Hortonworks

Slider Makes Running Applications on YARN a Breeze

Ted Yu

Agenda

Introduction to Slider

Authoring Application Packages

Ambari View for Slider

Docker based app packaging

Fun operational problems

Placement: where to run?

Installation

Configuration & Binding

Client configuration

Lifecycle

- Start/Stop
- Reconfigure
- Scale up/down
- Rolling-restart
- Decommission/Recommission

Failure handing and recovery

Logging

Upgrading

Metrics & Monitoring

Hortonworks

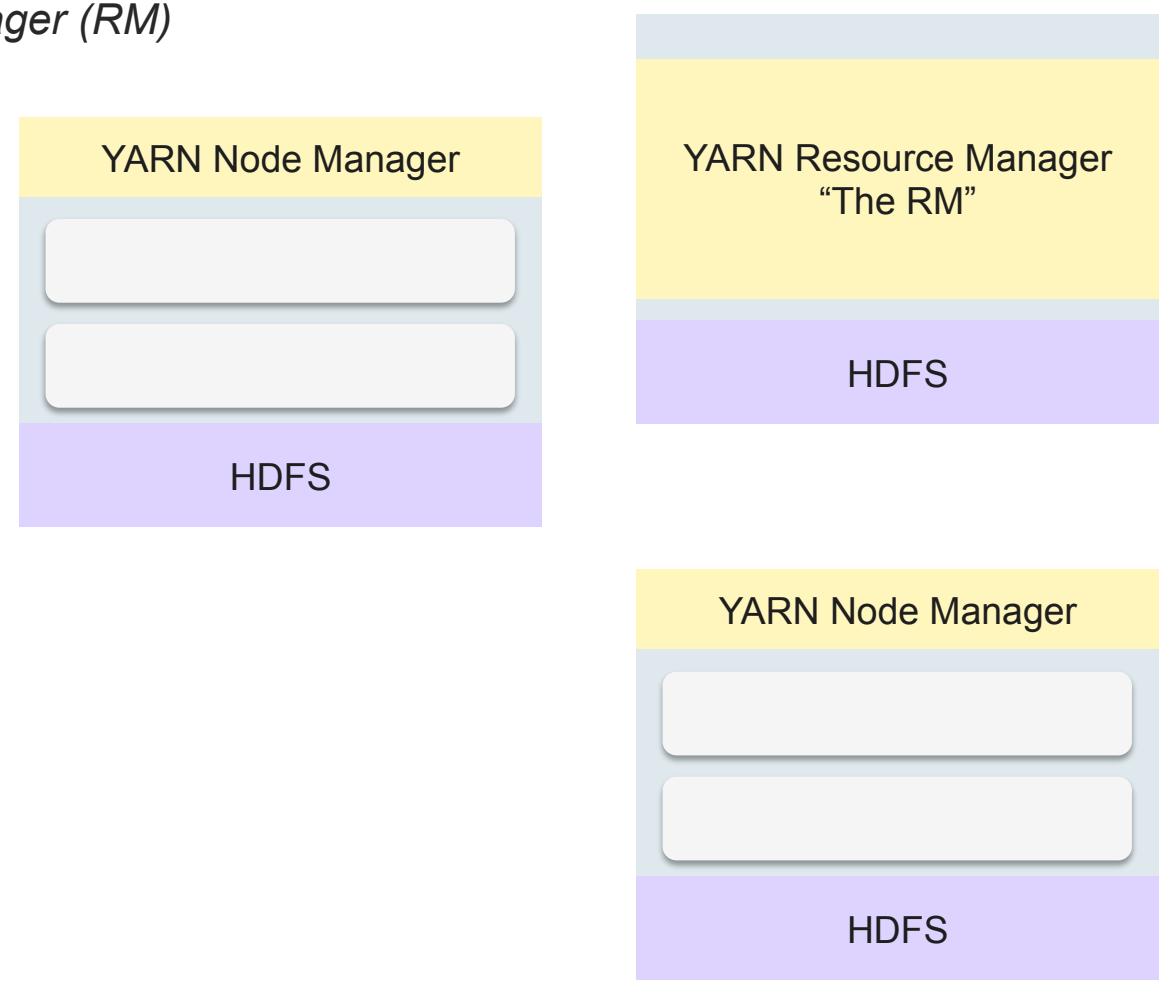
Apache Slider

