

# Proven Tools to Simplify Hadoop Environments



The power to do more



[ DELL Korea | Next Generation Compute Solutions | 최주열 부장 ]

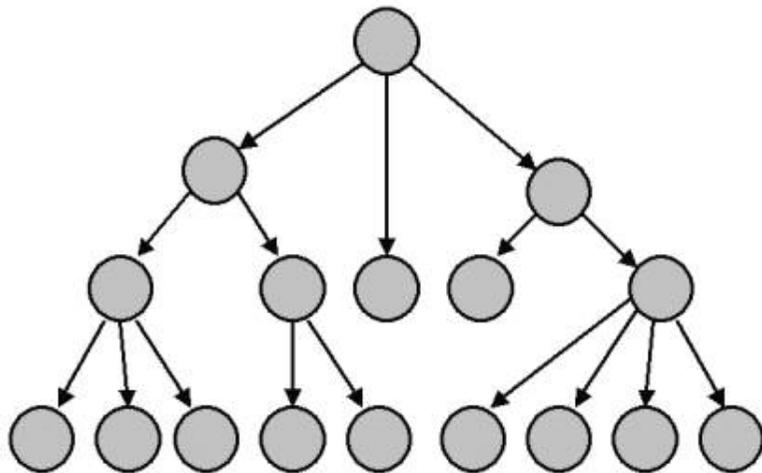


# Hadoop Operational Models



## Traditional Datacenter

- Assigned Servers
- Rigid Policies
- Tiered Software



## Cloud Infrastructure

- Elastic Resources
- Services (APIs)
- Distributed Software



# Operational Challenges



The power to do more

- Deployment
  - Complex because of scale (60 nodes to 1000 nodes)
  - Cumbersome because of high-touch processes
- Configuration & Tuning
  - Error-prone configuration management
  - State management
- Monitoring and Management
  - Complex troubleshooting and diagnostics
  - No proactive notification of problems
- Performance Optimization
  - Limitations of traditional tools



