CloudRank-D： A Benchmark Suite for Private Cloud Systems

Jing Quan

Institute of Computing Technology, Chinese Academy of Sciences and University of Science and Technology of China

HVC tutorial
in conjunction with The 19th IEEE International Symposium on High Performance Computer Architecture (HPCA 2013)

Contents

- Background & Motivation
- Introduction of CloudRank-D
- Use cases

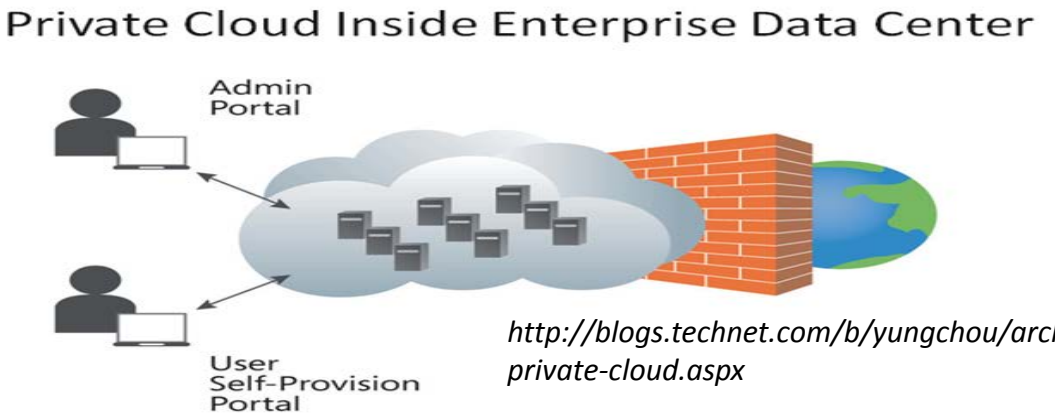
Contents

- Background & Motivation
- Introduction of CloudRank-D
- Use cases

What is Private Cloud ?

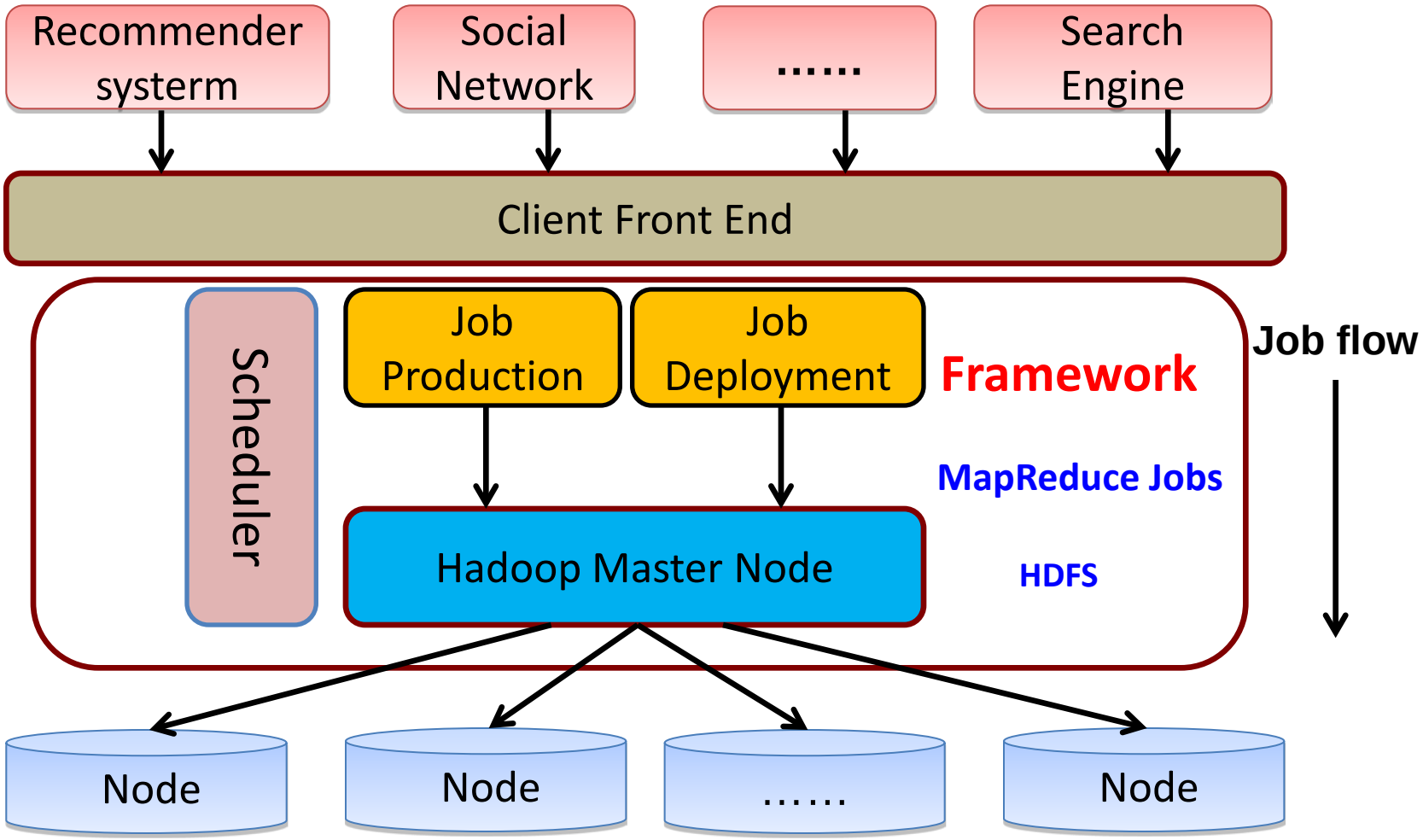
- *Private Cloud*
  - The cloud infrastructure is provisioned for exclusive use by a single organization comprising multiple consumers (e.g., business units). It may be owned, managed, and operated by the organization, a third party, or some combination of them, and it may exist on or off premises.

