Application-Aware Storage

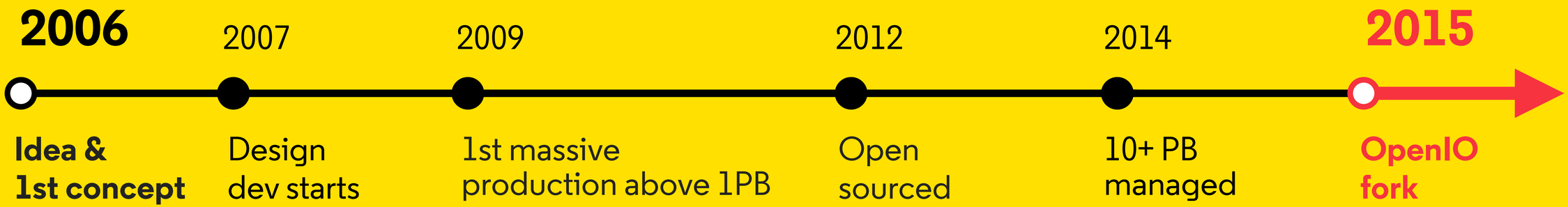
Agenda

1. Intro
2. OpenIO's missions
3. How it works
4. Demo
5. Next steps


OpenIO is a pure software object store.

