Aeromancer: A Workflow Manager for Large-Scale MapReduce-Based Scientific Workflows

Presented by Sarunya Pumma

Supervisors: Dr. Wu-chun Feng, Dr. Mark Gardner, and Dr. Hao Wang

VirginiaTech
Invent the Future

SyNeRG
synergy.cs.vt.edu

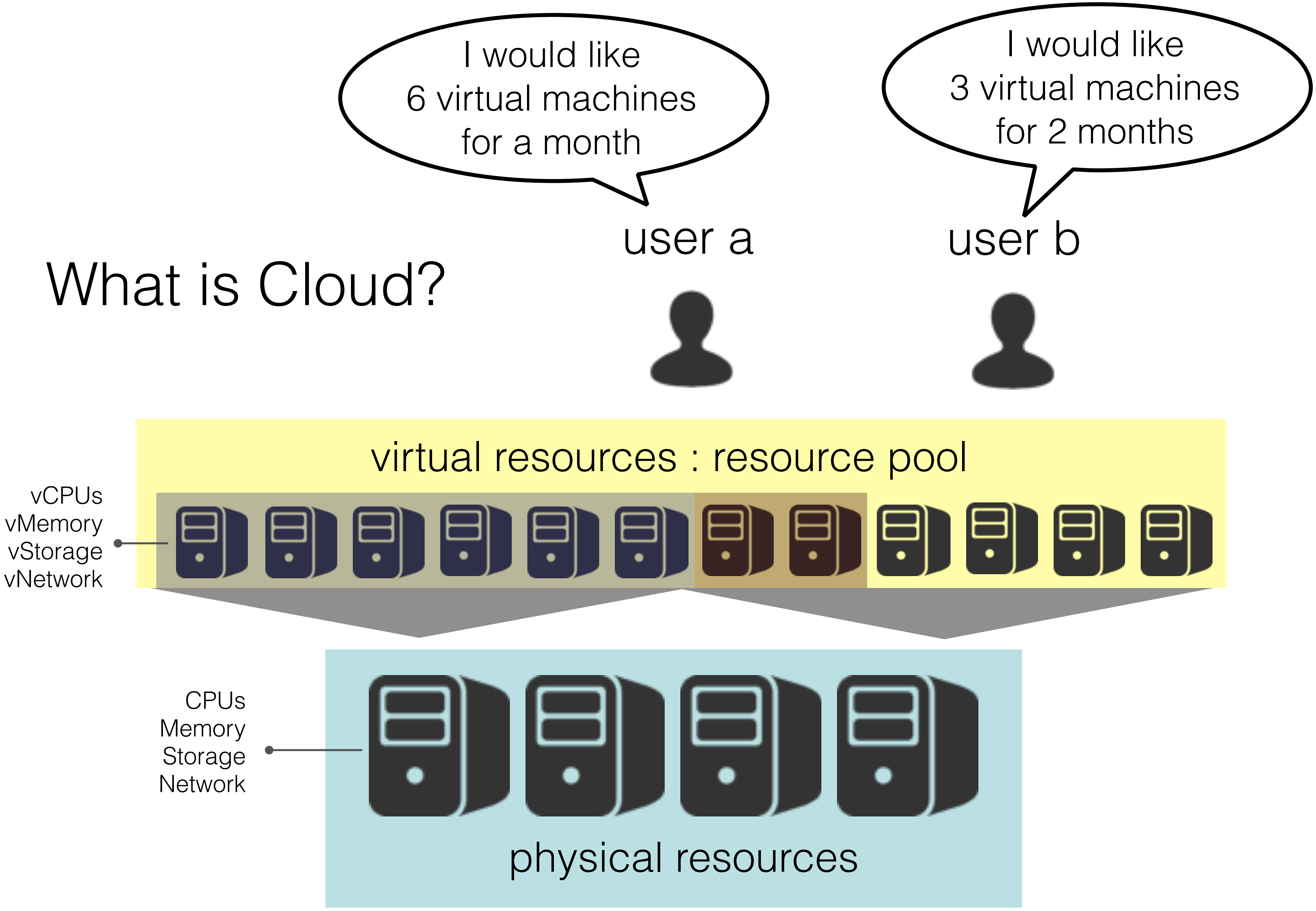
Outline

- Introduction & Motivation
  - Cloud computing and its benefits
  - SeqInCloud and scalability analysis
- Aeromancer
  - Usage scenario
  - Workflow and architecture
  - Discussion
- Experiments & Initial Results

Introduction & Motivation

What is Cloud?

What is Cloud?



I would like
6 virtual machines
for a month

I would like
3 virtual machines
for 2 months

user a

user b

virtual resources : resource pool

vCPUs
vMemory
vStorage
vNetwork

CPUs
Memory
Storage
Network

physical resources

How is Cloud beneficial for Us?

How is Cloud beneficial for Us?

Next-generation sequencing (NGS) produces a huge amount of DNA sequence data

{ 2k sequencers produce 15 PB genetic data/year }

Cost of sequencing a human-size genome decreases over time

2011	→	$95M
2012	→	$6.5k
2014	→	$1k

How is Cloud beneficial for Us?

Big Data

How is Cloud beneficial for Us?

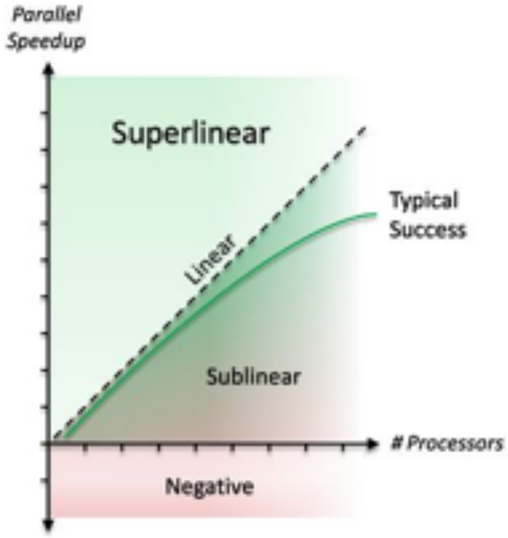
Cloud can accelerate our ability
to identify the cure for cancer

Big Data Challenge

Provisioning compute and storage resources to analyze the exponentially increasing genomic data