"The NIST Definition of Cloud Computing" National Institute of Standards and Technology. Retrieved 24 July 2011

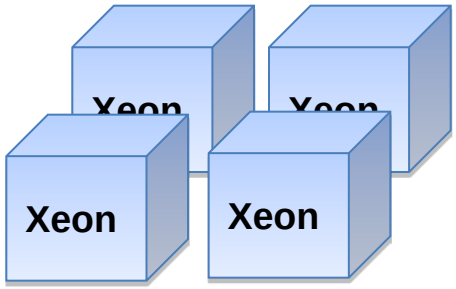


http://blogs.technet.com/b/yungchou/archive/2011/03/21/what-is-private-cloud.aspx

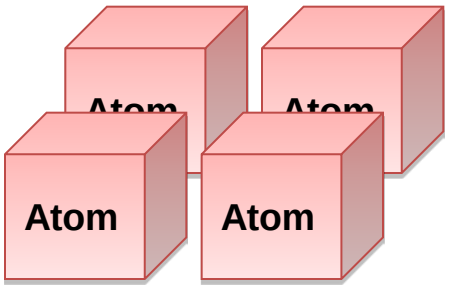
Typical Data Processing Application



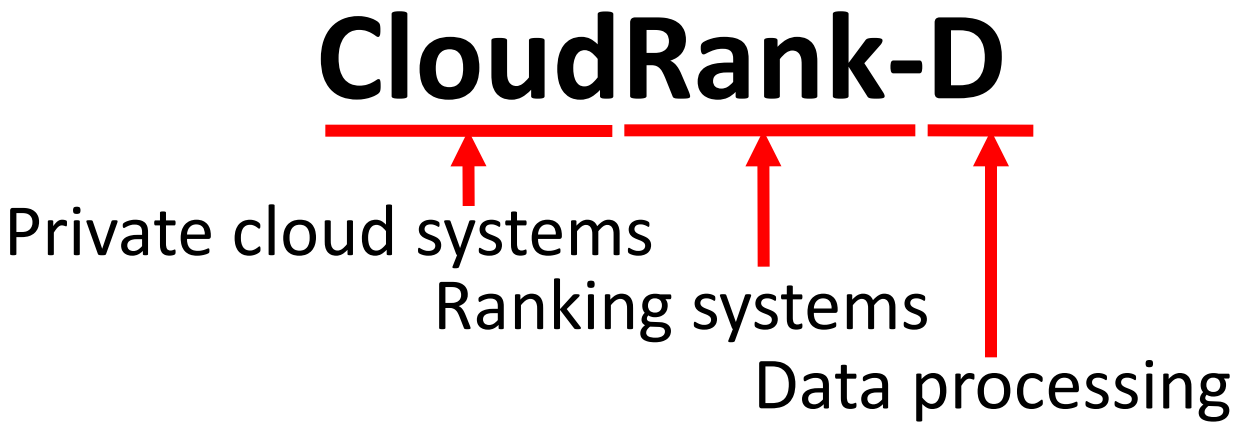
User Concerns



How to quantitatively measure systems?
Which one is better (ranking systems)?
How to guide optimization?