About

Object storage pioneers



San Francisco | **Lille** (FR) | **Montreal** | **Tokyo**

Use cases

→ Email Platforms

→ Video Streaming

→ Enterprise Object and File storage

- Storage-as-a-Service
- Compute + Storage Platform
- On-premise Data

SFR

numericable

Scalair

RELEX

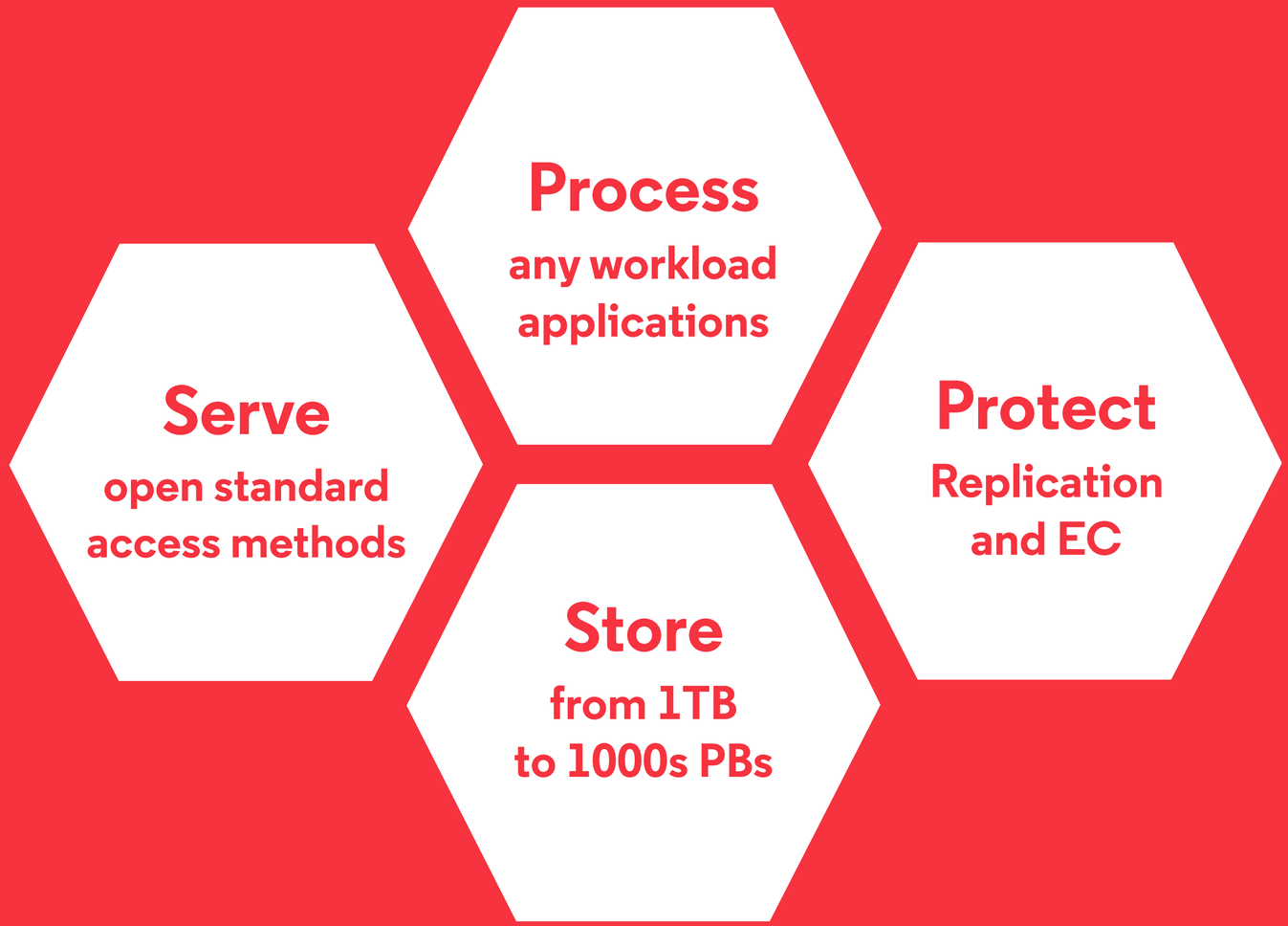
aqua ray

VadeSecure
Advanced Email Protection

IIJ
Internet Initiative Japan