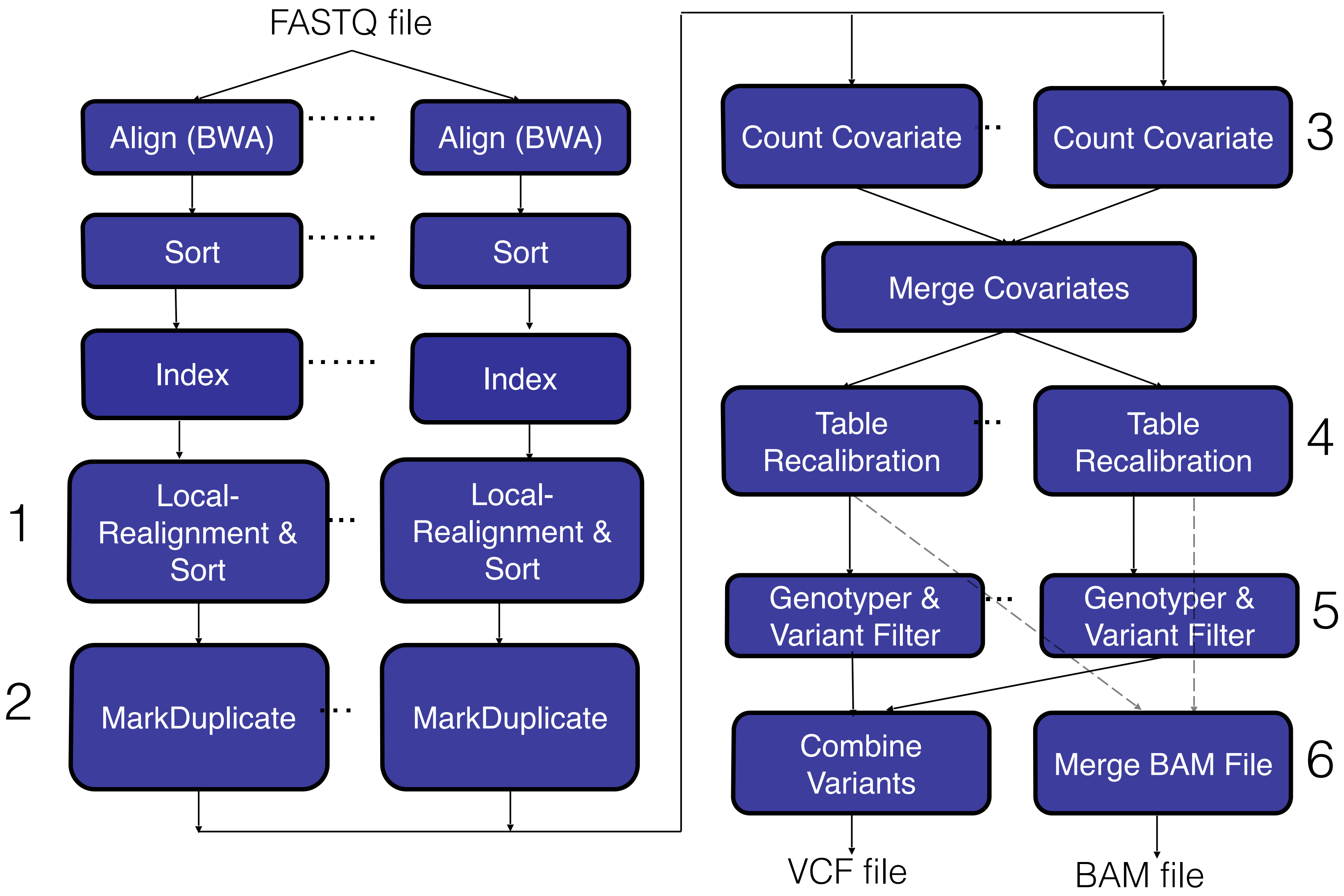
SeqInCloud

- It is implemented based on GATK (Genome Analysis Toolkit)
- All six stages have to be executed in order
- Each stage is implemented as a small separated Hadoop program
  - It can be broken down into small jobs and run on multiple computers/processors
  - We expect a linear speed-up on every stage

