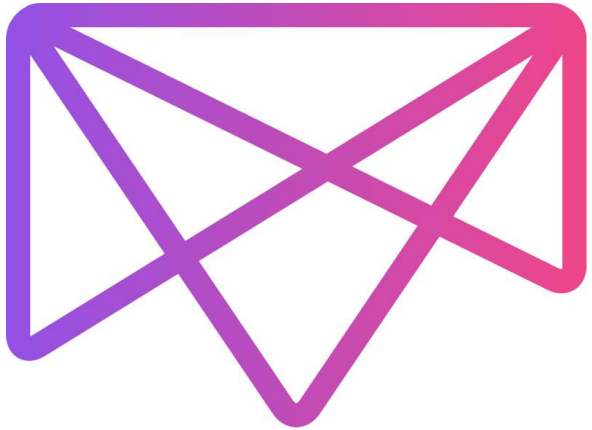
Prague Containers Meetup

Introduction to DC/OS

Matt Jarvis
@mattj_io

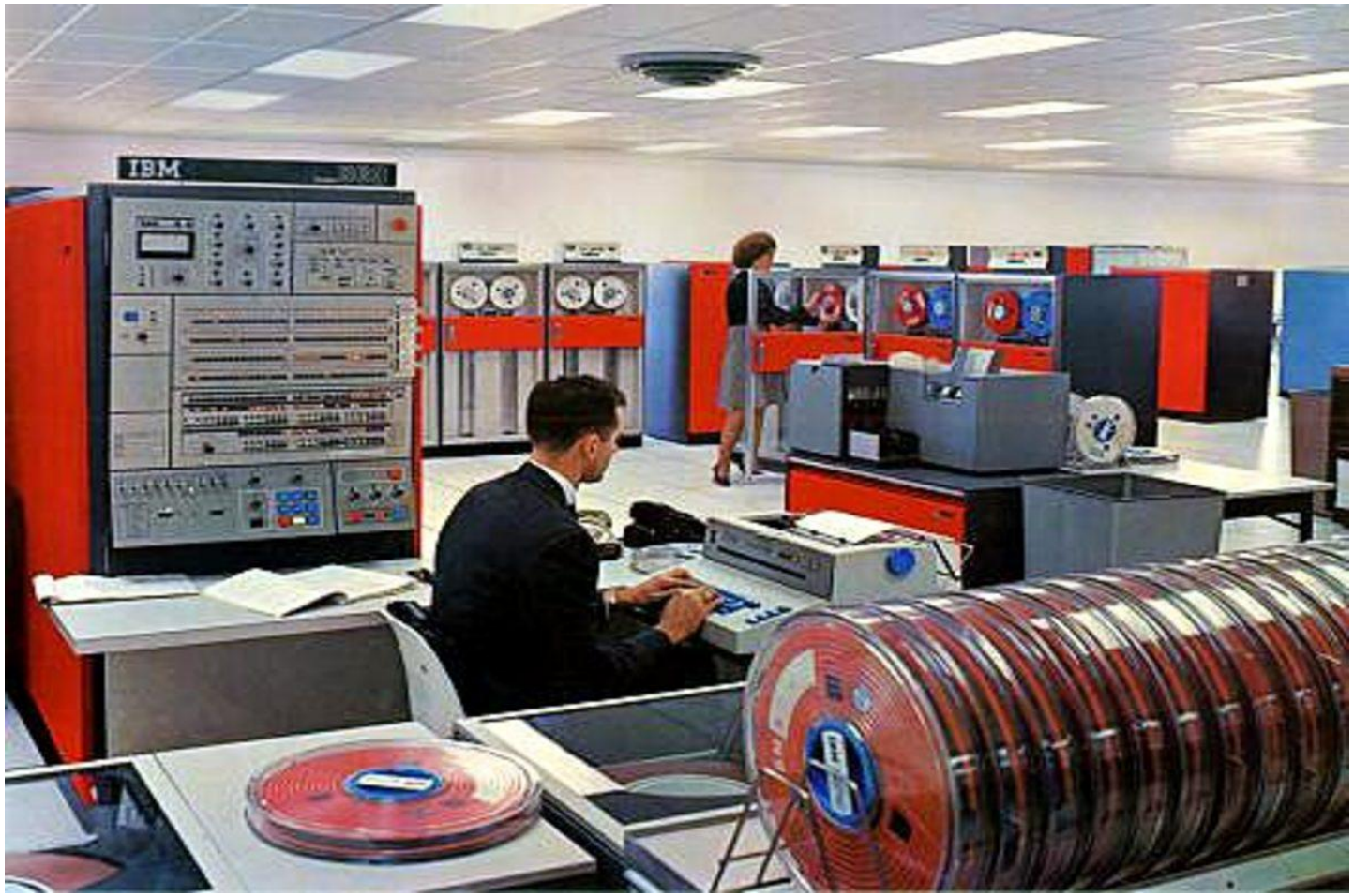
Matt Jarvis, Head of Community and Evangelism

- Head of Community and Evangelism at Mesosphere
- Building stuff with open source software for 15+ years
- Ops, Dev and Dev/Ops

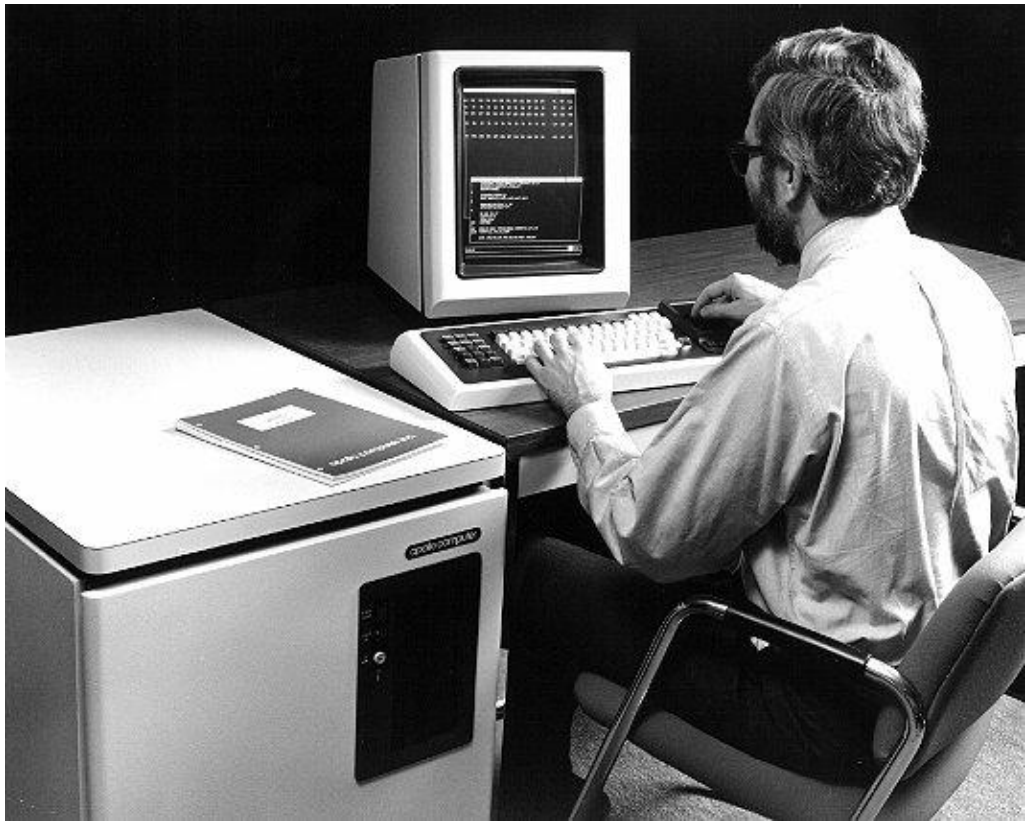


@mattj_io

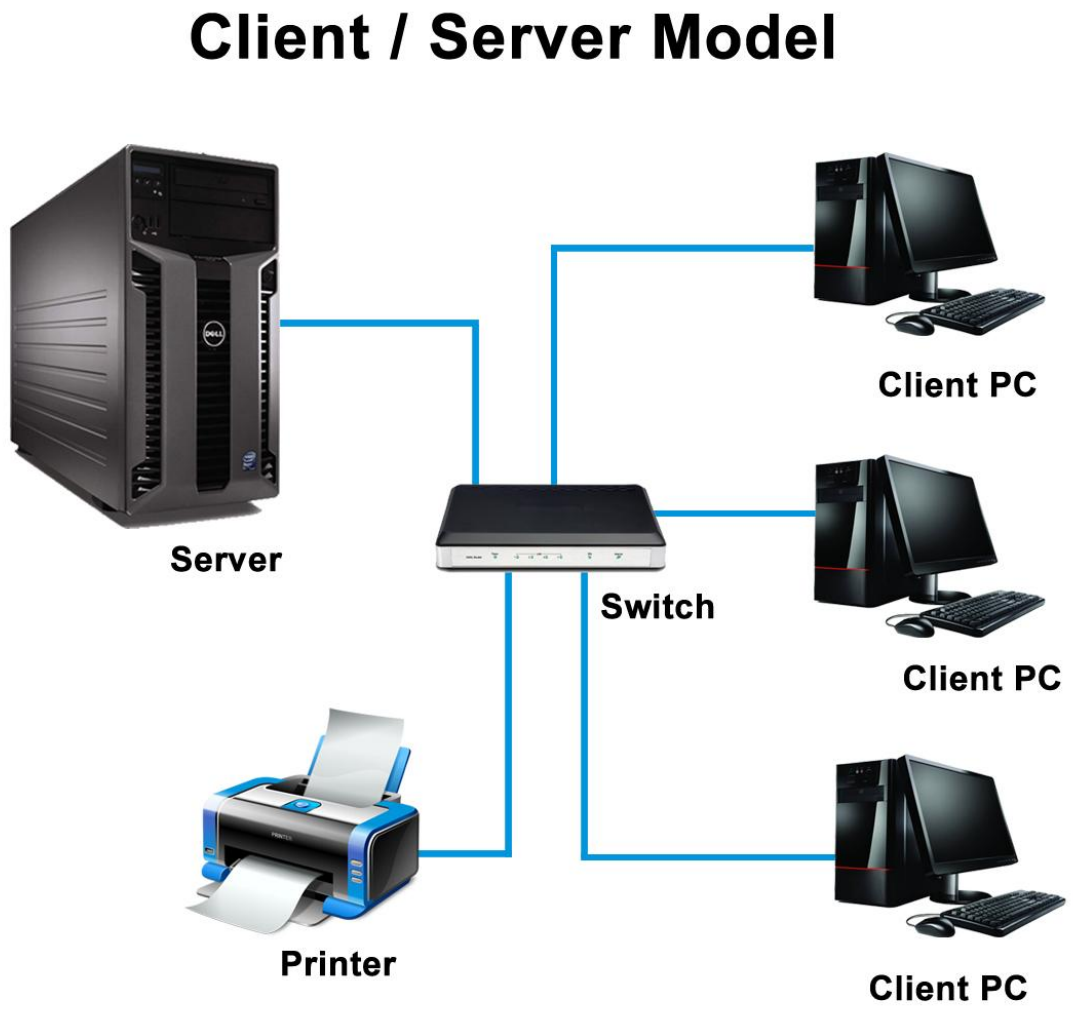
In the beginning there was ....