# CloudOps Framework



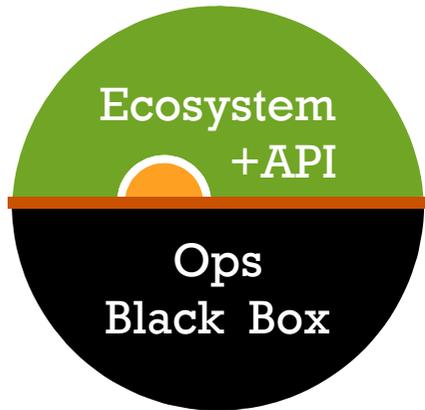
The power to do more



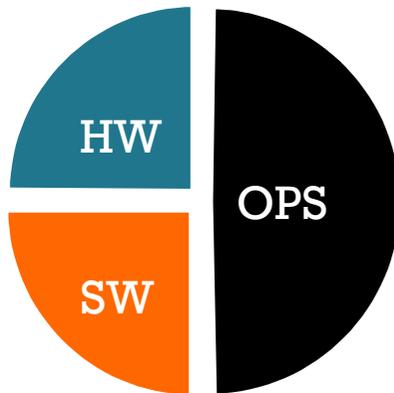
# Three aspects of revolutionary clouds



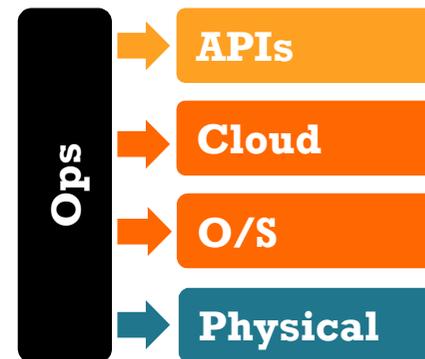
## Two Sides of Cloud



Cloud = Operations

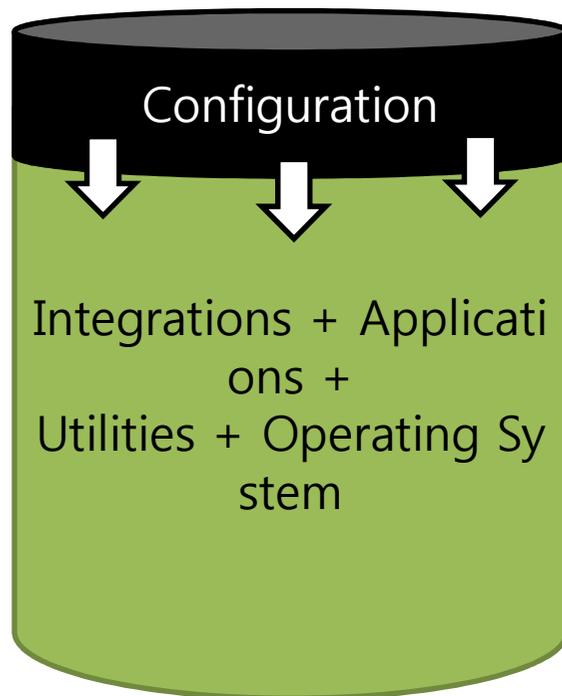


CloudOps

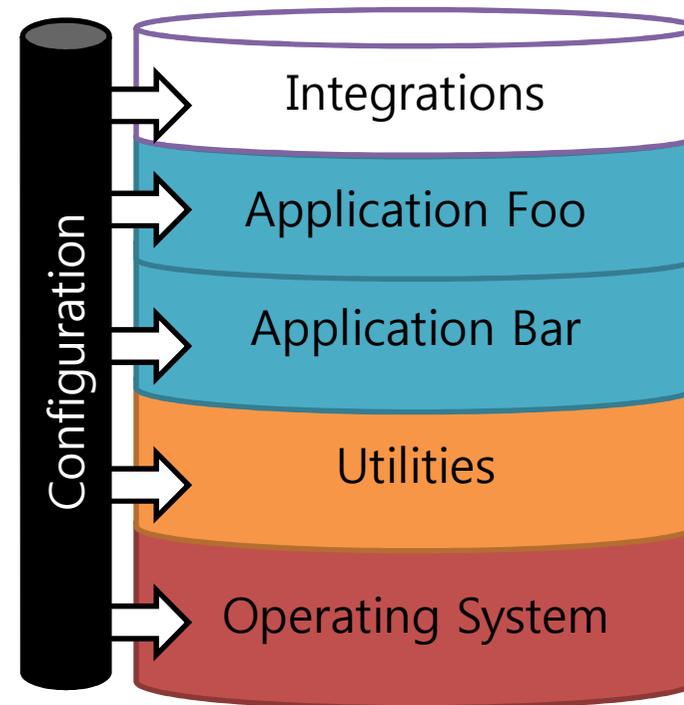


# Images v.s. Layers: Overview

- Images: Single Unit

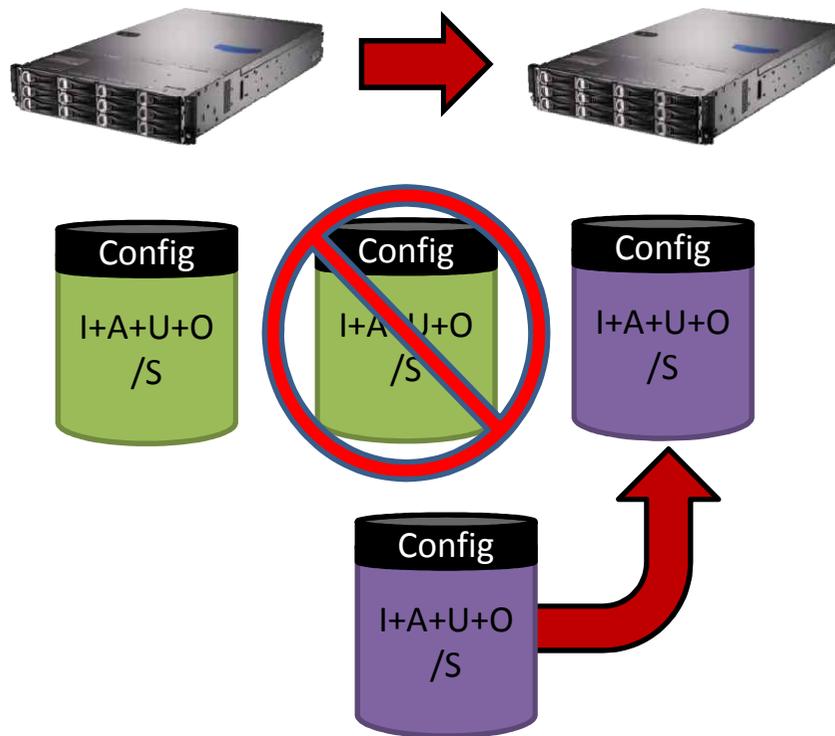


- Layers: Stacked Pieces