Deploying and managing
applications on Apache Hadoop YARN

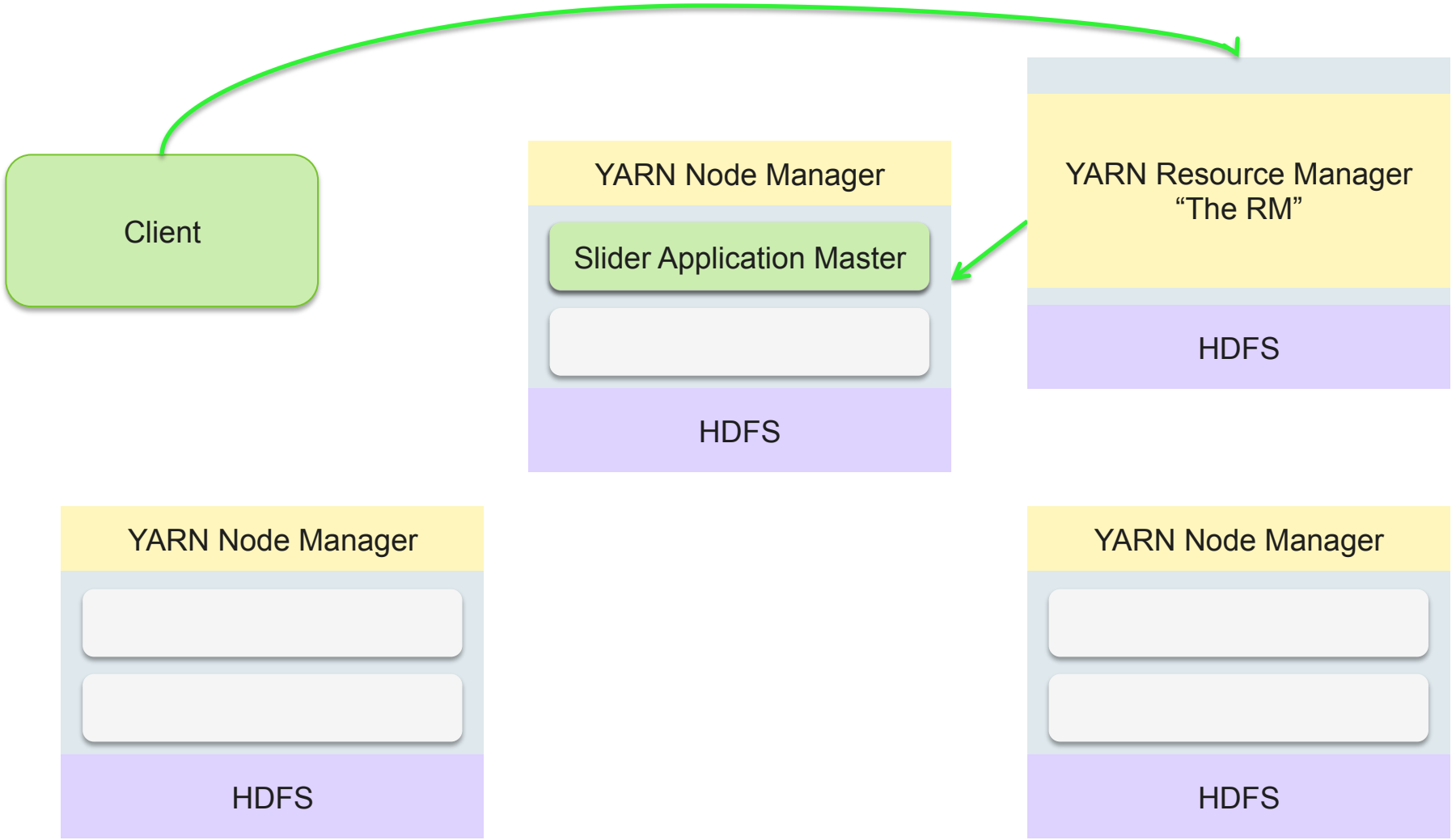
http://slider.incubator.apache.org/

YARN schedules work

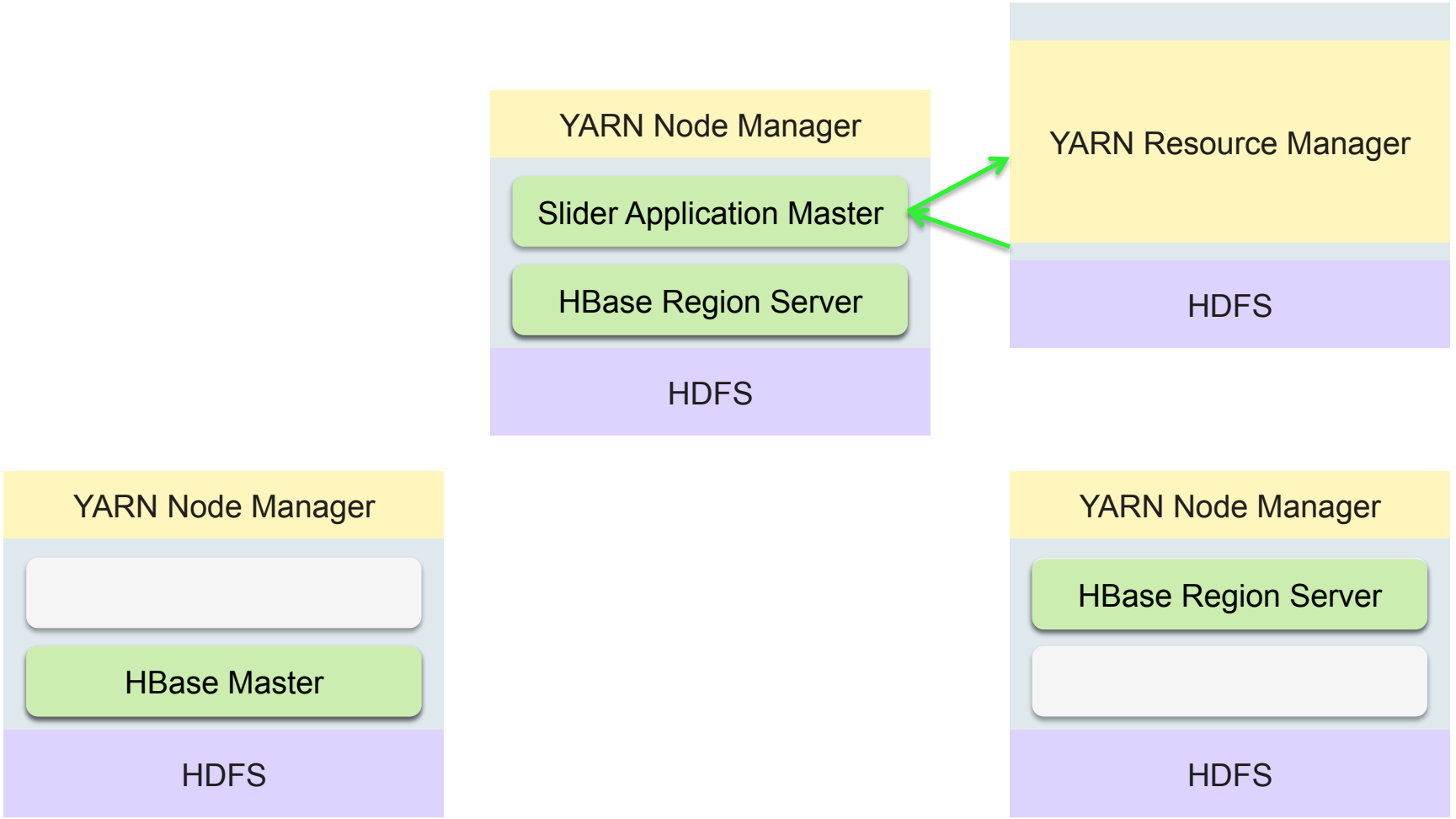
- Servers run YARN *Node Managers (NM)*
- NM's heartbeat to *Resource Manager (RM)*
- RM schedules work over cluster
- RM allocates containers to apps
- NMs start containers
- NMs report container health



Client creates Slider App Master



Slider App Master runs the application



An application consists of

"Application Package"

