SNIA **DSI** '16
DATA STORAGE INNOVATION
CONFERENCE

Innovation in Storage Products,
Services, and Solutions

June 13-15, 2016 | Marriott San Mateo | San Mateo, CA

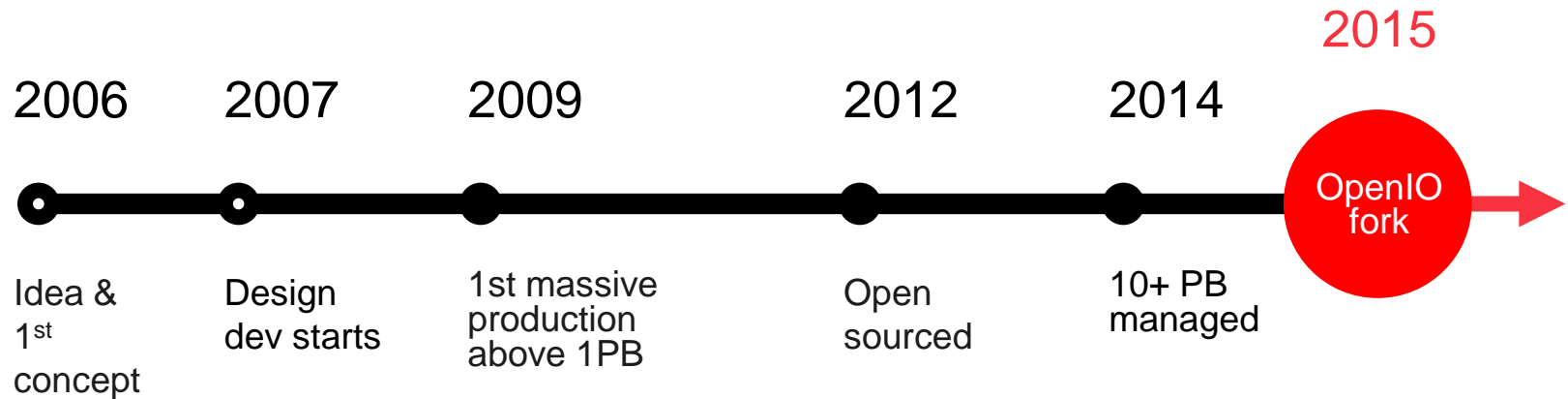# New Fresh Storage Approach for New IT Challenges

Laurent Denel – Philippe Nicolas

OpenIO

Agenda

- ☐ Company profile and background
- ☐ Business and Users needs
- ☐ OpenIO approach
- ☐ Competition
- ☐ Conclusion

Company profile and background

2006	2007	2009	2012	2014	2015

2015

OpenIO fork

| Idea & 1st concept | Design dev starts | 1st massive production above 1PB | Open sourced | 10+ PB managed | |

- Project started in 2006, Company launched in June 2015
- Privately held
- 27 people
- Lille (FR), San Francisco, Montreal, Tokyo
- 50 millions end-users (Email Storage)

SNIA™ openstack OPEN Compute Project LINUX FOUNDATION KINETIC Open Storage Project

