





Resource Sharing Beyond Boundaries


Mohit Soni Santosh Marella

A Mesos framework for scaling a YARN cluster (Incubating into Apache)
<https://issues.apache.org/jira/browse/MYRIAD> — Edit

 262 commits

 8 branches

 0 releases

 15 contributors



Branch: **phase1** ▾

myriad / +

 Code

 Pull requests 7

 Wiki

Contributors

Traffic

Commits

Code frequency

Punch card

Network

Members



Contributors

Traffic

Commits

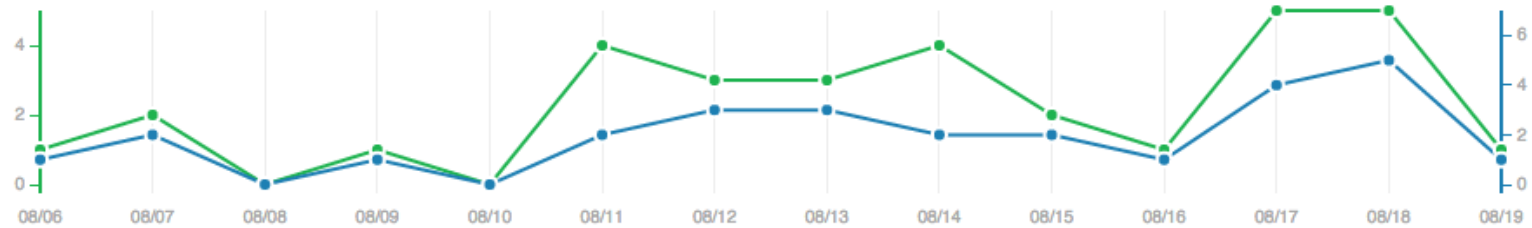
Code frequency

Punch card

Network

Members

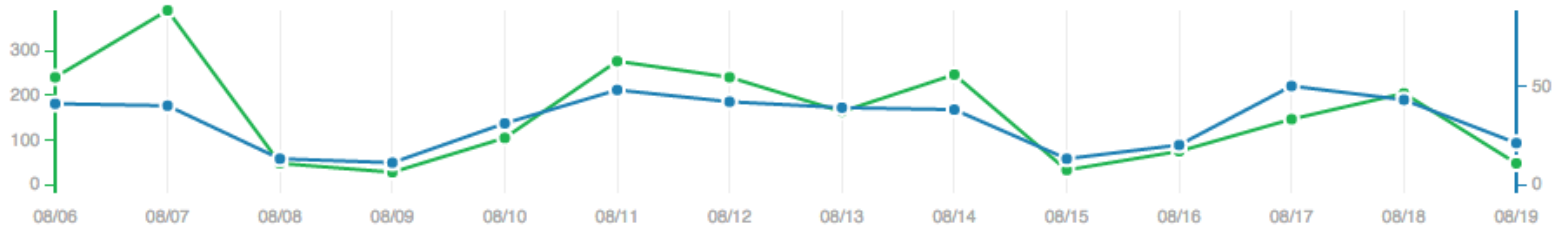
Git clones



33
Clones

25
Unique cloners

Visitors



2,391
Views

344
Unique visitors



Adam Bordelon

Brandon Gulla

Jim Klucar

Luciano Resende

Renan DelValle

Swapnil Daingade

Yuliya Feldman

Anoop Dawar

Danese Cooper

Kannan Rajah

Meghdoot Bhattacharya

Ruth Harris

Ted Dunning

Zhongyue Luo

Ben Hindman

Darin Johnson

Ken Sipe

Paul Reed

Shingo Omura

Will Ochandarena

Agenda

- What's up with Datacenters these days?
- Apache Mesos vs. Apache Hadoop/YARN?
- Why would you want/need both?
- Resource Sharing with Apache Myriad

What's running on your datacenter?

- Tier 1 services
- Tier 2 services
- High Priority Batch
- Best Effort, backfill







Requirements

- Programming models based on resources, not machines
- Custom resource types
- Custom scheduling algorithms:
Fast vs. careful/slow
- Lightweight executors, fast task launch time
- Multi-tenancy, utilization, strong isolation

Hadoop and More

- Support Hadoop/BigData ecosystem
- Support arbitrary (legacy) processes/containers
- Connect Big Data to non-Hadoop apps, share data, resources

Mesos from 10,000 feet

-  Open Source Apache project
-  Cluster Resource Manager
-  Scalable to 10,000s of nodes
-  Fault-tolerant, no SPOF
-  Multi-tenancy, Resource Isolation
-  Improved resource utilization

Mesos is more than Yet Another Resource Negotiator

- ≡ Long-running services; real-time jobs
- ▣ Native Docker; cgroups for years;
Isolate cpu/mem/disk/net/other
- 📄 Distributed systems SDK;
~200 loc for a new app
- </> Core written in C++ for performance,
Apps in any language



airbnb™

ebay

vimeo



NETFLIX



HubSpot



Time
Warner
Cable®



ignidata
igniting business with data

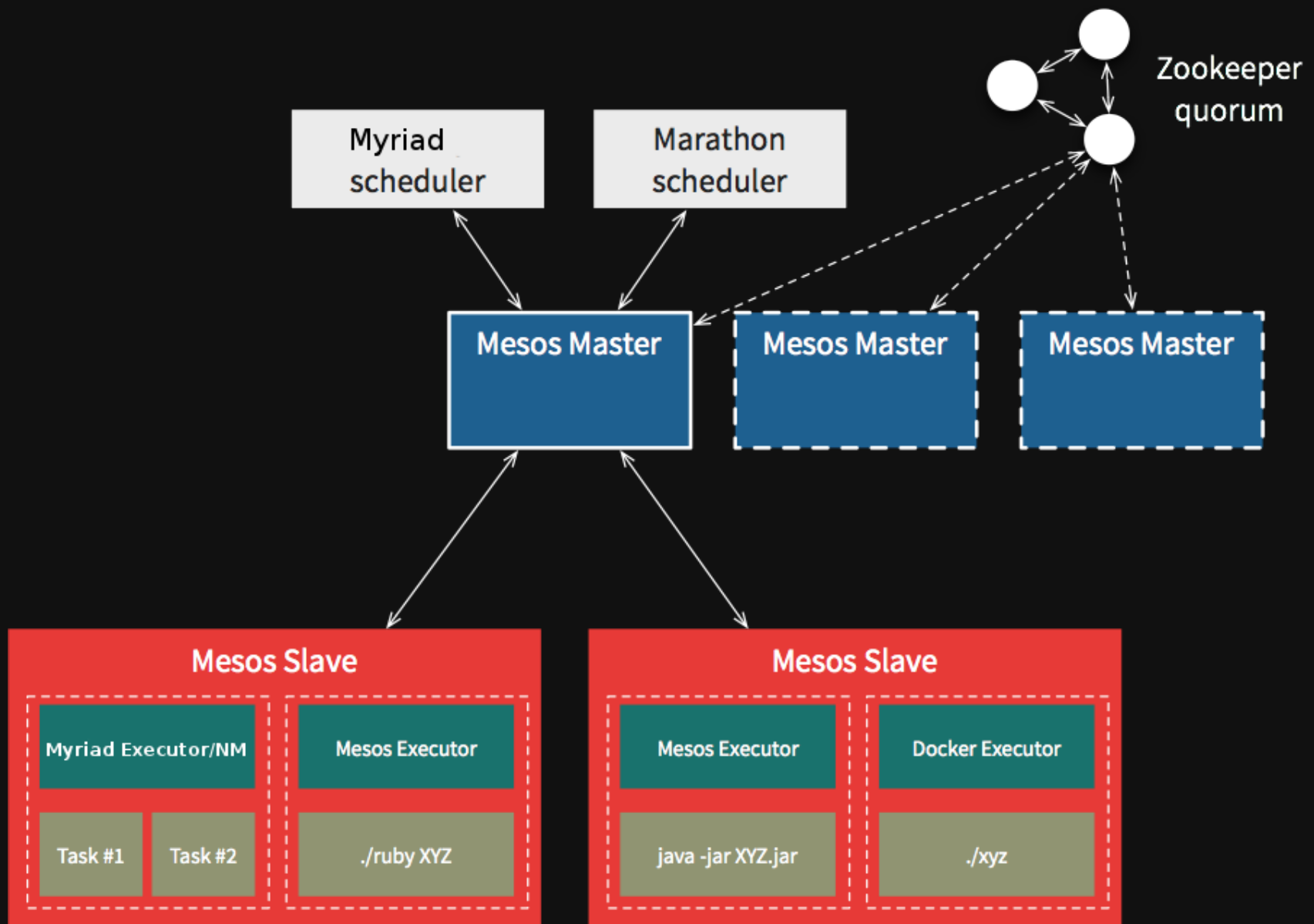
xogito
...radical thinking...