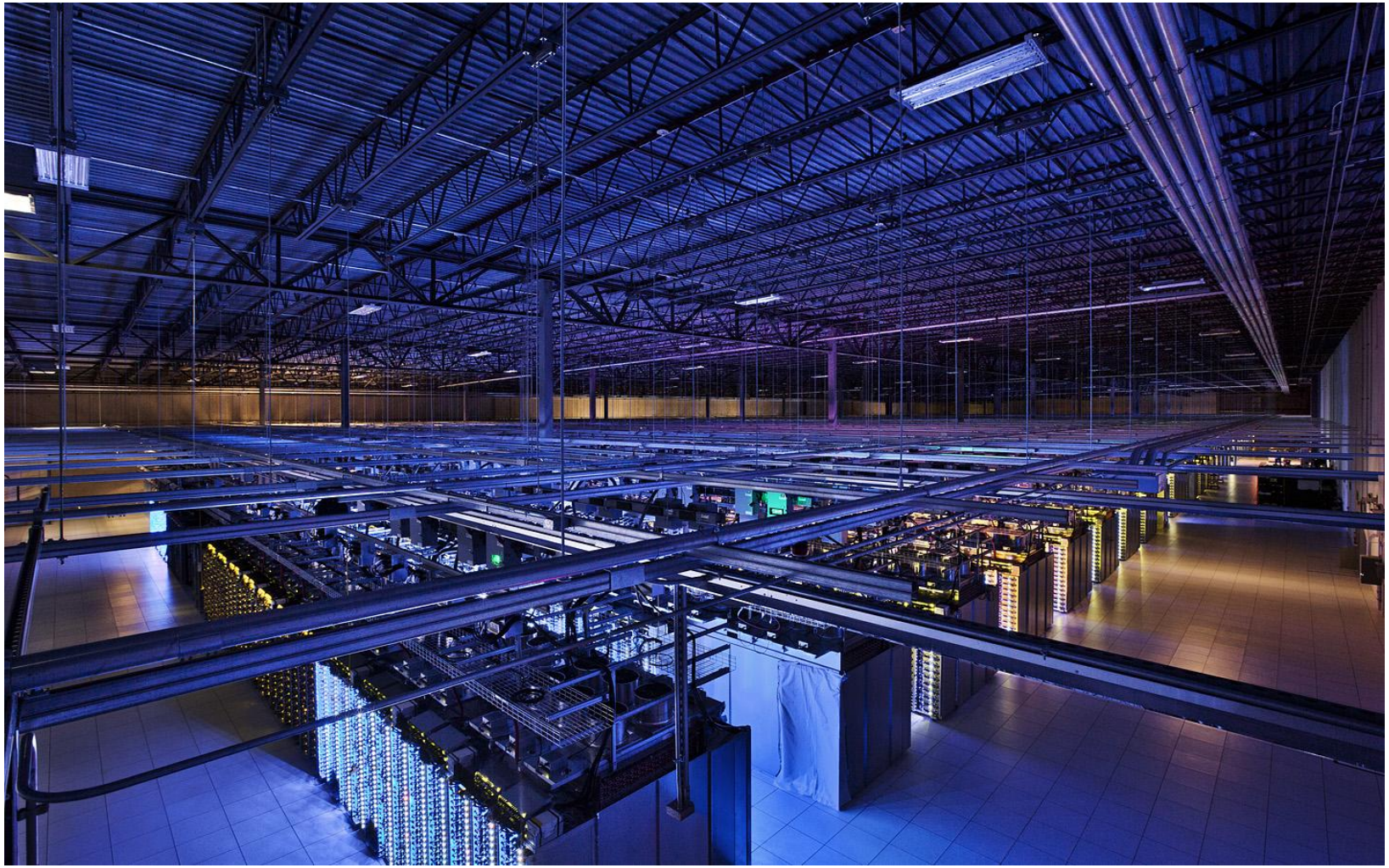
Things get smaller



Connect some smaller computers to the big one …



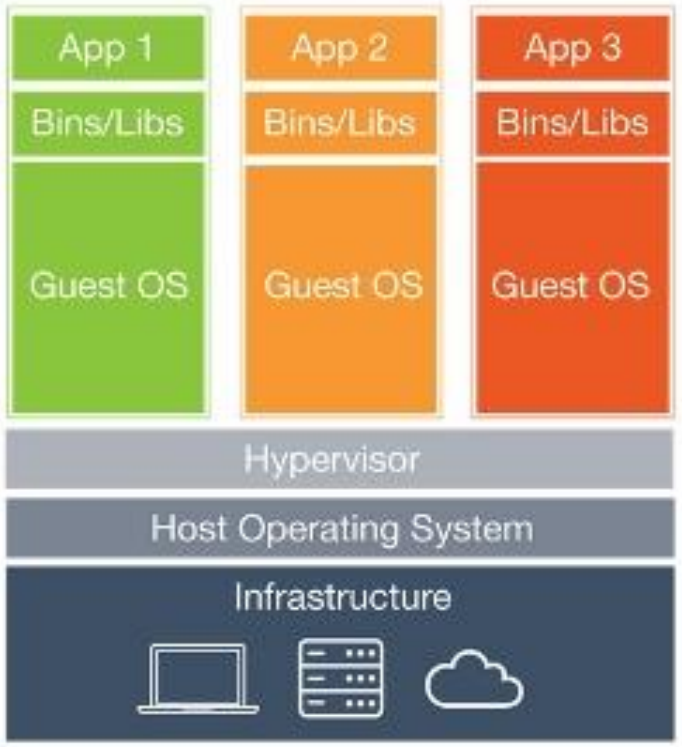
Scaling ...



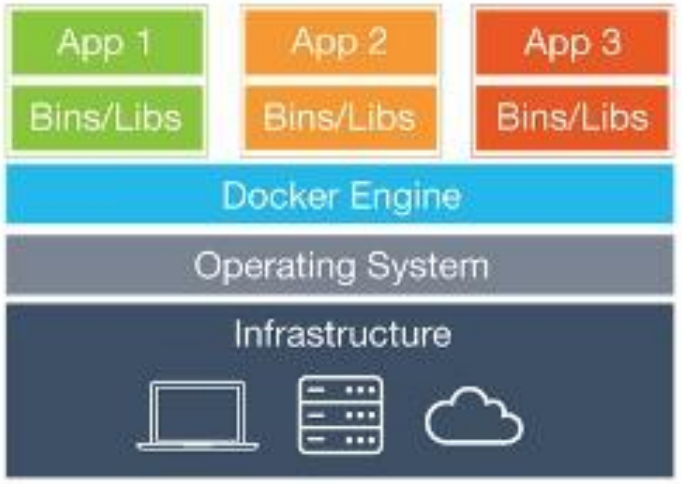
MapReduce is crunching data ..

First generation creates processing silos

Subdivision of infrastructure

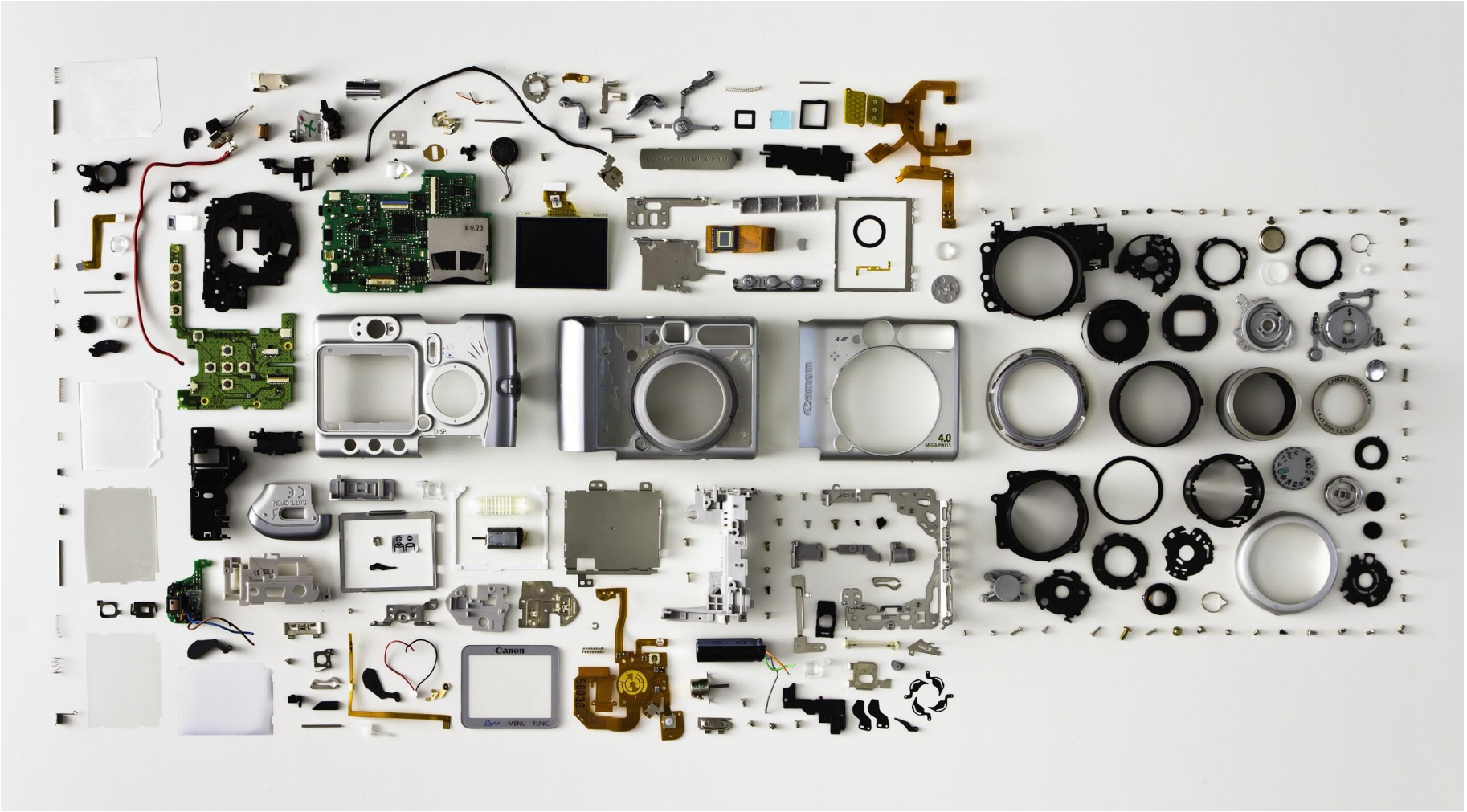


Virtual Machines



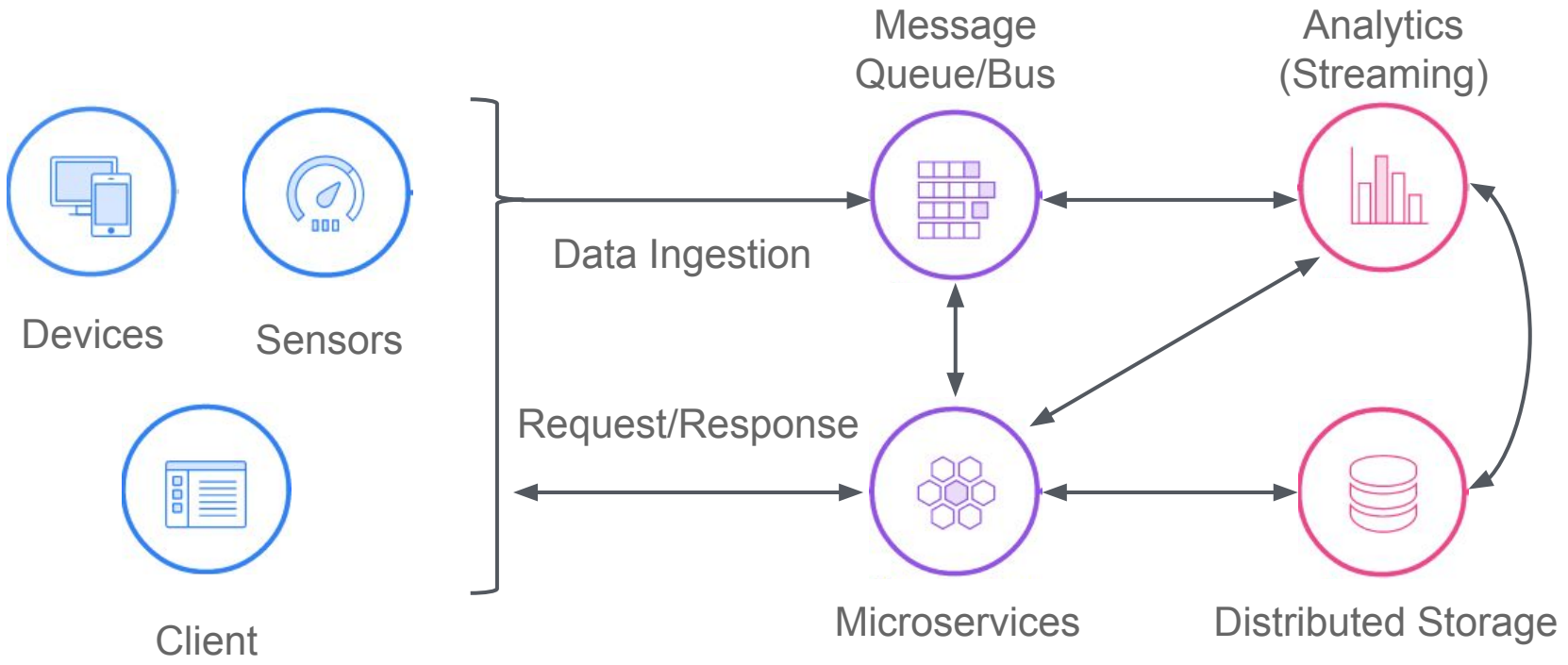
Containers

Splitting applications up



We need to turn faster ..