JSON configuration files: YARN resources + config

Data persisted in HDFS

Placement history in HDFS

resources.json

```
{
  "schema": "http://example.org/specification/v2.0.0",
  "global": {
     "yarn.memory": "512"
  },
  "components": {
    "HBASE_MASTER": {
      "yarn.role.priority": "1",
      "yarn.component.instances": "1",
      "yarn.vcores": "1",
    },
    "HBASE_REGIONSERVER": {
      "yarn.role.priority": "2",
      "yarn.component.instances": "1"
    }
  }
 }
```

configure: app_config.json

```
{
    "application.def": "/slider/hbase_v096.zip",

    "site.global.app_log_dir": "${AGENT_LOG_ROOT}/app/log",
    "site.global.app_pid_dir": "${AGENT_WORK_ROOT}/app/run",

    "site.global.hbase_master_heapsize": "1024m",
    "site.global.ganglia_server_host": "${NN_HOST}",
    "site.global.ganglia_server_port": "8667",
    "site.global.ganglia_server_id": "Application1",

    "site.hbase-site.hbase.tmp.dir": "${AGENT_WORK_ROOT}/work/app/tmp",
    "site.hbase-site.hbase.master.info.port": "${HBASE_MASTER.ALLOCAT
    "site.hbase-site.hbase.regionserver.port": "0",
    "site.hbase-site.hbase.zookeeper.quorum": "${ZK_HOST}",

    "site.core-site.fs.defaultFS": "${NN_URI}",
}
```

Configurations needed by Slider

Named variables

Variables for the application scripts

Site variables for application

Allocate and advertise

Named variables for cluster details

create, start, stop, destroy

```
$ slider create hbase1 --resources resources.json --template config.json

$ slider list

$ slider status hbase1

$ slider stop hbase1

$ slider start hbase1

$ slider destroy hbase1
```

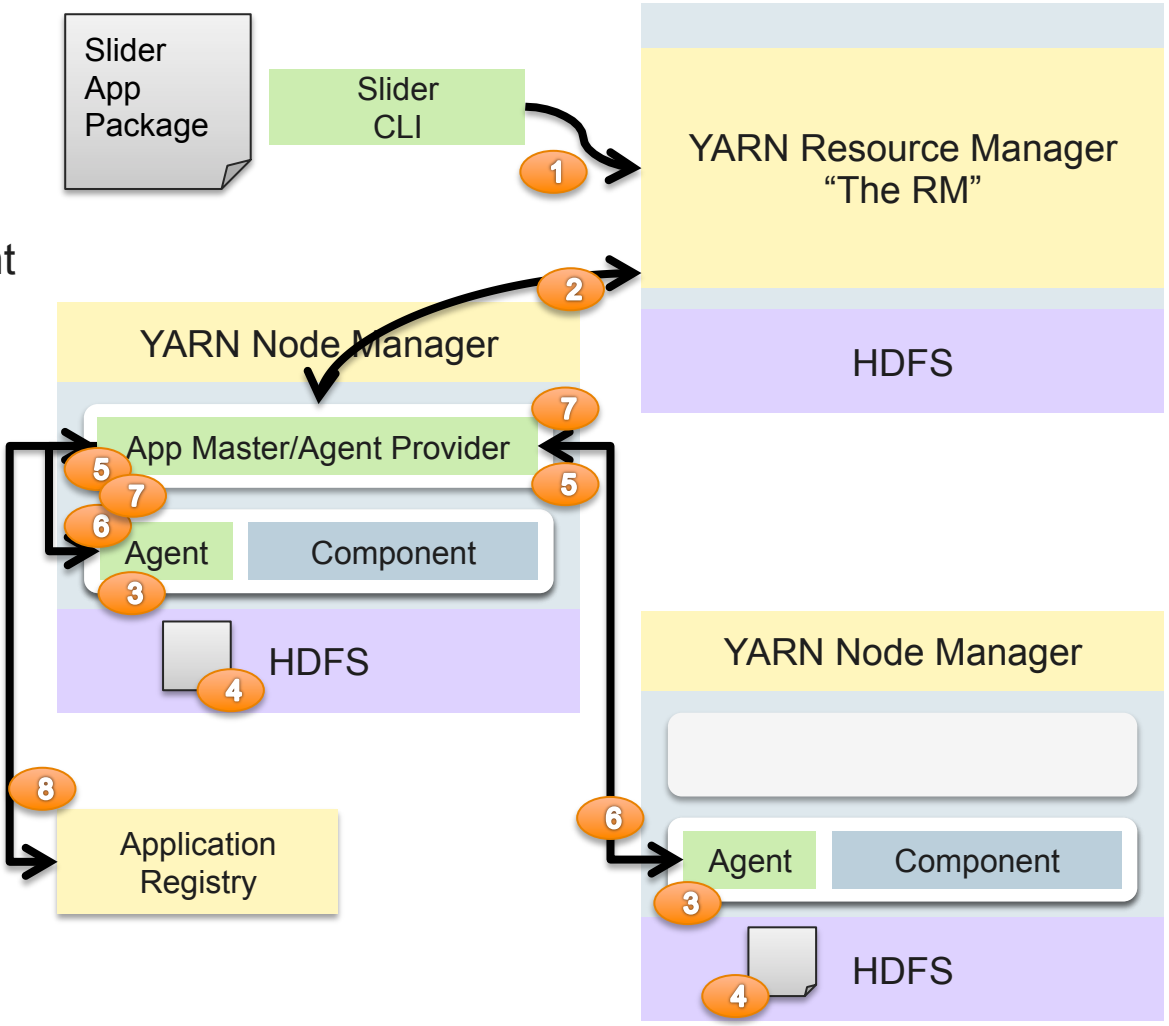
Dynamic application resize

```
slider flex hbase1
  --component HBASE_REGIONSERVER 2
```