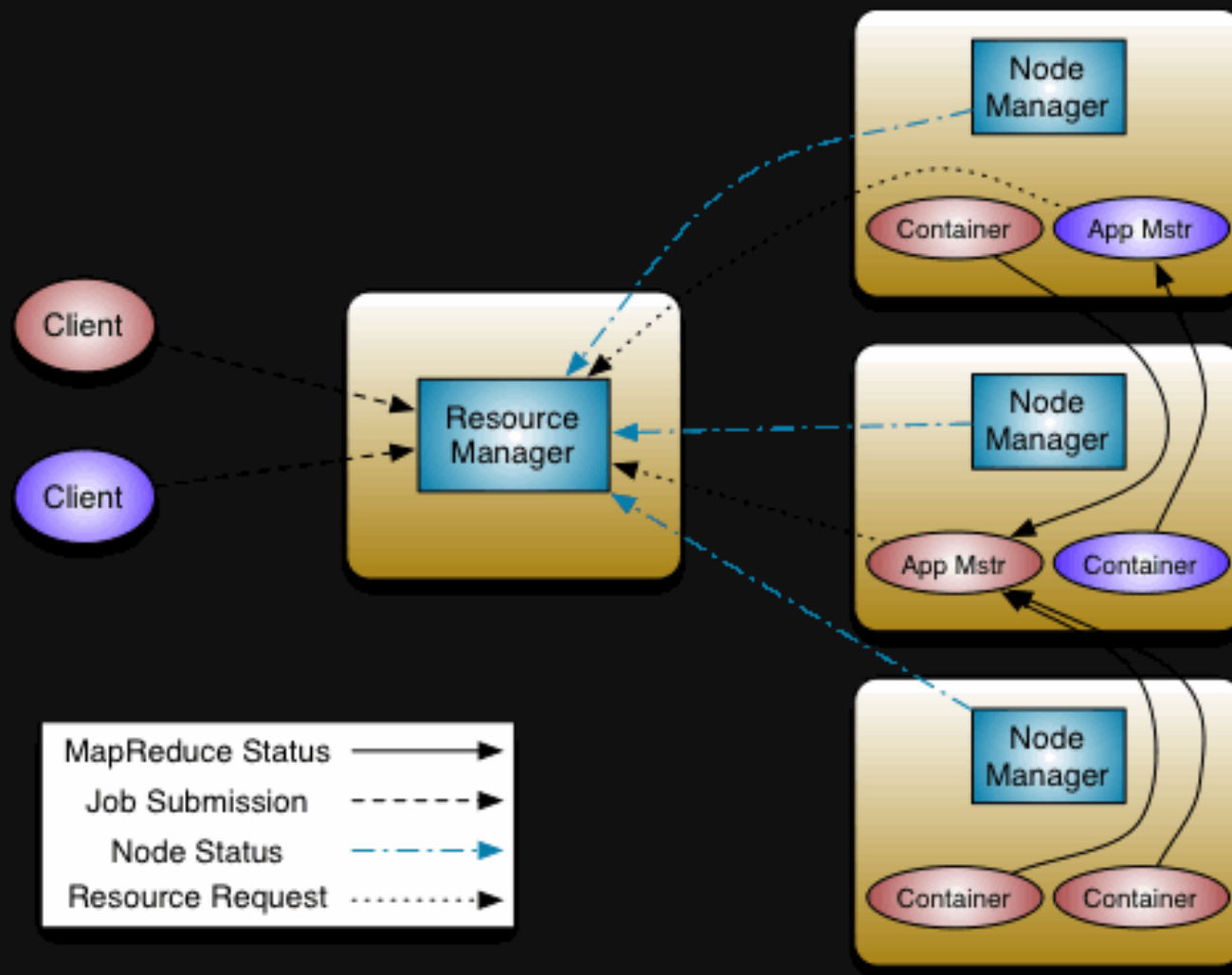
PayPal™



Atlassian







Why two resource managers?

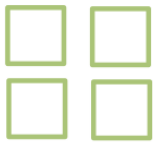
Mesos

Framework
+
master

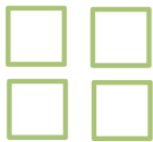
YARN

Resource
Manager

Node



Node



Node



Node



Node



Node



Node



Node



Node



Node



Node

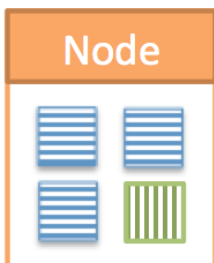
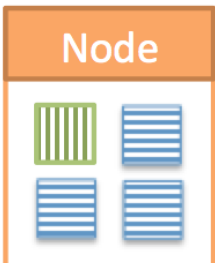
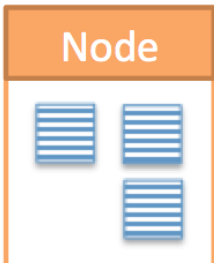
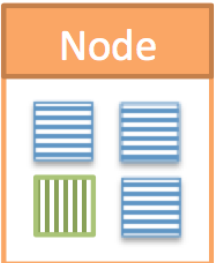
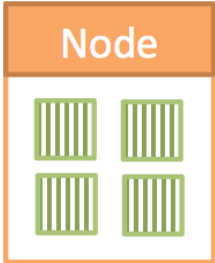
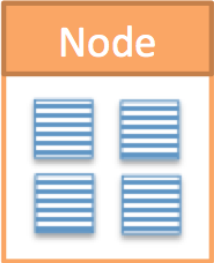
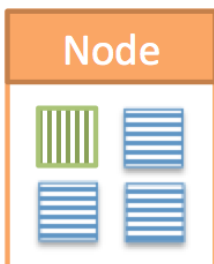
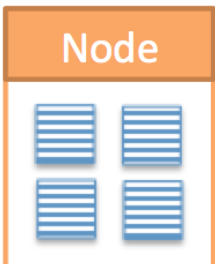
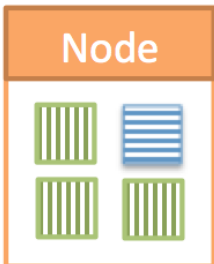
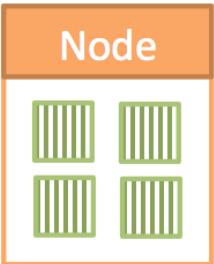
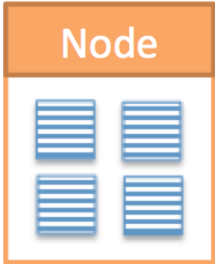
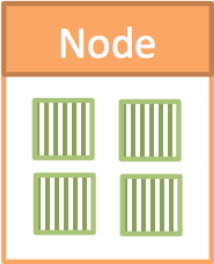
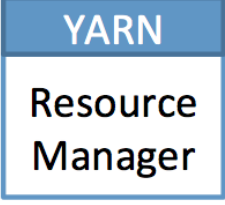
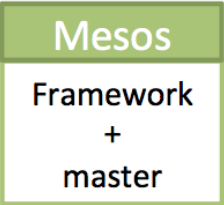


Node



Static Partitioning sucks

- Hadoop teams fine with isolated clusters, but Ops team unhappy; slow to provision
- Resource silos, no elasticity
- Want to run Hadoop on the same infrastructure, without interrupting Tier-1 services
- Want multi-tenancy, resource sharing/isolation