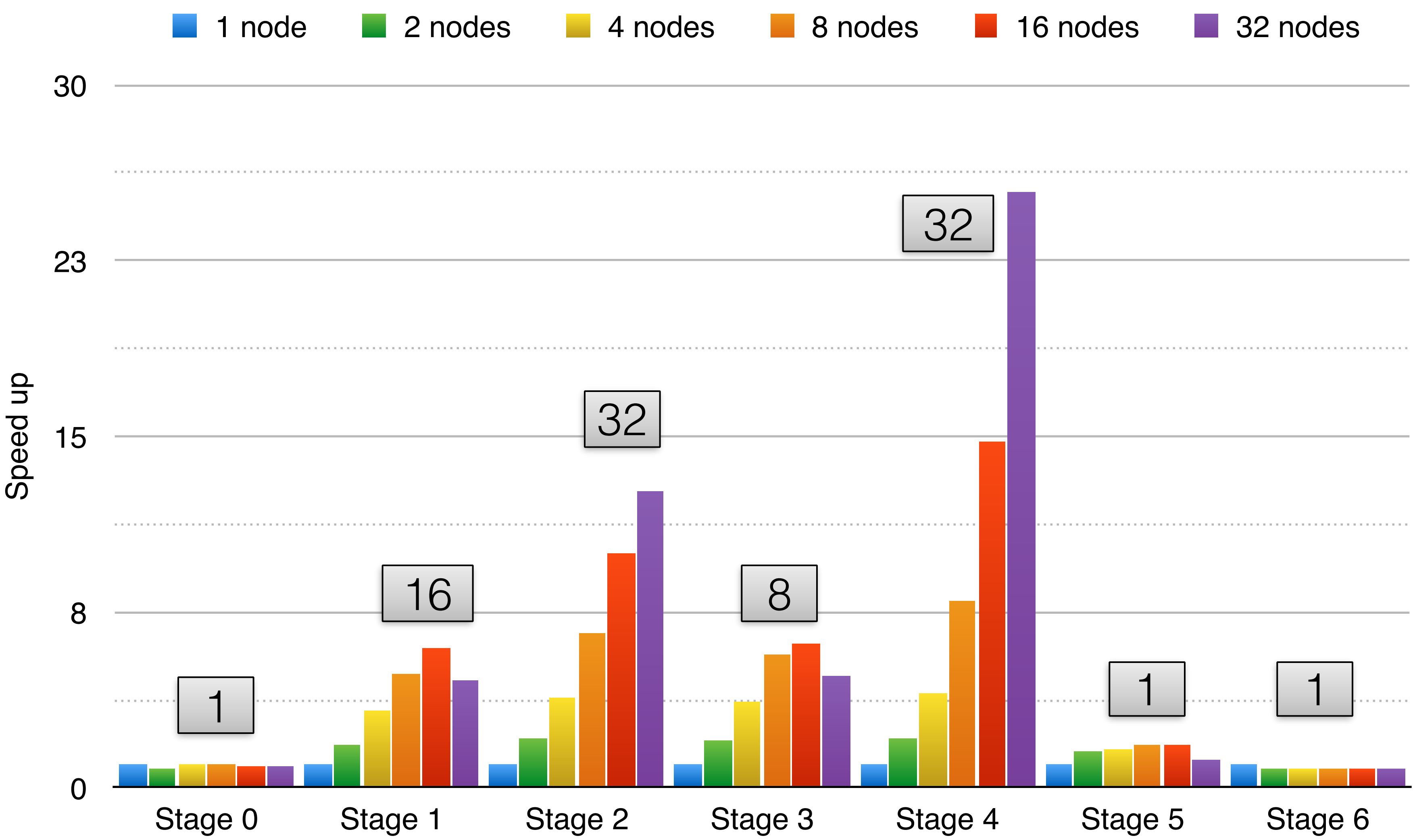


SeqInCloud

Genome variant analysis pipeline



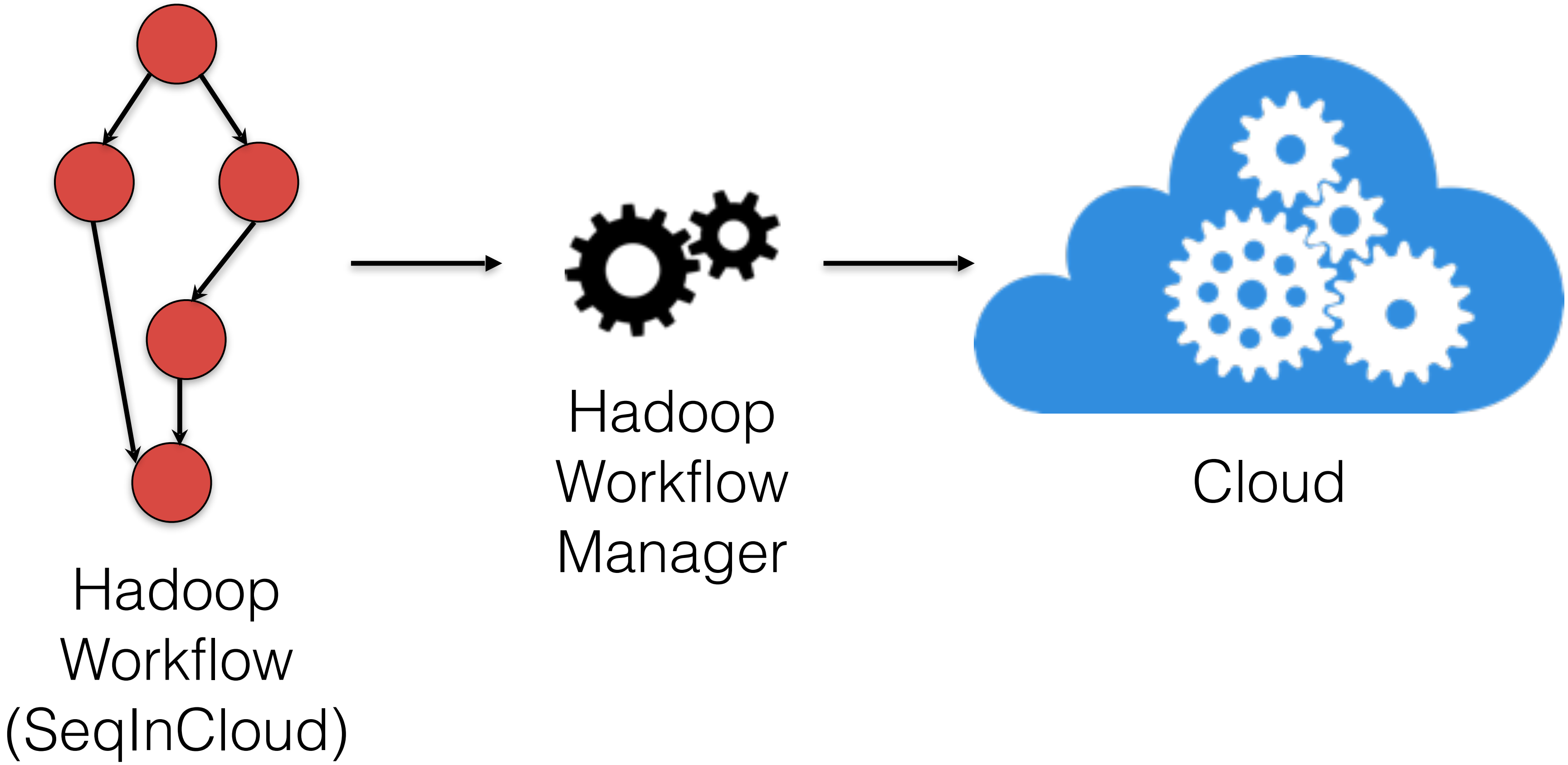
SeqInCloud Scalability Analysis



SeqInCloud Scalability Analysis

- To run SeqInCloud efficiently on Hadoop
  - Number of compute nodes must be varied based on the scalability of each stage
- Challenges
  1. How many nodes do we need for each stage?
  2. How to rapidly adjust a number of nodes with negligible overhead?

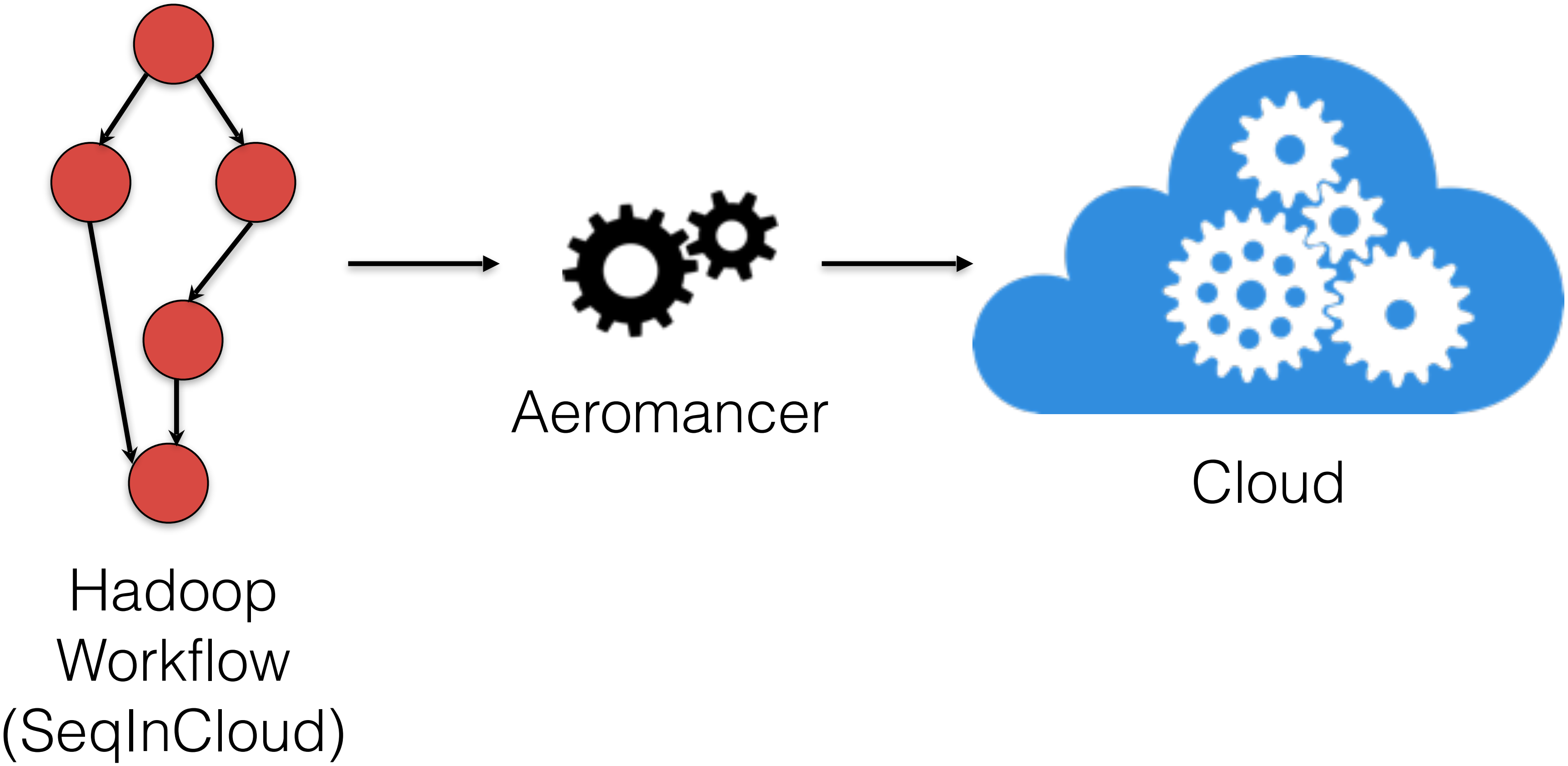
An automatic Hadoop workflow execution is needed!