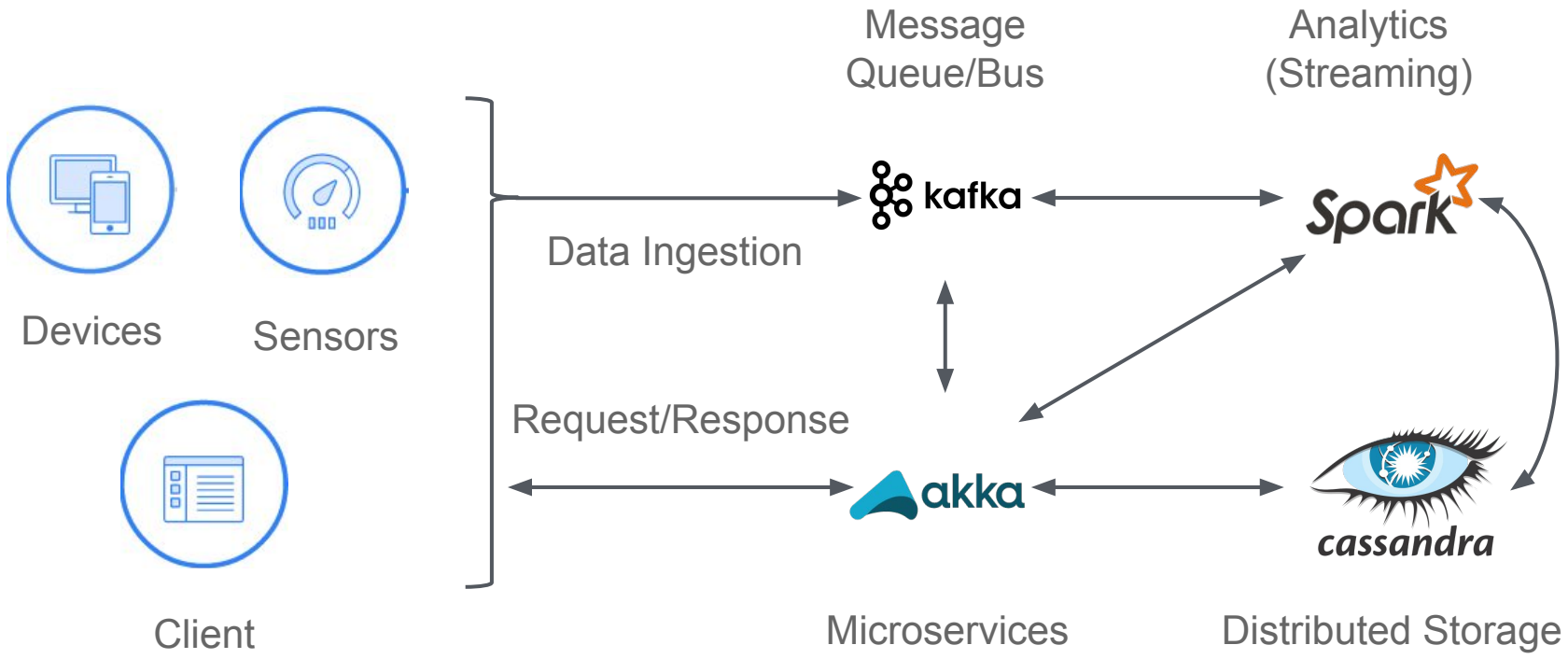
MODERN APPLICATION -> FAST DATA BUILT-IN



Use Cases:

- Anomaly detection
- Personalization
- IoT Applications
- Predictive Analytics
- Machine Learning

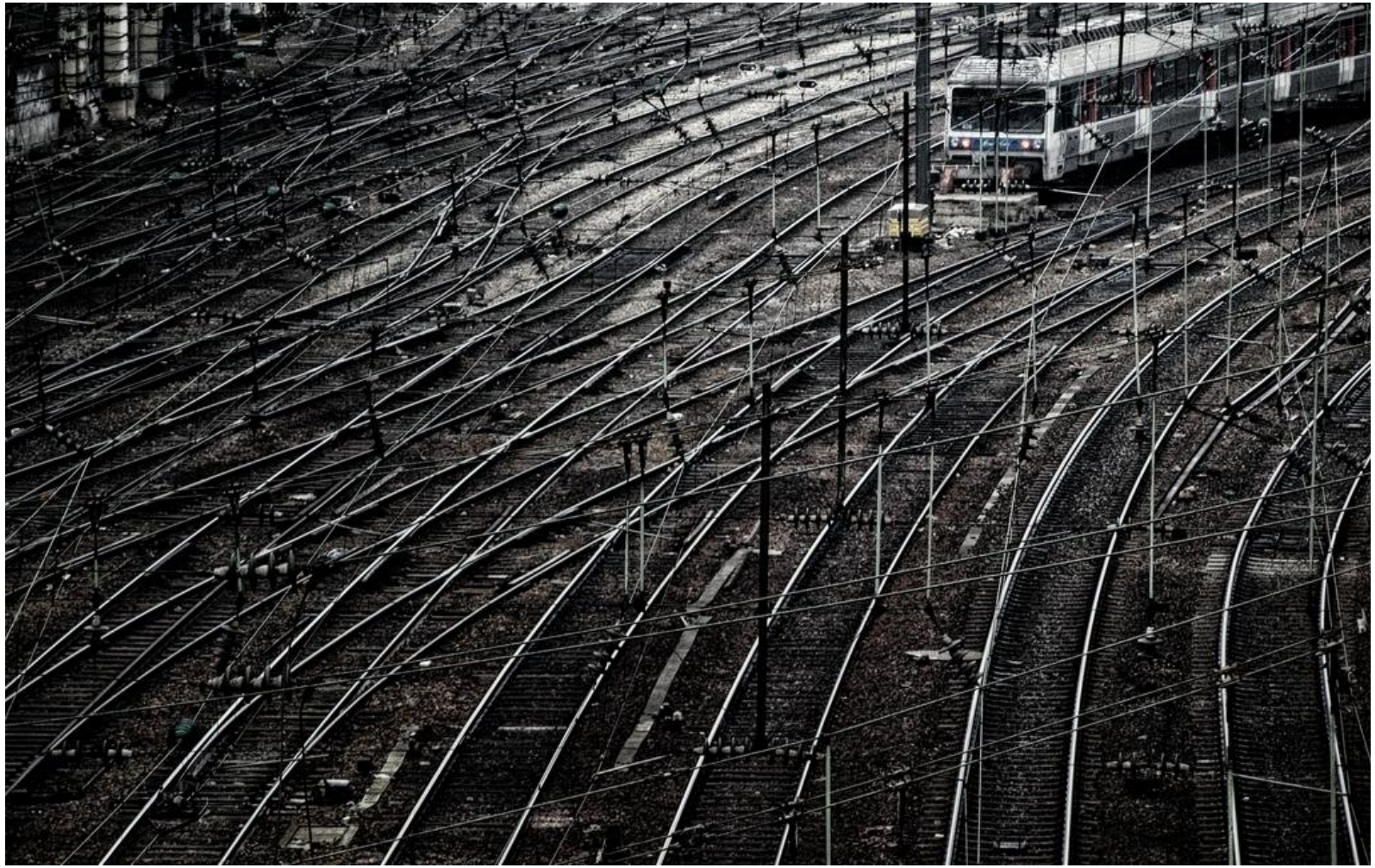
The SMACK Stack

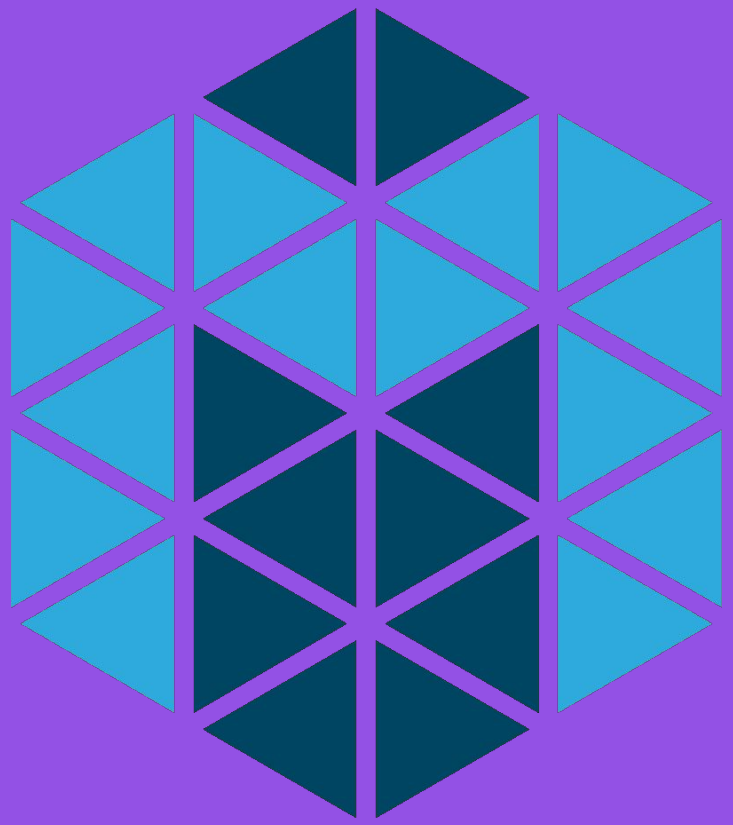


Use Cases:

- Anomaly detection
- Personalization
- IoT Applications
- Predictive Analytics
- Machine Learning