Mesos task



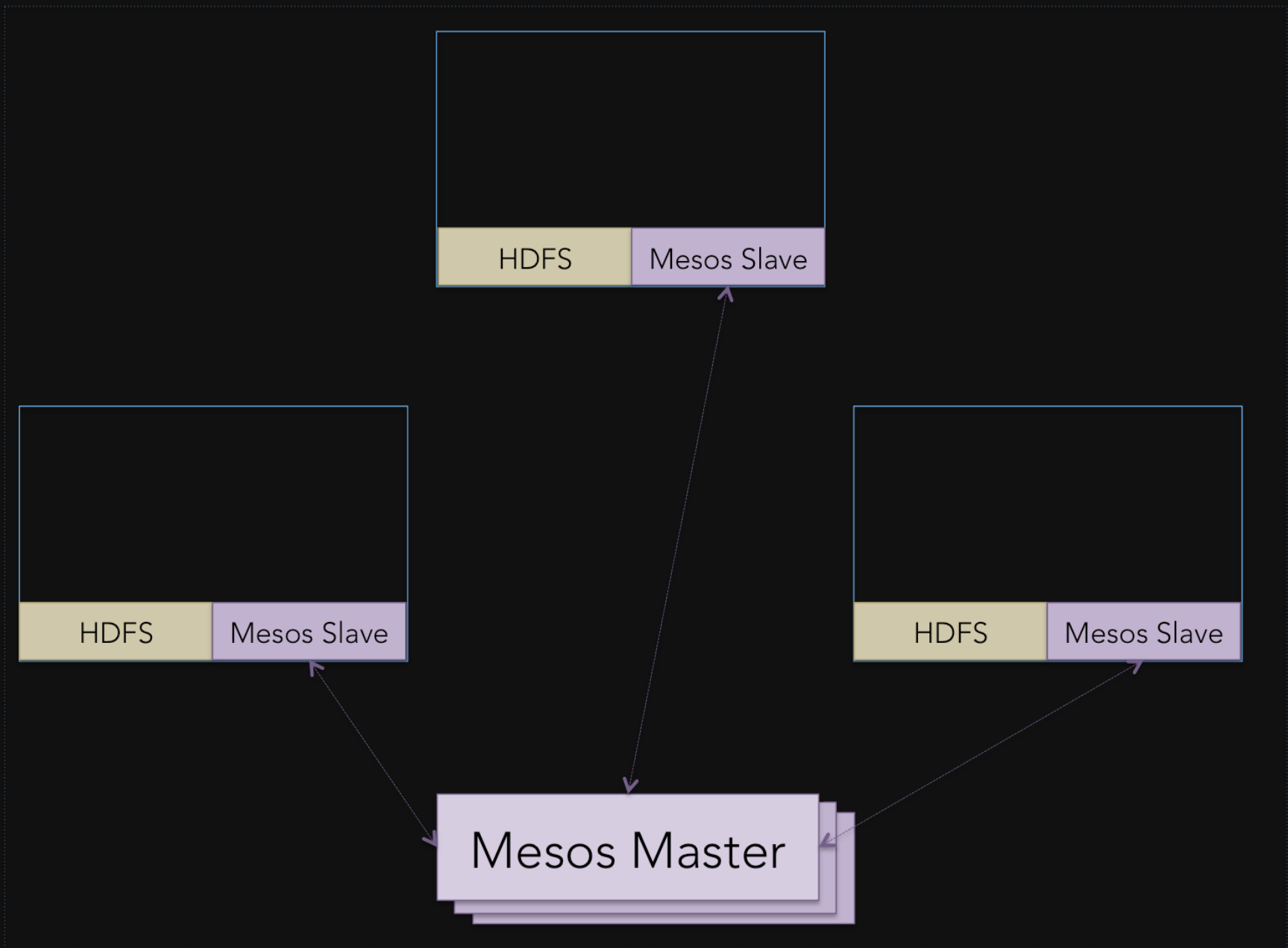
YARN task

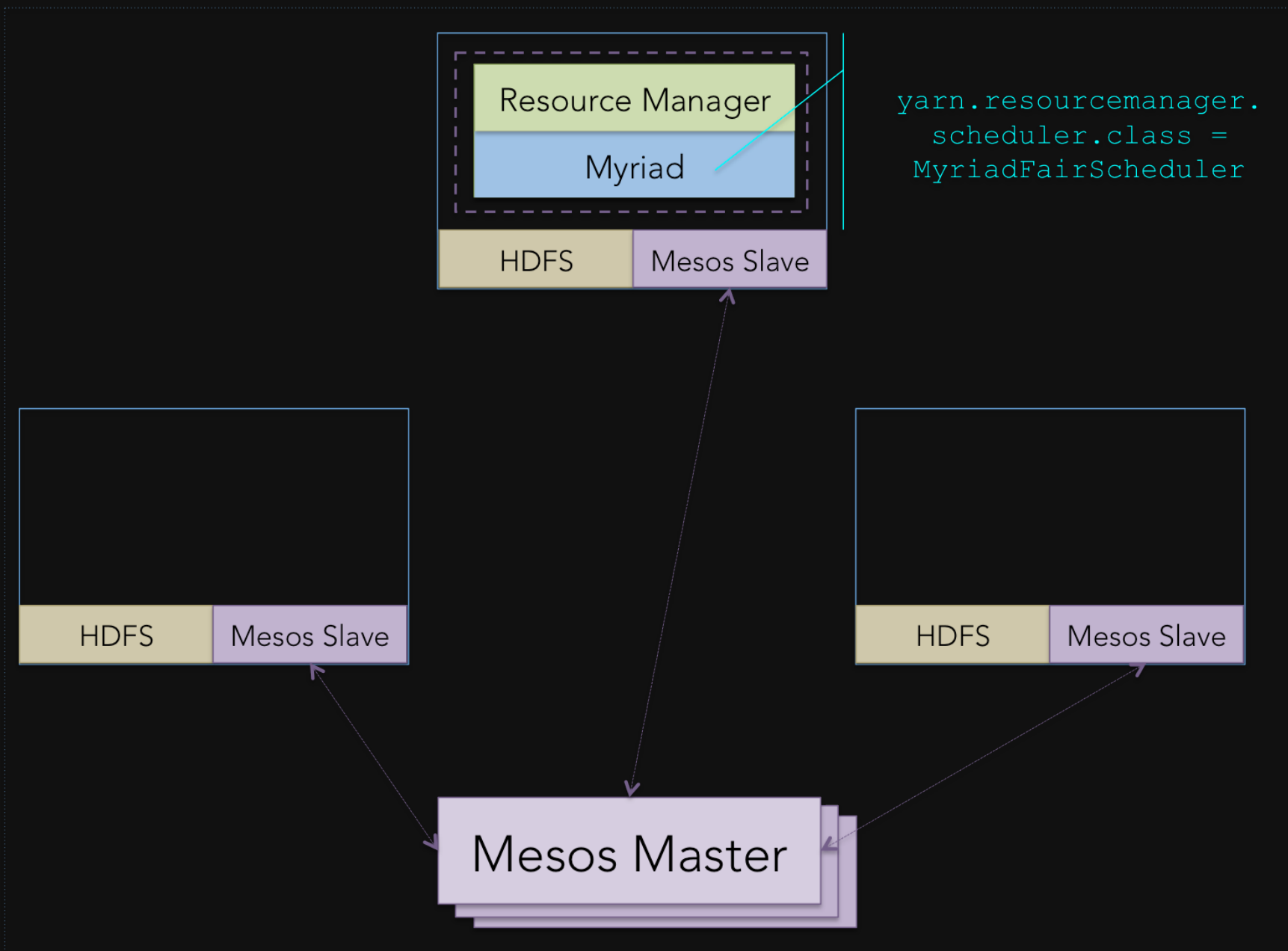
Introducing Myriad

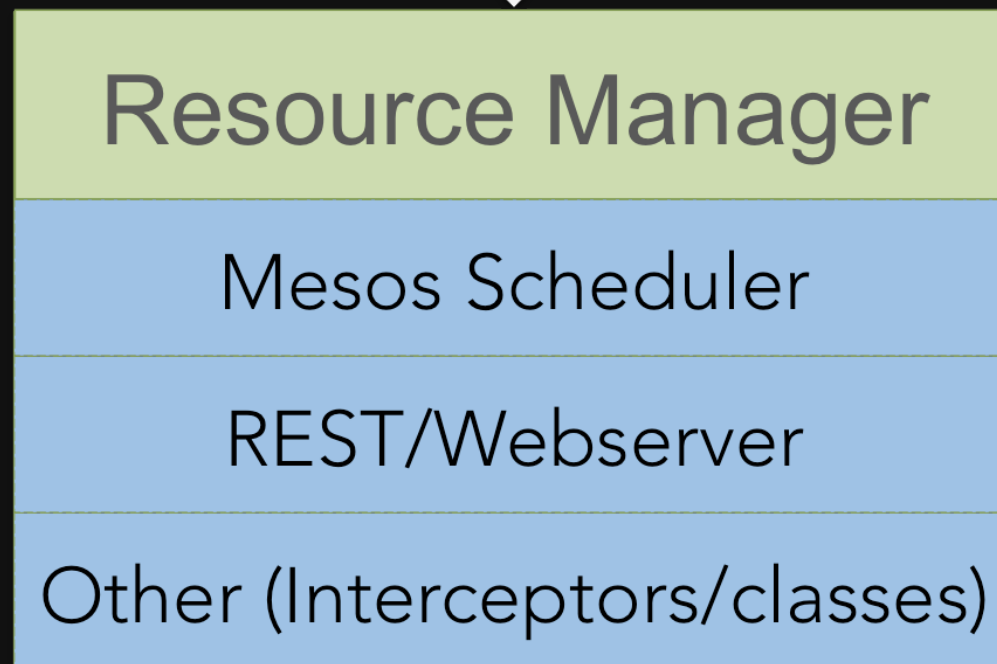
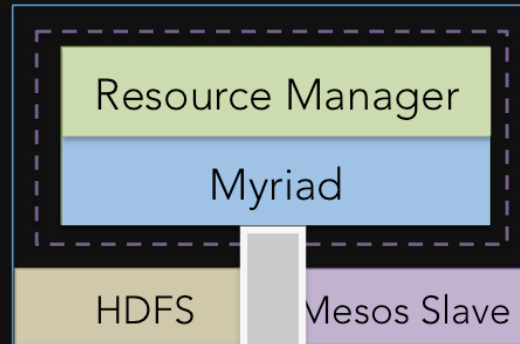
Myriad Overview

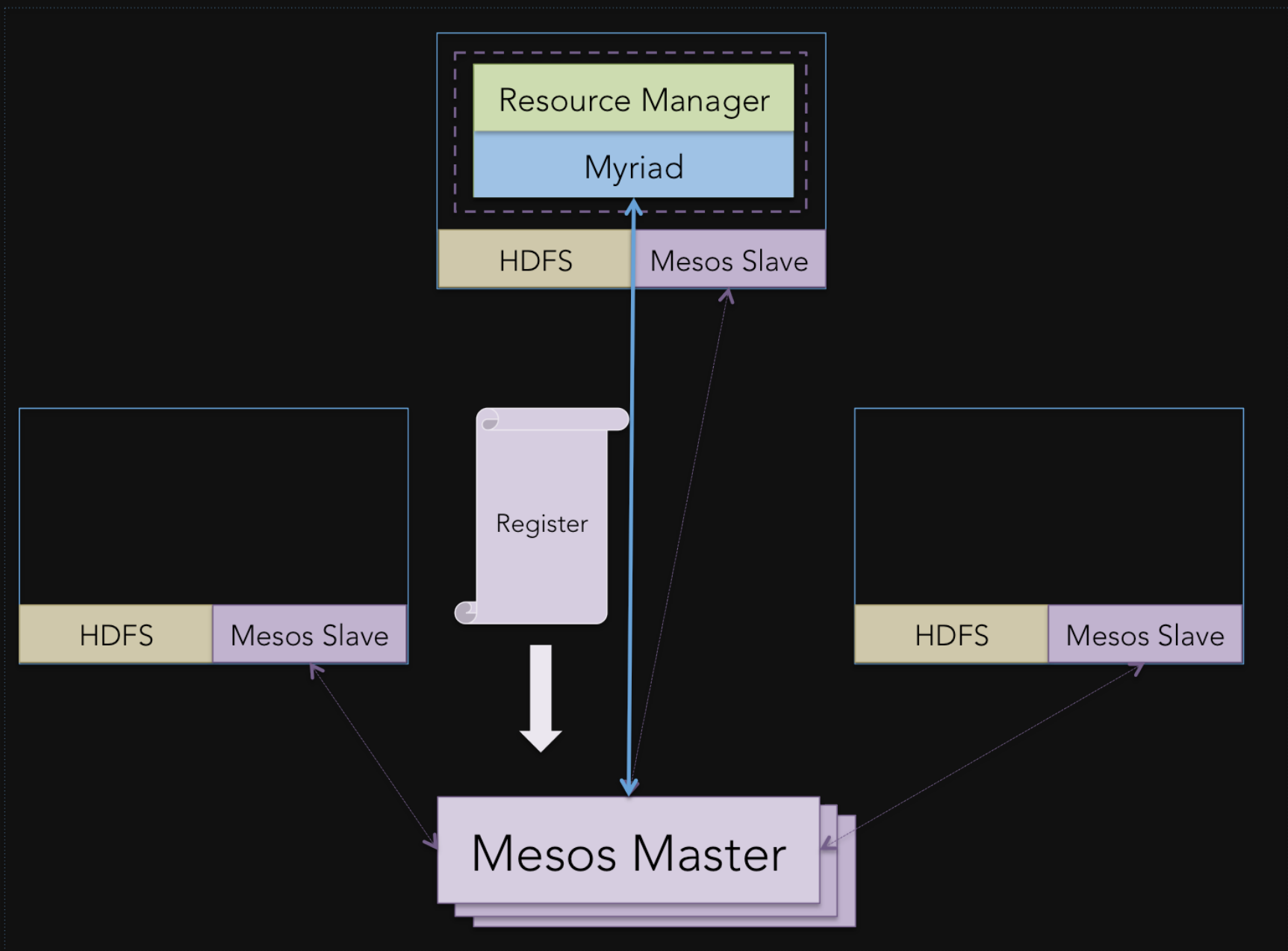
- Mesos Framework for Apache YARN
- Mesos manages DC, YARN manages Hadoop
- Coarse and fine grained resource sharing