Hadoop Workflow (SeqInCloud)

Hadoop Workflow Manager

Cloud



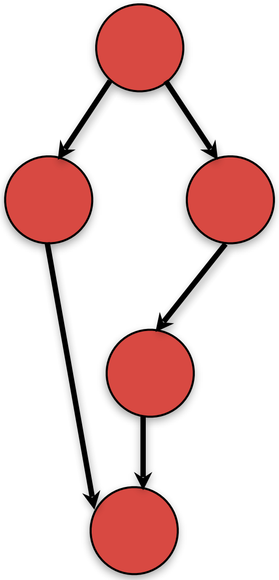
Hadoop Workflow (SeqInCloud)

Aeromancer

Cloud

Aeromancer

- Platform for executing the scientific Hadoop workflows
  - Based on Hadoop
    - MapReduce
    - HDFS
  - Build on top of CloudGene
  - Workflow (DAG)
    - Node/Stage
    - Edge (represent dependency between two stages)

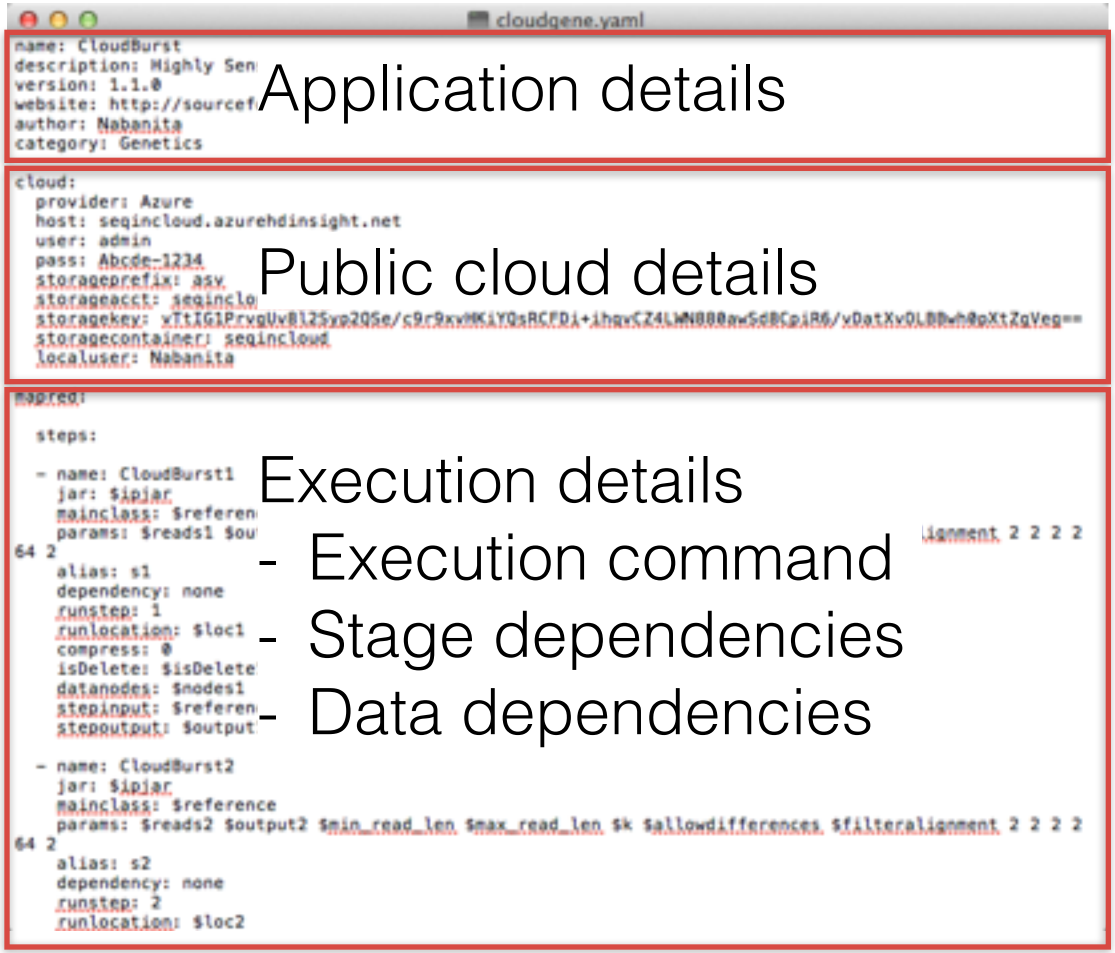


Aeromancer Usage Scenario

1

Add a new application to Aeromancer

Aeromancer Usage Scenario



```
name: CloudBurst
description: Highly Sen
version: 1.1.0
website: http://sourcef
author: Nabanita
category: Genetics
cloud:
  provider: Azure
  host: seqincloud.azurehdinsight.net
  user: admin
  pass: Abcde-1234
  storageprefix: asv
  storageacct: seqinclo
  storagekey: vTtIG1PrvgUvBl2Syp2QSe/c9r9xvHKiYQsRCFDi+ihgvCZ4LWN880awSdBCpiR6/vDatXvDLB0wh0pXtZgVeg==
  storagecontainer: seqincloud
  localuser: Nabanita
mapred:

  steps:

  - name: CloudBurst1
    jar: $ipjar
    mainclass: $referen
    params: $reads1 $ou                                                    lignment 2 2 2 2
64 2
    alias: s1
    dependency: none
    runstep: 1
    runlocation: $loc1
    compress: 0
    isDelete: $isDelete
    datanodes: $nodes1
    stepinput: $referen
    stepoutput: $output

  - name: CloudBurst2
    jar: $ipjar
    mainclass: $reference
    params: $reads2 $output2 $min_read_len $max_read_len $k $allowdifferences $filteralignment 2 2 2 2
64 2
    alias: s2
    dependency: none
    runstep: 2
    runlocation: $loc2
```