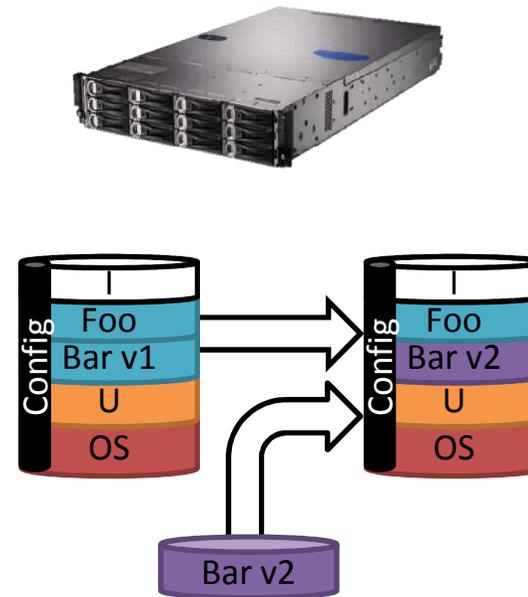


# Images v.s. Layers: Lifecycle

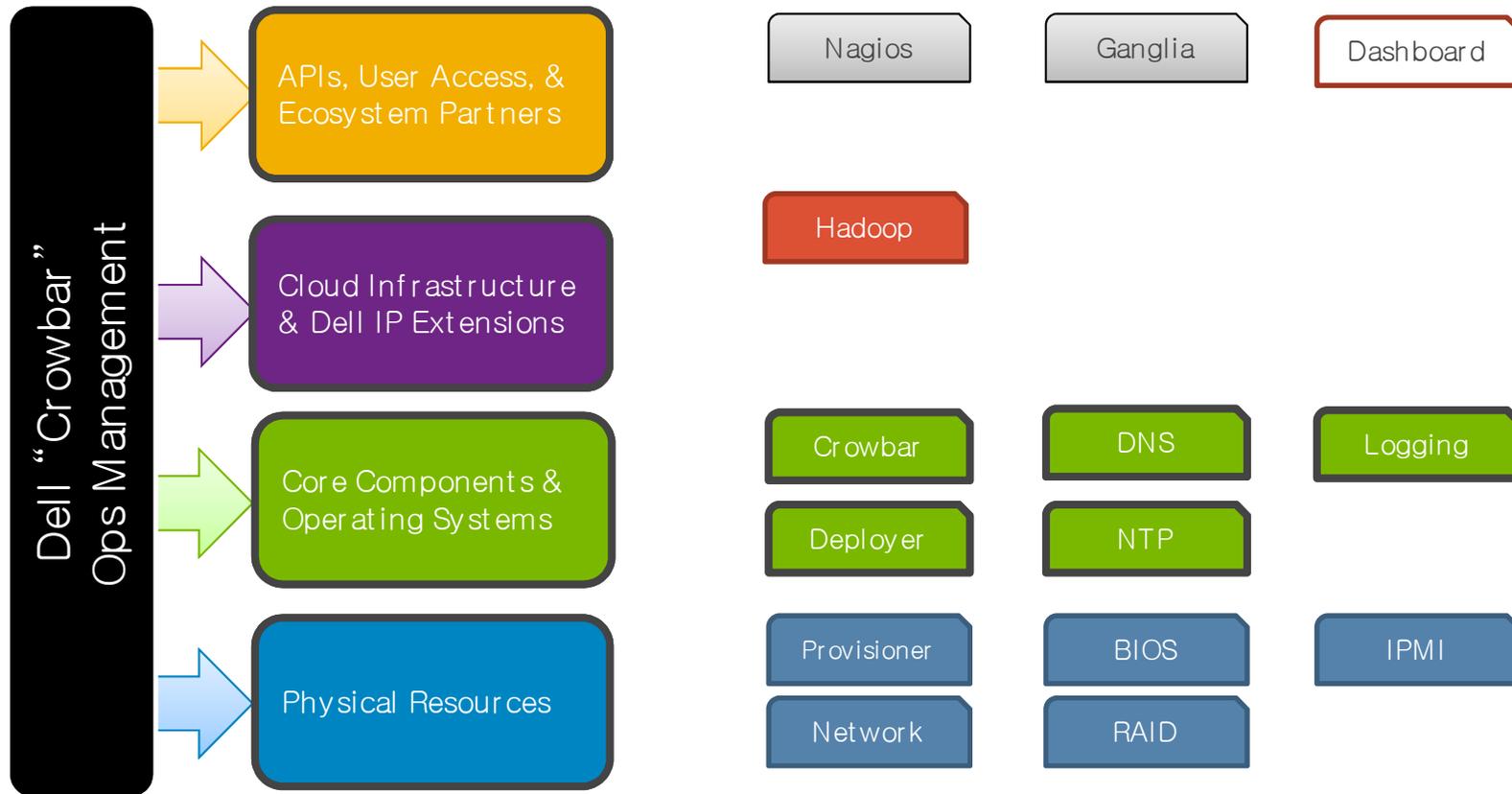
## Images: Replacement



## Layers: Upgrade



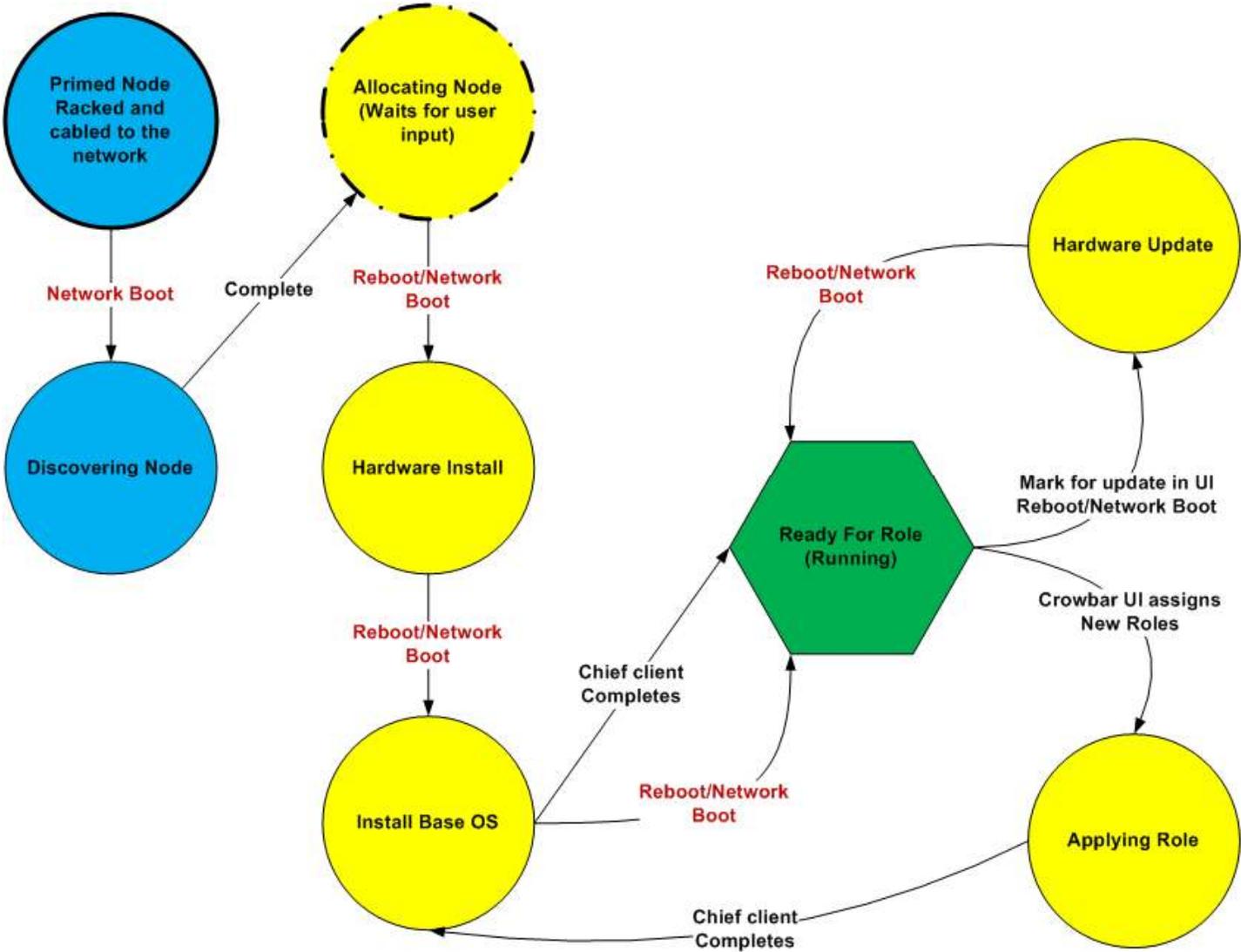
# Modular Design: Bar clamps



# Crowbar = Install State Machine



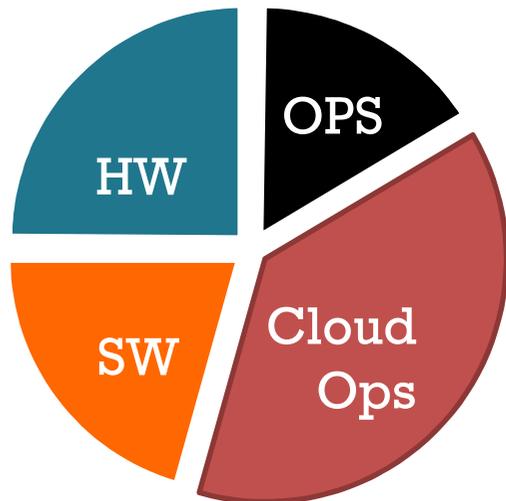
The power to do more



# Cloud = Ops



We have capable hardware & software, the *real question* is how are we going to operate it as a service?



This is *CloudOps*

Software mindset to infrastructure