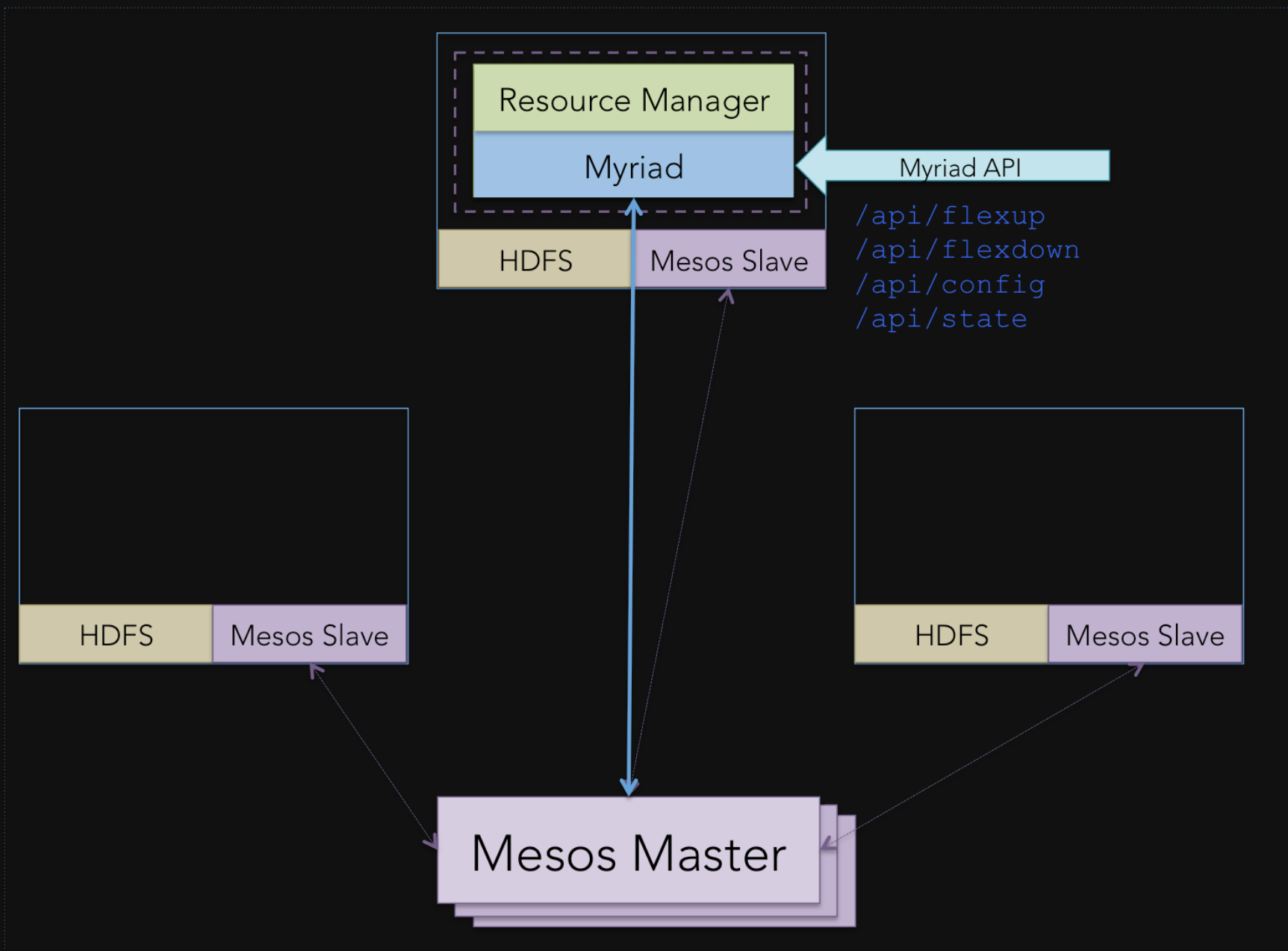
Resource Sharing

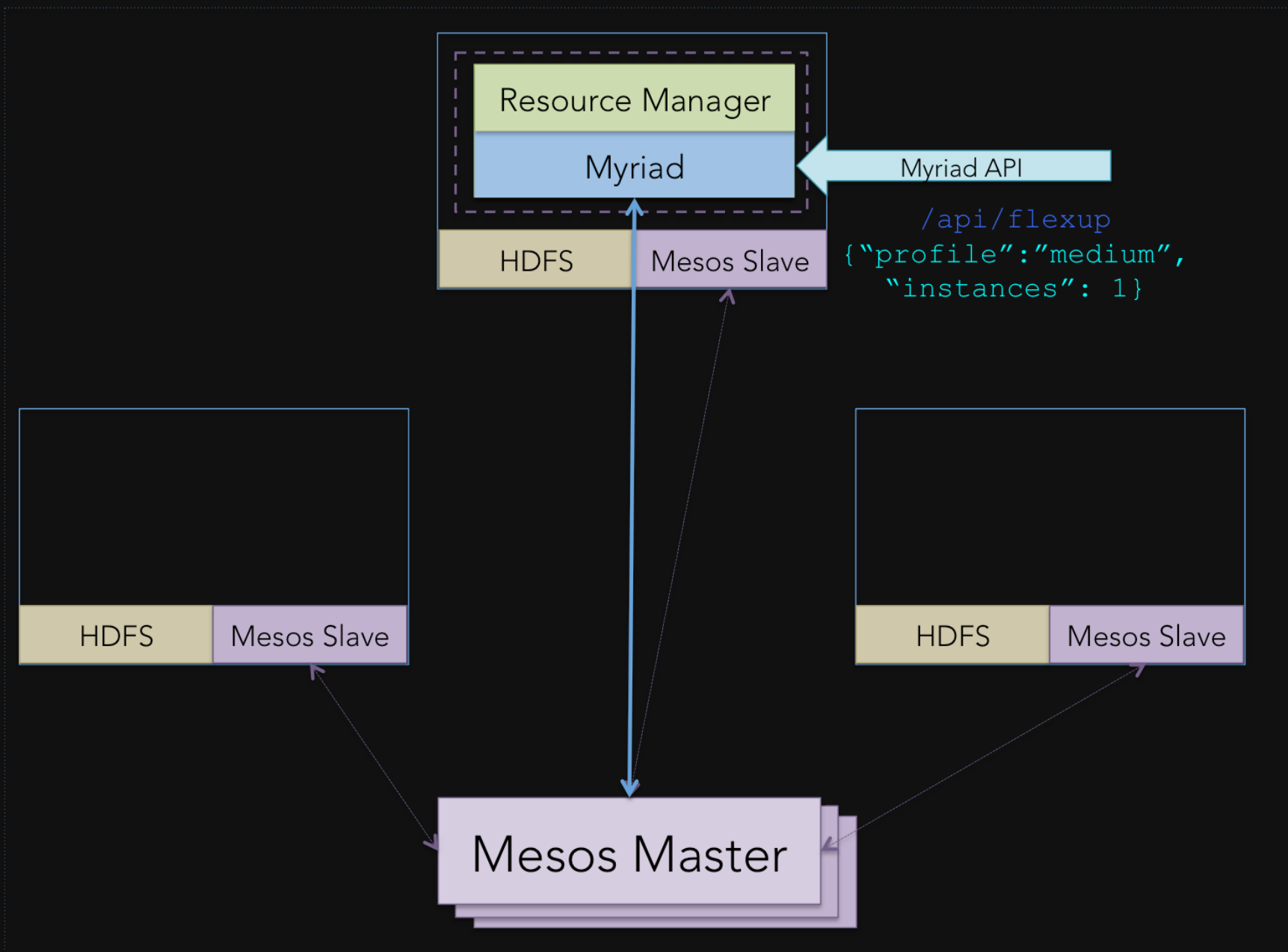


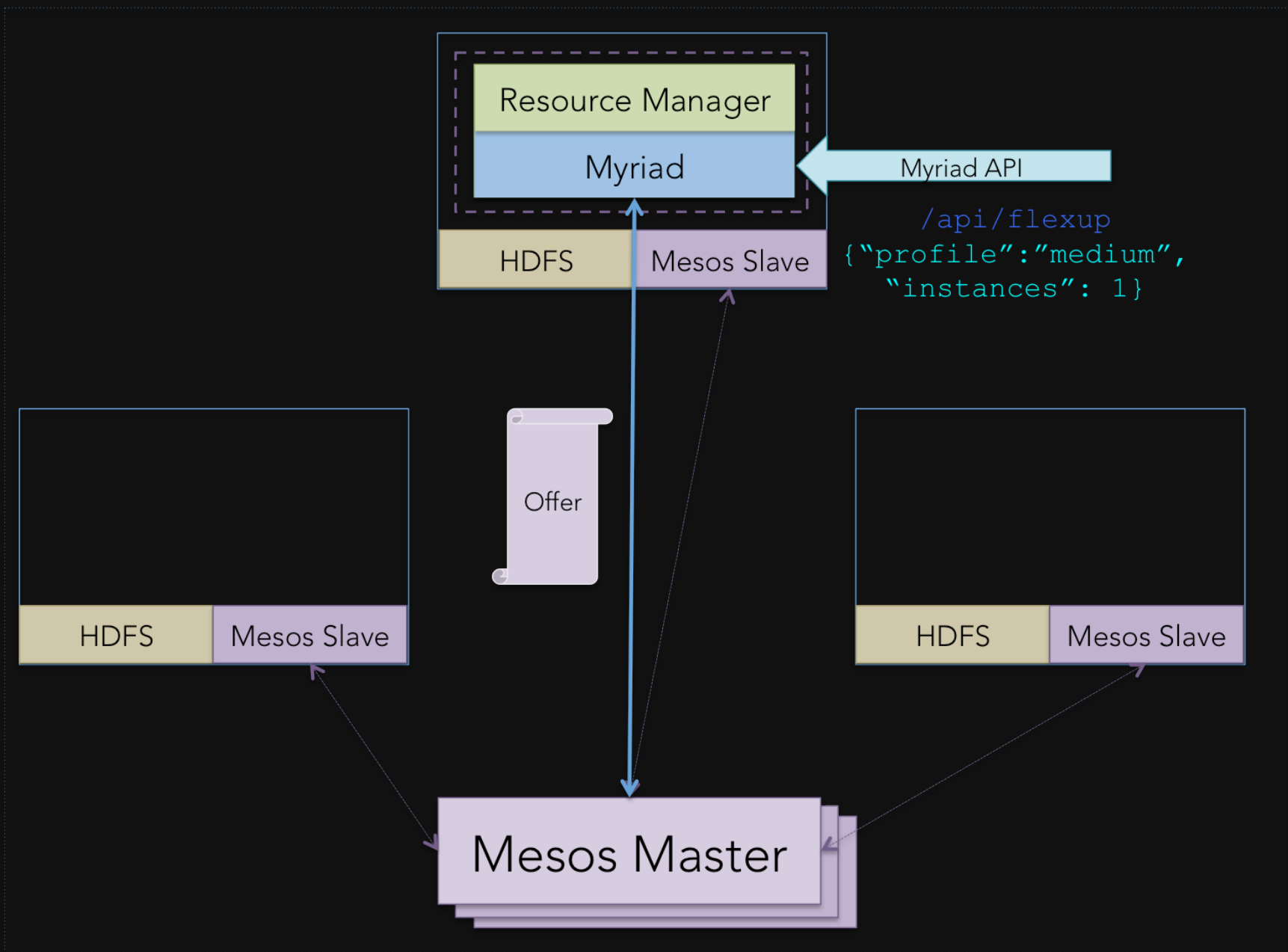


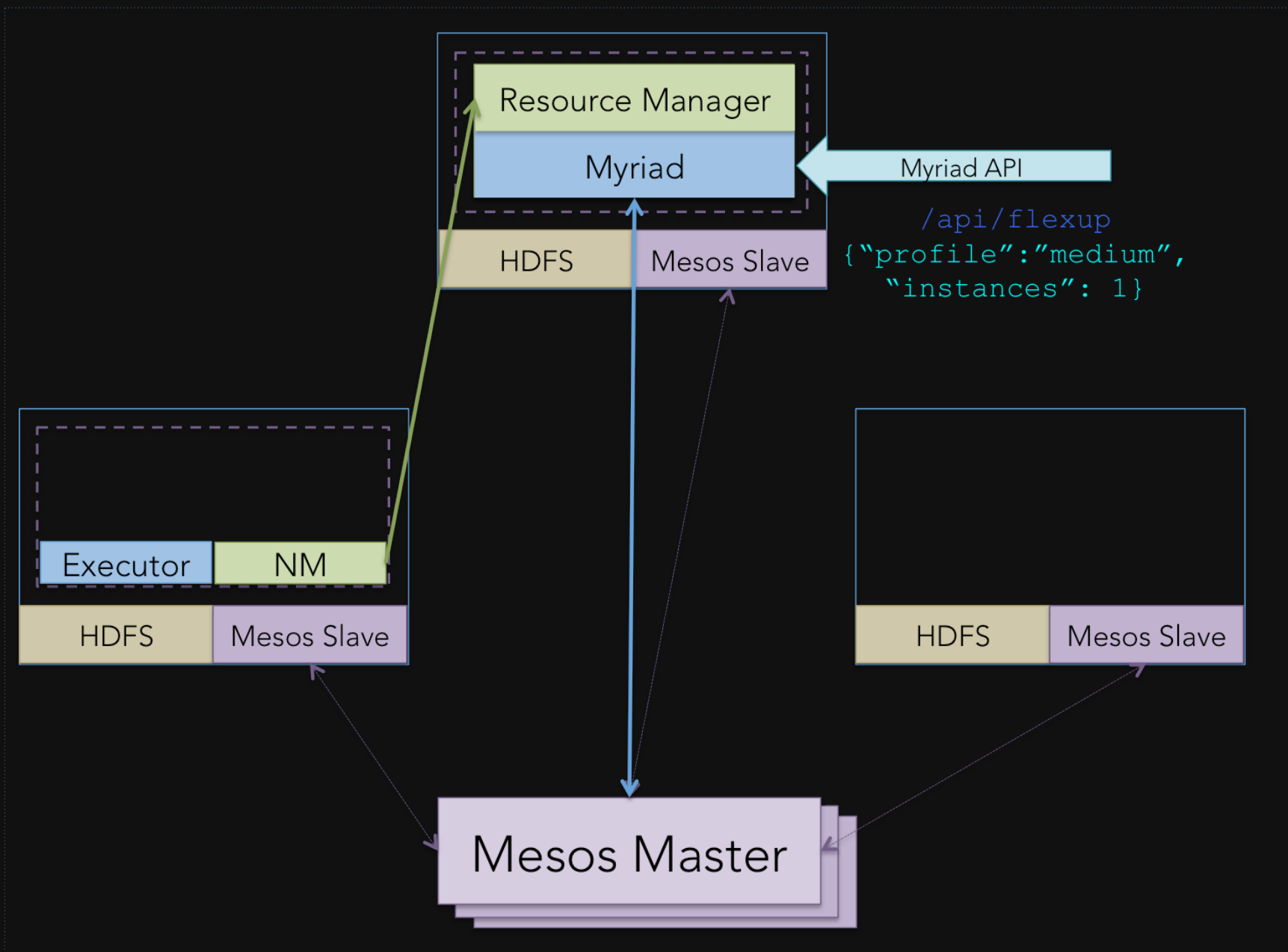


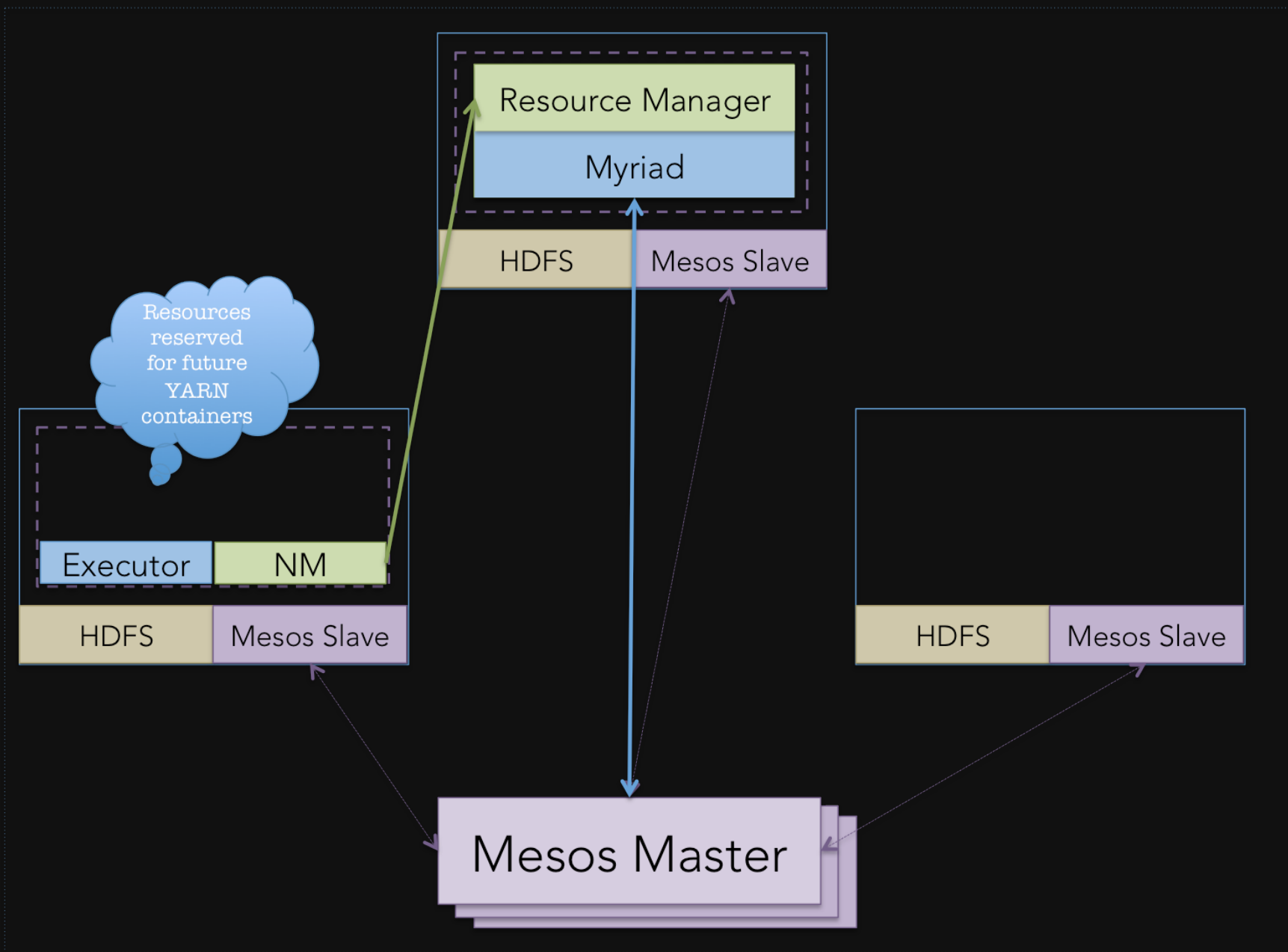


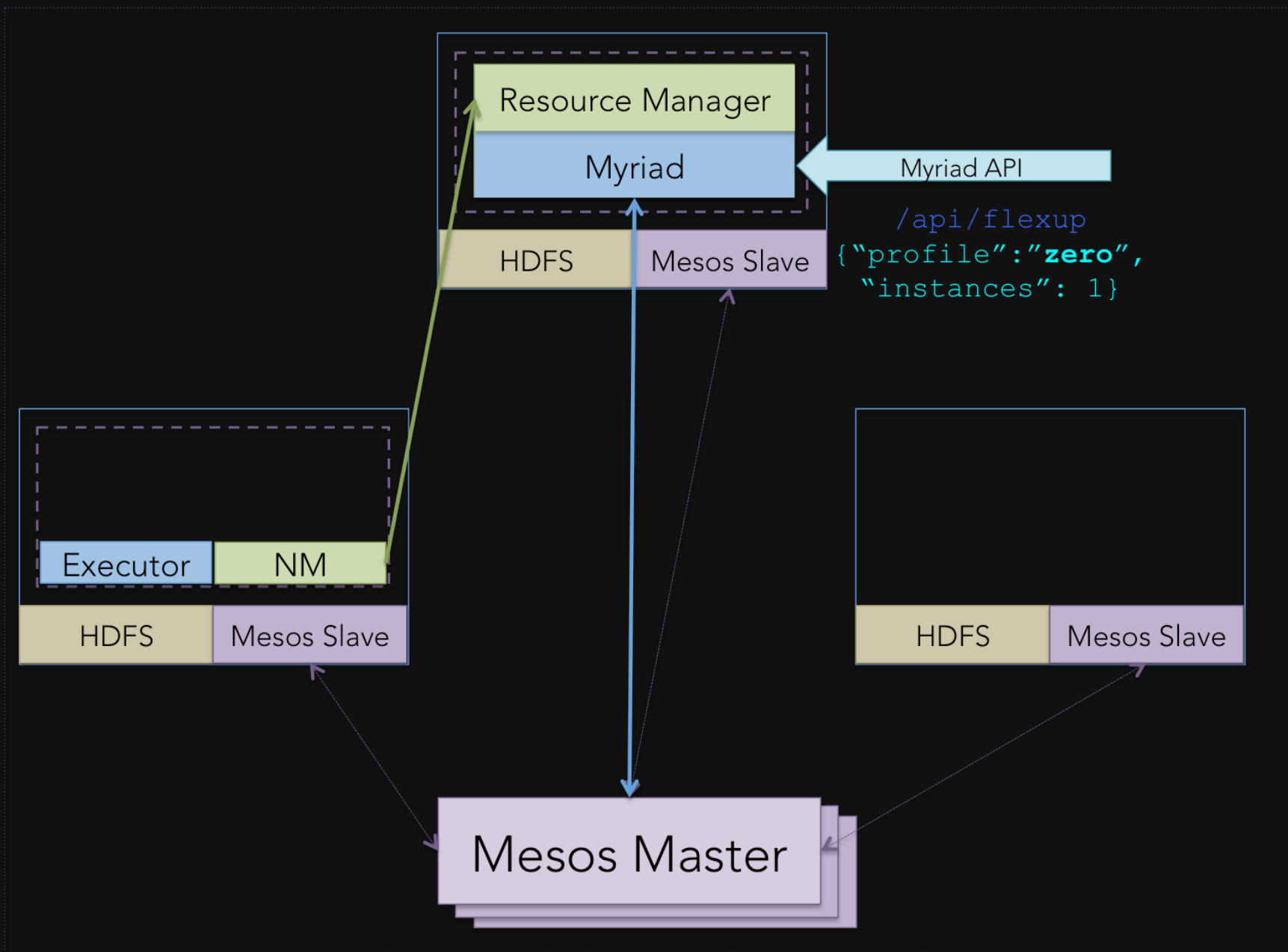


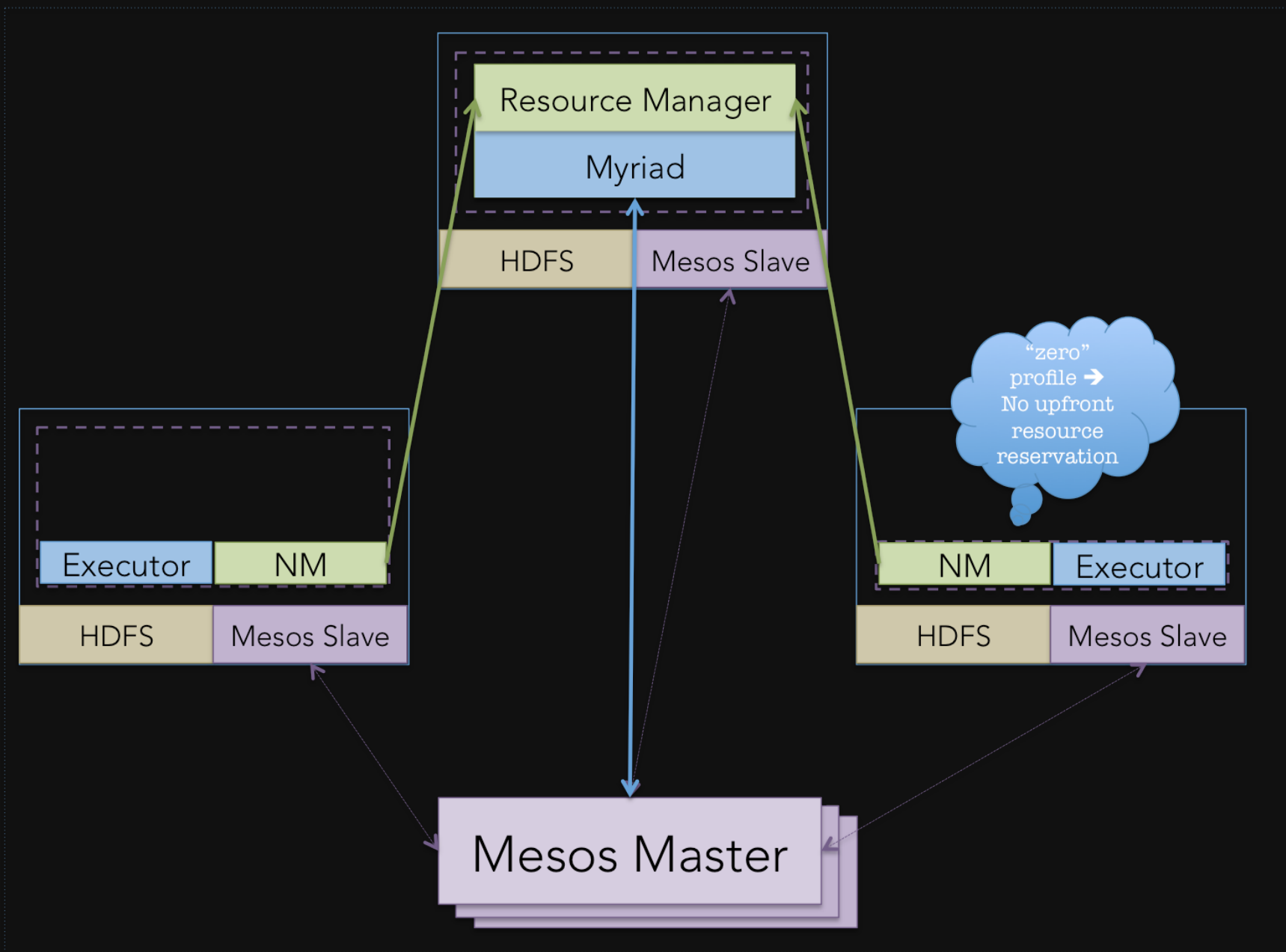


