Game Changer: Software-Defined Storage and Container Schedulers

David vonThenen
{code} by Dell EMC
@dvonthenen
dvonthenen.com
github.com/dvonthenen

Agenda

- Container Schedulers
- Containers In Production
- Software-Defined Storage (SDS)
- Schedulers + SDS = Game Changing
- Demo

Schedulers

What is a Scheduler?

- Fair and efficient workload placement
- Adhering to a set of constraints
- Quickly (and deterministically) dispatching jobs
- Robust and tolerates errors

Apache MESOS™

kubernetes

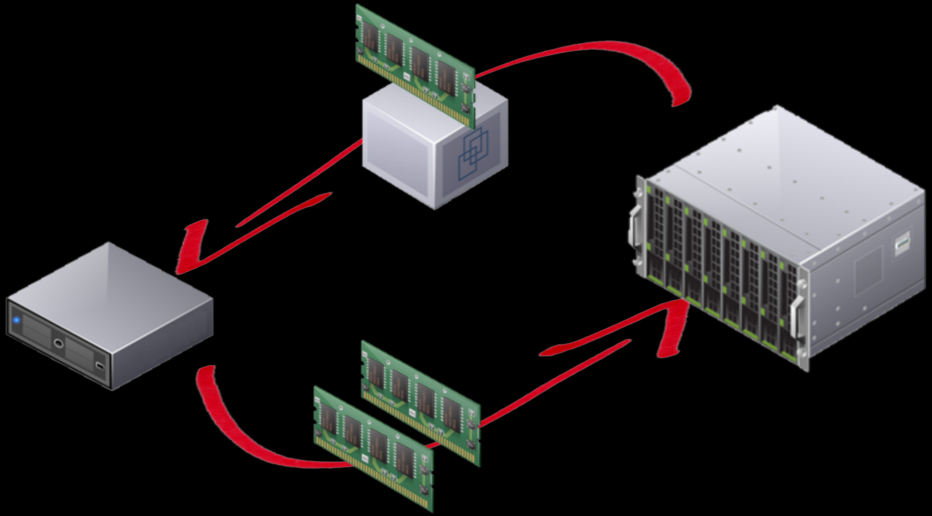
Pivotal Cloud Foundry®

Let’s take a look: Apache Mesos

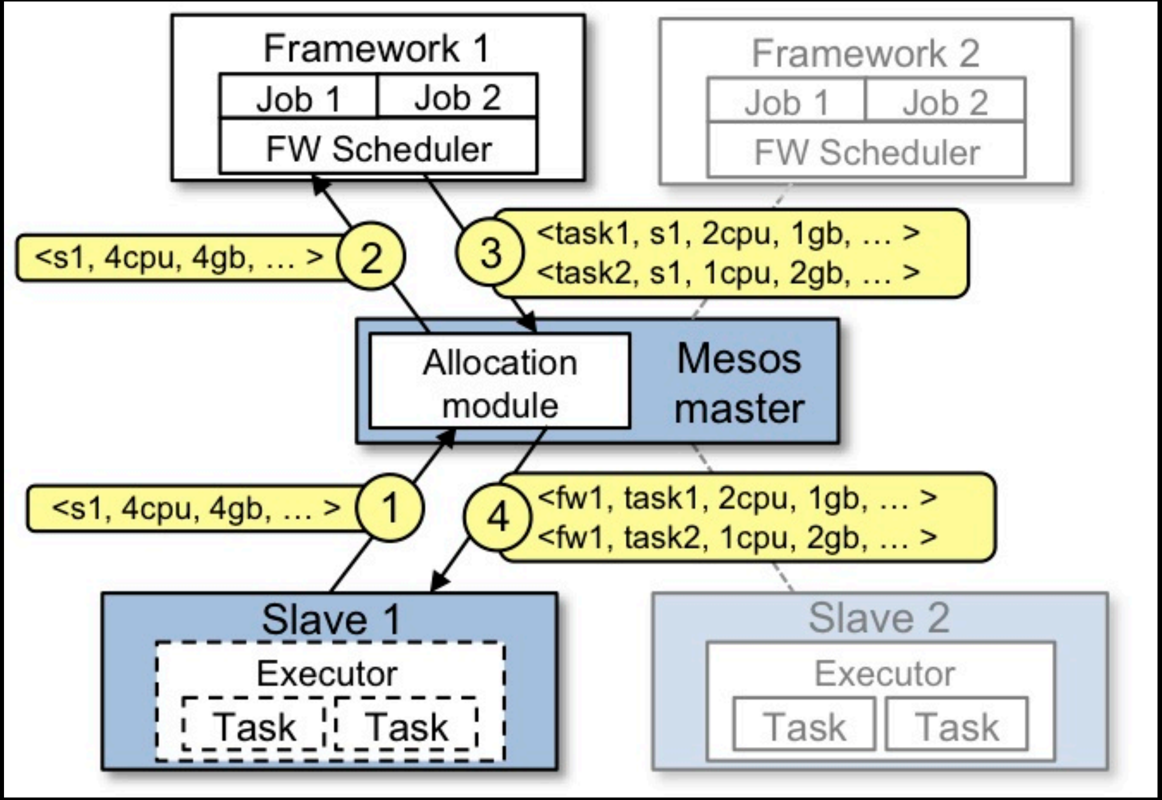
- Is a Container Scheduler
  - Docker
  - Unified Containerizer
- Cluster Manager
- Task placement based on CPU, Memory, and Disk
- User defined constraints
- 2 Layer Scheduler – Offer/Accept Model