What is CloudRank-D?



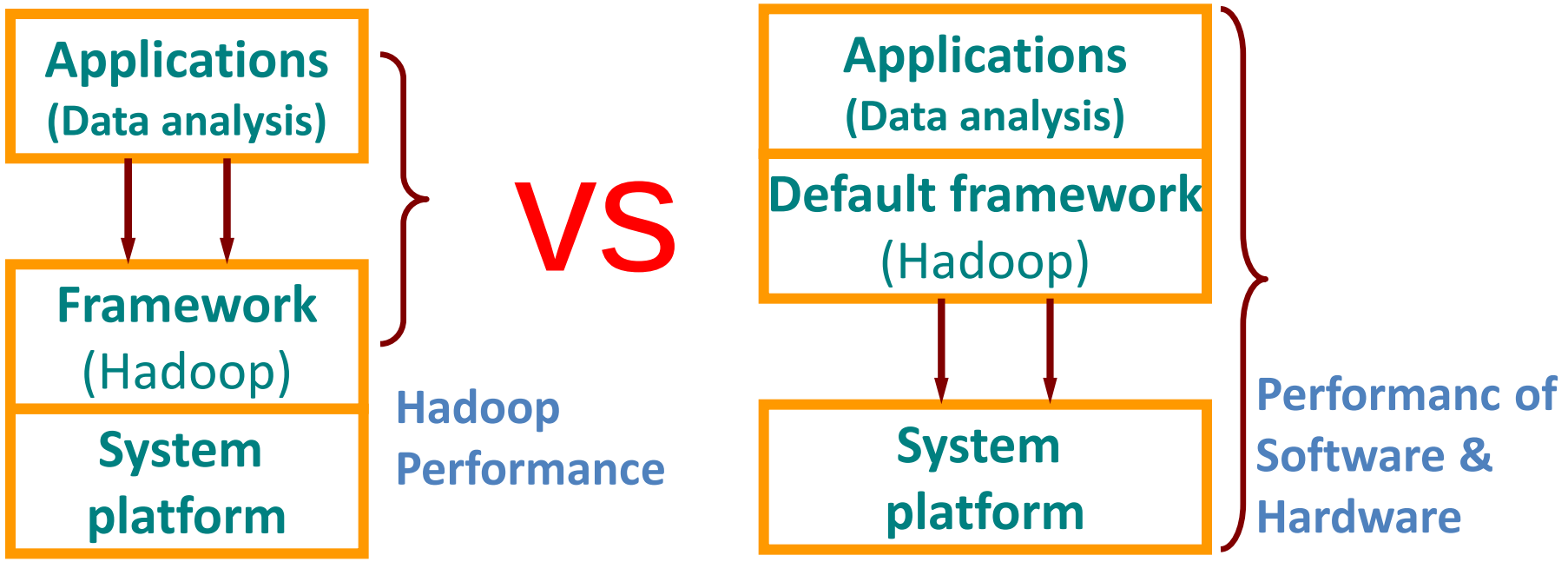
General Description

CloudRank-D is a benchmark suite, used to evaluate private cloud systems that is shared for running data processing applications.

Why CloudRank-D?

Benchmark	Target of Evaluation
MineBench	Data mining algorithms
GridMix	Hadoop framework
HiBench	Hadoop framework
WL suite	Hadoop framework
CloudRank-D	The whole system

Our Focus: Evaluating the Whole System



GridMix etc.

CloudRank-D

Comparison of Different Benchmarks Suites

		Mine-Bench	Grid-Mix	HiBench	WL suite	CloudSuite	CloudRank-D
Representa-tive applications	Basic operations	n	y	y	y	n	y
	Classification	y	n	y	n	y	y
	Clustering	y	n	y	n	n	y
	Recommend-ation	n	n	n	n	n	y
	Sequence learning	y	n	n	n	n	y
	Association rule mining	y	n	n	n	n	y
	Data warehouse operations	n	n	n	y	n	y