DE LA RECHERCHE À L'INDUSTRIE
cea

OpenIO

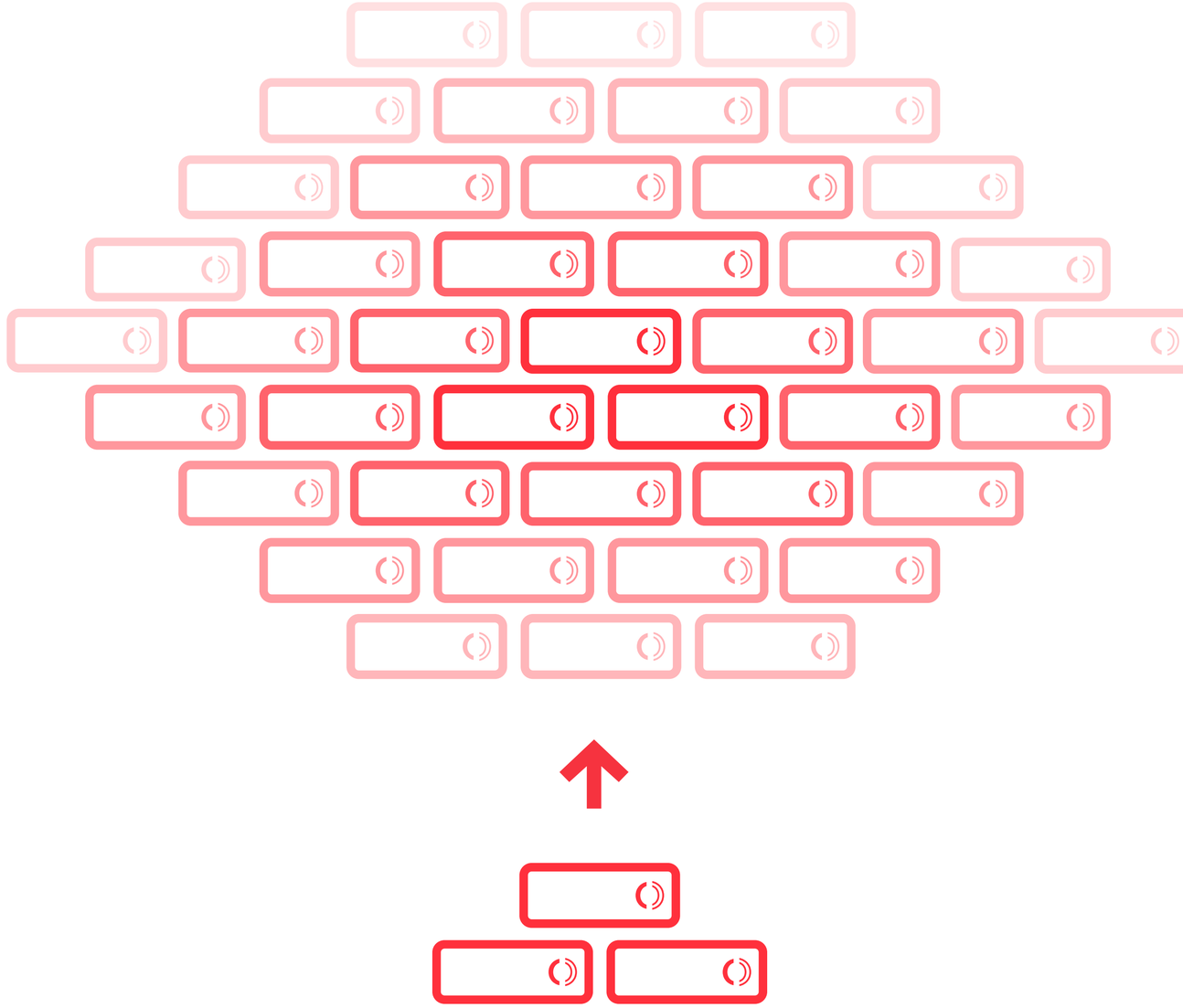


on a single platform

Store

Hyper Scalable Storage

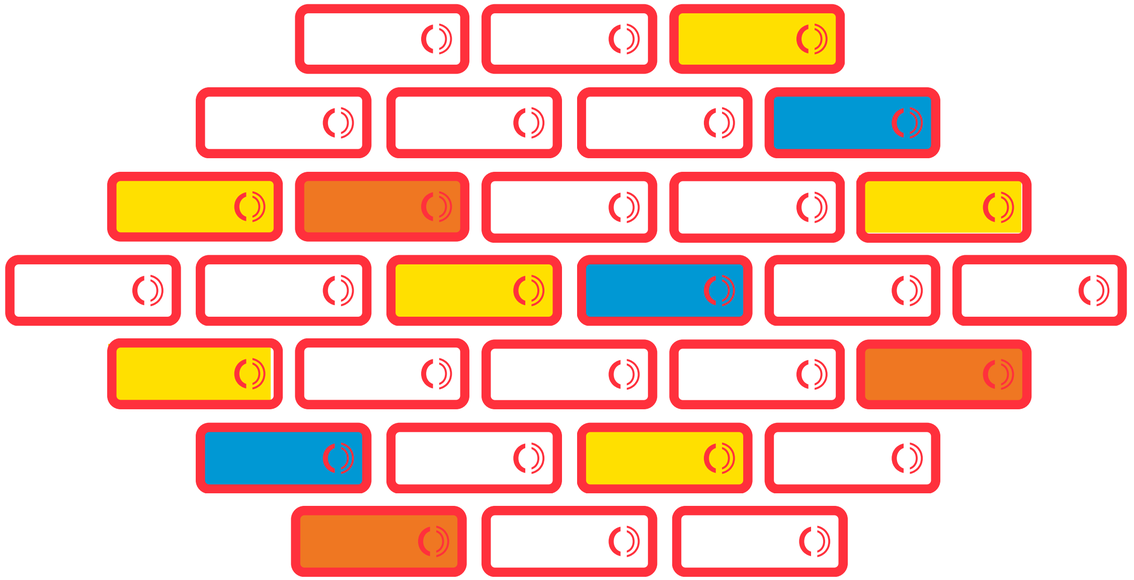
- Scale-out by nature with **shared-nothing** model to aggregate storage capacity from independent x86 servers
- **Limitless** storage based on **open source** object storage technology
- Store **thousands of PBs** of data and **billions of objects**