Mesos Frameworks

- Ability to sub-schedule tasks based on Application needs
- Framework implements a Scheduler and Executor
  - Scheduler – Accepts/Denies resources
  - Executor – Application
- Multiple Frameworks run within the cluster



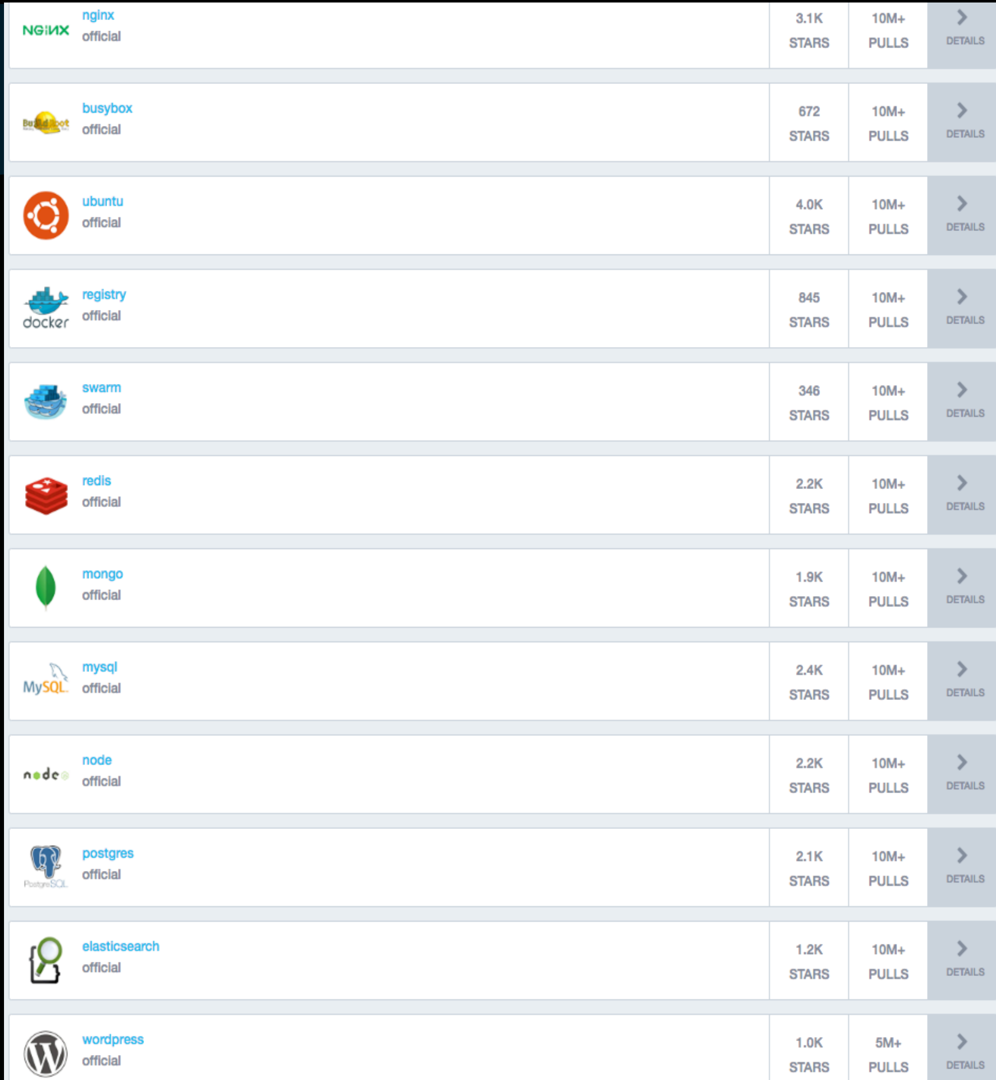