Complexity increases





Apache Mesos:
The datacenter kernel

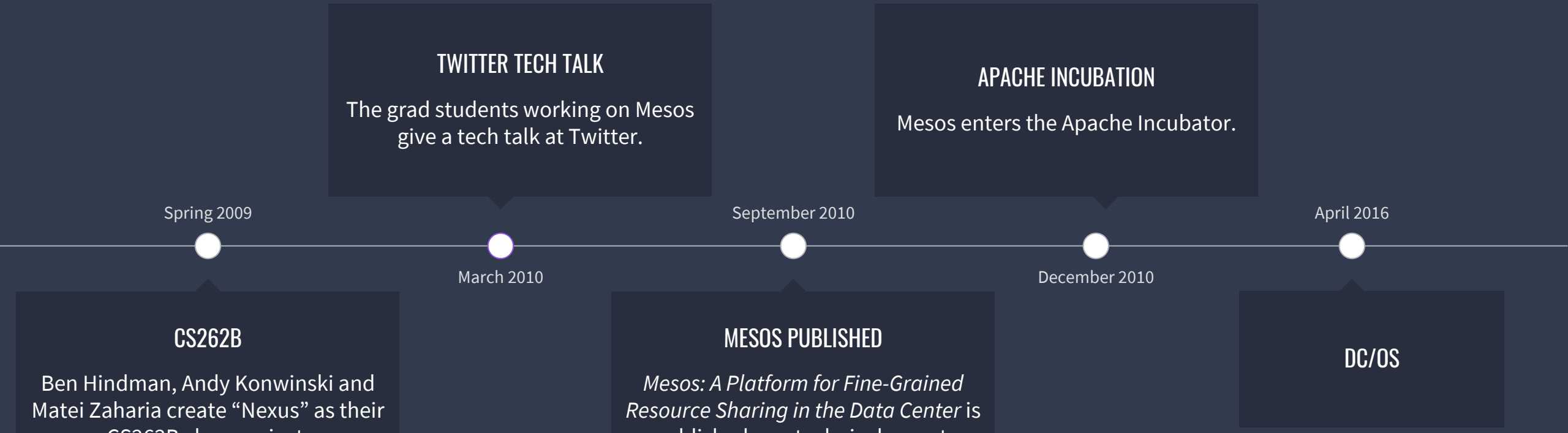
http://mesos.apache.org/

Building block of the modern internet

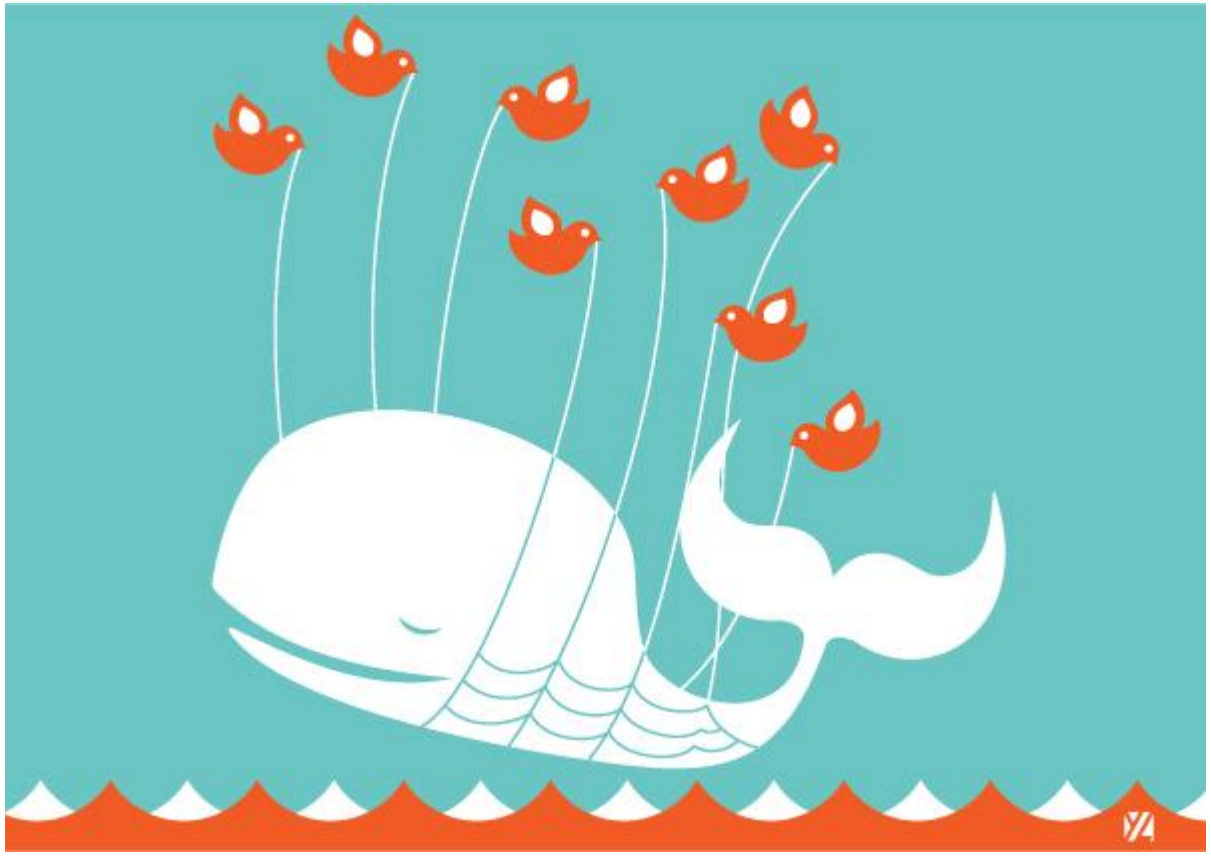
- A cluster resource negotiator
- A top-level Apache project
- Scalable to 10,000s of nodes
- Fault-tolerant, battle-tested
- An SDK for distributed apps
- Native Docker support

http://mesos.apache.org/documentation/latest/powered-by-mesos/

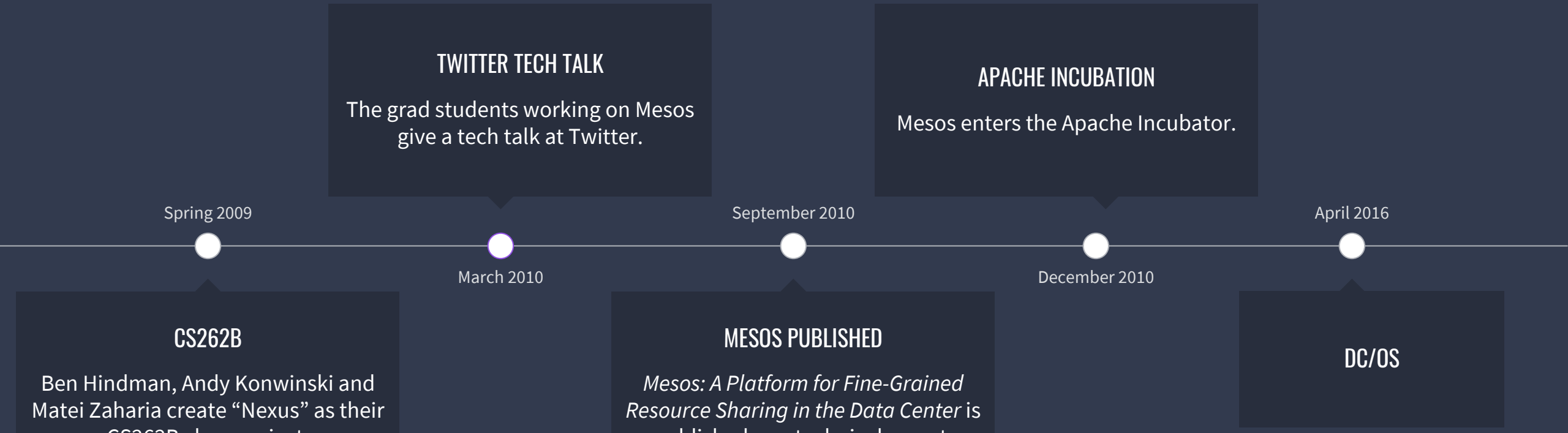
THE BIRTH OF MESOS



Solving the Fail Whale



THE BIRTH OF MESOS



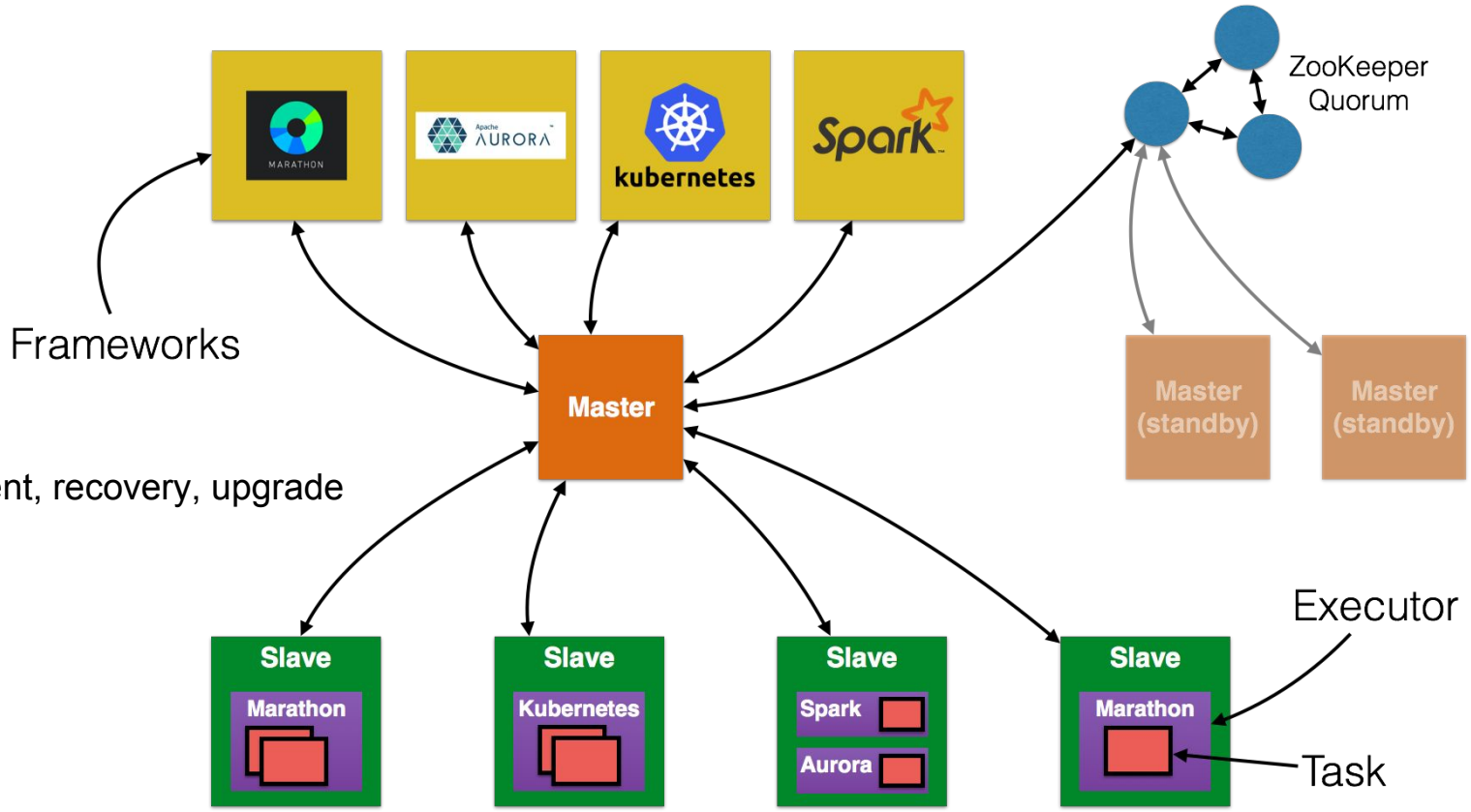
Two level scheduling

Mesos Master and Agents

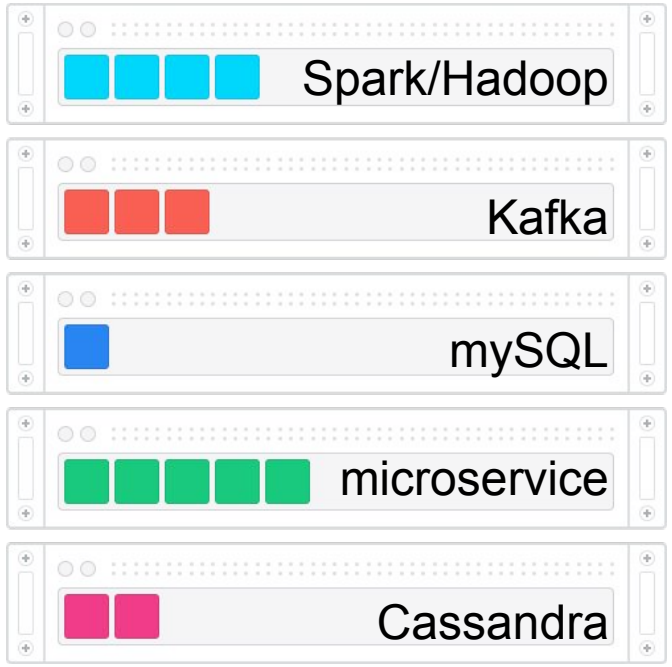
- Abstract resources into single pool
- Offers and tracks resources
- Guarantees isolation
- Handles workload restart on failure