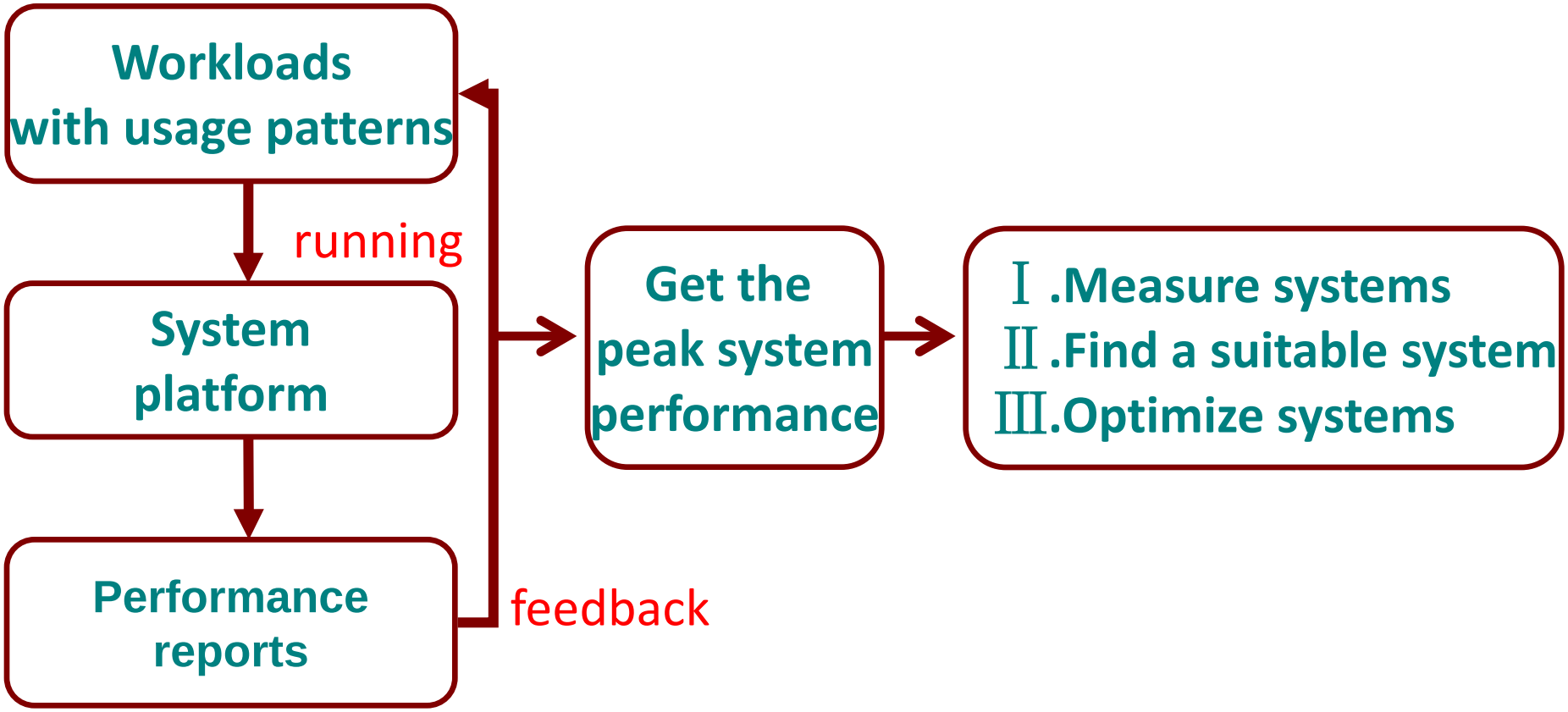
Comparison of Different Benchmarks Suites(Cont')

		MineBench	GridMix	HiBench	WL suite	CloudSuite	CloudRank-D
Workloads description	Submission pattern	n	n	n	y	n	y
	Scheduling strategies	n	n	n	n	n	y
	System software configuration	n	n	n	n	n	y
	Data models	n	n	n	n	n	y
	Data semantics	n	n	n	n	n	y
	Scalable data size	y	y	n	y	n	y
	Category of datacentric computation	n	n	n	y	n	y

Contents

- Background & Motivation
- Introduction of CloudRank-D
  - Methodology
- Use cases