Framework / Offer Mechanism



Containers

Containers Today

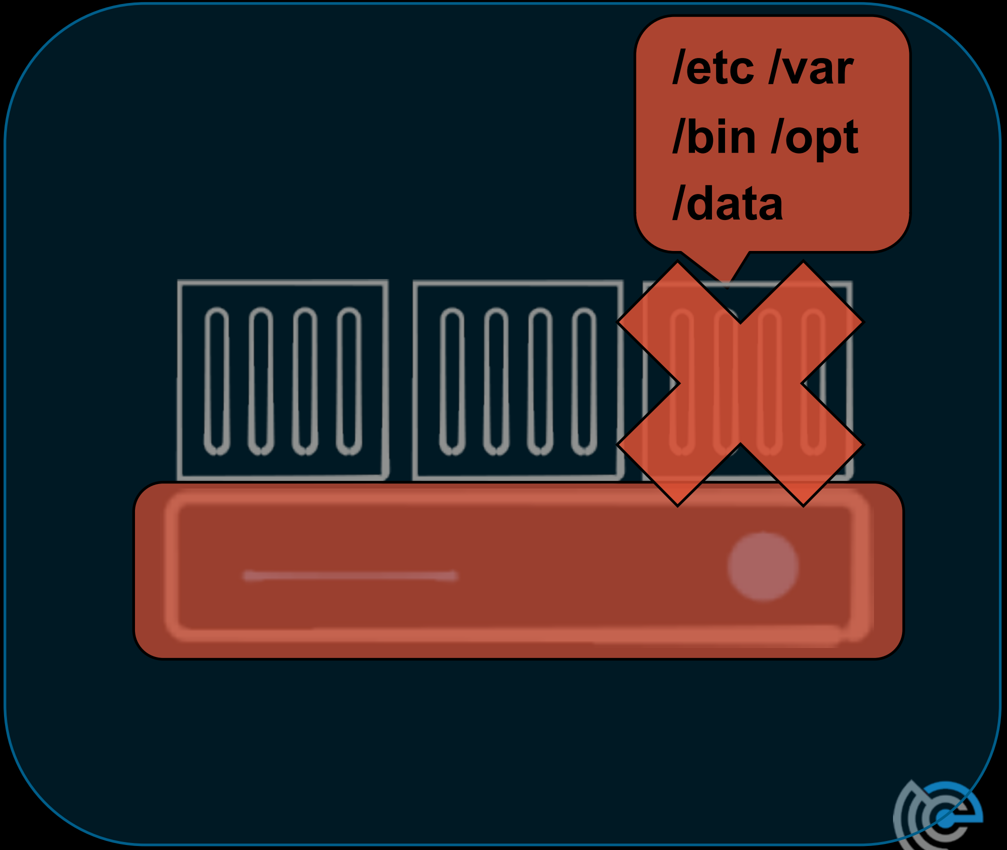
- Many container workloads are long running
- Many have state: user data, configuration, and etc
- Top 7 of 12 Apps in Docker Hub are persistent applications