Application by Slider

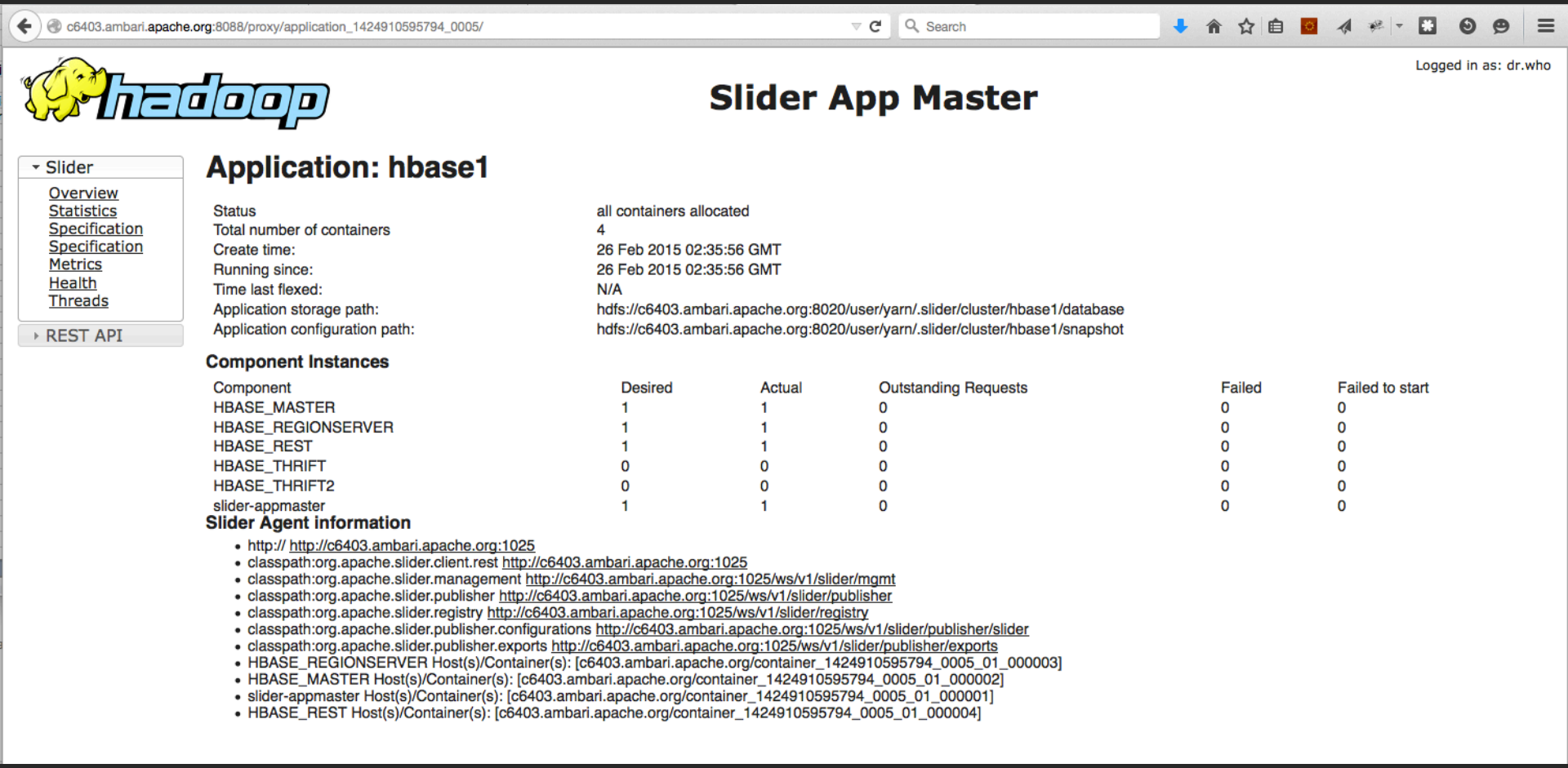
Similar to any YARN application

1. CLI starts an instance of the AM
2. AM requests containers
3. Containers activate with an Agent
4. Agent gets application definition
5. Agent registers with AM
6. AM issues commands
7. Agent reports back, status, configuration, etc.
8. AM publishes endpoints, configurations