Mesos Framework

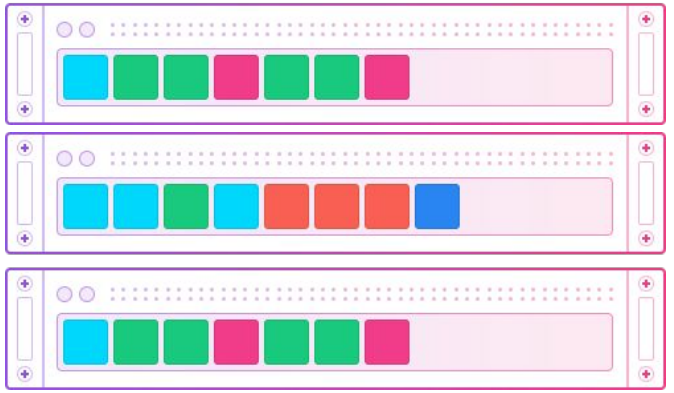
- Consumes resources
- Deploys tasks
- Provides application specific logic for deployment, recovery, upgrade



MULTIPLEXING OF DATA, SERVICES, USERS, ENVIRONMENTS



Typical Datacenter
siloed, over-provisioned servers, low utilization



Apache Mesos
automated schedulers, workload multiplexing onto the same machines

Apache MESOS | Frameworks | Agents | Roles | Offers | Maintenance | ejoseph-te4msh6

Master 62dff48e-dfaa-4309-94f0-73d5e94ab01e

Cluster: ejoseph-te4msh6
Leader: 10.0.5.237:5050
Version: 1.4.0
Built: 5 days ago by
Started: 53 minutes ago
Elected: 53 minutes ago

LOG

Agents

Activated	5
Deactivated	0
Unreachable	0

Tasks

Staging	0
Starting	0
Running	11
Unreachable	0
Killing	0
Finished	1
Killed	0

Active Tasks

Find...

Framework ID	Task ID	Task Name	Role	State	Started ▼	Host	
62dff48e-dfaa-4309-94f0-73d5e94ab01e-0001	bus-demo_dashboard.37943816-8677-11e7-b432-425ffcbc45b8	dashboard.bus-demo	slave_public	RUNNING	a minute ago	10.0.5.101	Sandbox
62dff48e-dfaa-4309-94f0-73d5e94ab01e-0001	bus-demo_ingest.0999da65-8676-11e7-b432-425ffcbc45b8	ingest.bus-demo	slave_public	RUNNING	9 minutes ago	10.0.1.204	Sandbox
62dff48e-dfaa-4309-94f0-73d5e94ab01e-0004	broker-2__581647a0-6953-4cfe-af96-356d04535c38	broker-2	kafka-role	RUNNING	12 minutes ago	10.0.3.240	Sandbox
62dff48e-dfaa-4309-94f0-73d5e94ab01e-0004	broker-1__d24b1885-860b-4ae9-9feb-502ffcded5fe	broker-1	kafka-role	RUNNING	13 minutes ago	10.0.3.7	Sandbox
62dff48e-dfaa-4309-94f0-73d5e94ab01e-0004	broker-0__eb077cd0-f416-4918-9cbd-1f5b1ea8c10d	broker-0	kafka-role	RUNNING	13 minutes ago	10.0.1.204	Sandbox
62dff48e-dfaa-4309-94f0-73d5e94ab01e-0001	kafka.8a668774-8675-11e7-b432-425ffcbc45b8	kafka	slave_public	RUNNING	13 minutes ago	10.0.0.68	Sandbox
62dff48e-dfaa-4309-94f0-73d5e94ab01e-0003	node-2__a9c29921-d7c1-4a32-8eb5-4fd37b25665d	node-2	cassandra-role	RUNNING	14 minutes ago	10.0.3.7	Sandbox

DC/OS brings it all together

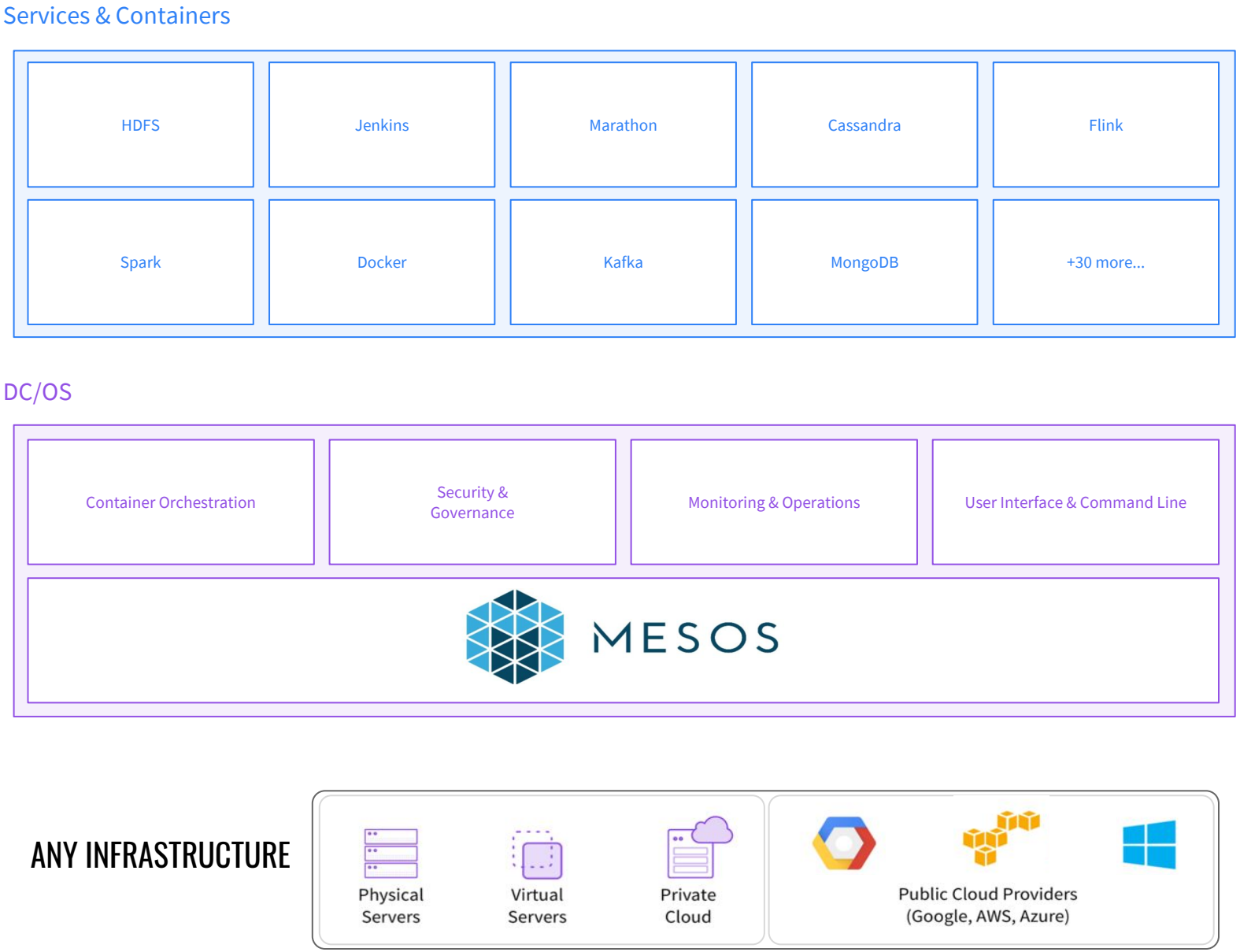
- Service Discovery
- Load Balancing
- Security
- Ease of installation
- Comprehensive tooling for operations
- Built in frameworks for long running and scheduled jobs
- Catalog of pre-configured apps (including Apache Spark, Apache Kafka…), browse at http://universe.dcos.io/
- And much more https://dcos.io/

DC/OS

DC/OS is ...

- 100% open source (ASL2.0)
  - A big, diverse community
- An umbrella for ~30 OSS projects
  - Roadmap and designs
  - Documentation and tutorials
- Not limited in any way

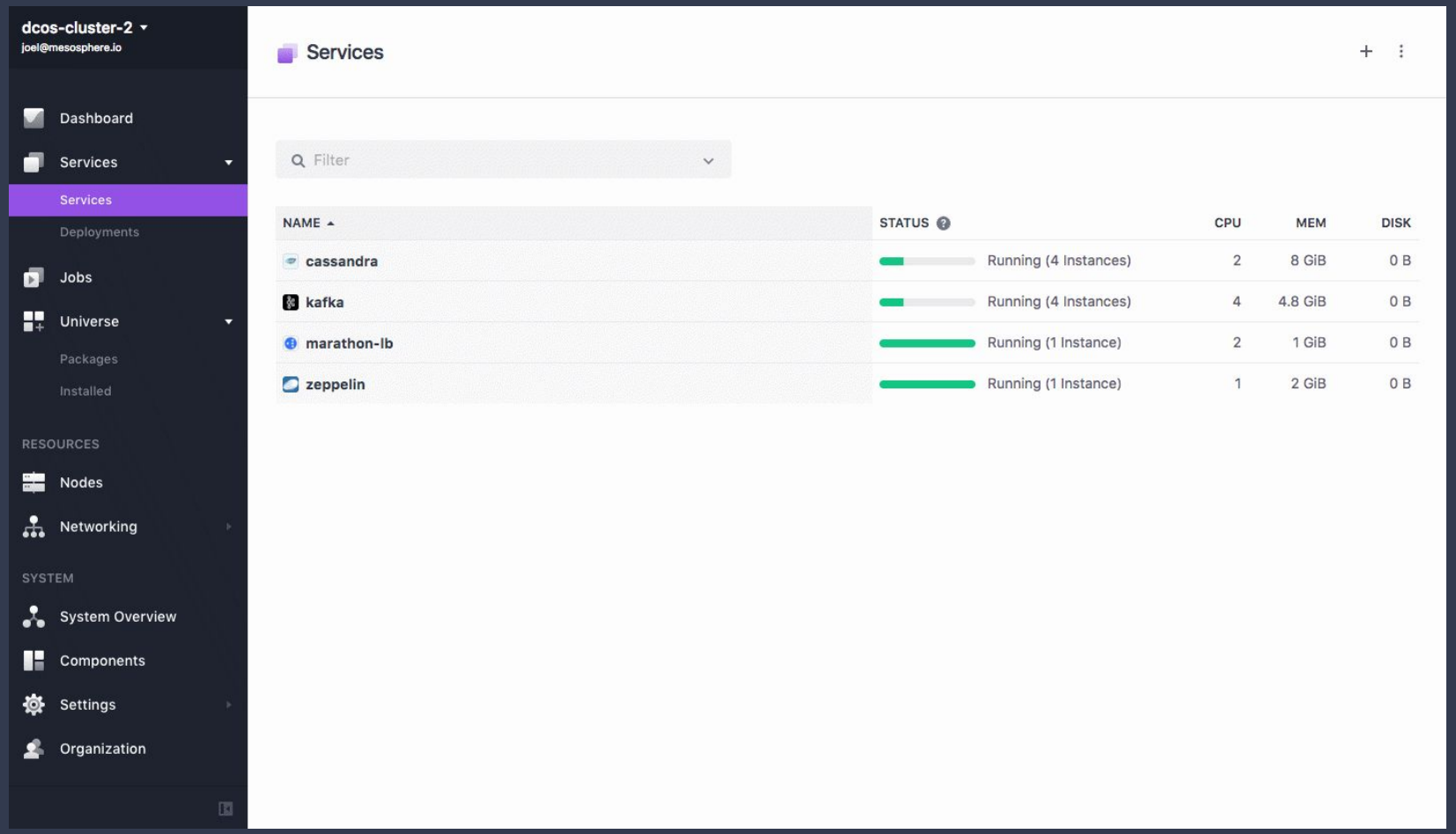
DC/OS Architecture Overview



Interact with DC/OS (1/2)