IT Challenges Impact Business

☐ Explosion of Data Volume

☐ IT Infrastructure Cost
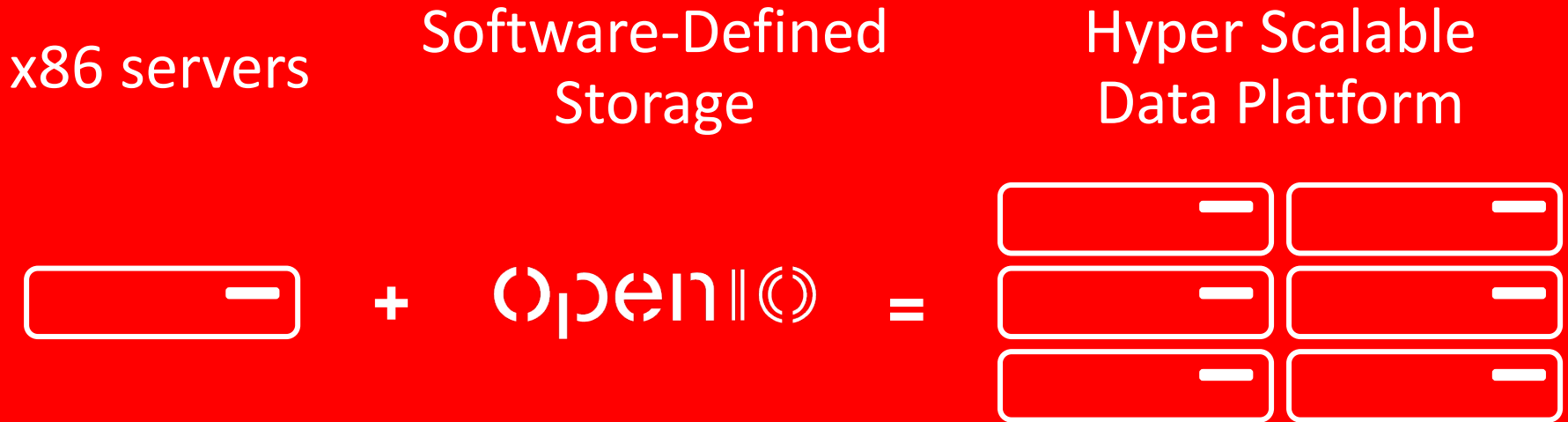
☐ Complexity – Silos

The Compute and Storage dilemma

Silos of Independent Compute and Storage resources

Over-provision to keep up with SLA and QoS

Explosion of Costs and Complexity

OpenIO's Answer

x86 servers

Software-Defined Storage

Hyper Scalable Data Platform

$+$ OpenIO $=$

OpenIO transforms a rack of x86 servers into a large storage and compute pool

OpenIO Democratizes Large Data Platform

Internet Giant
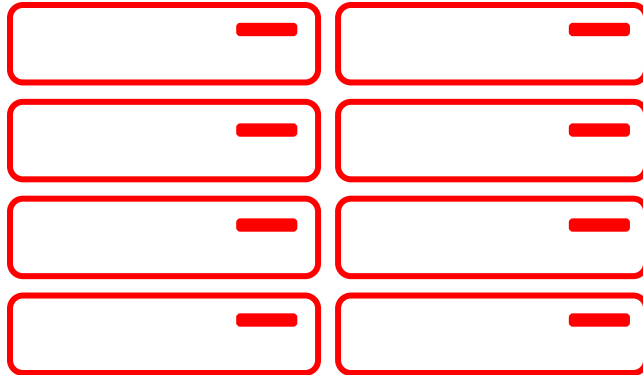have initiated the wave
and proven the technology

openIO

Same model "On-Premise"

Simple
Elastic
Flexible
Cost effective
On-demand

Data Storage and
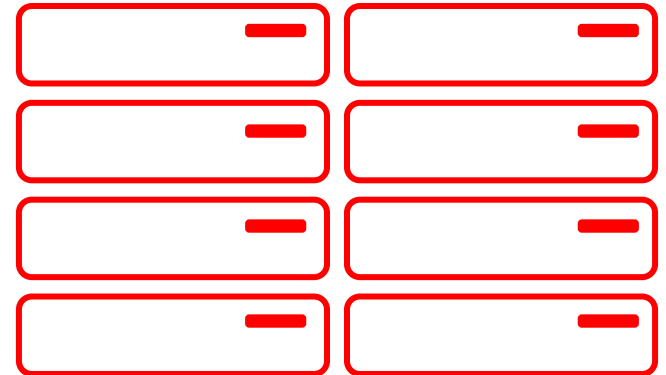Compute Platform

OpenIO's Approach

- Object Storage to scale and store Thousands of PBs of data, Billions of objects