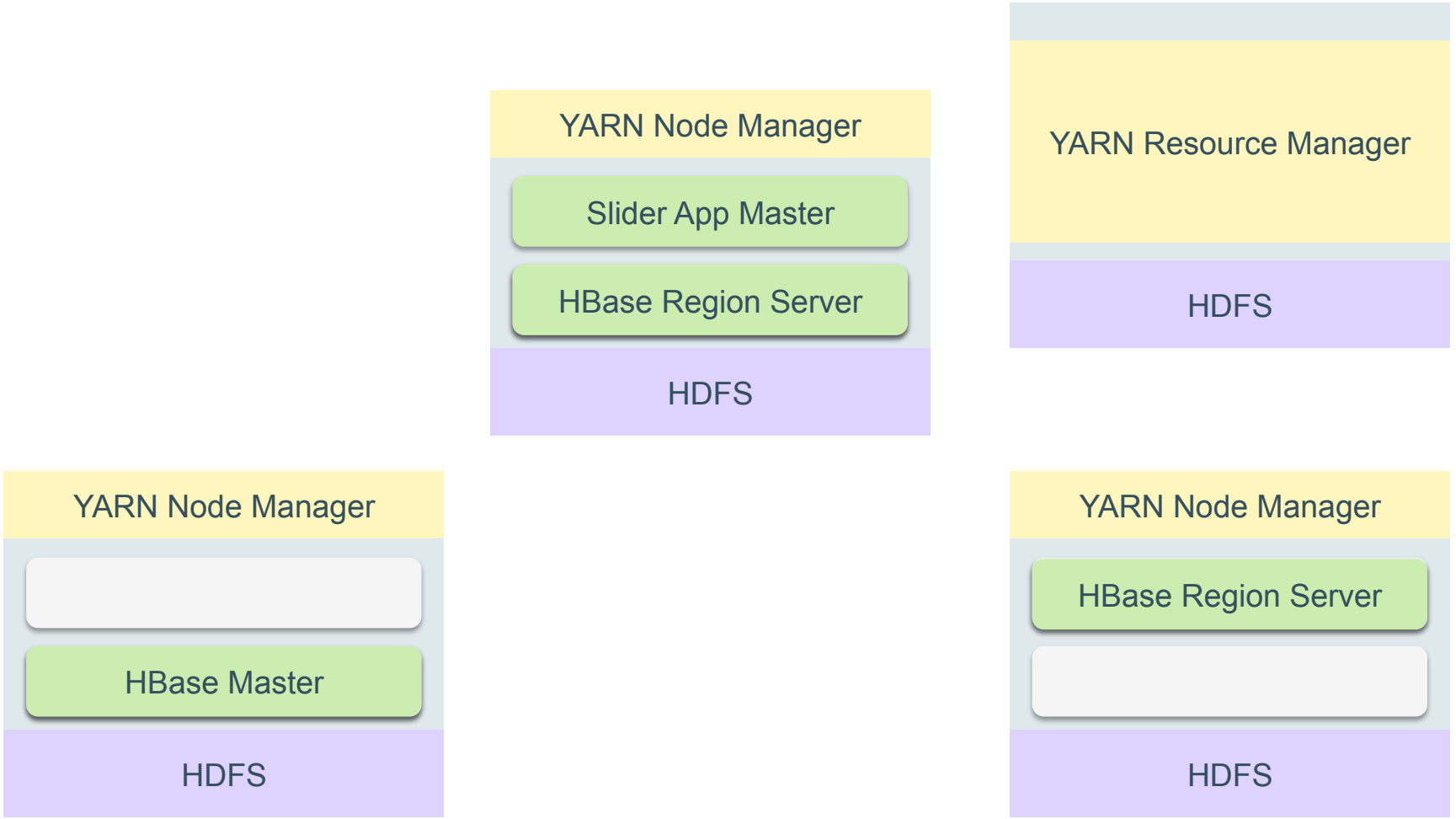


Slider AppMaster/Agent/Client

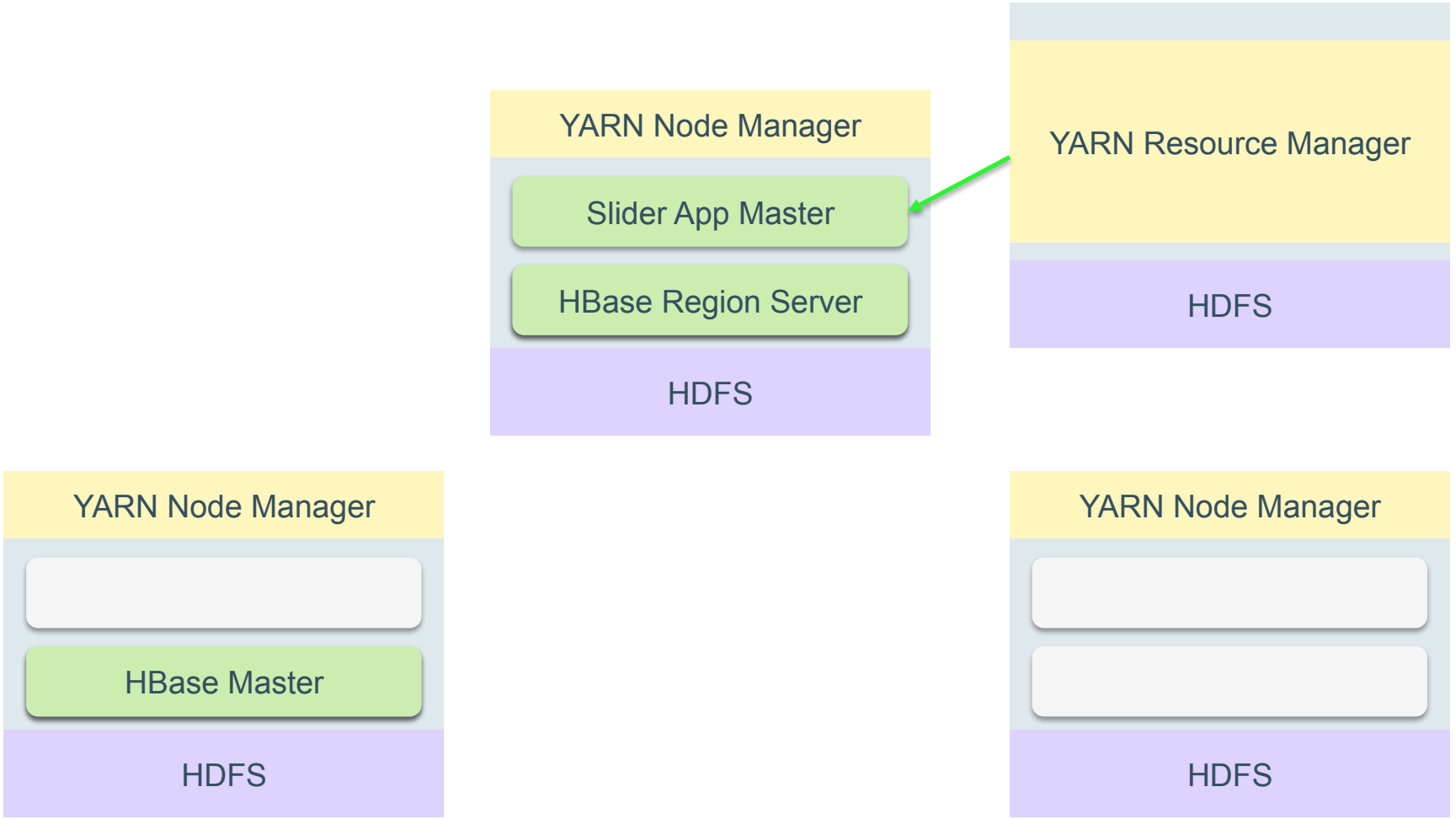
- AppMaster
  - Common YARN interactions
  - Common *-client interactions
  - Publishing needs
- Agent
  - Configure and start
  - Re-configure and restart
  - Heartbeats & failure detection
  - Port allocations and publishing
  - Custom commands if any (e.g. graceful-stop)
- Client
  - App life cycle commands (flex, status, …)
  - Package installation


c6403.ambari.apache.org:8088/proxy/application_1424910595794_0005/
Search
Logged in as: dr.who
hadoop
Slider App Master
Slider
Overview
Statistics
Specification
Specification
Metrics
Health
Threads
REST API
Application: hbase1
Status all containers allocated
Total number of containers 4
Create time: 26 Feb 2015 02:35:56 GMT
Running since: 26 Feb 2015 02:35:56 GMT
Time last flexed: N/A
Application storage path: hdfs://c6403.ambari.apache.org:8020/user/yarn/.slider/cluster/hbase1/database
Application configuration path: hdfs://c6403.ambari.apache.org:8020/user/yarn/.slider/cluster/hbase1/snapshot
Component Instances
Component Desired Actual Outstanding Requests Failed Failed to start
HBASE_MASTER 1 1 0 0 0
HBASE_REGIONSERVER 1 1 0 0 0
HBASE_REST 1 1 0 0 0
HBASE_THRIFT 0 0 0 0 0
HBASE_THRIFT2 0 0 0 0 0
slider-appmaster 1 1 0 0 0
Slider Agent information
http:// http://c6403.ambari.apache.org:1025
classpath:org.apache.slider.client.rest http://c6403.ambari.apache.org:1025
classpath:org.apache.slider.management http://c6403.ambari.apache.org:1025/ws/v1/slider/mgmt
classpath:org.apache.slider.publisher http://c6403.ambari.apache.org:1025/ws/v1/slider/publisher
classpath:org.apache.slider.registry http://c6403.ambari.apache.org:1025/ws/v1/slider/registry
classpath:org.apache.slider.publisher.configurations http://c6403.ambari.apache.org:1025/ws/v1/slider/publisher/slider
classpath:org.apache.slider.publisher.exports http://c6403.ambari.apache.org:1025/ws/v1/slider/publisher/exports
HBASE_REGIONSERVER Host(s)/Container(s): [c6403.ambari.apache.org/container_1424910595794_0005_01_000003]
HBASE_MASTER Host(s)/Container(s): [c6403.ambari.apache.org/container_1424910595794_0005_01_000002]
slider-appmaster Host(s)/Container(s): [c6403.ambari.apache.org/container_1424910595794_0005_01_000001]
HBASE_REST Host(s)/Container(s): [c6403.ambari.apache.org/container_1424910595794_0005_01_000004]