nginx official	3.1K STARS	10M+ PULLS	DETAILS
busybox official	672 STARS	10M+ PULLS	DETAILS
ubuntu official	4.0K STARS	10M+ PULLS	DETAILS
registry official	845 STARS	10M+ PULLS	DETAILS
swarm official	346 STARS	10M+ PULLS	DETAILS
redis official	2.2K STARS	10M+ PULLS	DETAILS
mongo official	1.9K STARS	10M+ PULLS	DETAILS
mysql official	2.4K STARS	10M+ PULLS	DETAILS
node official	2.2K STARS	10M+ PULLS	DETAILS
postgres official	2.1K STARS	10M+ PULLS	DETAILS
elasticsearch official	1.2K STARS	10M+ PULLS	DETAILS
wordpress official	1.0K STARS	5M+ PULLS	DETAILS

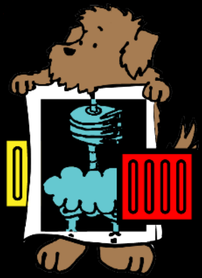
Death of a Container

- Where does my data go?
- Turned to the compute node's local disk to store data
- What happens on a node failure?
- Production applications require high availability
- External Storage!

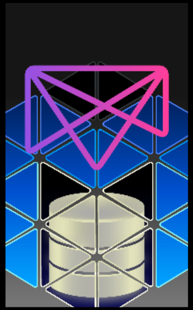


How do we achieve this?

- REX-Ray
  - Vendor agnostic storage orchestration engine
  - AWS, GCE, ScaleIO, VirtualBox, many more
  - GitHub: https://github.com/emccode/rexray



- mesos-module-dvdi
  - Provides hooks to Mesos agent nodes to manage external storage
  - GitHub: https://github.com/emccode/mesos-module-dvdi



Enablement is Out-Of-Band

- The “glue” that combines compute to external storage is add-on to the resource manager
- Obvious but easily dismissive answer: DevOps
  - Software upgrades? On all nodes…
  - Maintenance? Infrastructure, Storage Platform, etc
  - Changes to Container Scheduler? Behaviors, APIs, etc
- Just make it happen!
- Almost 100% of the way there...

Got to be an easier way…

Software-Defined Storage

What are they?

- Software-Defined Storage (SDS) serve as abstraction layer above underlying storage
- Provides a (programmatic) mechanism to provision storage
- Varying degrees of SDS: NFS, VMware Virtual Volumes

What makes them unique?

- Manage provisioning and data independent of underlying hardware (operational)
- Abstract consumed logical storage from underlying physical storage (physical)
- Automation of policy driven SLAs both external (users) and internal (platform)