- Open Source and Commodity x86 servers to reduce cost and TCO

- Build Storage with x86 servers and Run applications on same infrastructure
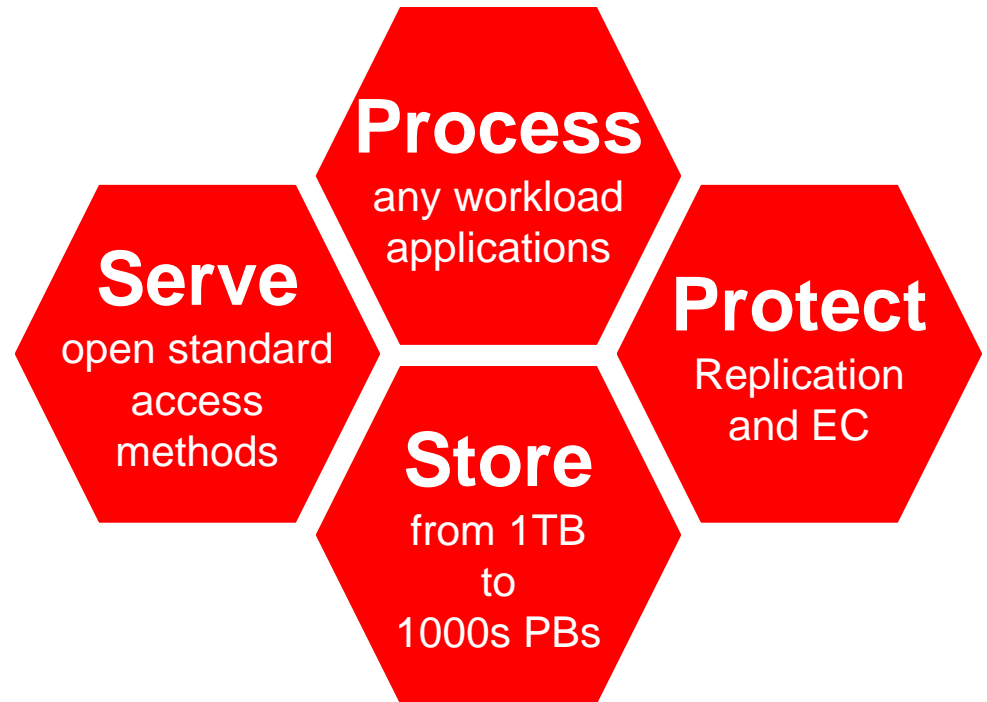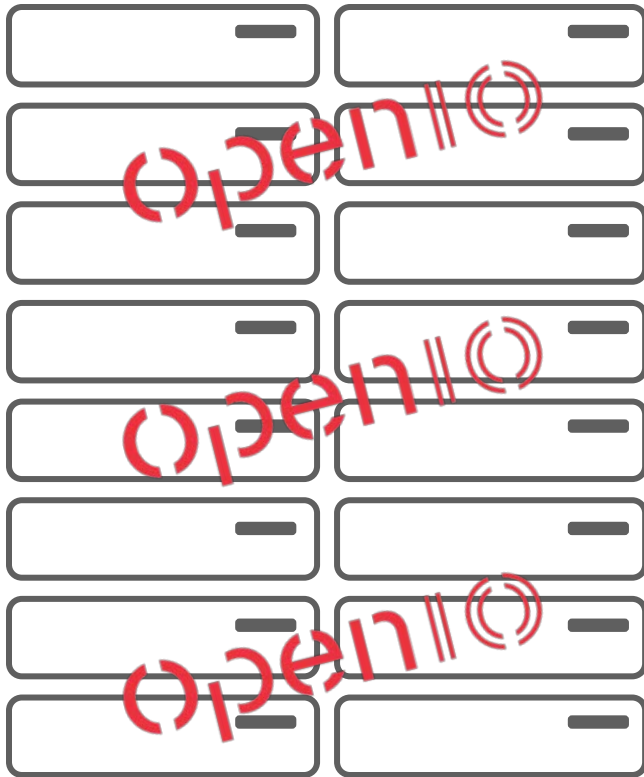
Crush the Stack with OpenIO