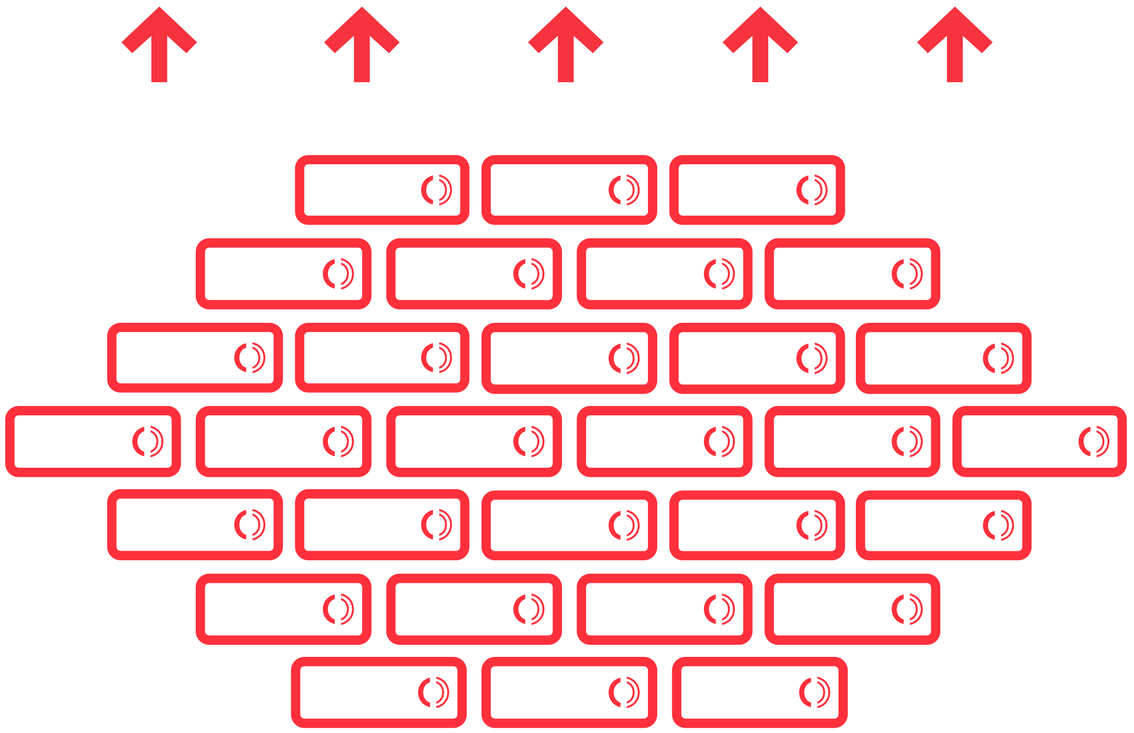
Protect

Super Data Resiliency

- **Data Replication** with multiple copies
- **Erasure coding** based on Reed-Solomon
- Various **topologies** from 1 Data Center to multiple or stretched cluster across geos
- Synchronous or Asynchronous geo-replication



Serve

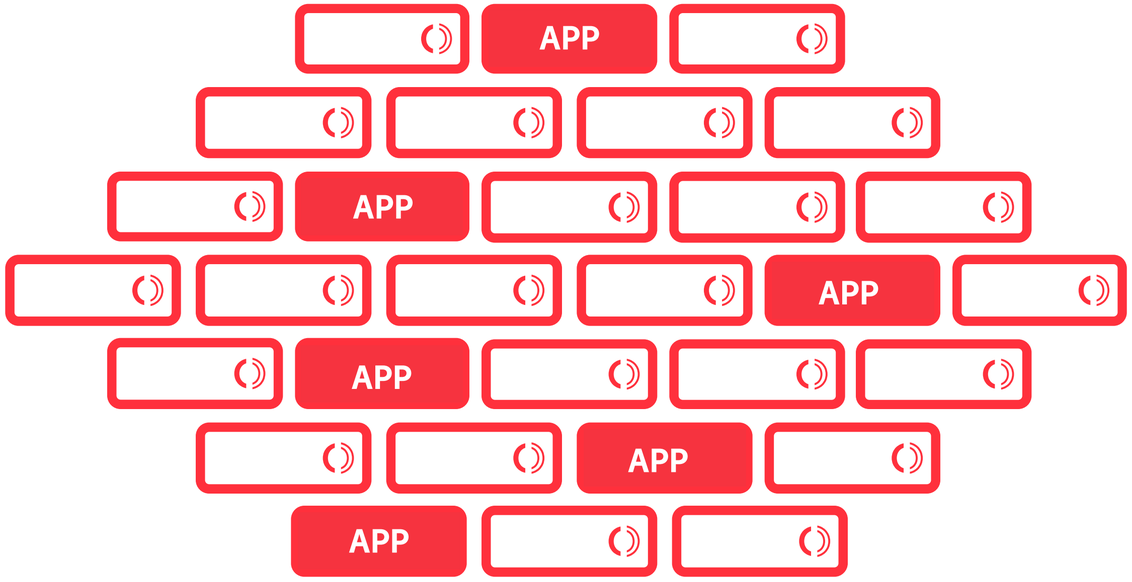


Open Data Services to Applications

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

Compute Data “Locally”

- Move and **run applications where data resides**
- Consolidate the app tier and the storage tier
- Better applications SLAs with data locality
- Flexible application support with **APIs, file-based access or SDKs**
- Drastic TCO Reduction