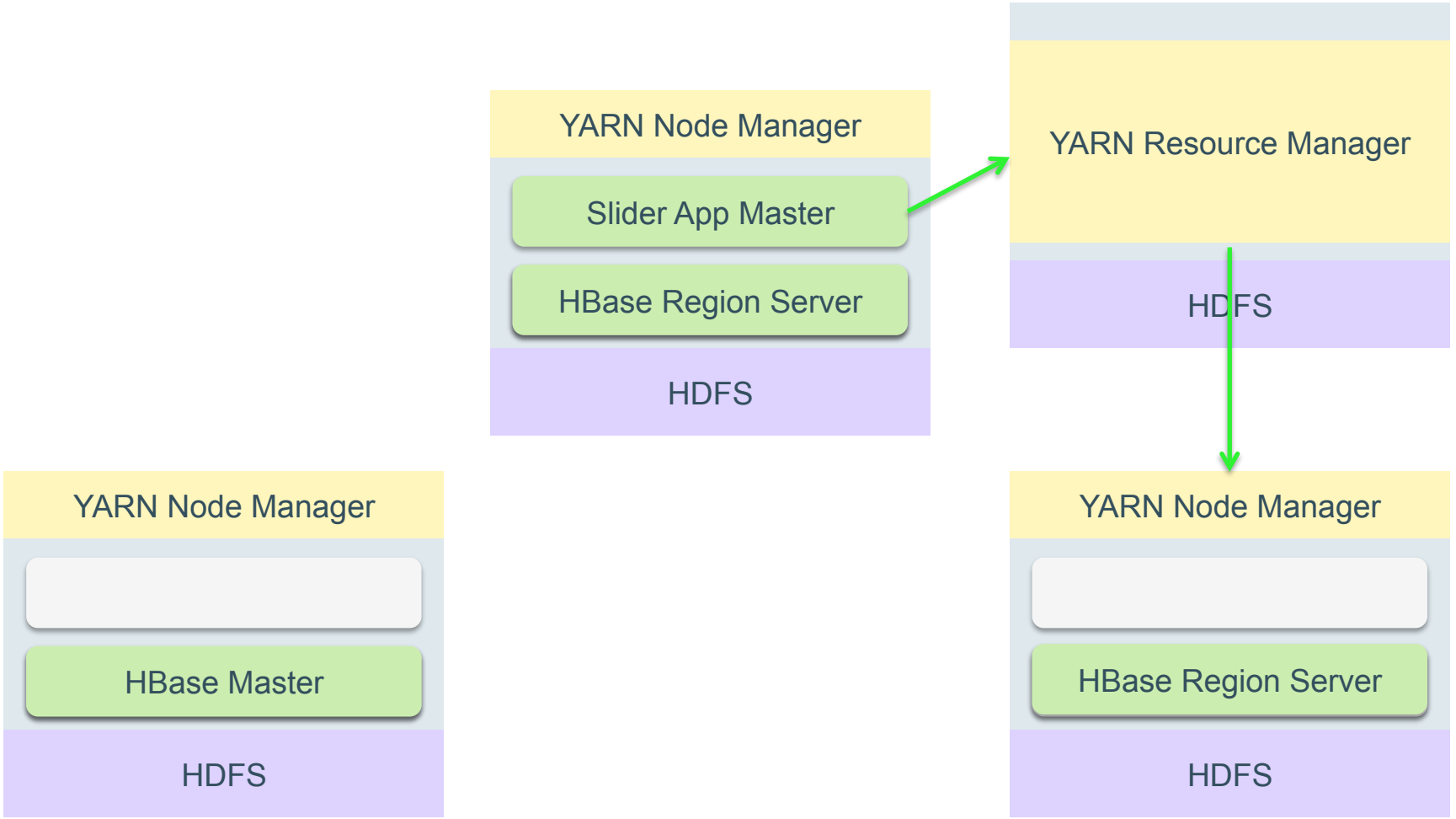
Failures



YARN notifies AM



AM requests replacement



Registration and Discovery

- Application must declare itself
  - URLs
  - Host/port
  - Config (client config)
- Application must be discoverable
  - Registry
  - Name-based lookups
  - Regularly updated
- Client support
  - Callback if “data” changes; thick clients
  - Configurable gateway; thin clients

Application Registry

- A common problem (not specific to Slider)
  - https://issues.apache.org/jira/browse/YARN-913
- Currently,
  - Apache Curator based
  - Register URLs pointing to actual data
  - AM doubles up as a webserver for published data
- Plan
  - Registry should be stand-alone
  - Slider is a consumer as well as publisher
  - Slider focuses on declarative solution for Applications to publish data
  - Allows integration of Applications independent of how they are hosted

App Packaging Capabilities

- Dynamic port allocation and sharing
- Inter-component dependency
  - Specify the start order of components
- Exports
  - Construct arbitrary name value pairs
  - E.g. URLs (org.apache.slider.monitor: **http://${HBASE_MASTER_HOST}:${site.hbase-site.hbase.master.info.port}/master-status**)
- Default HDFS and ZK isolation