Compute + Storage running on same infrastructure

OpenIO Data Storage and Compute Platform

**Serve**
open standard access methods

Process
any workload applications

Store
from 1TB to 1000s PBs

Protect
Replication and EC

on a **Single Platform**

SNIA DSI '16
DATA STORAGE INNOVATION
CONFERENCE

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
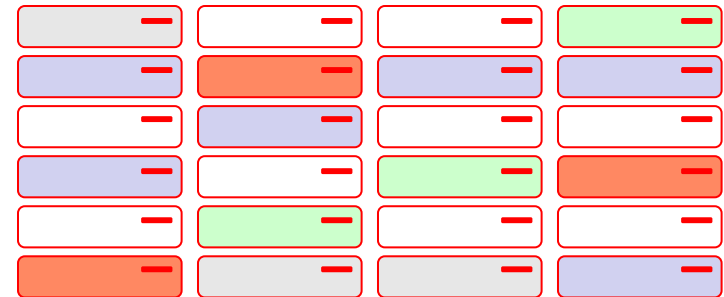
- Store Thousands of PBs of data and Billions of objects
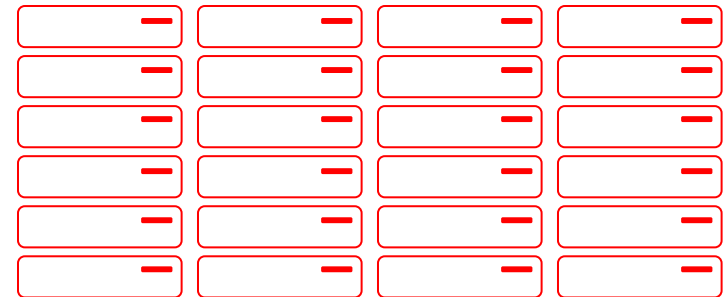
Scale in any dimension

High Data Durability

- Data Replication with multiple copies

- Erasure Coding based on Reed-Solomon

- Various topologies from 1 Data Center to Multiple or stretched cluster across geos

Multi-copies Replication
Erasure Coding based on Reed-Solomon

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, AFP and FTP + FUSE

Objects APIs / File Sharing Protocols
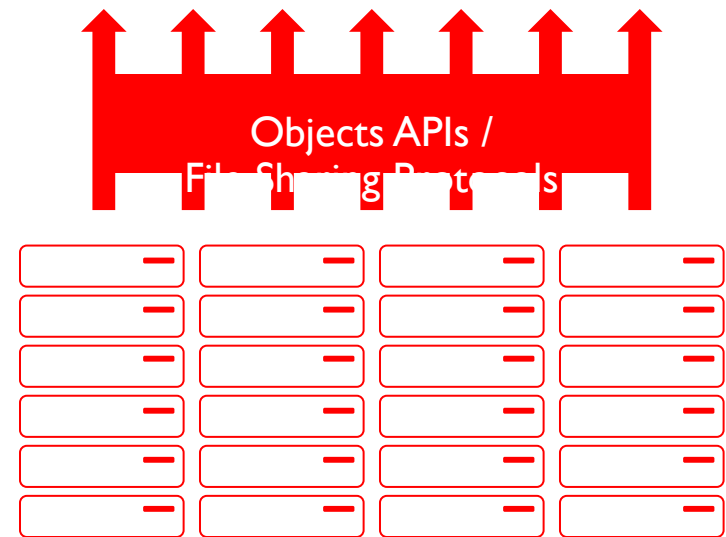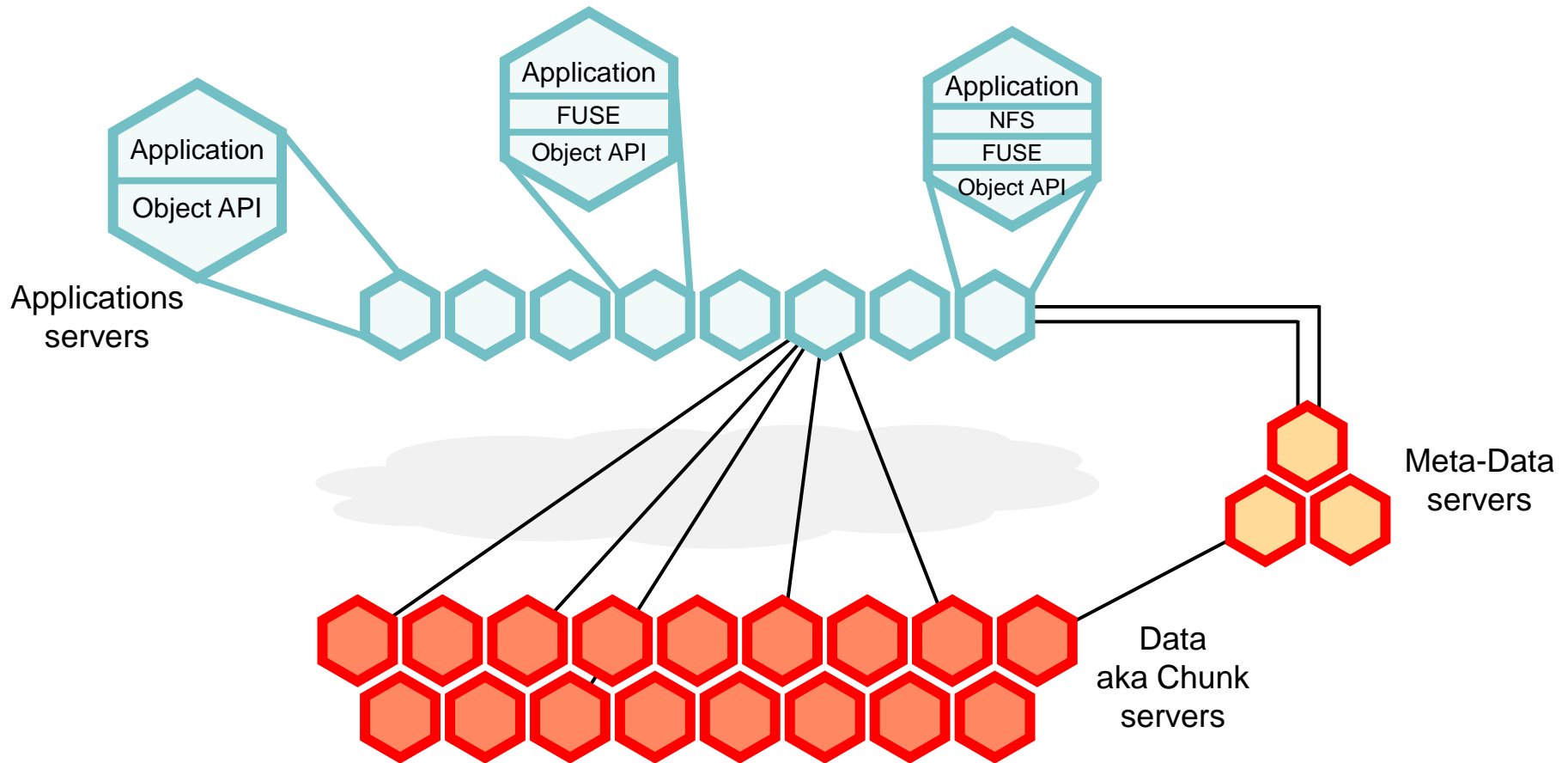
Compute Data "locally"