Application details

Public cloud details

Execution details
- Execution command
- Stage dependencies
- Data dependencies

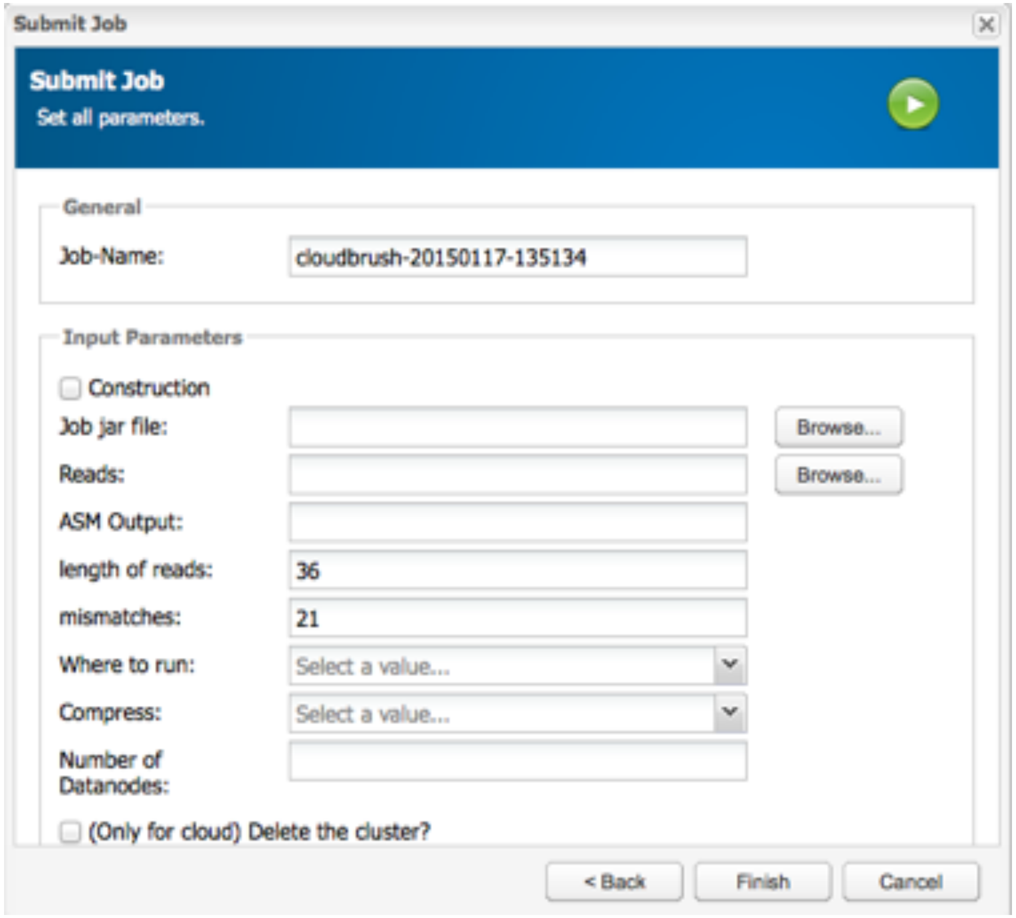
YAML file

Aeromancer Usage Scenario

1 Add a new application to Aeromancer

2 Submit a new job

Aeromancer Usage Scenario



Submit Job

Set all parameters.

General

Job-Name: cloudbrush-20150117-135134

Input Parameters

□ Construction

Job jar file:

Reads:

ASM Output:

length of reads: 36

mismatches: 21

Where to run: Select a value...

Compress: Select a value...

Number of Datanodes:

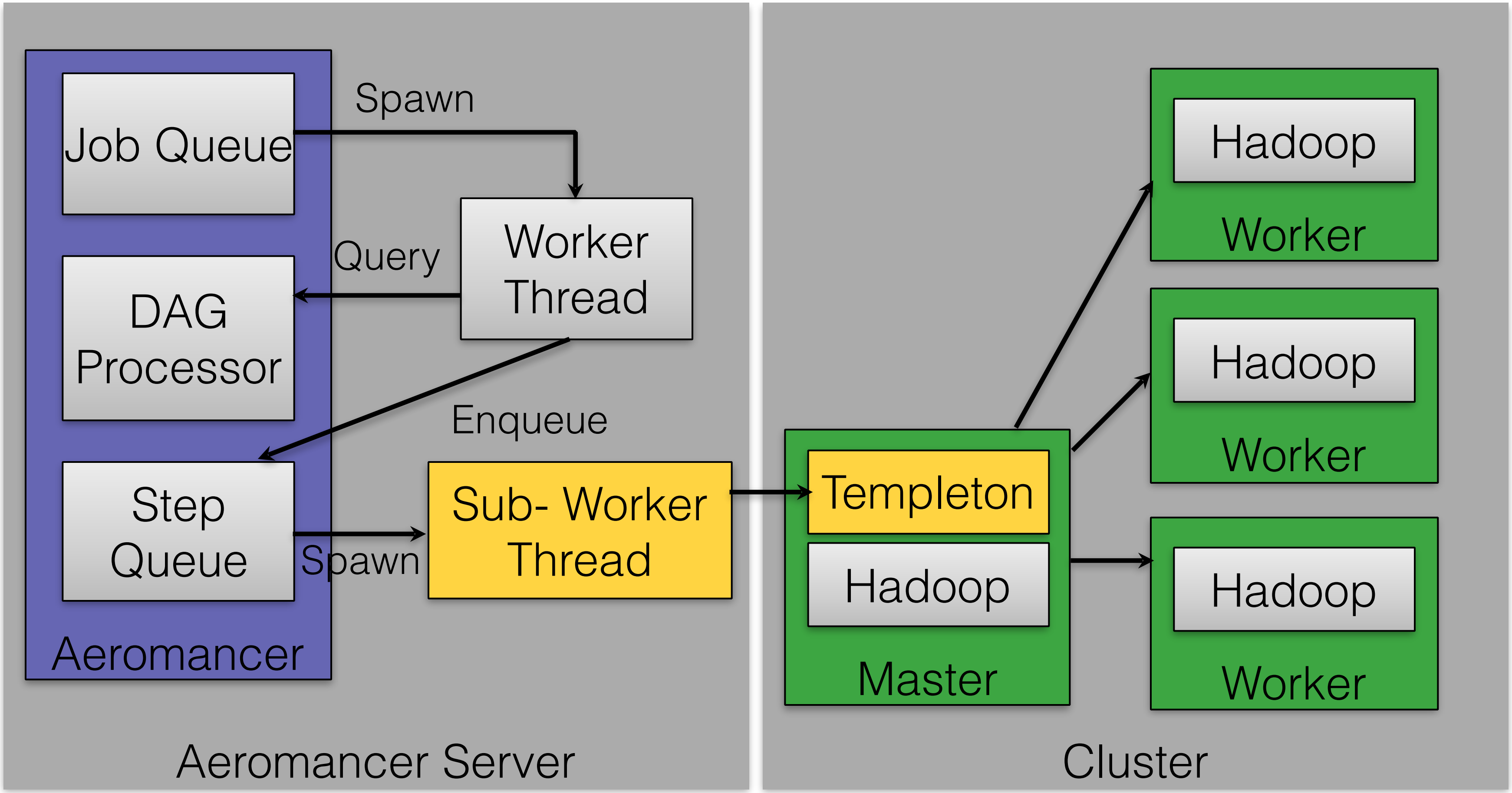
□ (Only for cloud) Delete the cluster?

User Portal

Aeromancer Usage Scenario

1. Add a new application to Aeromancer
2. Submit a new job
3. Wait for the execution to complete

Aeromancer Workflow & Architecture



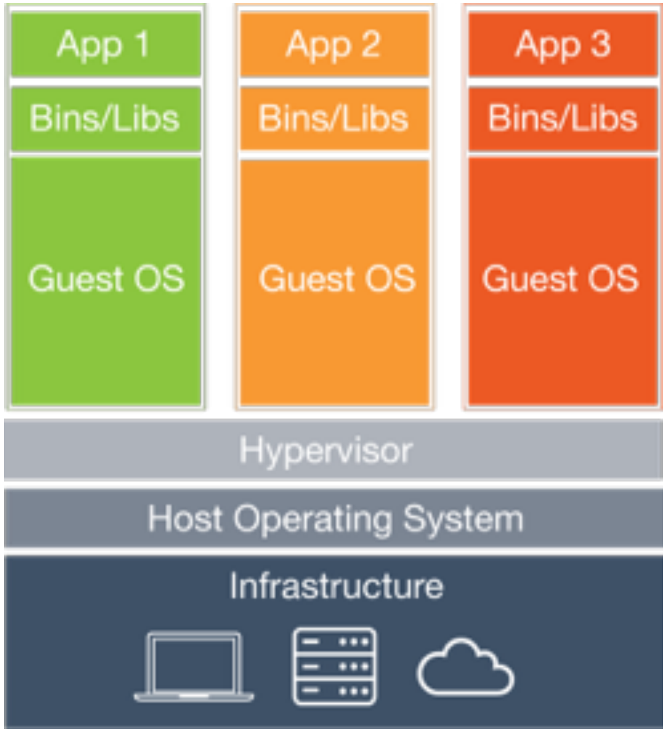
Aeromancer Discussion

- Currently, Aeromancer worker nodes are **virtual machines** in the cloud
- Pro
  - Number of worker nodes are flexibly adjustable upon demand
- Con
  - The overhead of creating/deleting worker nodes is high
  - Execution speed drops significantly