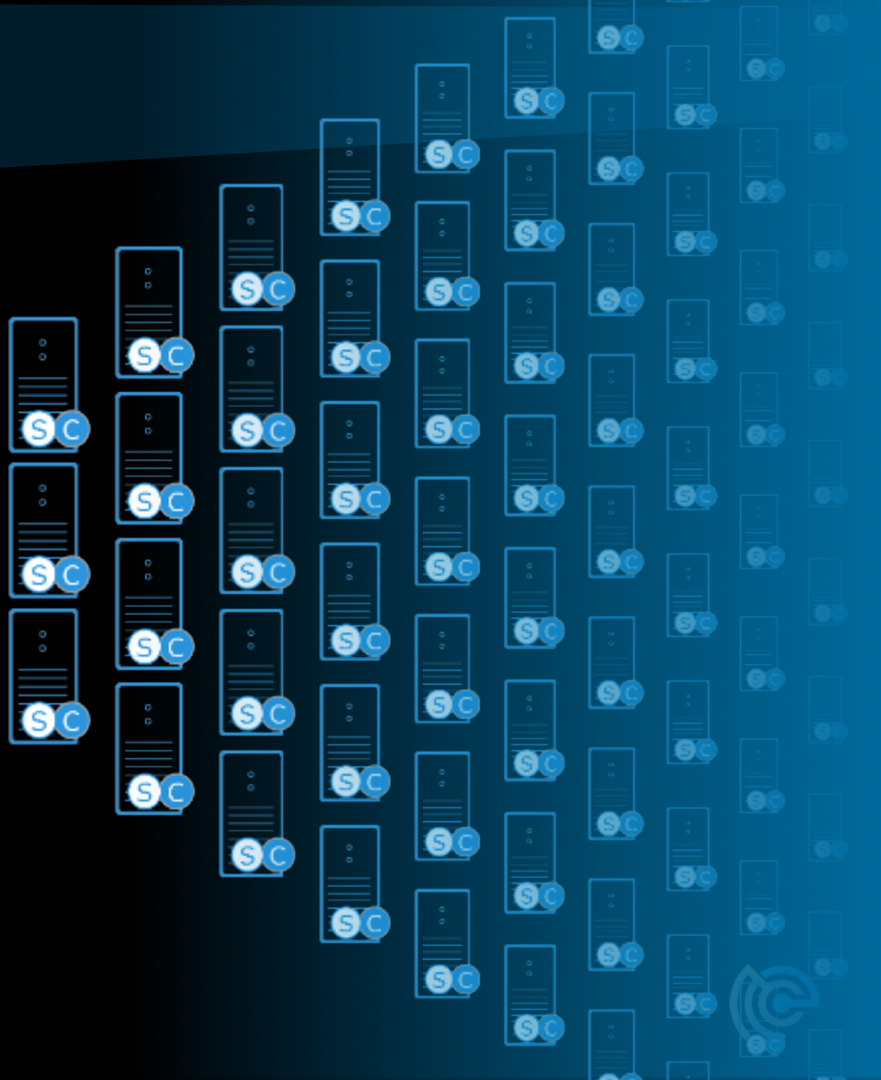
Let’s take a look: ScaleIO

- Scale-out block storage
- Linear performance
- Elastic architecture
- Infrastructure agnostic
- Try ScaleIO as a free download:

https://www.emc.com/products-solutions/trial-software-download/scaleio.htm

Scale-out Block Storage

- Scale from 3 nodes to 1000s of nodes
- Add storage services and servers on the fly to increase capacity and performance
- Storage growth always automatically aligned with application needs