- Move and Run Applications where Data reside
- Consolidate the App Tier and the Storage Tier
- Better applications SLAs with data locality
- Flexible application support with APIs, File-based access or SDKs
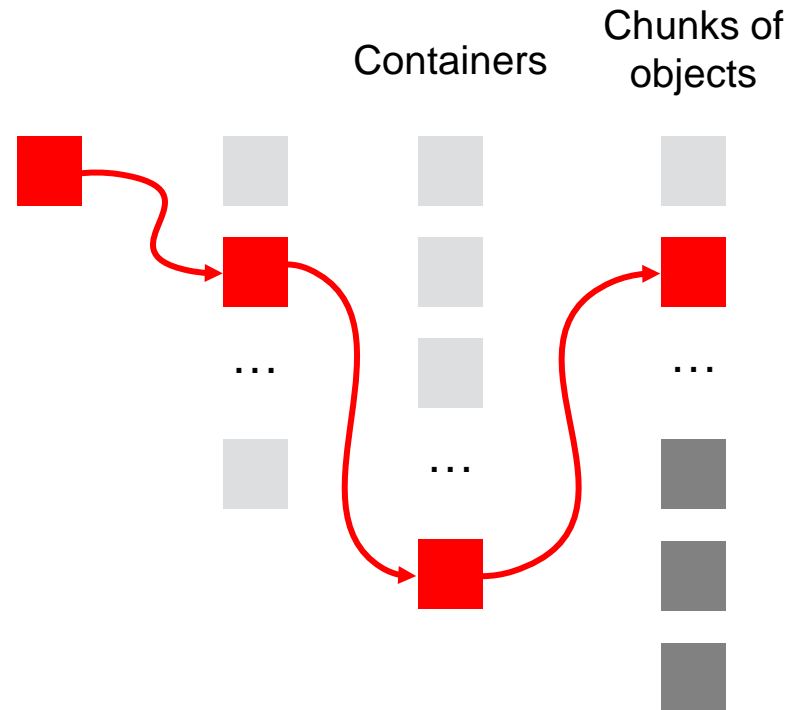- Drastic TCO Reduction