Web-based GUI

https://dcos.io/docs/latest/usage/webinterface/



Interact with DC/OS (2/2)

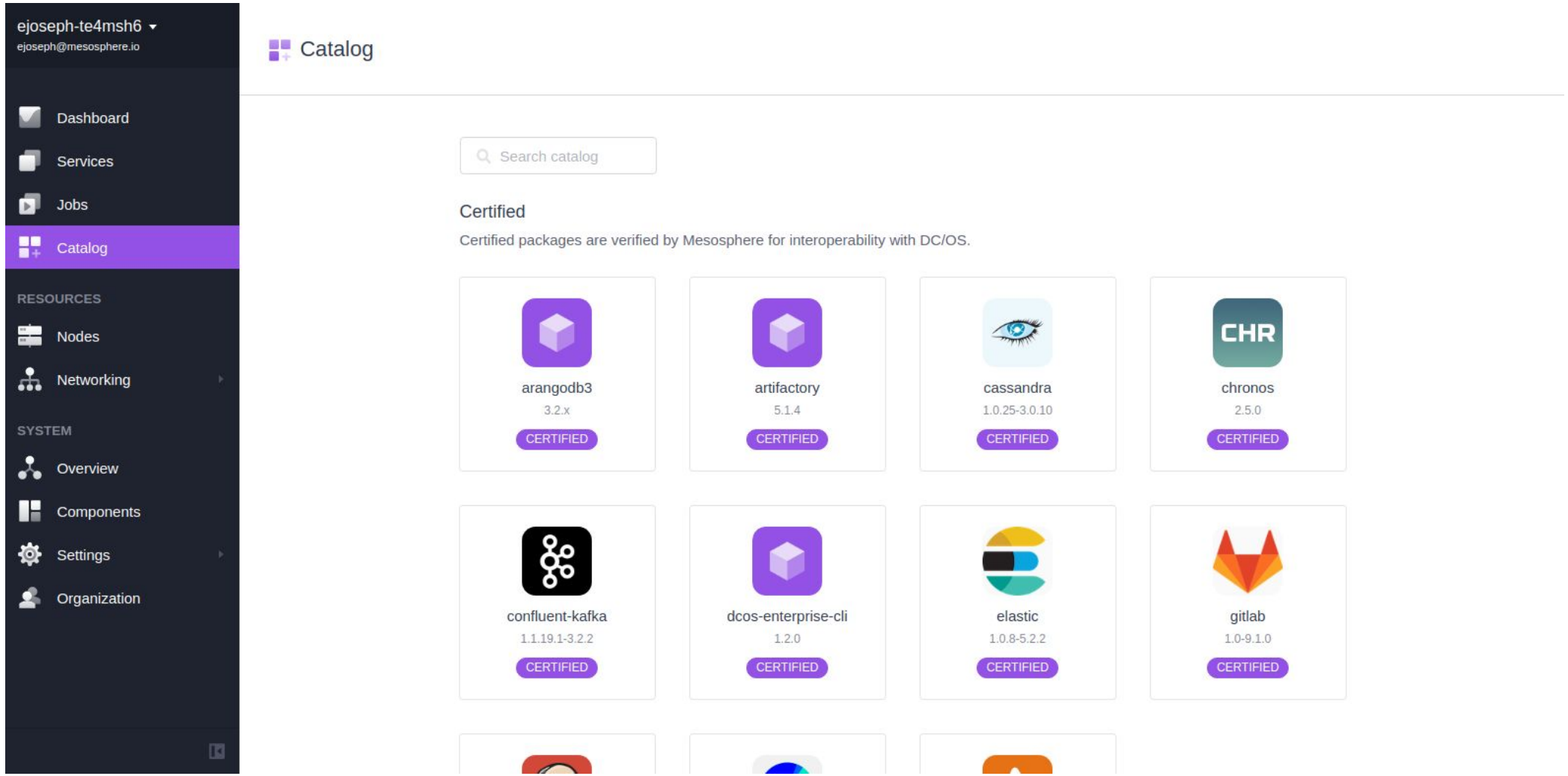
CLI tool

https://dcos.io/docs/latest/usage/cli/

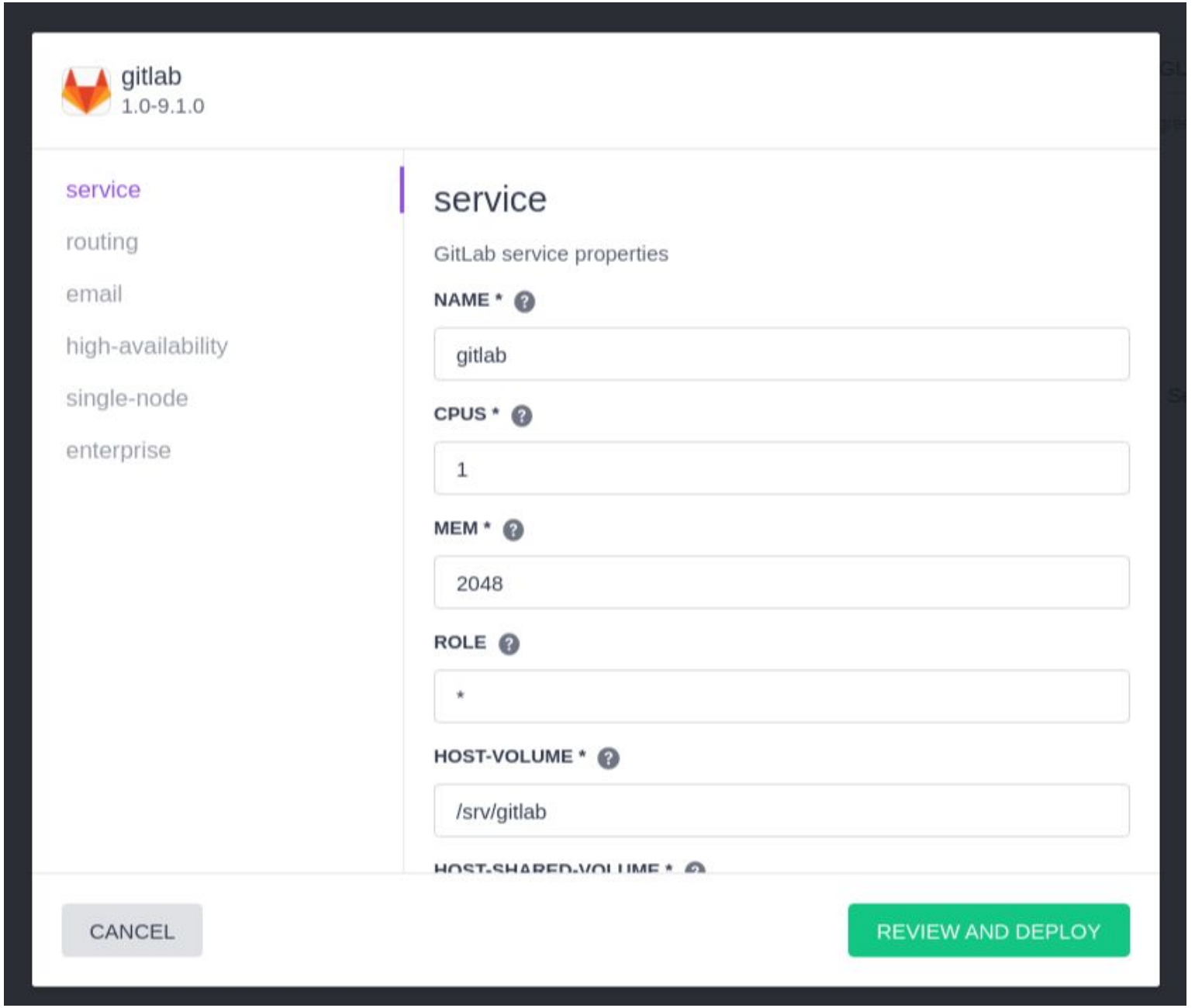
API

https://dcos.io/docs/latest/api/

Catalog of Applications



Install an Application



Application JSON

```
{
  "service": {
    "name": "gitlab",
    "cpus": 1,
    "mem": 2048,
    "role": "*",
    "host-volume": "/srv/gitlab",
    "host-shared-volume": "/srv/gitlab-data"
  },
  "routing": {
    "https-redirect": false,
    "ssh-port": 22222,
    "registry-port": 50000
  },
  "email": {
    "enabled": false,
    "port": 25,
    "authentication": "login",
    "enable-starttls-auto": true,
    "openssl-verify-mode": "peer",
    "tls": false
  },
  "high-availability": {
    "enabled": false,
    "postgres": {},
    "redis": {}
  },
  "single-node": {
    "local-volumes": {},
    "external-volumes": {
      "enabled": false
    }
  },
  "enterprise": {
    "enterprise-edition": false
  }
}
                                                          1,1           All
```

Service Discovery

Critical to distributed systems since container can be spawned anywhere

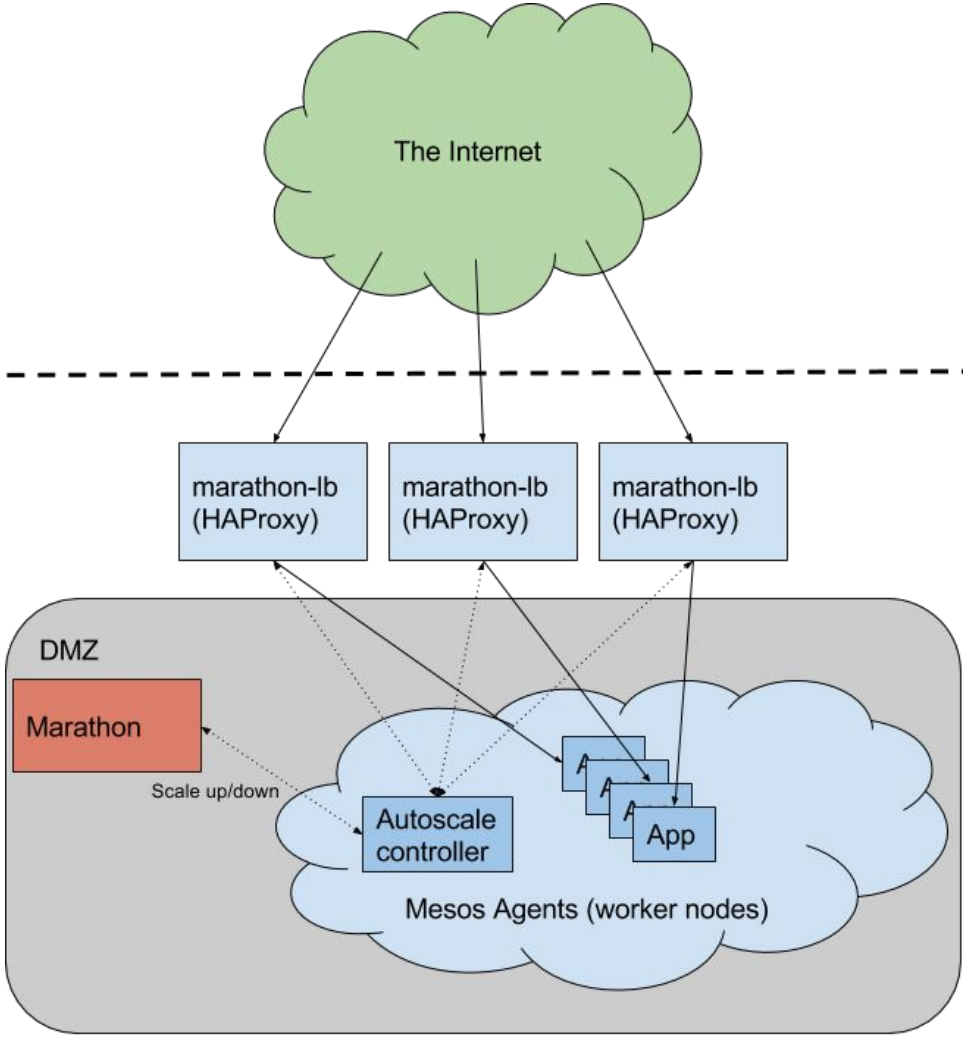
Mesos DNS

- Each instance of a service given a DNS entry in pattern task.scheduler.mesos eg. myapp.marathon.mesos
- Basic round robin load balancing
- Requires an A and SRV lookup