Elastic Architecture

Add, remove, re-allocate, on the fly, without stopping IO

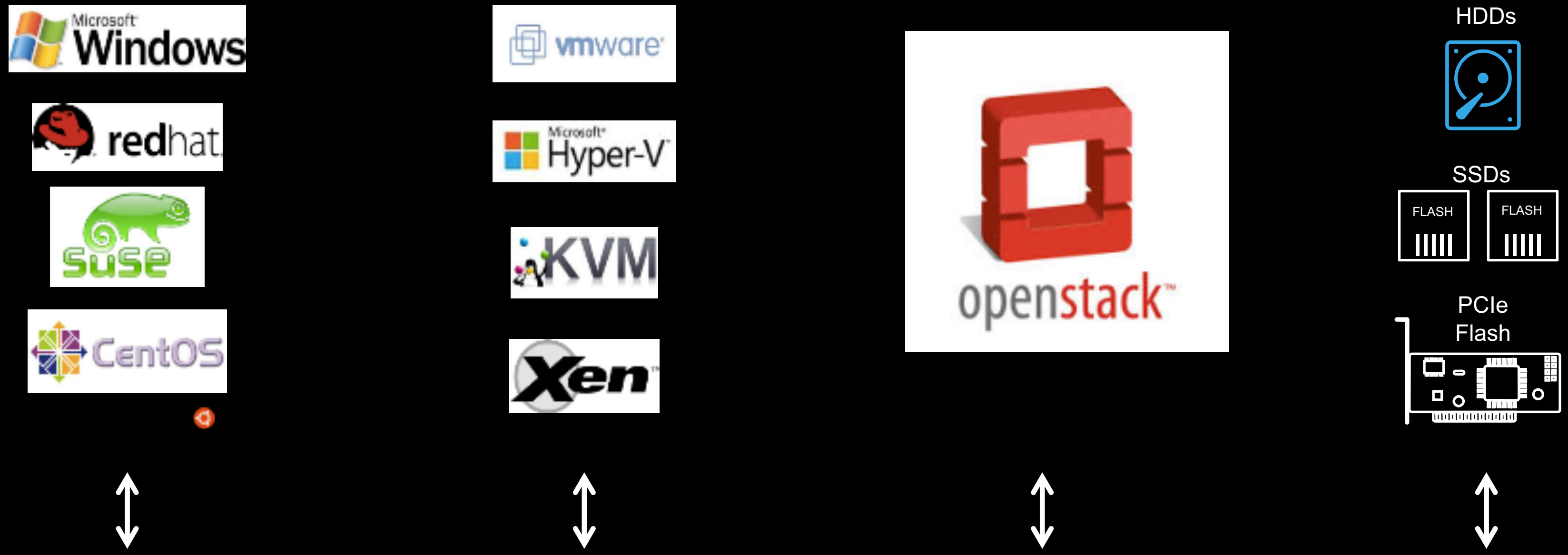
AUTO-REBALANCE
when resources are added

AUTO-REBUILD
when resources fail or removed

NO CAPACITY PLANNING OR MIGRATION!

Infrastructure Agnostic

Mix and match OS, hypervisors, platforms, media in the same ScaleIO system



HDDs

SSDs

PCIe Flash

OPERATING SYSTEM

HYPERVISORS

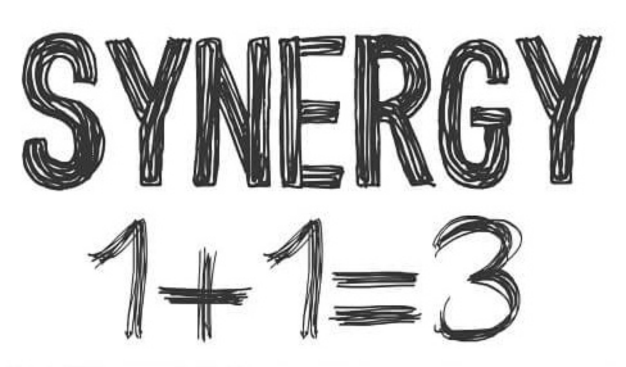
CLOUD

MEDIA

Game Changer

Let's Review…

- Container Schedulers:
  - Great platform for container management
  - Needs persistent storage for production Apps
  - Adding persistent storage out-of-band presents challenges
- Software-Defined Storage:
  - Scale-out storage
  - Elastic architecture
  - Infrastructure agnostic



Schedulers + SDS = ????

Better than the Sum of Our Parts

- Let's create a Software-Defined Storage Framework
- ScaleIO + Mesos Framework = Awesome Sauce!
- https://github.com/codedellemc/scaleio-framework

SDS Framework = Mind Blown

- External persistent storage **native** to scheduling platform
- Globally accessible storage
- Storage array? Reduce complexity
- Reduces maintenance
- Deploy Anywhere!

What this Means for your Apps

- No data loss on infrastructure failure
- Insulates changes with cluster manager (APIs, etc)
- Highly Available containers and Apps!
- Production ready!
- Tolerates failures