Containerization may be able to
improve Aeromancer’s performance

Virtualization

KVM



App 1 | App 2 | App 3

Bins/Libs | Bins/Libs | Bins/Libs

Guest OS | Guest OS | Guest OS

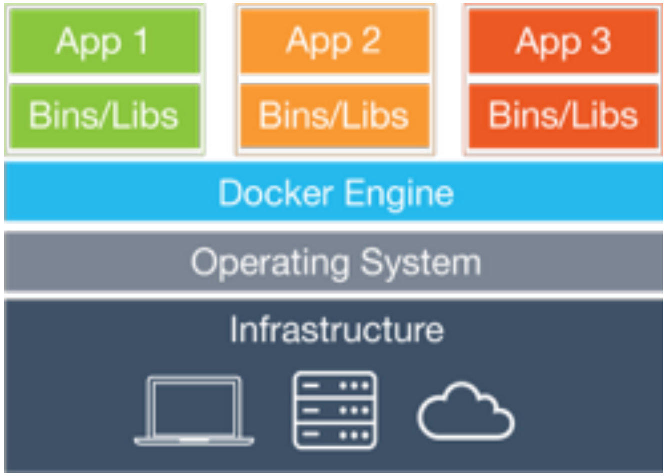
Hypervisor

Host Operating System

Infrastructure

Containerization

docker



App 1 | App 2 | App 3

Bins/Libs | Bins/Libs | Bins/Libs

Docker Engine

Operating System

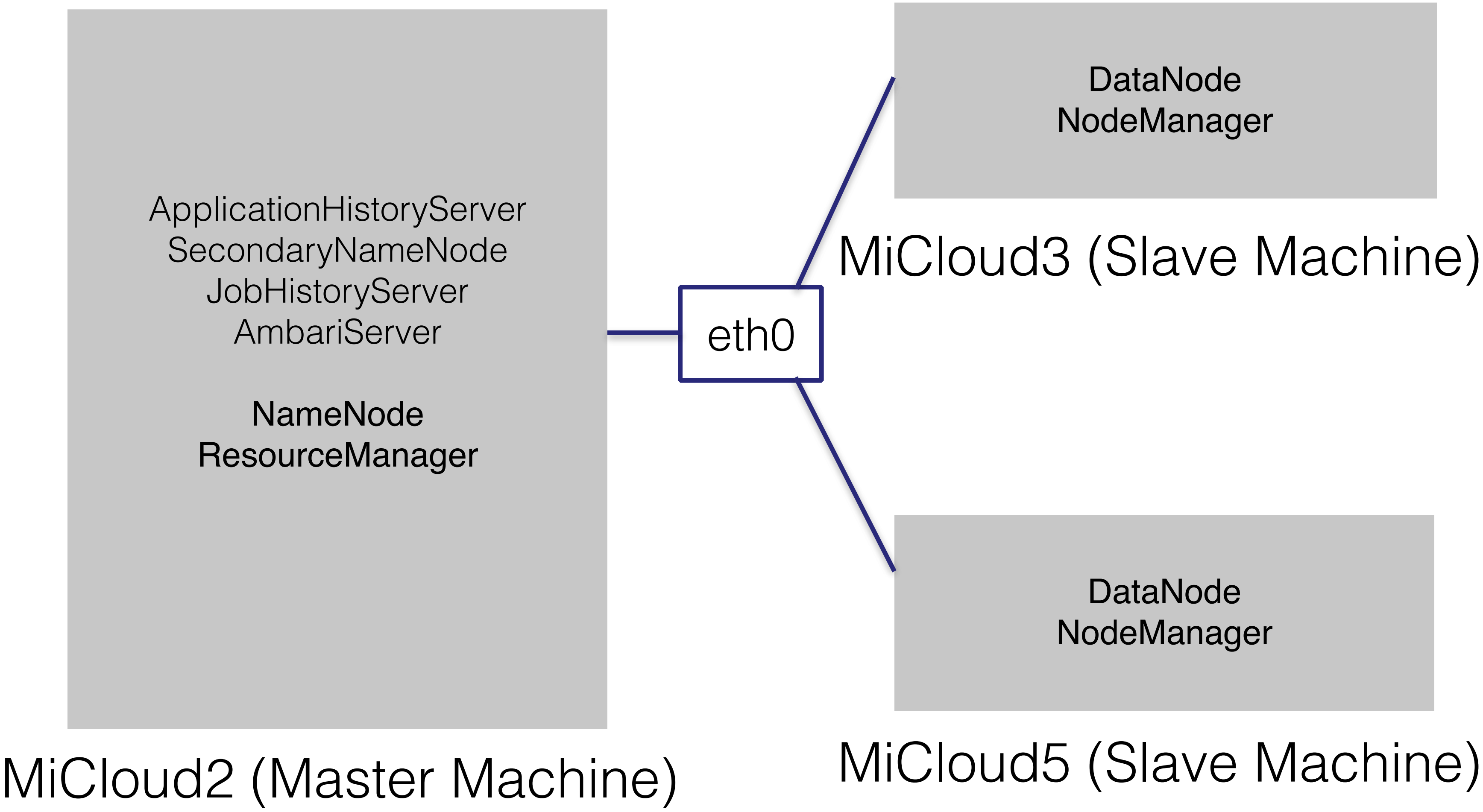
Infrastructure

Reference: https://www.docker.com/whatisdocker

Experiment

- Platforms
    - Bare metal machines
    - KVM
    - Docker
- Benchmarks (HiBench)
    - Sort
    - WordCount
    - TeraSort
    - Nutch Indexing
    - PageRank
    - EnhancedDFSIO
- Performance
    - Speed (execution time)
    - HDFS throughput

Experiment:
Hadoop on Bare Metal Environment



ApplicationHistoryServer
SecondaryNameNode
JobHistoryServer
AmbariServer

NameNode
ResourceManager

MiCloud2 (Master Machine)