- Software is constantly changing
- Fluid resources instead of servers
- Manual touch is unacceptable



# Second Act



The power to do more



# Platform Selection



## Dell PowerEdge C2100

- Designed with big data in mind
- Compact 2U form factor
- 2-socket 6-core
- Intel® Xeon® 5620 processor
- High performance memory system
- Expansive disk storage



## Recommended Configuration

- Intel Xeon Processor 5600 series
- 4-6 1TB or 2TB 7200 RPM SATA SSD
- 12-24GB DDR3 R-ECC RAM
- 1-2 dual-port 1GigE
- Linux kernel 2.6.30 or later
- Sun Java 6u14 or later
- Hadoop version 0.20.x or later

Intel Whitepaper: "Optimizing Hadoop Deployments" (<http://software.intel.com/file/31124>)

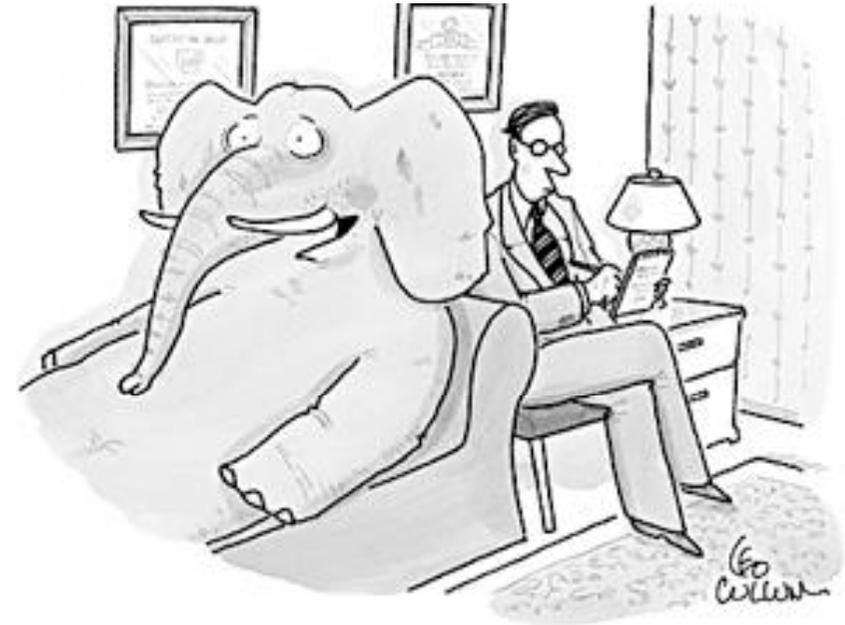


# So what seems to be the problem?



The power to do more

- Data flow and high level abstraction make it difficult to understand runtime behaviors
- Large distributed system makes it difficult to correlate concurrent performance-related activities