Surprising Combination

Demo

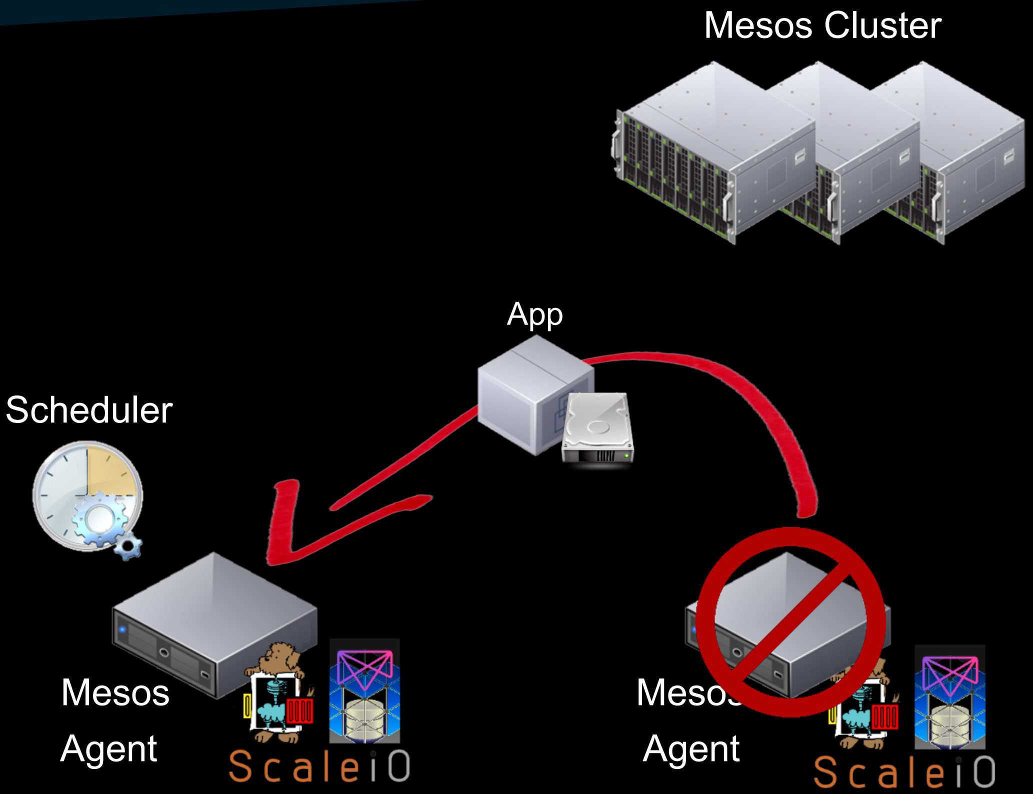
Configuration

- 3 Node Mesos Cluster (Management)
- 2 Mesos Agent nodes (Compute)
  - Initially the first node online
  - Second node will be onboarded or introduced later
- ScaleIO Cluster (Scale-out storage)
  - 3 management nodes
  - 180 GB local disks on each management node to comprise this storage pool

Configuration (Cont.)

- ScaleIO Framework
  - GitHub: https://github.com/codedellemc/scaleio-framework
- Persistent External Storage
  - Using REX-Ray
    - GitHub: https://github.com/emccode/rexray
  - Using mesos-module-dvdi
    - GitHub: https://github.com/emccode/mesos-module-dvdi

The Moving Parts



Mesos Cluster

App

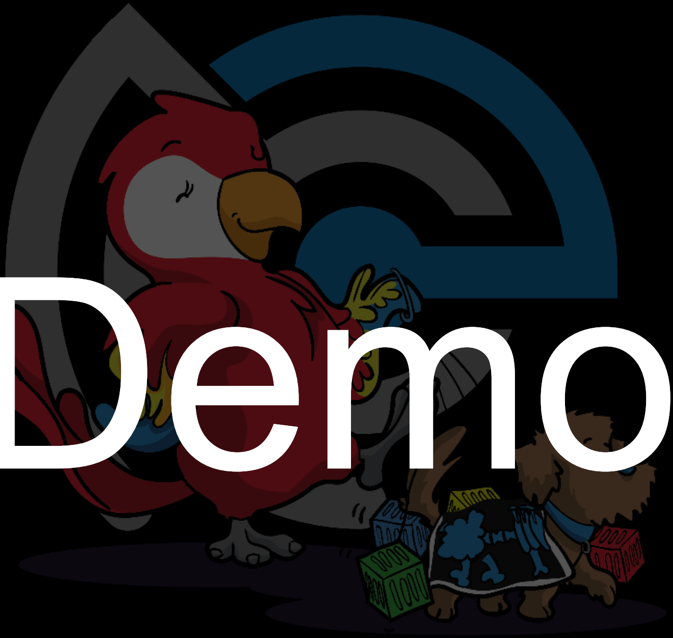
Scheduler

Mesos Agent

Mesos Agent

#CodeOpen

Demo



codedellemc.com

#CodeOpen

Thank you

codedellemc.com

#CodeOpen

codedellemc.com

DELL EMC