eth0

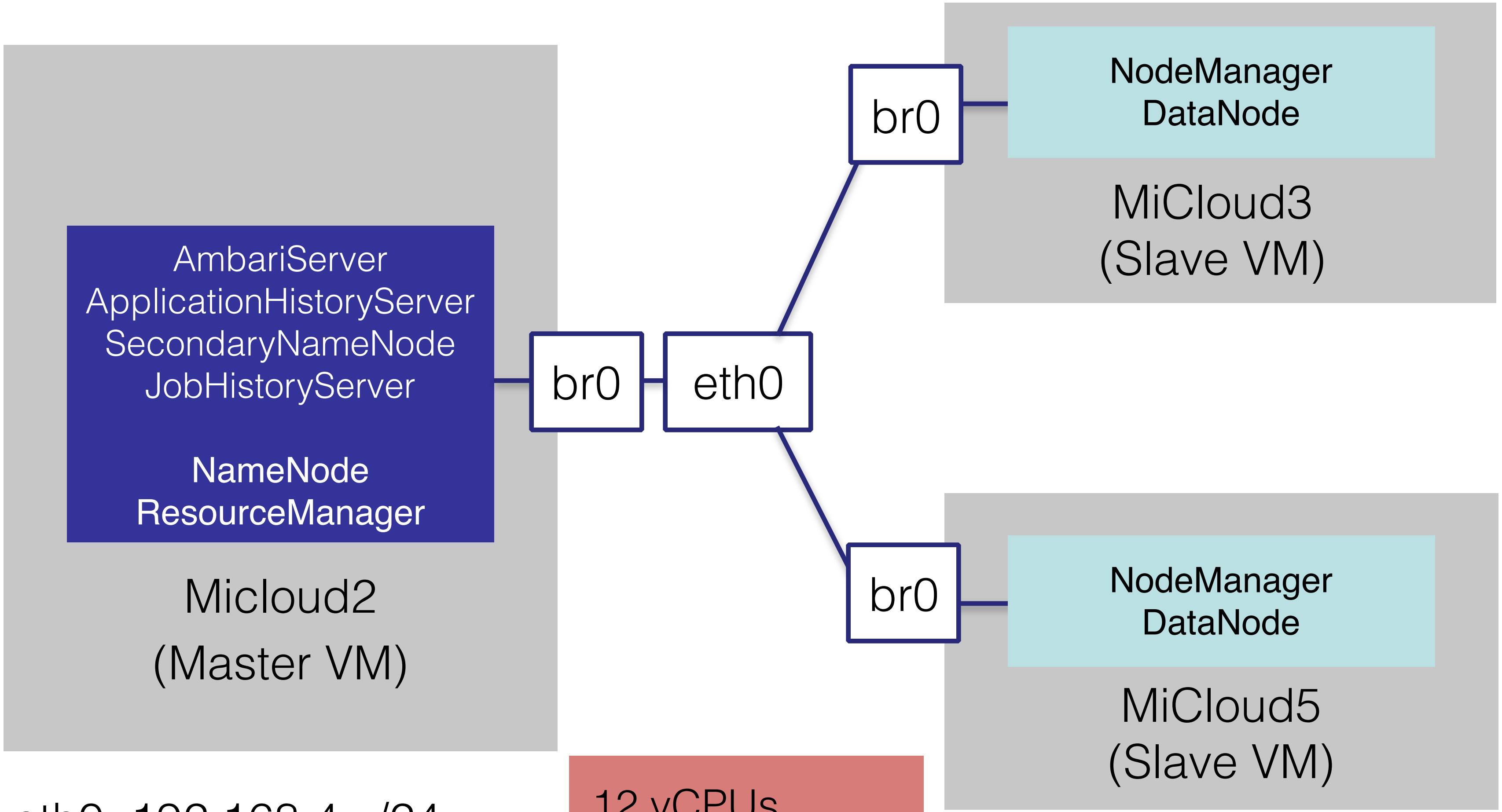
DataNode
NodeManager

MiCloud3 (Slave Machine)

DataNode
NodeManager

MiCloud5 (Slave Machine)

Experiment:
Hadoop on Virtual Machine Environment



AmbariServer
ApplicationHistoryServer
SecondaryNameNode
JobHistoryServer

NameNode
ResourceManager

Micloud2
(Master VM)

br0

eth0

br0

NodeManager
DataNode

MiCloud3
(Slave VM)

br0

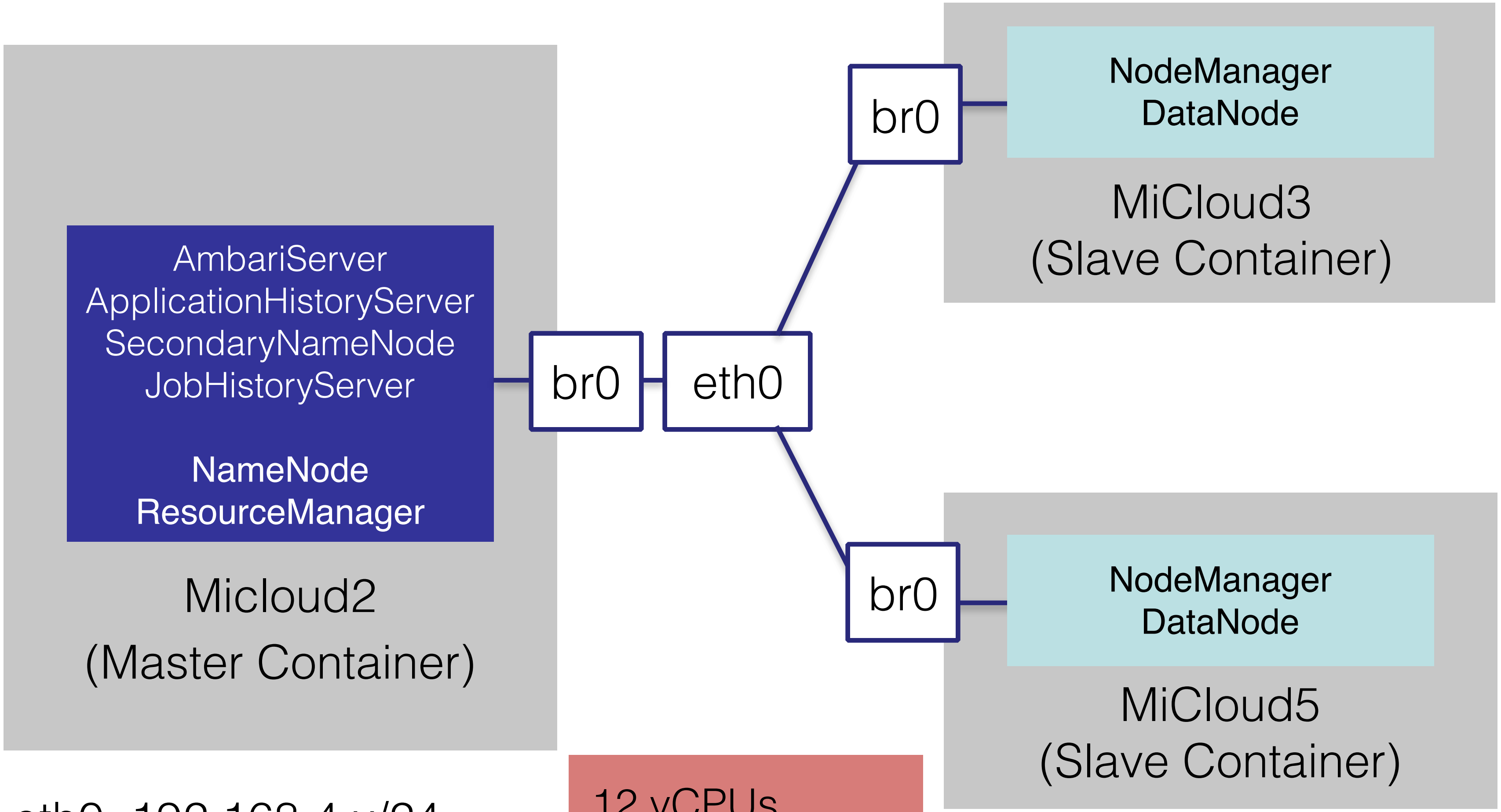
NodeManager
DataNode

MiCloud5
(Slave VM)