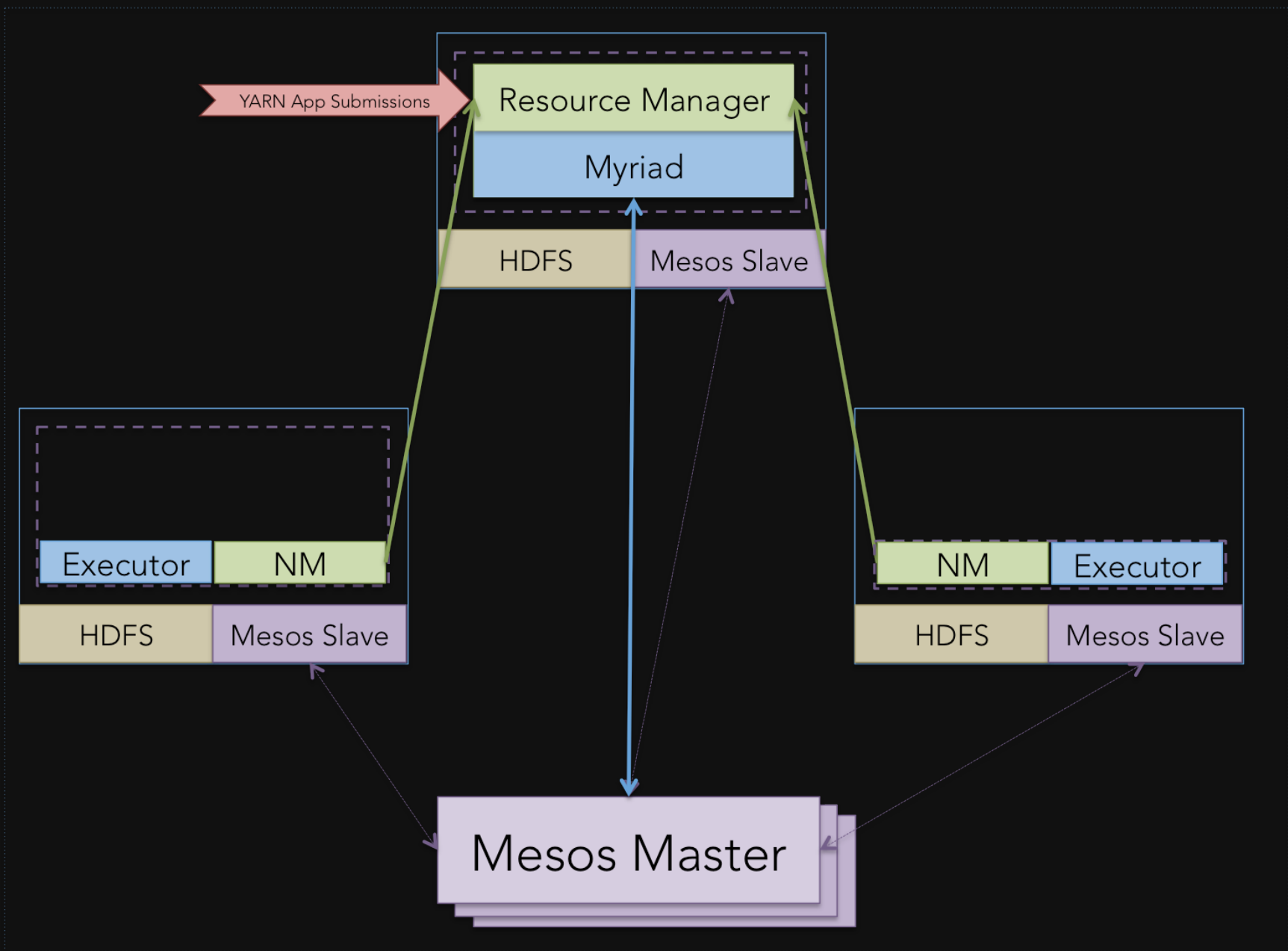


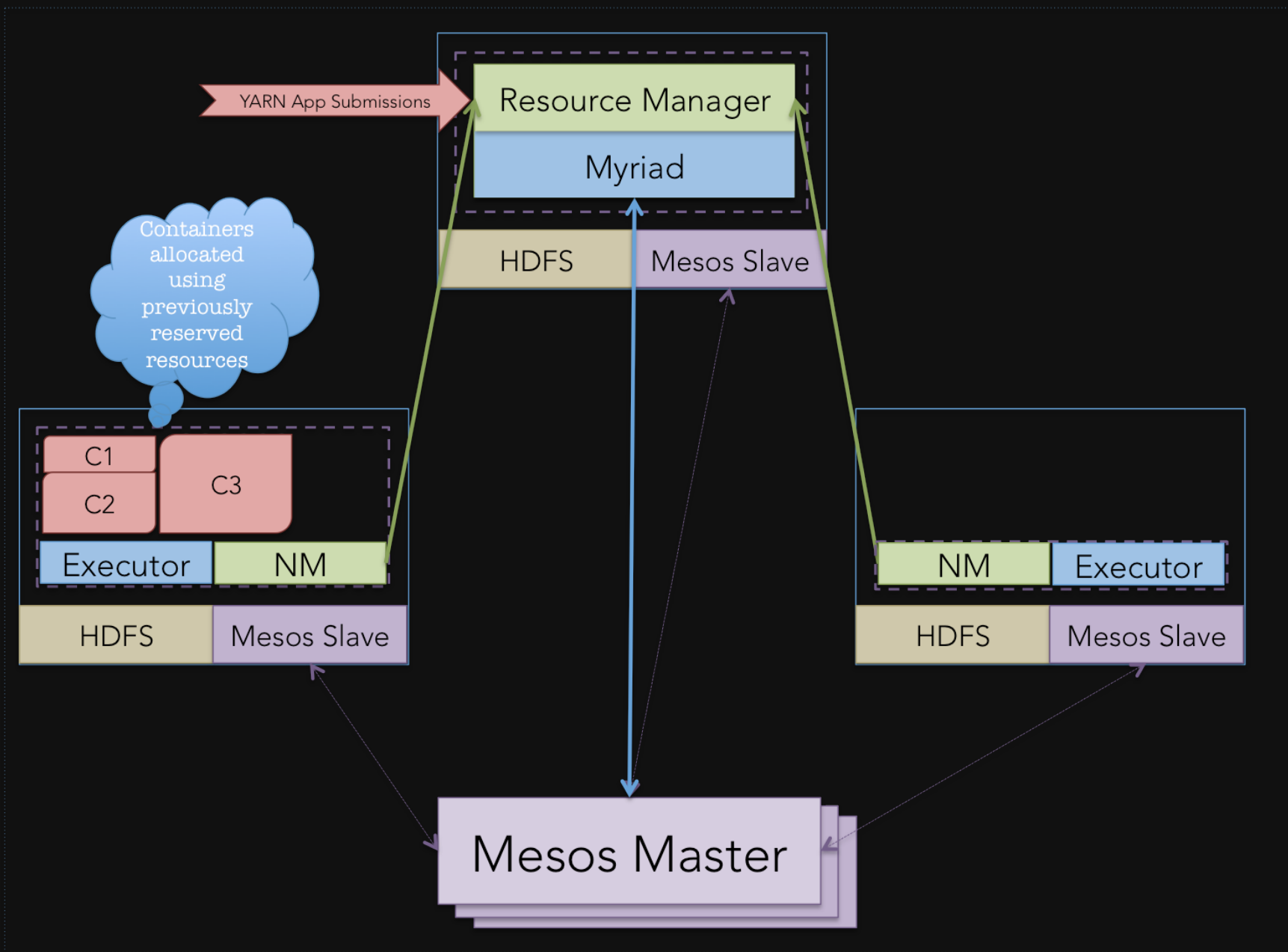


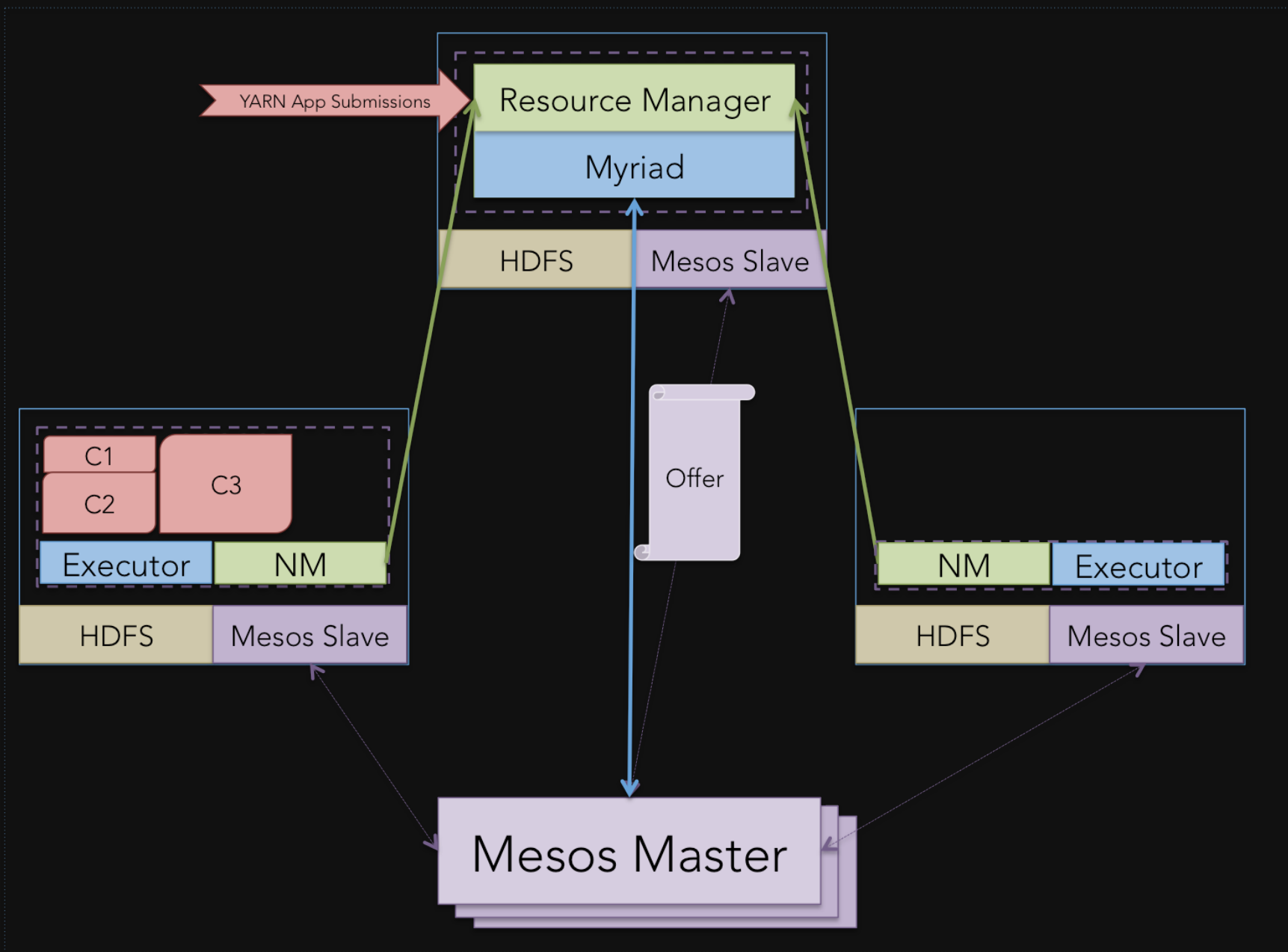


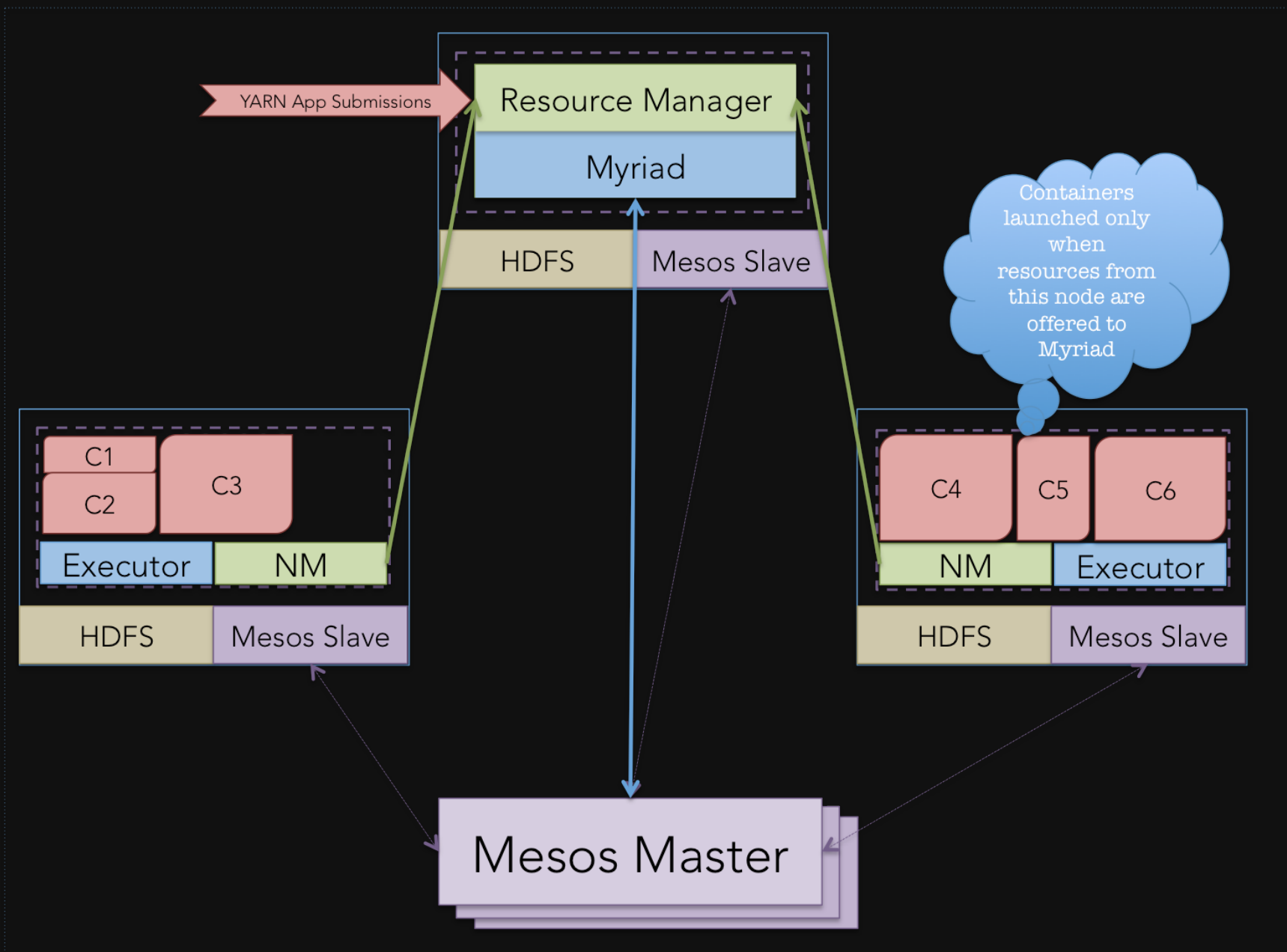


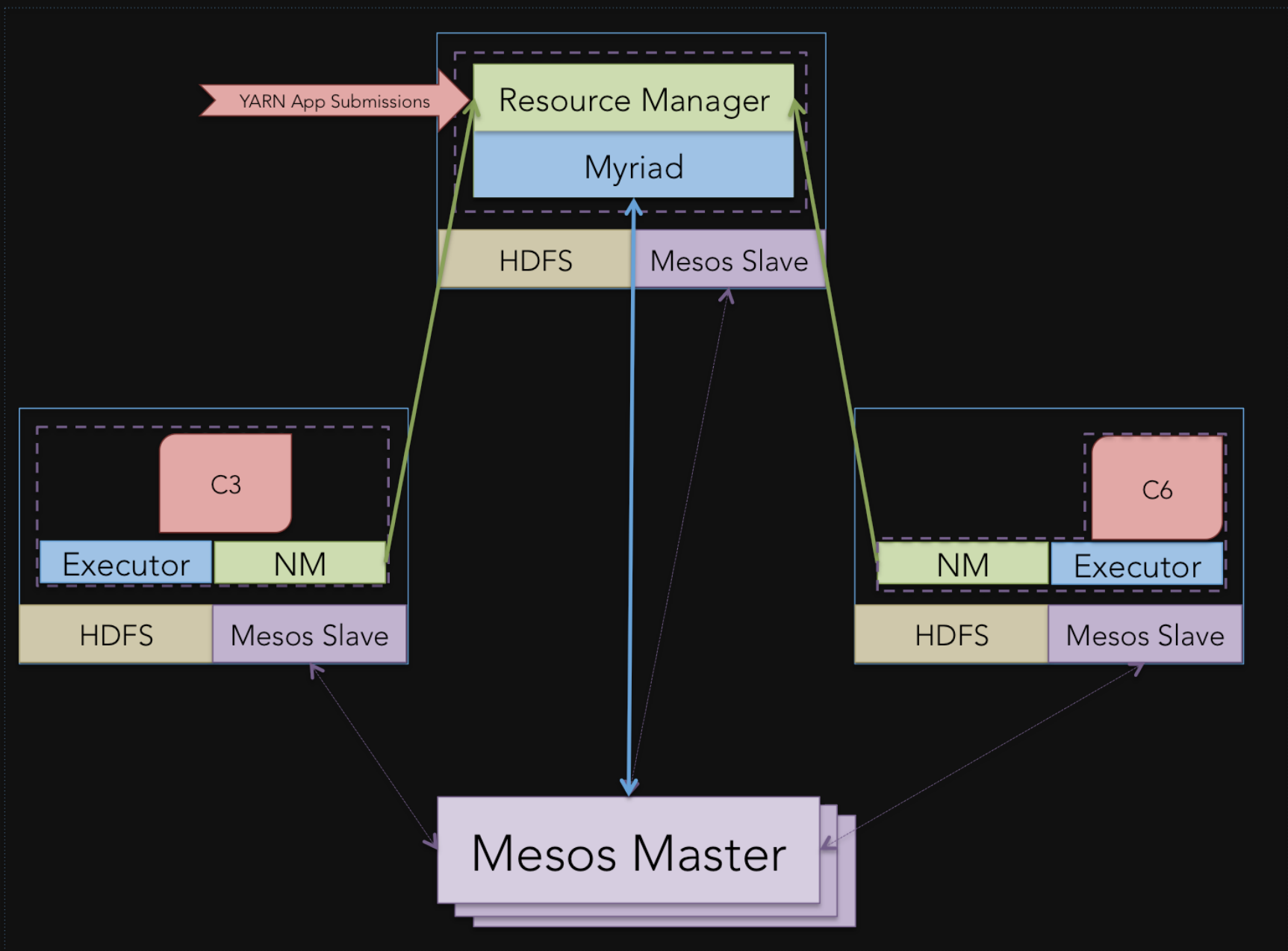


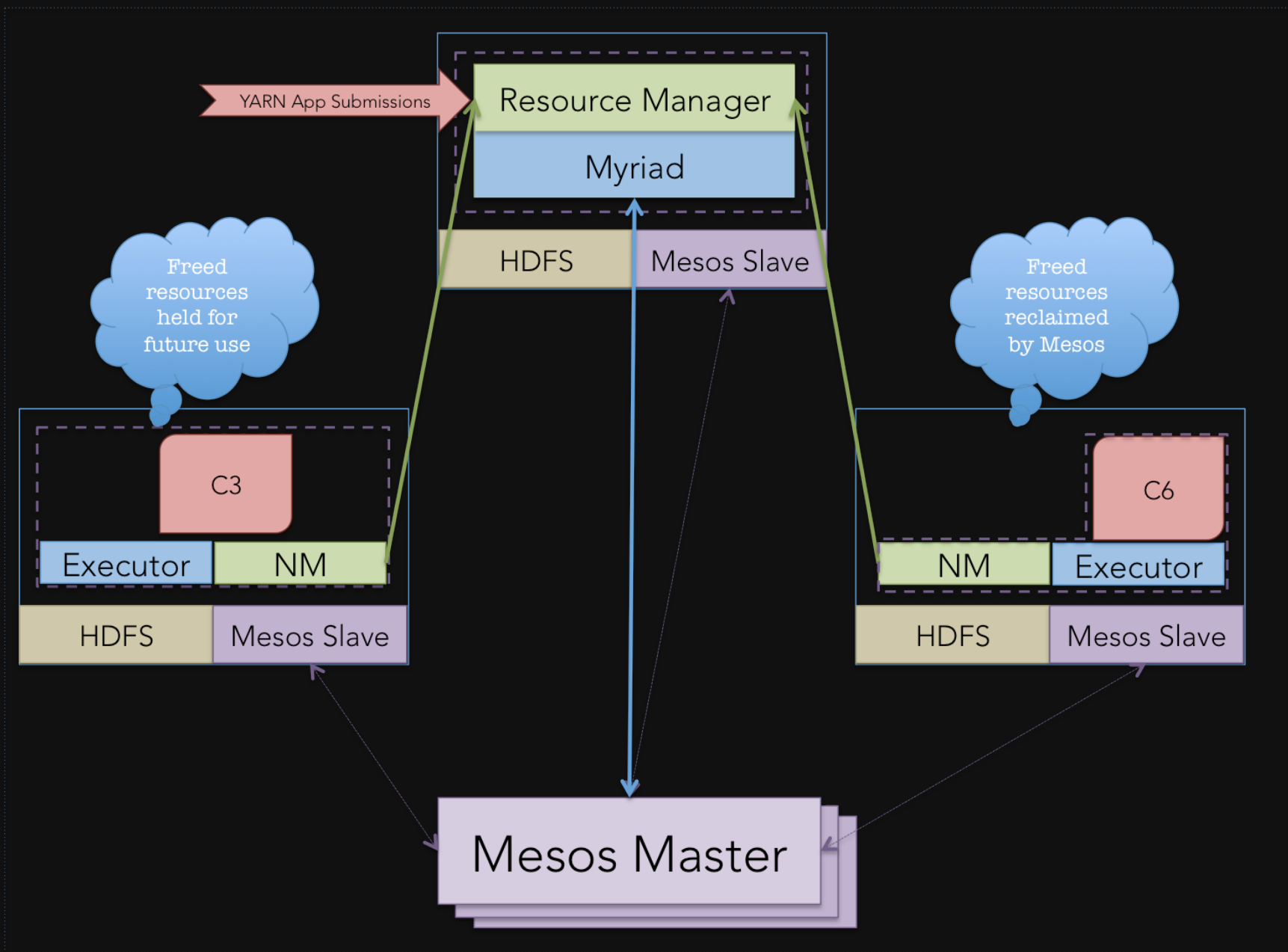


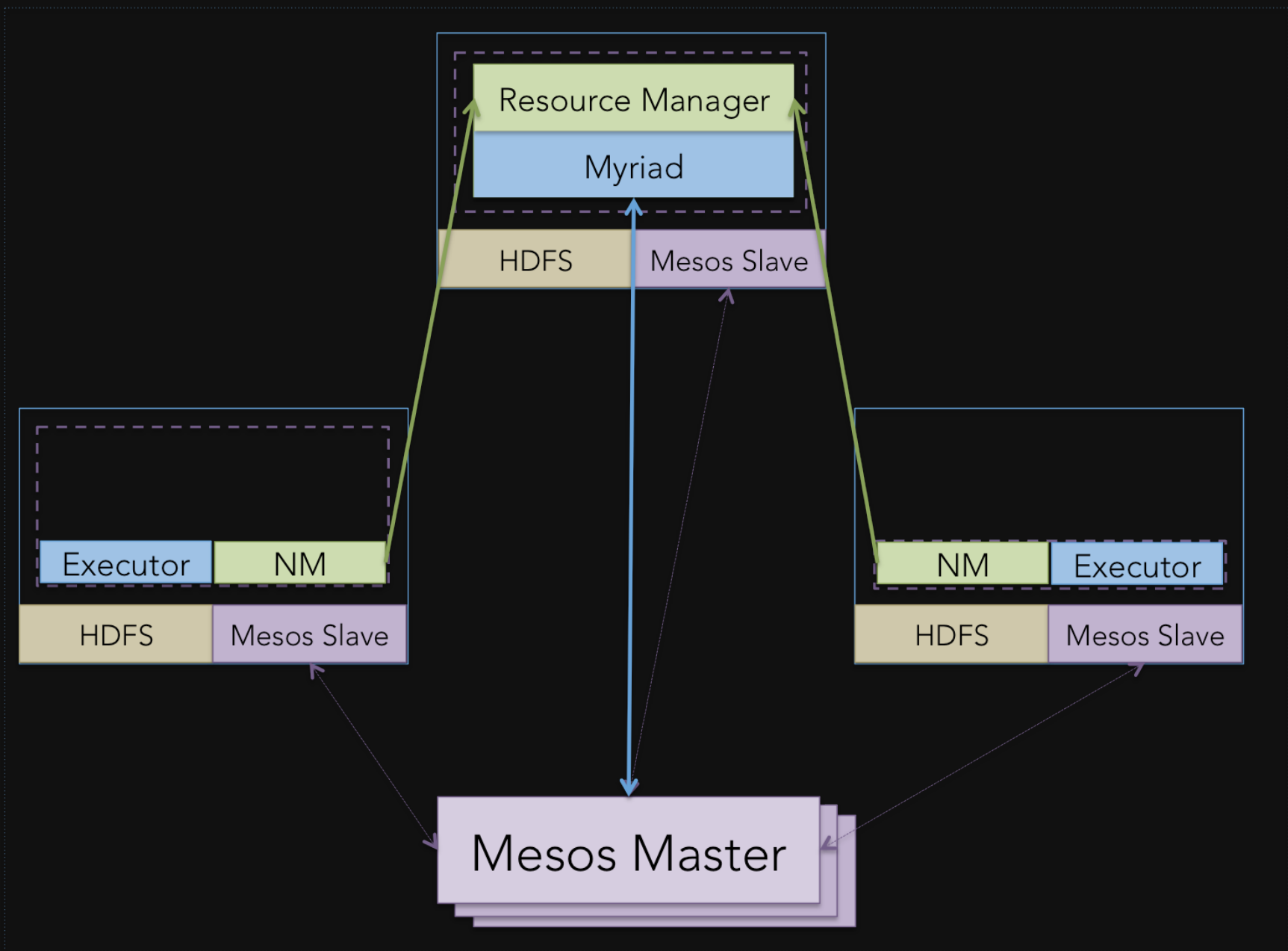














Demo

Myriad improves Mesos

- ⚙️ Tighter integration with Hadoop frameworks like HBase, Hive, Pig
- ⚙️ Borrow resources from Hadoop when traffic spikes for tier-1 services
- ⚙️ Backfill unused resource capacity with best-effort Hadoop jobs
- 🔗 No Mesos code changes necessary

Myriad improves Hadoop

- ..|| Elastic scaling
- 💣 Fault-tolerant: Maintain NM capacity
- ≡ Share resources with other workloads, improve resource utilization
- 🕒 High SLA hadoop jobs unaffected
- 🔗 No YARN/Hadoop code changes

Other Features

- RM failover/discovery using Marathon/Mesos-DNS
- Distribution of hadoop binaries
- Web Interface
- Myriad scheduler HA, task reconciliation (in progress)
- Ability to launch Job History Server (in progress)
- Your favorite feature here!

Learn More!

<https://github.com/mesos/myriad>
dev@myriad.incubator.apache.org

[MYRIAD JIRA](#)

[Apache Myriad Incubator Proposal](#)

[Apache Myriad Incubator Status Page](#)