Run Application
directly on storage nodes



HOW

Directory with indirections

→ **Track containers, not objects**

Namespace

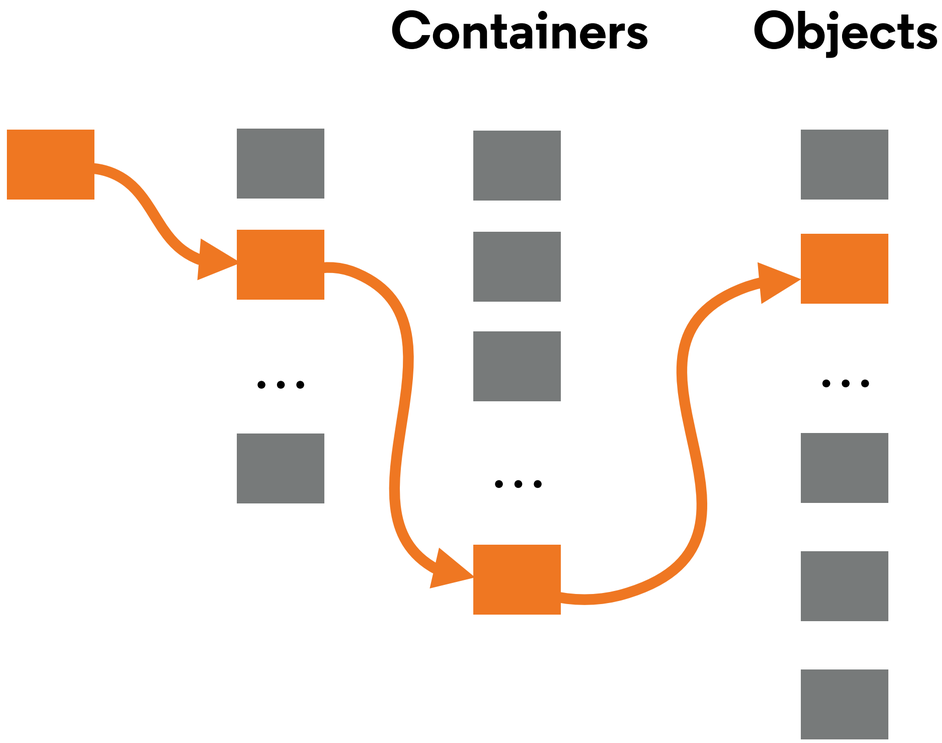
Account

Container

Object

Multi tenancy

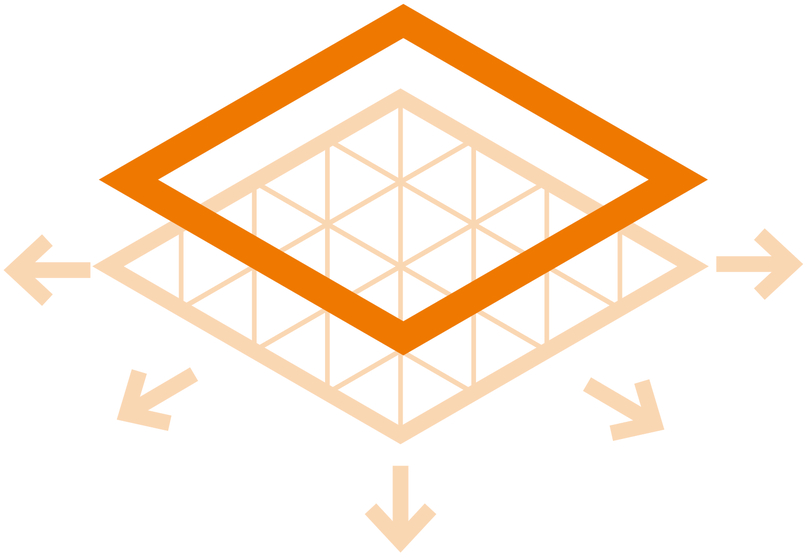
Flat structure

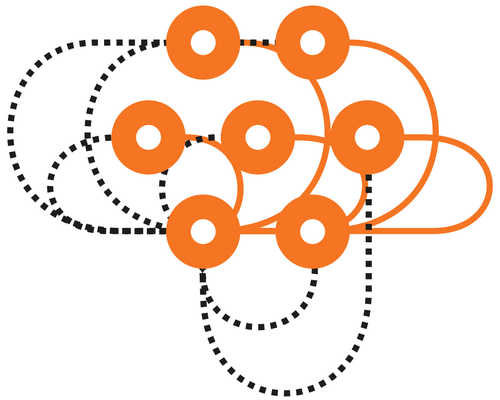


grid://namespace/account/container/object

Conscience

→ Realtime load balancing for optimal data placement





1. **Collects metrics from the services of each node**
2. **Computes a score for each service**
3. **Distributes scores to every nodes and clients**
4. **On the fly best match making for each request**

