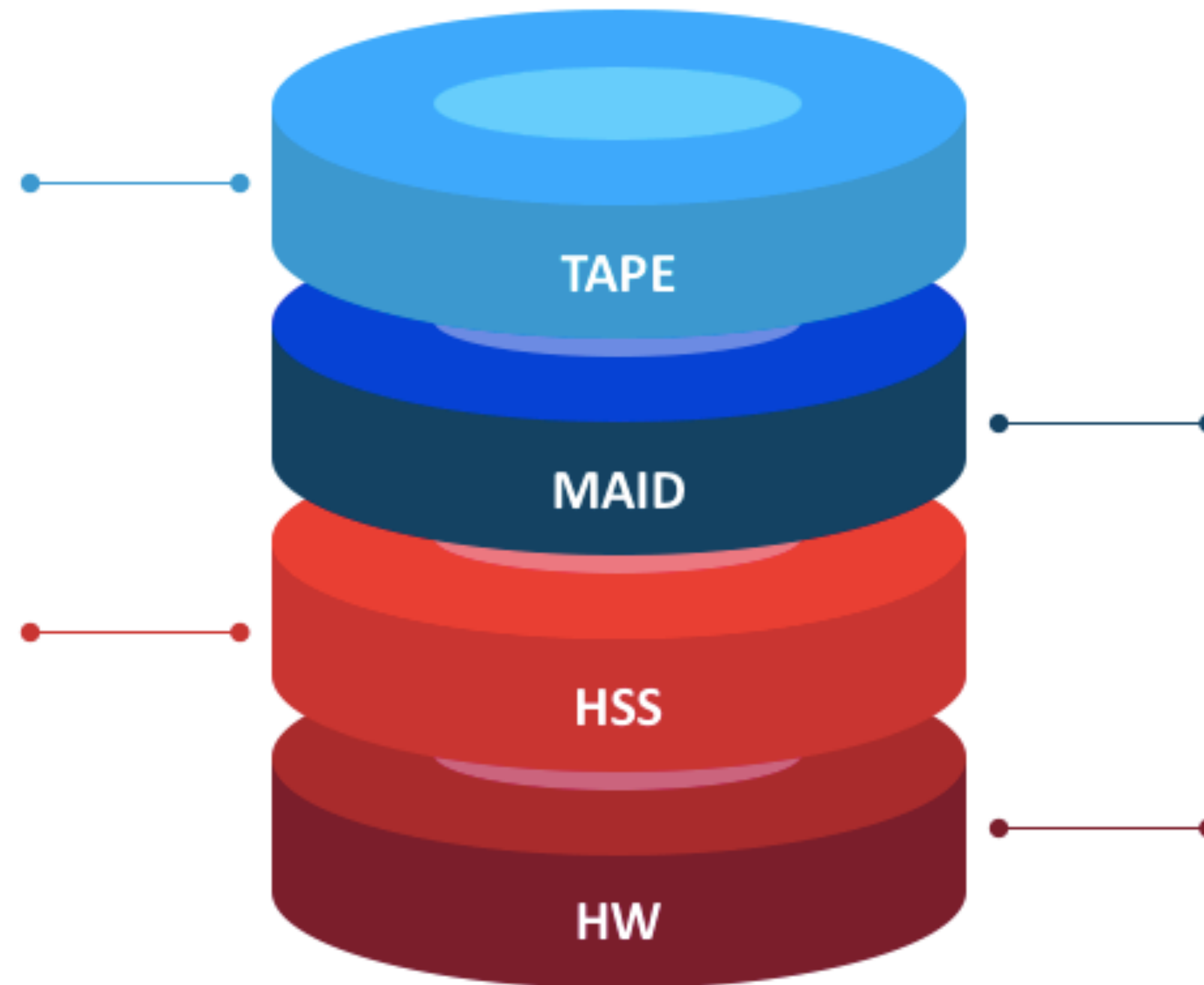
Storage

Tape

Long term storage. HIPAA retention requirements: 10 years.
LTO-7 (6PB total capacity) but cloud storage under study.

HSS

High Speed Storage, composed of *Spectrum Scale* (GPFS) and *Intel Lustre*. Absorb data traffic from dedicated hardware acquisition such MRI, microscope, ...
Connected with every satellite component around (compute node, databases, desktop, dedicated applications).