App.   App.   App.   App.
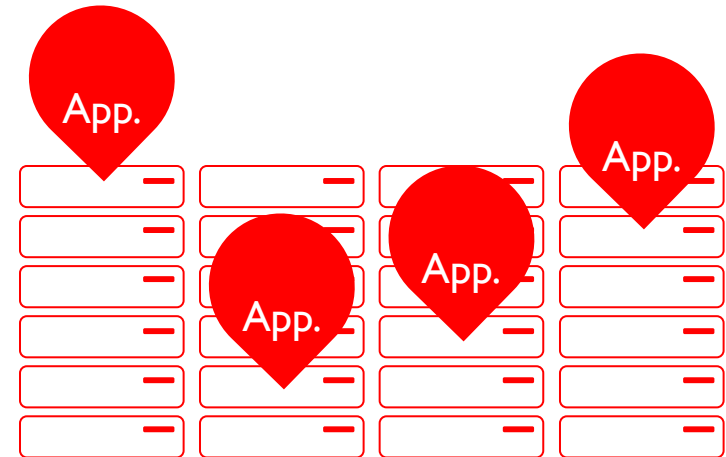
Run Application directly on storage nodes

Logical Architecture

Meta Directories

- Directories with Indirections
- Track Containers not Objects
- 2 Hops Max. to reach any Object
- Multi-Tenancy at the Account level
- Flat structure at the Container level

Containers
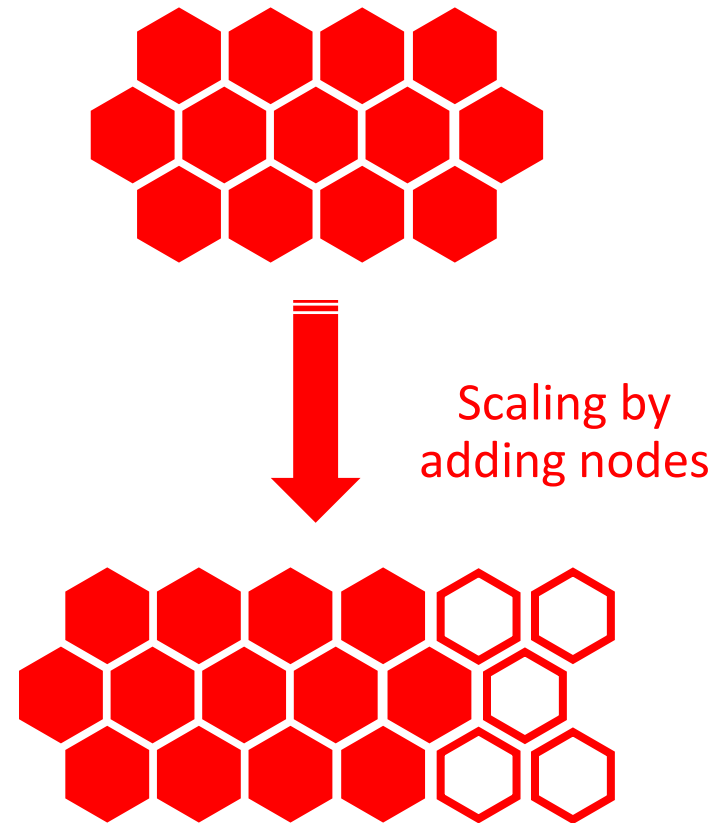
Chunks of objects

...

...