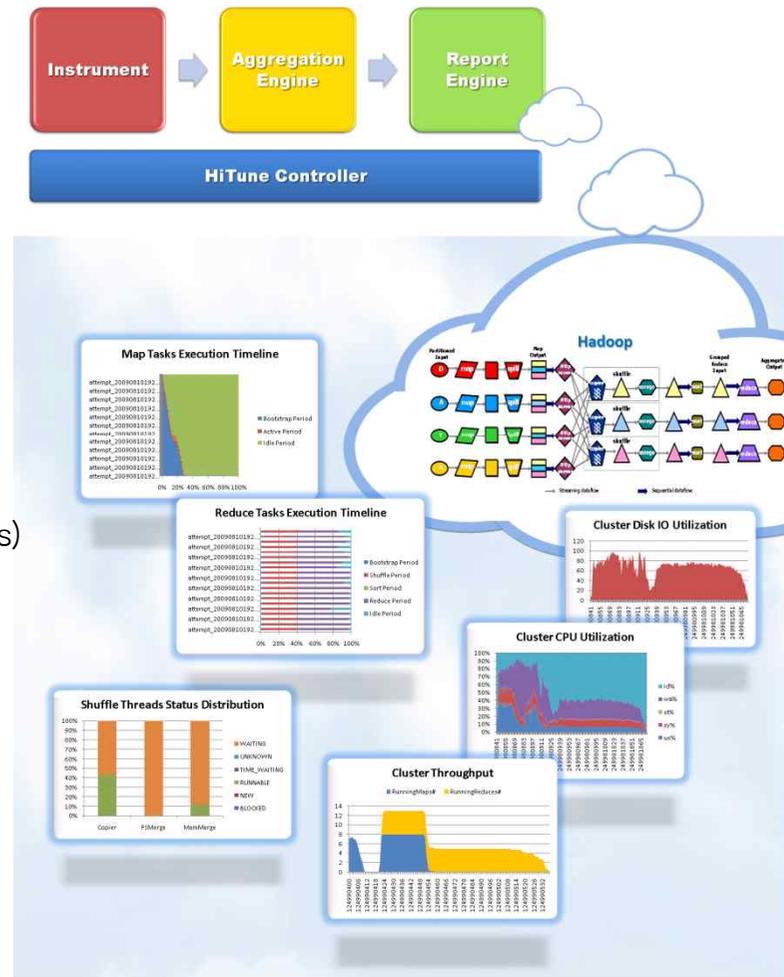


*"I'm right there in the room, and no one even acknowledges me."*

# HiTune: Hadoop Performance Analyzer



- Collects metrics from each node
  - Aggregates data using Chukwa
  - Analyzes results using Hadoop
  - Generates reports for visualization
- 
- System metric (CPU, Disk I/O, Network IO, Memory)
  - Hadoop metrics (NameNode, DataNode, JobTracker, TaskTracker, JVM metrics)
  - Dataflow based statistics (Job, MapTasks, Reduce Tasks, Threaddump for M/R)
  - Summary view of a single job
  - Summary view by comparing multiple jobs