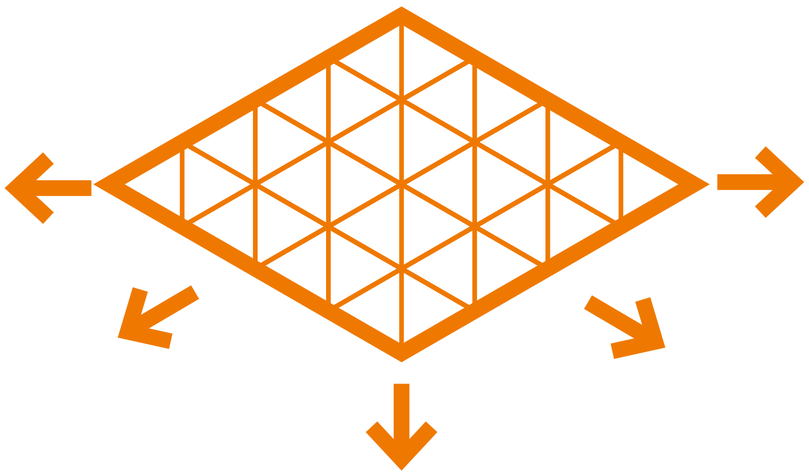
The score is computed with a configurable formula usually based on:
capacity, io performance, CPU

Grid of nodes with no consistent hashing

→ **Never rebalance**

No consistent hashing algorithm == no recalculation of the key space

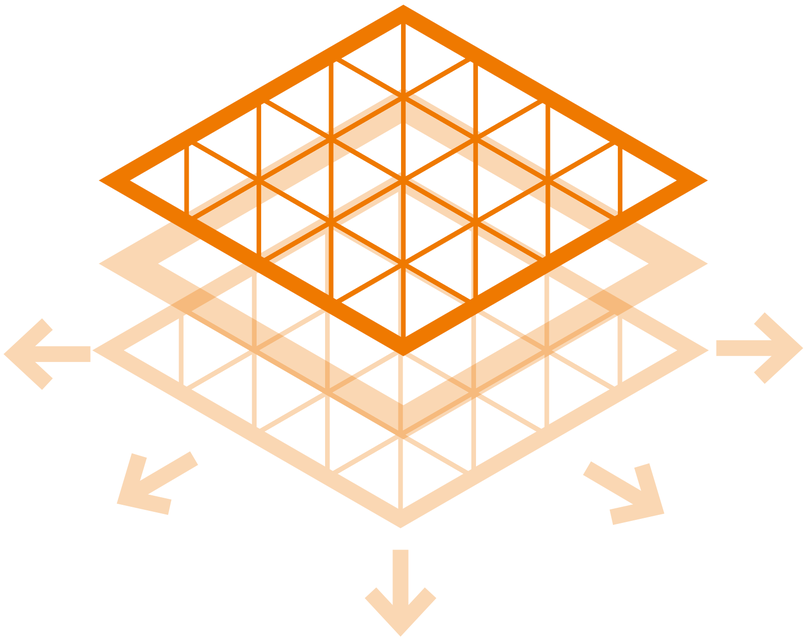
New nodes are automatically discovered and immediately available