...

grid://namespace/account/container/object

SNIA | DSI-'16-
DATA STORAGE INNOVATION
CONFERENCE

Grid of nodes

- No Consistent Hashing –
  No Key Space to update –
  Never Rebalance

- Transparent Elasticity – No
  Impact on production

- Immediate availability of
  new nodes for storage and
  compute

Scaling by
adding nodes

Conscience

- Real-time load balancing for optimal data placement
- Collects metrics from the services of each node
- Computes a score for each service
- Distributes scores to every nodes and clients
- On the fly best match making for each request
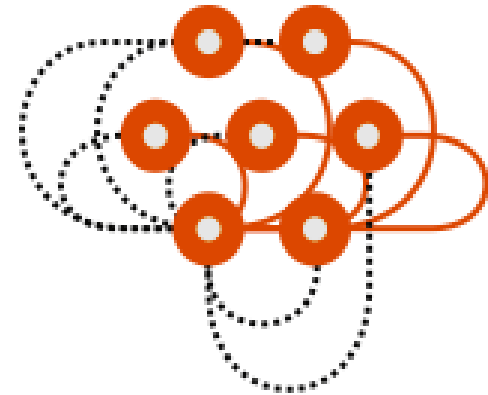
The score is computed
with a configurable formula
usually based on:
capacity, io performance, CPU

Grid for Apps – Data Access

- 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

- Optimized native Object APIs
  - OpenIO, Amazon S3, OpenStack Swift
  - C, Python, Java, Go
- Specific Application Connectors / Editions
  - Email, Video/Media and Enterprise Storage
- Command Line Interface
- SDK Soon

Easy Deployment and Management

- Full Operational Control
- Ubiquitous Management via Web GUI
- CLI available
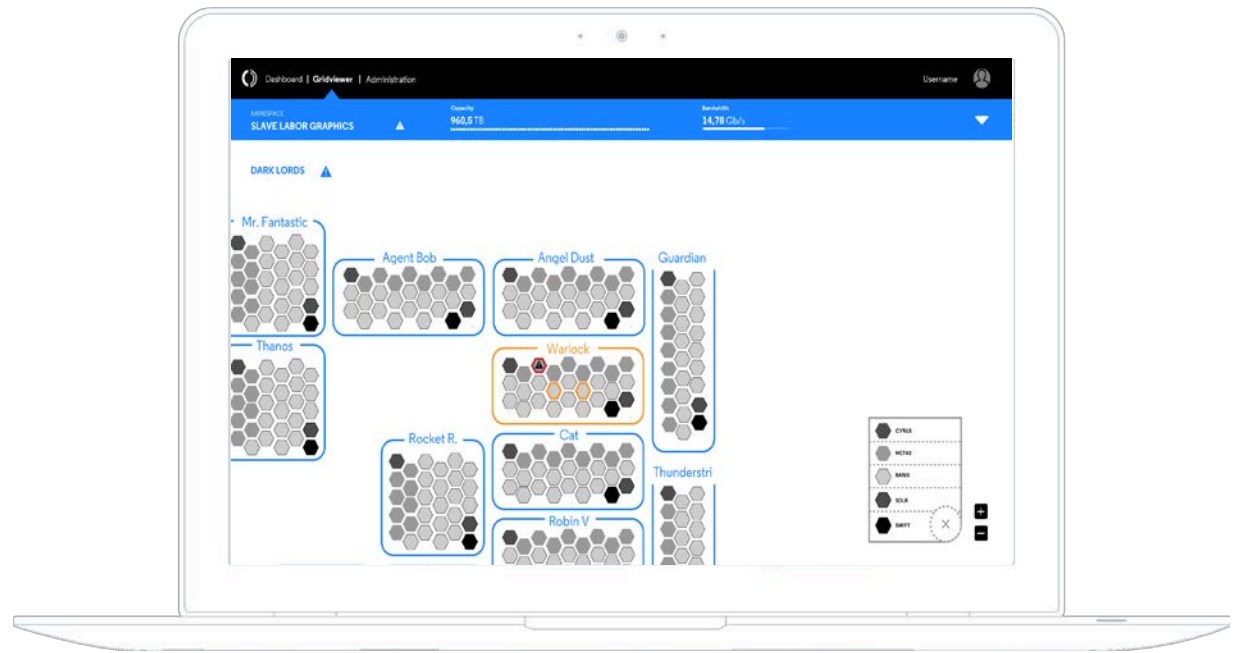
Editions and Pricing Model

- Software
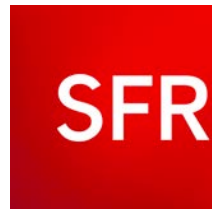 - Core $0
  - Editions $30k/Edition/Year
- Support
  - Standard $0.08/actual GB/Yr
  - Premium Support $60k/Yr
  - Platinum $150k/Yr (premium support + All Editions)