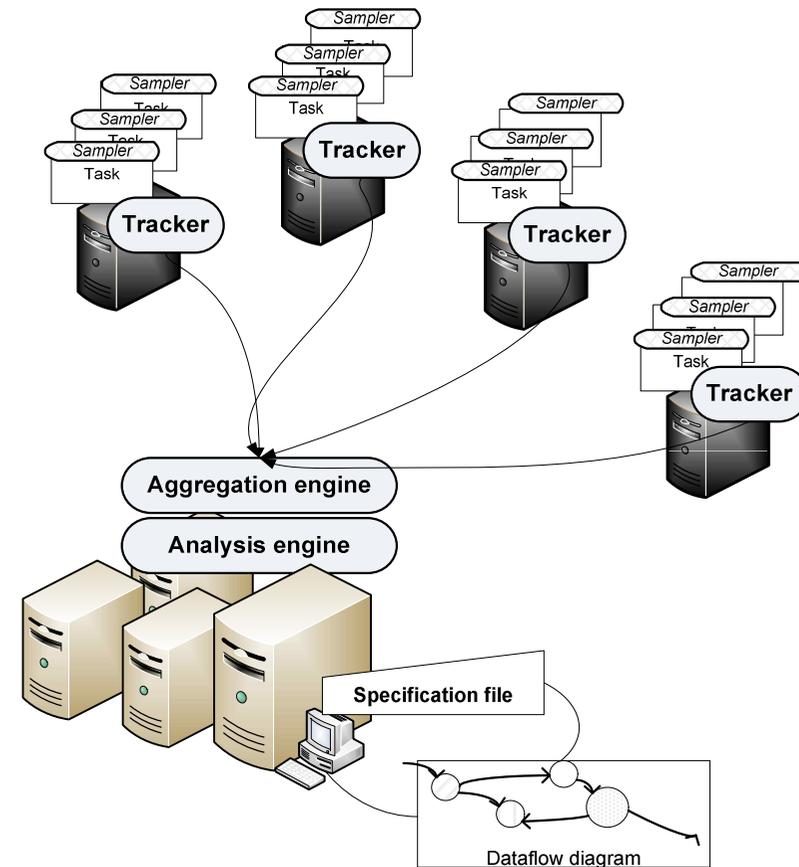


# HiTune Architecture

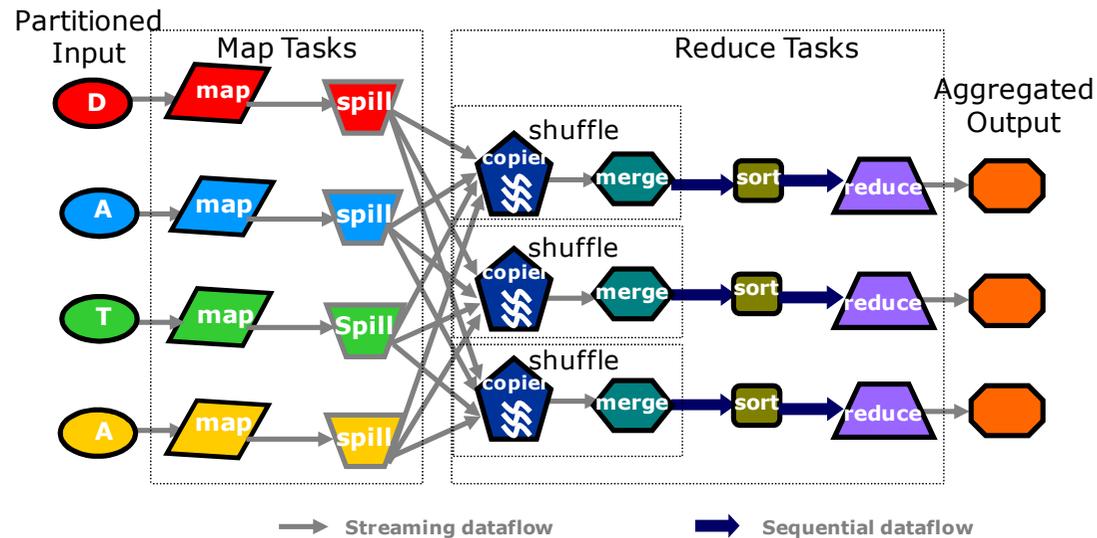


The power to do more

- Tracker
  - Lightweight agent running on each node in the Hadoop cluster
    - Sysstat, Hadoop logs and metrics, Java instrumentation
- Aggregation engine
  - Merges the results of all the trackers in a distributed fashion
- Analysis engine
  - Generates reports based on data flow model



# Case Study



## Terasort with zlib

- Large gap between end of map and end of shuffle
- No CPU, I/O, or network bandwidth bottlenecks
- Adding copiers does not change “shuffle fetchers busy percent” = 100

## Terasort with LZ0

- Copier threads idle 80% waiting for memory merge threads
- Memory merge threads busy mostly due to compression
- Changing compression codec to LZ0 closes the gap
- Improves job running time by 2.3x



# Have at it



The power to do more

- Pull Crowbar
  - <https://github.com/dellcloudedge/crowbar>
- Pull HiTune
  - <https://github.com/HiTune/HiTune>





The power to do more