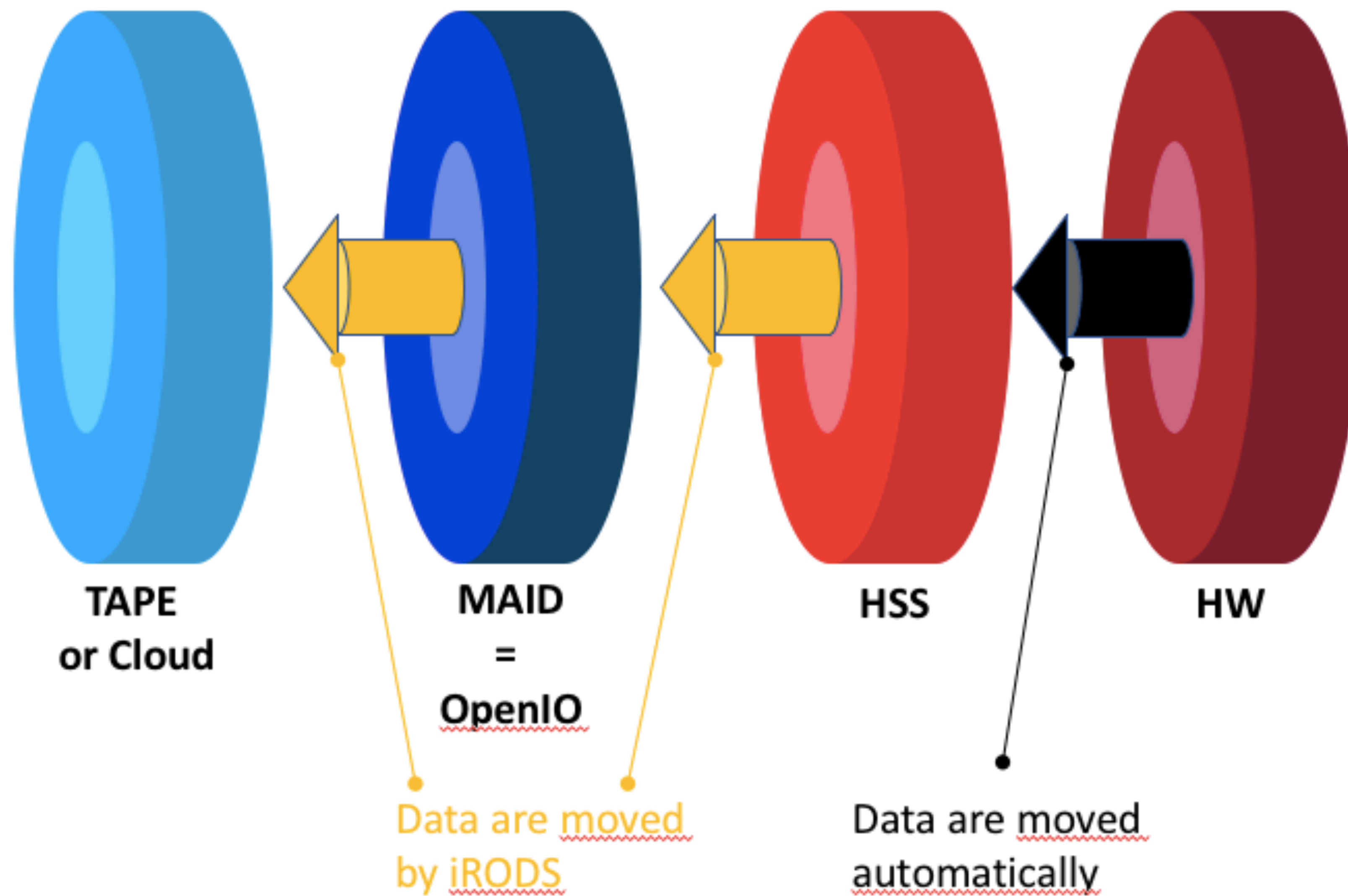
MAID

Massive Array of Idle Disk. Give access to large and low cost but slower storage. Mainly to backup or archive. RAW data are automatically moved to this storage and have only read-only access (immutable + RO).

HW

Dedicated acquisition hardware. Almost each equipment have flash storage (NVMe, PCIe or SATA). Data are automatically moved to the high speed storage.

ICM





github.com/open-io



[@OpenIO](https://twitter.com/OpenIO)



[OpenIO](https://www.linkedin.com/company/open-io)

Guillaume Delaporte
guillaume@openio.io

openio.io