Use Cases & Customers' Stories

- ❐ Email Storage
- ❐ Storage-as-a-Service
- ❐ Media Content Processing & Delivery
- ❐ Compute + Storage Platform
- ❐ Long-term Data Archive
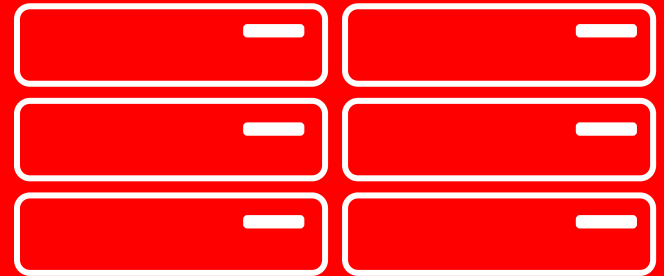
What you can do in 5 minutes

Drink
an espresso

Stretch
your body

Deploy*
an OpenIO cluster

https://github.com/open-io/oio-sds

OpenIO

Competition with Object Storage Solutions

	Open Source Object Storage	Commercial Object Storage	OpenIO
HW Agnostic	Mostly	Mostly if SW, Absolutely not if Appliance	**Any mixed HW**
Open Source	Yes	No by definition	**Yes**
TCO	$	$$$	**$**
Scalability	Webscale	Webscale	**Webscale**
Compute + Storage	No just Storage	No just Storage	**Yes (Grid for Apps)**
Data Access Methods	Objects APIs, Block GW	Objects APIs, some offer file access or partner w/ GW vendor	**Objects APIs, File sharing protocols and Editions (Mail, Video…)**
Data Protection	Replication and EC	Replication and EC	**Replication and EC**

SNIA DSI '16 DATA STORAGE INNOVATION CONFERENCE

Competition with Data Center Platforms

	Object Storage	Big Data Hadoop	Hyper-Converged	OpenIO
HW Agnostic	Mostly	Mostly	Appliance based, HW restrictions	**Any mixed HW**
Compute + Storage	Storage only	Yes but dedicated	Yes but limited	**Yes (Grid for Apps)**
Open Source	Only 4: OpenIO, Minio, Ceph and Swift	Yes	Mostly Commercial	**Yes**
TCO	$$	$$	$$$	**$**
Webscale	Yes	Yes at Internet giant sites	No	**Yes**

OpenIO Key Takeaways

OpenIO

Open Source SW
HW Agnostic
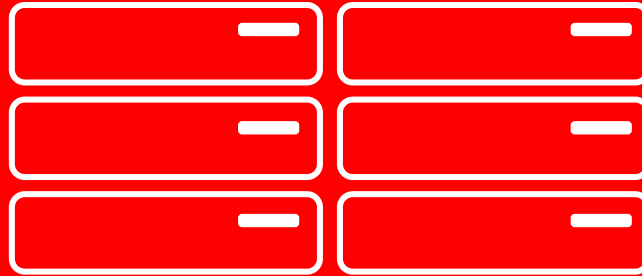
Distributed
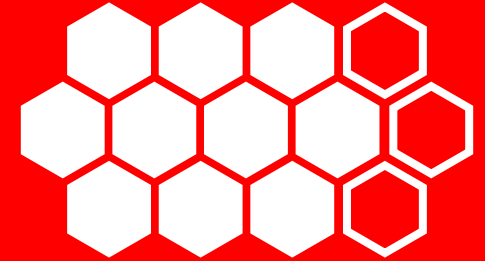
No Rebalance

Conscience
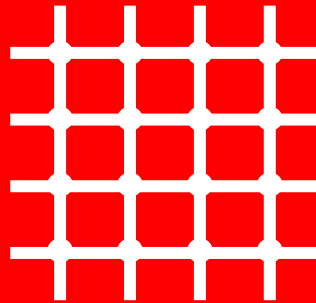
Grid for Apps

Webscale

openIO

CRUSH
THE STACK

More info on openio.io

New Fresh Storage Approach for New IT Challenges

Laurent Denel – Philippe Nicolas

OpenIO