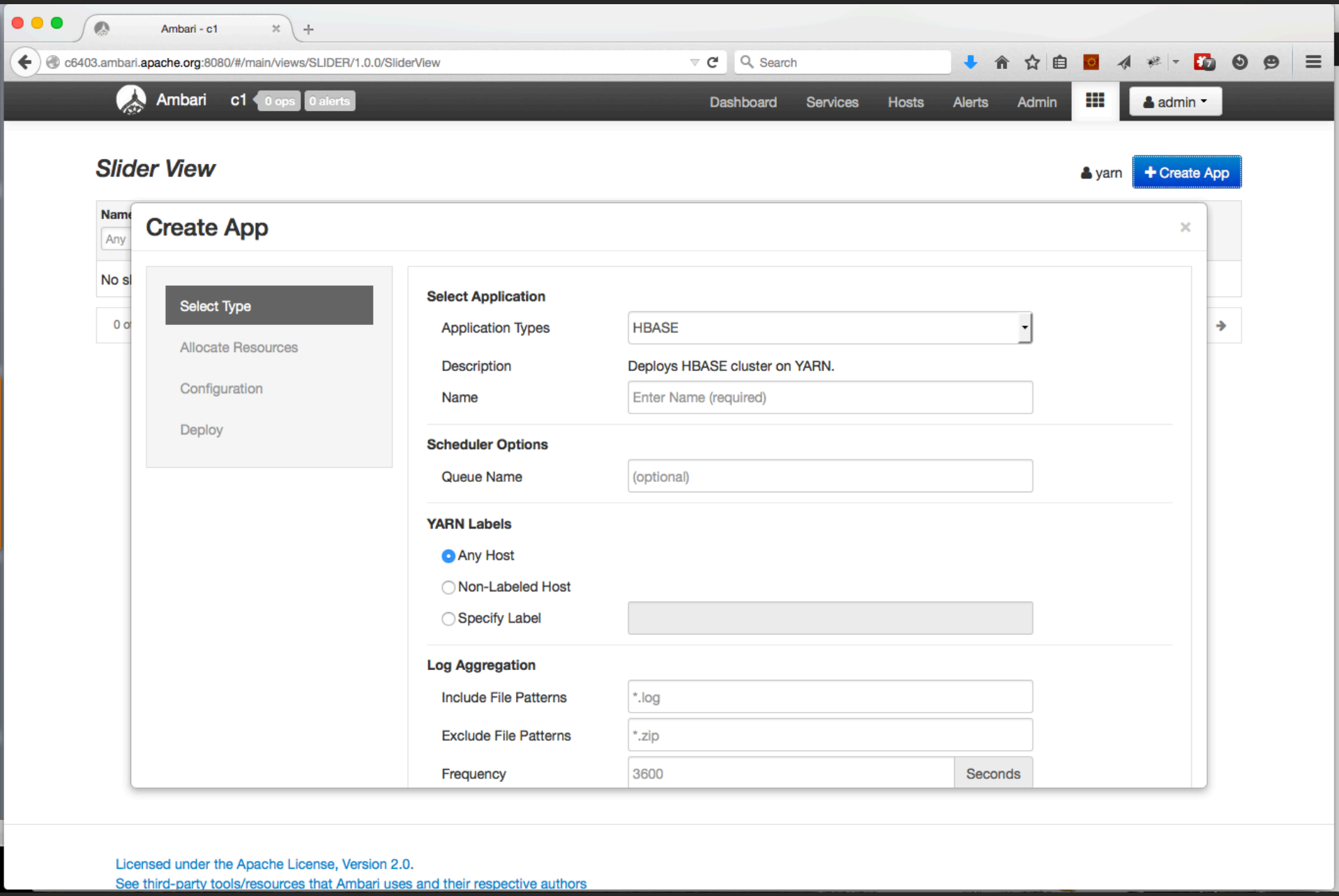
In addition to …

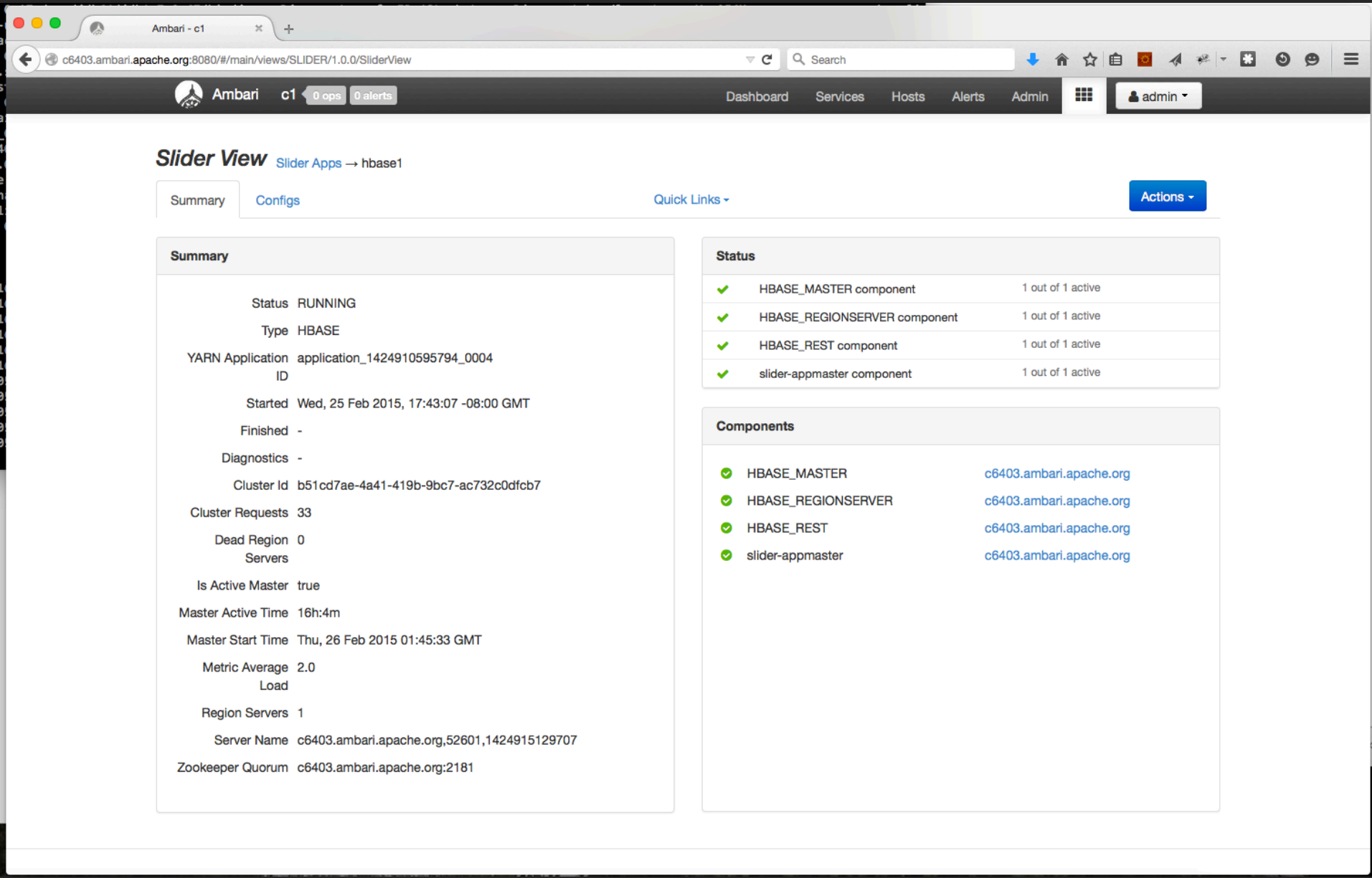
- Security
  - Configured for security
  - Token renewal and/or keytabs
- High Availability
  - On a highly available cluster (NN, RM HA)
  - Itself highly available (multi-master)
- Packaging
- Configurability
- …