Named VIPS

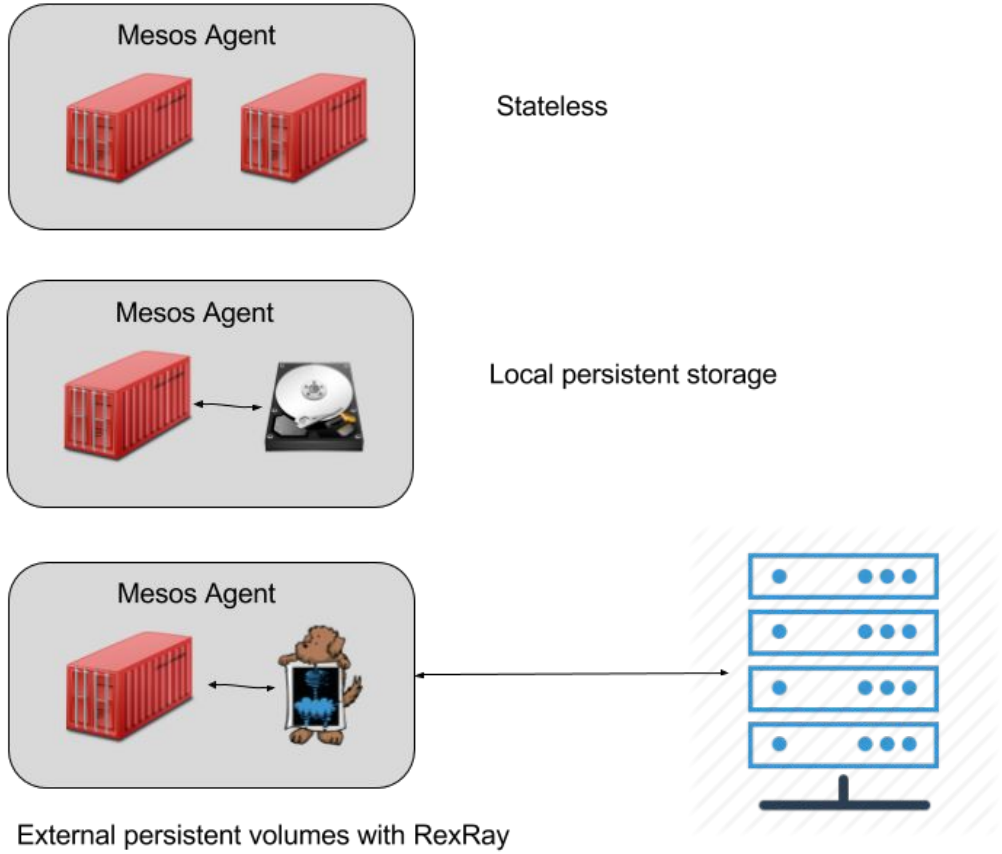
- Service allocated name based virtual IP *eg.* redis.marathon.l4lb.thisdcos.directory:6379
- Very high performance - integrates with connection tracking table in kernel for real address resolution
- Also provides low cost East/West load balancing
- Uses gossip protocol to propagate between nodes
- ~100ms update times

Load balancing - MarathonLB



- Based on HAProxy
- Ingests state of running applications
- Regenerates HAProxy configuration
- Can be North/South or East/West

Integrated storage options



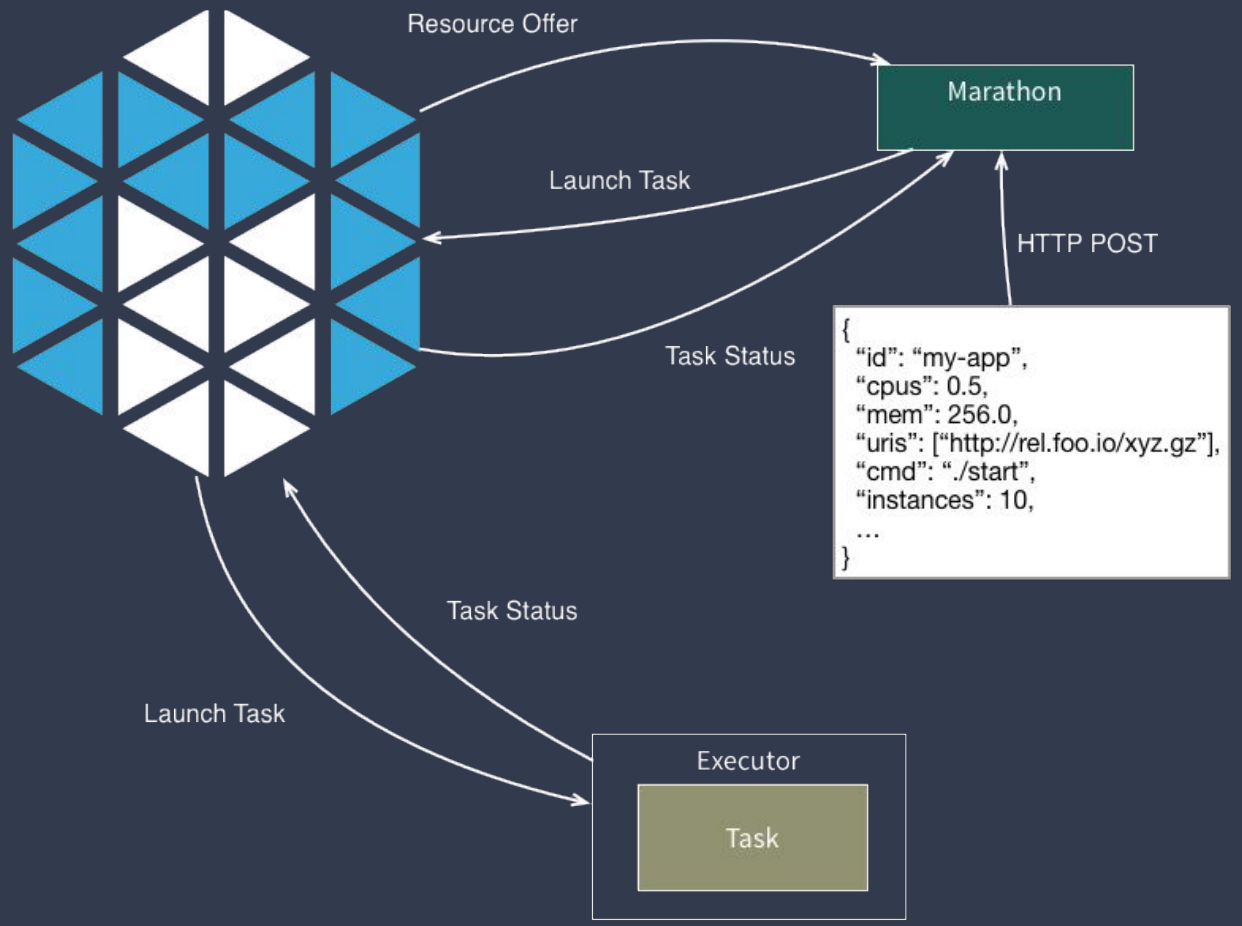
CONTAINER STORAGE INTERFACE

Networking

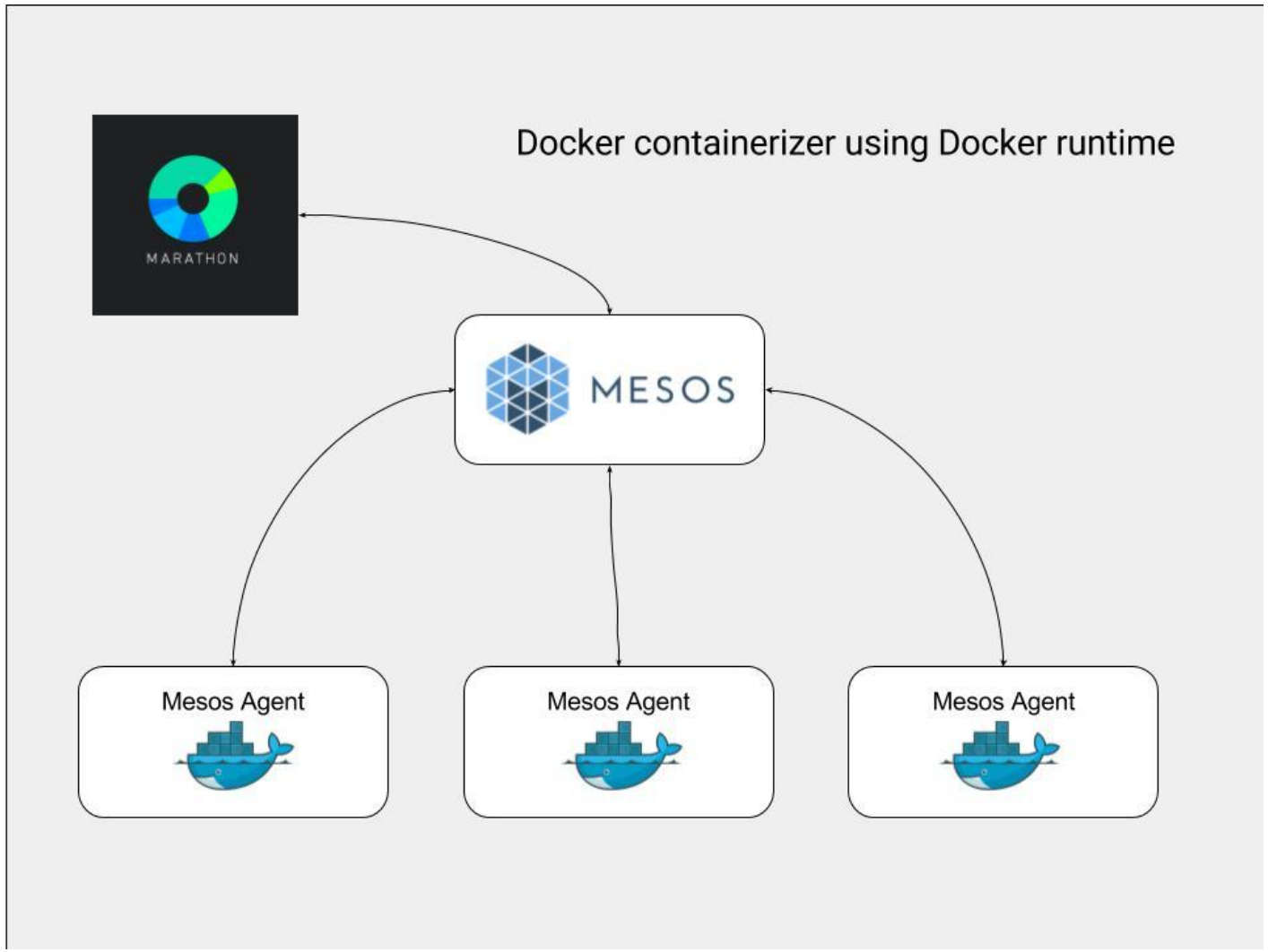
- Integrated VXLAN based virtual networks
- CNI compatible upstream integrations - Calico, OpenContrail etc.