Scaling

Grid for Apps

→ Data usage at the heart of the datacenter



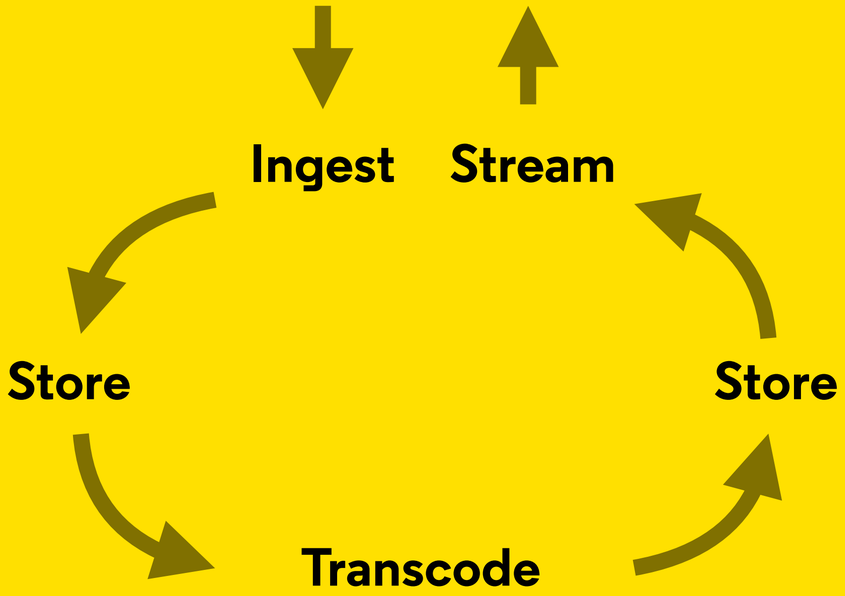
A data processing framework integrated inside OpenIO's Grid

Scale-out application back-ends can be built on the storage platform itself

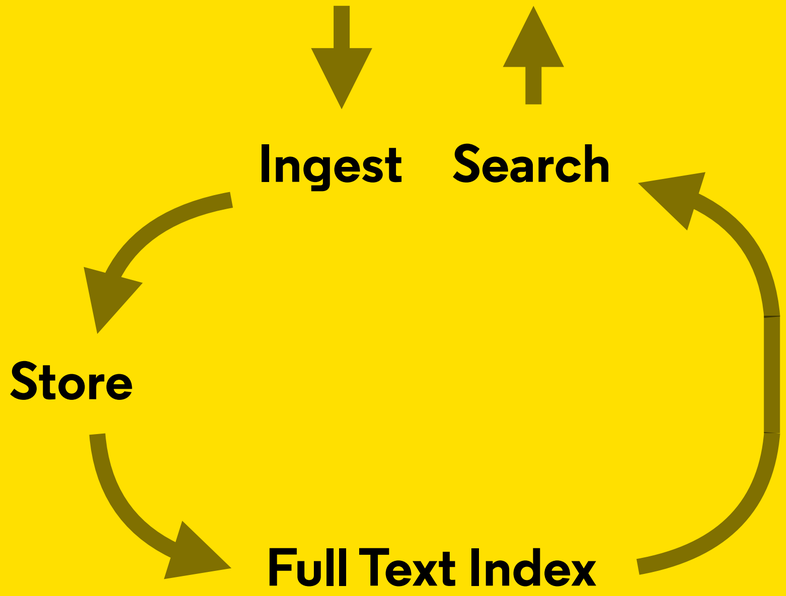
Avoid wasted resources and simplify load balancing for storage and processing

Grid for Apps

Scale-out real life use cases



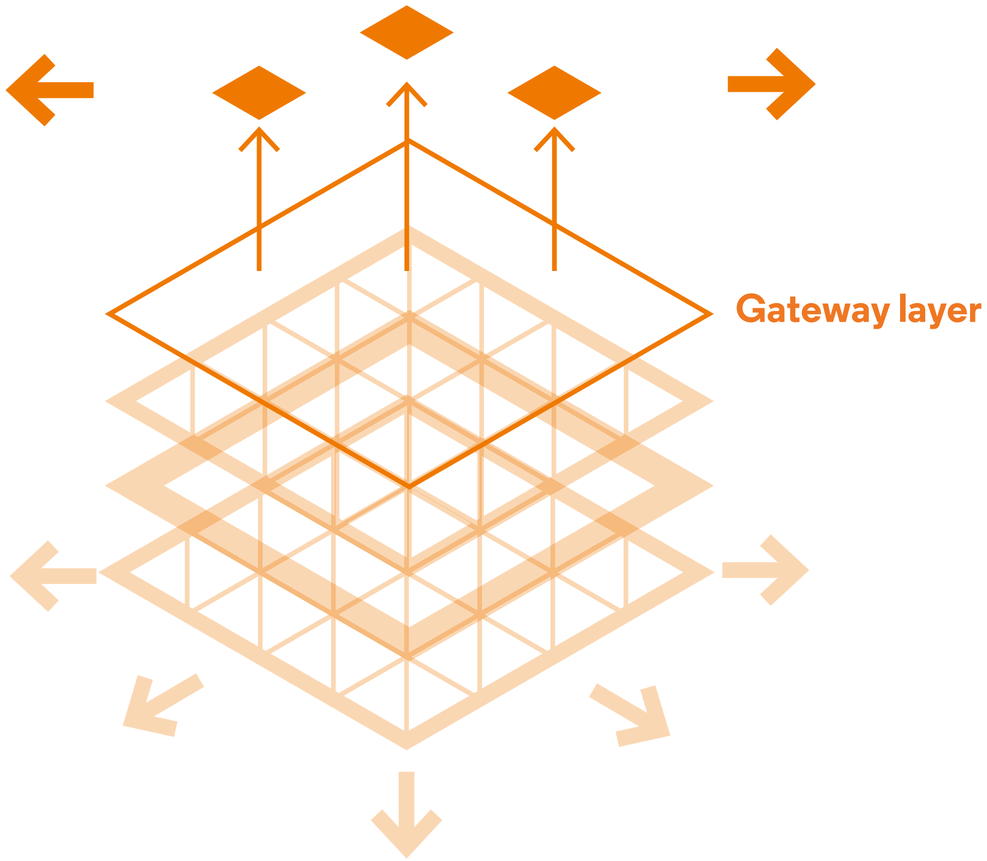
Video workflow
for user generated contents