Hortonworks

Slider View for Apache Ambari

http://ambari.apache.org/


Ambari - c1
c6403.ambari.apache.org:8080/#/main/views/SLIDER/1.0.0/SliderView
Search
Ambari
c1
0 ops
0 alerts
Dashboard
Services
Hosts
Alerts
Admin
admin
Slider View
yarn
+ Create App
Create App
Select Type
Allocate Resources
Configuration
Deploy
Select Application
Application Types
HBASE
Description
Deploys HBASE cluster on YARN.
Name
Enter Name (required)
Scheduler Options
Queue Name
(optional)
YARN Labels
Any Host
Non-Labeled Host
Specify Label
Log Aggregation
Include File Patterns
*.log
Exclude File Patterns
*.zip
Frequency
3600
Seconds
Licensed under the Apache License, Version 2.0.
See third-party tools/resources that Ambari uses and their respective authors


Ambari - c1
c6403.ambari.apache.org:8080/#/main/views/SLIDER/1.0.0/SliderView
Search
Ambari
c1
0 ops
0 alerts
Dashboard
Services
Hosts
Alerts
Admin
admin
Slider View
Slider Apps → hbase1
Summary
Configs
Quick Links
Actions
Summary
Status RUNNING
Type HBASE
YARN Application ID application_1424910595794_0004
Started Wed, 25 Feb 2015, 17:43:07 -08:00 GMT
Finished -
Diagnostics -
Cluster Id b51cd7ae-4a41-419b-9bc7-ac732c0dfcb7
Cluster Requests 33
Dead Region Servers 0
Is Active Master true
Master Active Time 16h:4m
Master Start Time Thu, 26 Feb 2015 01:45:33 GMT
Metric Average Load 2.0
Region Servers 1
Server Name c6403.ambari.apache.org,52601,1424915129707
Zookeeper Quorum c6403.ambari.apache.org:2181
Status
HBASE_MASTER component 1 out of 1 active
HBASE_REGIONSERVER component 1 out of 1 active
HBASE_REST component 1 out of 1 active
slider-appmaster component 1 out of 1 active
Components
HBASE_MASTER c6403.ambari.apache.org
HBASE_REGIONSERVER c6403.ambari.apache.org
HBASE_REST c6403.ambari.apache.org
slider-appmaster c6403.ambari.apache.org

Project Status

- In ASF incubation
- 0.70-incubating release - March
- Growing set of application packages
- Working in simplifying packaging
- HDP 2.2 added long-lived service support for Slider *and other services*

HDP2.2 features for Slider

- ✓ Labelled nodes & queues
- ✓ Log aggregation
- ✓ Container retention over Application Master restart
- ✓ Service registration & discovery
- ✓ Kerberos token renewal
- ✓ Windowed Application failure tracking

—essential for all long-lived YARN services

On-boarding Dockerized Application on YARN via Slider

- Application definition – metainfo.json (contains docker image name)
- Application instance configuration (if any) – app_config.json
- Application resource definition – resources.json
- Dockerized node.js app using dockerized redis for persistence

—No Application bundle and lifecycle scripts

metainfo.json – structure of the application

```
{
    "schemaVersion": "2.1",
    "application": {
            "name": "NODEJS-REDIS",
            "components": [
                {
                    "name": "NODEJS",
                    "type": "docker",
                    "dockerContainers": [
                        {
                            "name": "nodejs”,
                            "image": "rsahahw/centos-node-redis”,
                            "ports": [{"containerPort" : "8000” }]
                        }
                    ]
                },
                 {
                    "name": "REDIS",
                    "type": "docker",
                    "dockerContainers": [
                    {
                            "name": "redis",
                            "image": "tutum/redis",
                            "ports": [{
                                   "containerPort" : "6379",
                                   "hostPort": "6379”
                            ...
```

resources.json – resource requirement of the application

```
{
  "schema" : "http://example.org/specification/v2.0.0",
  "metadata" : {
  },
  "global" : {
  },
  "components": {

    "NODEJS": {
      "yarn.role.priority": "1",
      "yarn.component.instances": "1",
      "yarn.memory": "512"
     },
     "REDIS": {
      "yarn.role.priority": "2",
      "yarn.component.instances": "1",
      "yarn.memory": "512"
    }
  }
}
```

app_config.json – instance specific configuration of the application

```
{
  "schema": "http://example.org/specification/v2.0.0",
  "metadata": {
  },
  "global": {
  },
  "components": {
        "NODEJS": {
         "nodejs.options":"-d -e REDIS_HOST=${REDIS_HOST}",
              "nodejs.statusCommand":"/usr/bin/docker ps"
        },
        "REDIS": {
            "redis.options":"-d -e REDIS_PASS=**None**",
             "redis.statusCommand":"/usr/bin/docker ps"
        }
  }
}
```

Coming up

- SLIDER-799: AM-managed placement escalation

1. Redundant container request from slider may cause high load on busy cluster

- SLIDER-787 App Upgrade/Reconfig support in Slider

1. Allows rolling upgrade / downgrade
2. app packages are versioned

Hortonworks

http://slider.incubator.apache.org/