eth0: 192.168.4.x/24
br0: 192.168.254.x/24

12 vCPUs
16 GB-vMemory

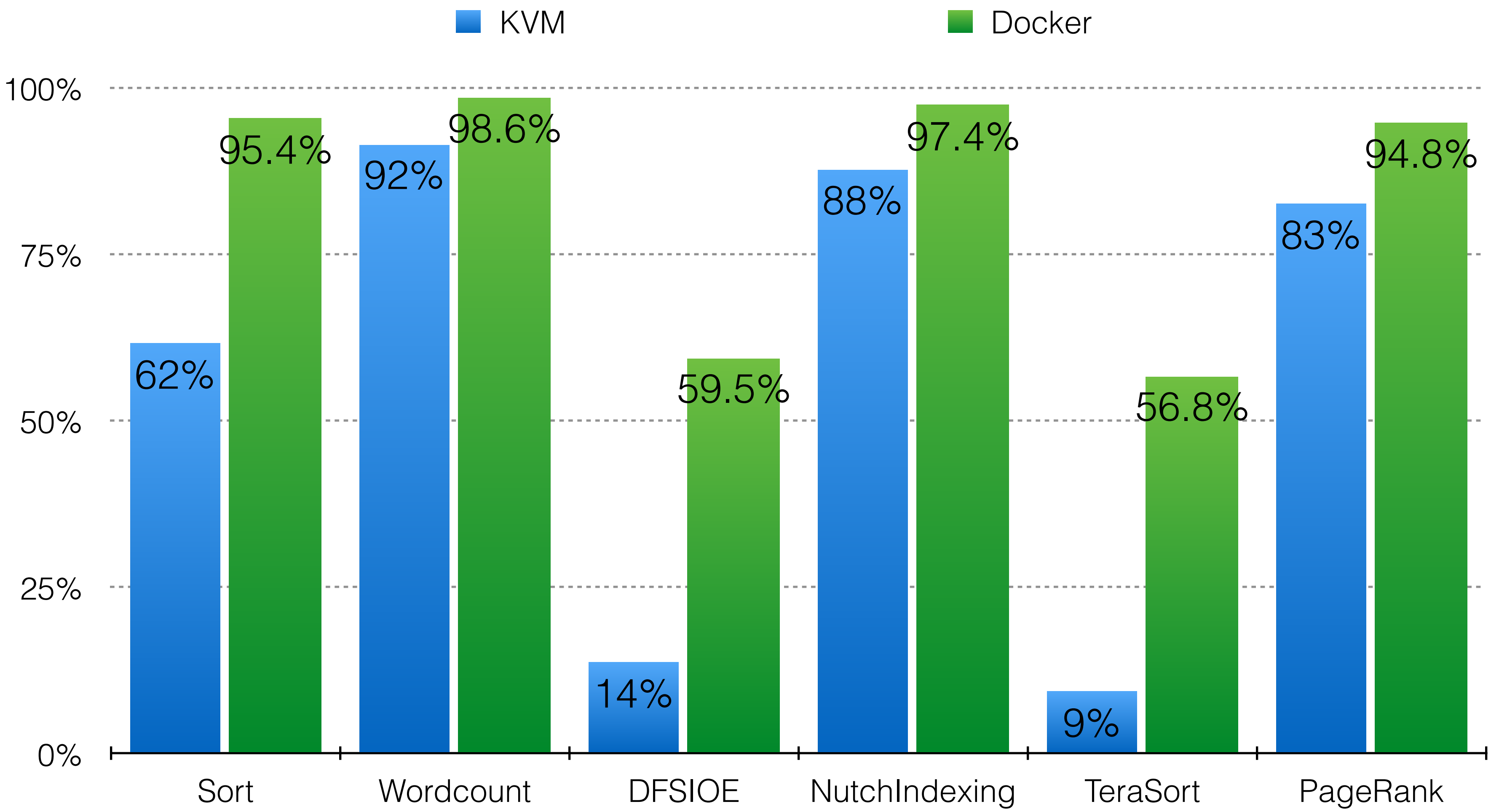
Experiment:
Hadoop on Docker



eth0: 192.168.4.x/24
br0: 192.168.6.x/24

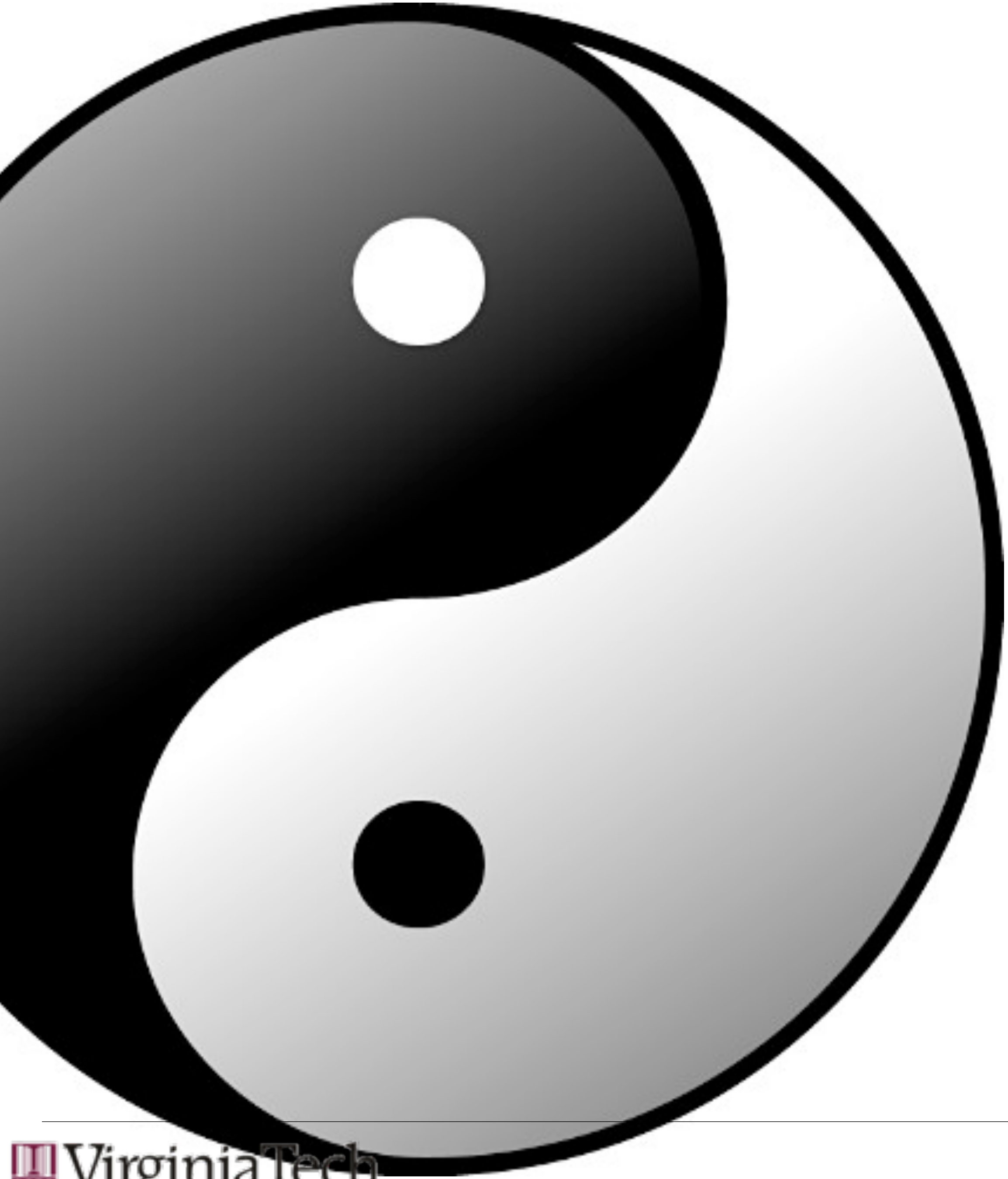
12 vCPUs
16 GB-vMemory

Initial Results: Slow Down



Future Work

- Automatic resources provisioning
  - Determining the number of computing nodes needed for each stage
    - Based on the scalability of each stage
    - Based on user’s constraints (budget/expected response time)
- Automatic task scheduling
  - Determining the optimal execution location for each stage (either client or cloud)



Thank you