Spam sample
long-term archiving and search

APIs and Specific App Connectors

→ Data at the heart of the datacenter



Optimized native APIs

- C
- Python
- Java
- Go

Specific App Connectors

- NFS/SMB/CIFS
- Dovecot/Zimbra/Cyrus
- Video streaming

Command line interface

REST APIs

- Amazon S3
- Openstack Swift

What makes us the most flexible Object Storage on the market ?

1.

Grid for Apps

Unifying Compute and Storage on a single platform

2.

0 TB → 1000+ PB

Start small and grow with your needs (from 3 nodes for a small config to thousands)

3.

Ubiquitous tiering

Built-in feature fully configurable and transparent

4.

Ease of use

Easy to test, deploy in production, manage and use

OpenIO tiering is obvious

1.

Built-in feature

Across storage devices managed by OIO or connected to an external pool

2.

Multicriteria settings

Name, date, size, patterns, type, container name, etc.

3.

Transparent

At upload, or with background jobs

Demo

1. Easy to deploy

   Less than 5’

2. Multi tenancy

   With flexibility

3. Easy to scale

   Easy to start with and scale later

4. Production ready

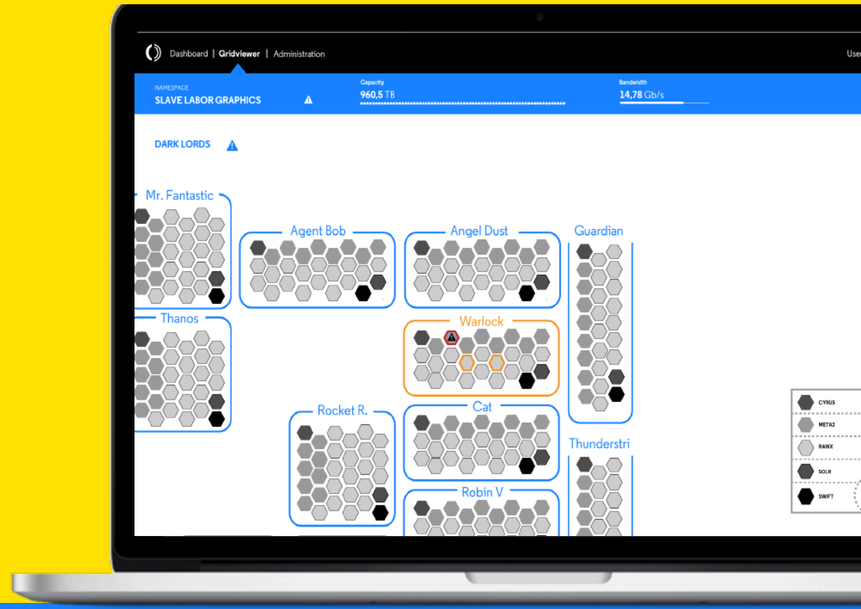
5. Let’s do it!

Easy
deployment and management

Full Operational Control

CLI available

Ubiquitous Management via Web GUI



Application-Aware Storage