Marathon

- Mesos can’t run applications on its own.
- A Mesos framework is a distributed system that has a scheduler.
- Schedulers like Marathon start and keep your applications running. A bit like a distributed init system.
- Learn more at https://mesosphere.github.io/marathon/

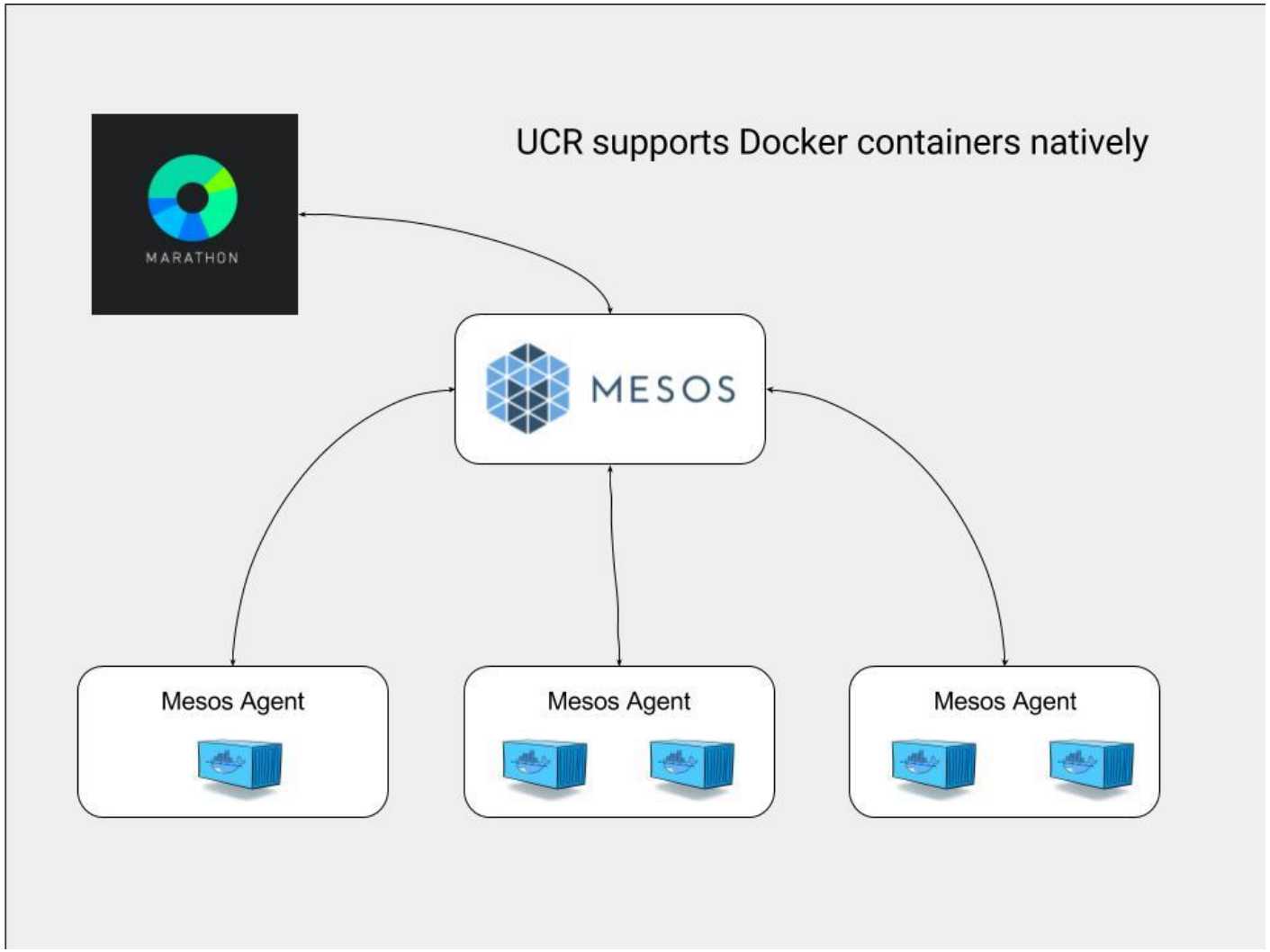


Mesos Docker containerizer



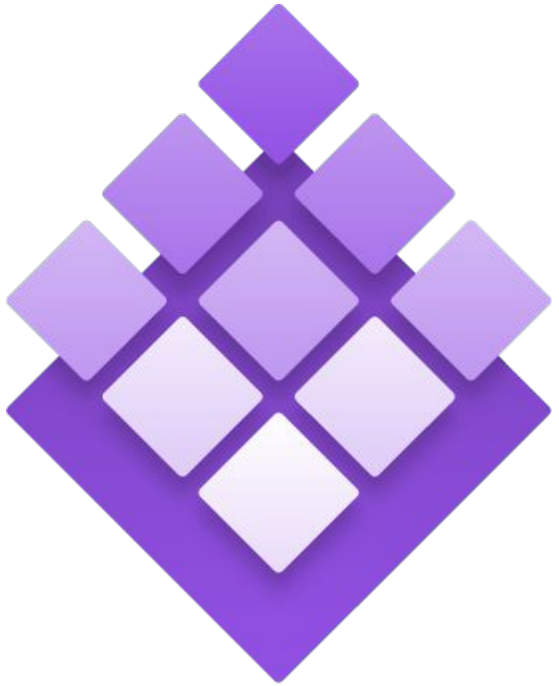
DC/OS

Mesos Universal Container Runtime - no Docker runtime



DC/OS

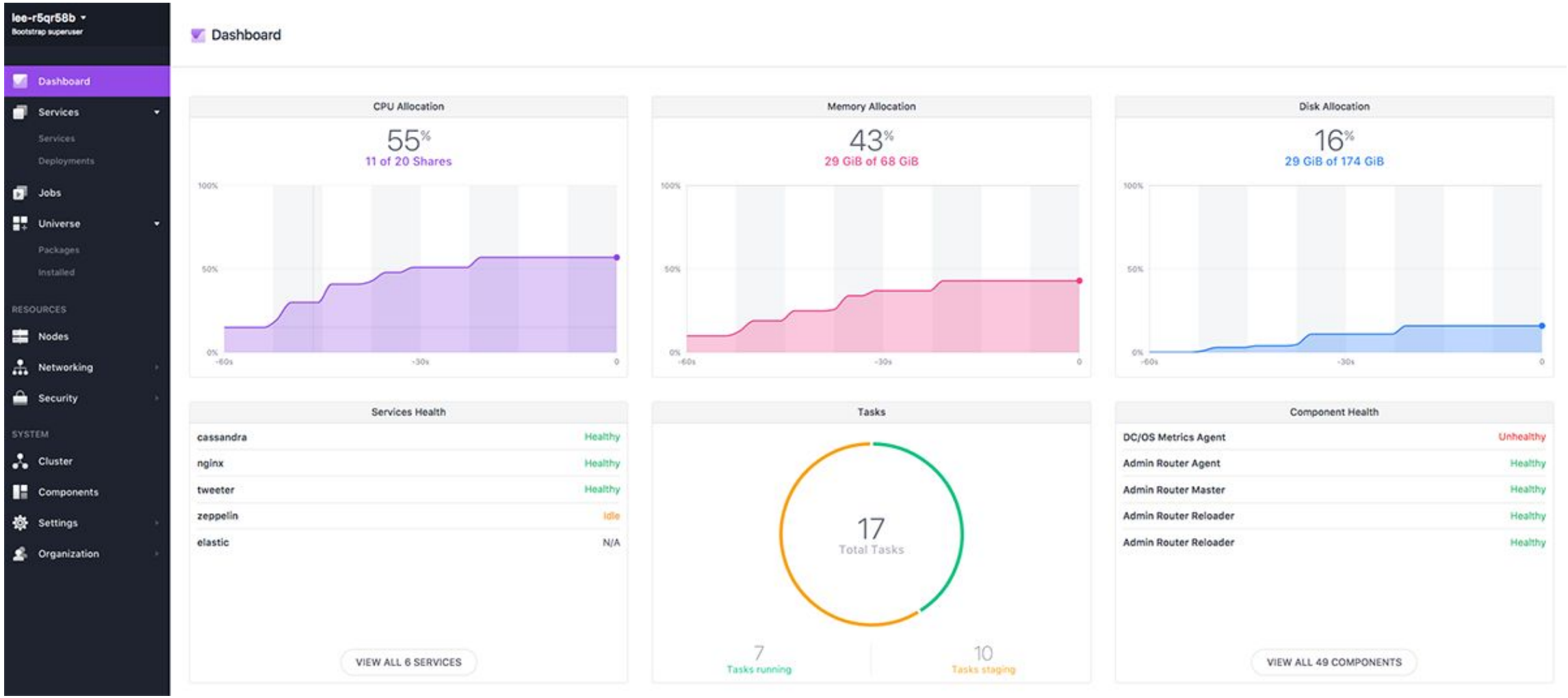
Kubernetes



- Beta framework for running Kubernetes clusters
- Multiple clusters within same DC/OS cluster
- Multiple versions within same DC/OS cluster

DC/OS

Centralised operations



Centralised operations

Framework integrations
- CLI extensions
- Integrated UI's

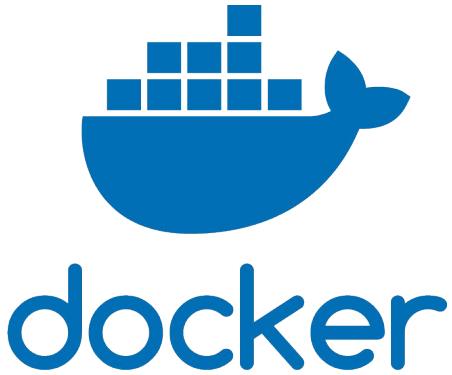
Logging aggregation
- Available through API, CLI and UI
- Integration with ELK and others

Metrics aggregation
- Available through UI, CLI and API
- Integrates with Prometheus, InfluxDB etc.

Single pane of glass across all frameworks

```
dcos kafka topic create topic1 --partitions 1 --replication 1
```

Try it out !



@dcos

chat.dcos.io

users@dcos.io

/dcos
/dcos/examples
/dcos/demos

Questions?

Matt Jarvis
Twitter: @mattj-io
Email: mjarvis@mesosphere.com
https://dcos.io

The SMACK Stack

Evolution of Data Analytics

Days *Hours* *Minutes* *Seconds* *Microseconds*

Batch → **Micro-Batch** → **Event Processing**

Reports what has happened using descriptive analytics

Solves problems using predictive and prescriptive analytics

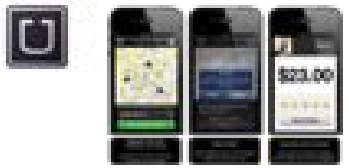
Billing, Chargeback

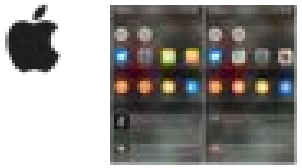
Product recommendations

Real-time Pricing and Routing

Real-time Advertising

Predictive User Interface





DEMO

Financial Transaction Processing using Apache Kafka and Apache Flink