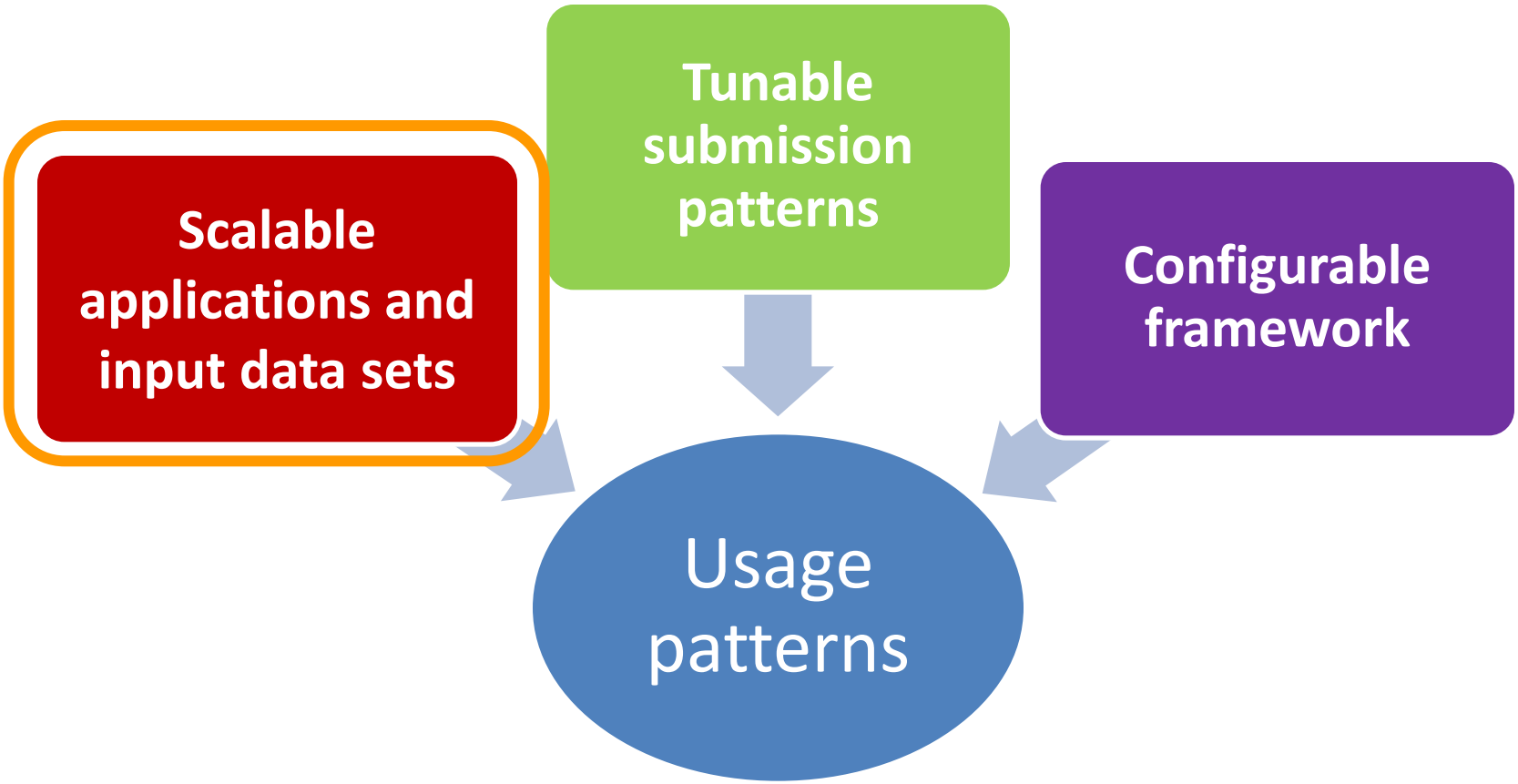
CloudRank-D Methodology



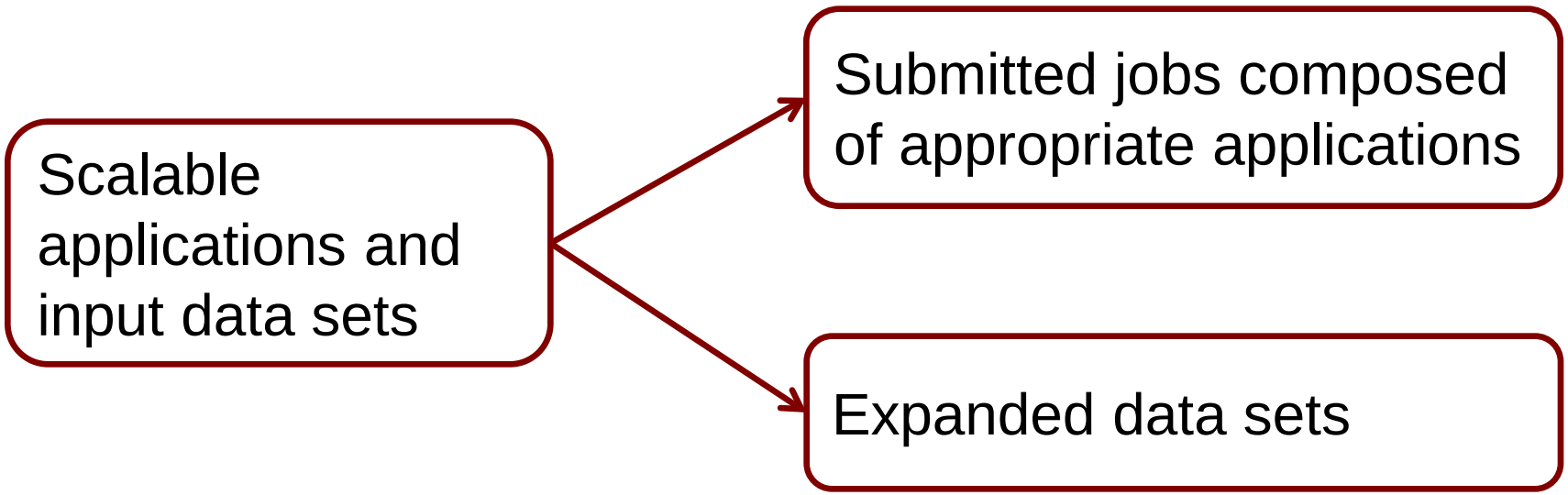
Configurable Workloads with Tunable Usage Patterns

Scalable applications and input datasets	Tunable submission patterns	Configurable runtime system
• Representive applications domains • User specific • Scalable data size	• Modeling production system logs	• Experiences from industry and academic

CloudRank-D Methodology: Workloads with Usage Patterns



Scalable Applications and Input Data Sets



Applications and Input Data Sets

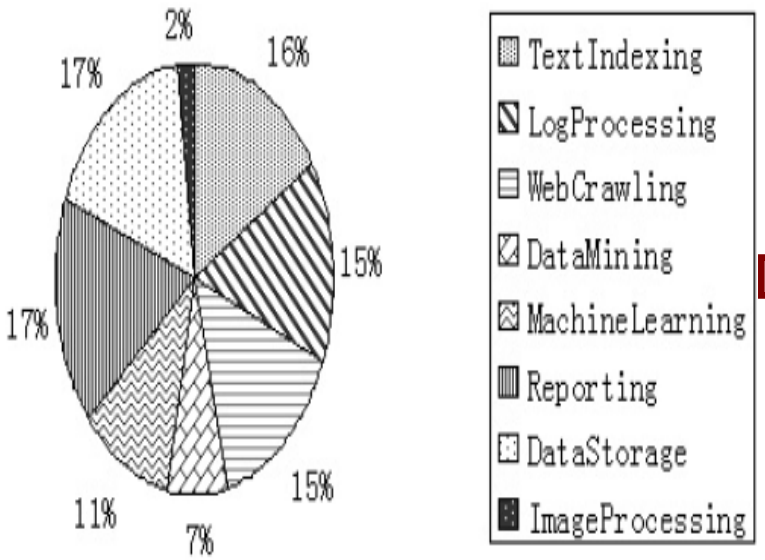
NO.	Category	Application	Data size	Data semantics
1	basic operation	sort	scalable (scale to 10PB)	automatically generated
2		word count		
3		grep		
4	classification	naive bayes		
5		support vector machine		Scientist Search
6	cluster	k-means	scalable	sougou corpus
7	recommendation	Item based collaborative filtering	scalable	ratings on movies

Applications and Input Data Sets (Cont')

NO.	Category	Application	Data size	Data semantics
8	association rule mining	frequent pattern growth	fixed	retail market basket data click-stream data , traffic accident data, collection of web html documents
9	sequence learning	hidden morkov model	scalable	Scientist Search
10	warehouse operation	grep select		automatically generated table
11		ranking select		
12		aggregation		
13		uservisits-ranking join		

You can add any applications you want !

Applications Combinations Demonstration



wiki.apache.org/hadoop/PoweredBy

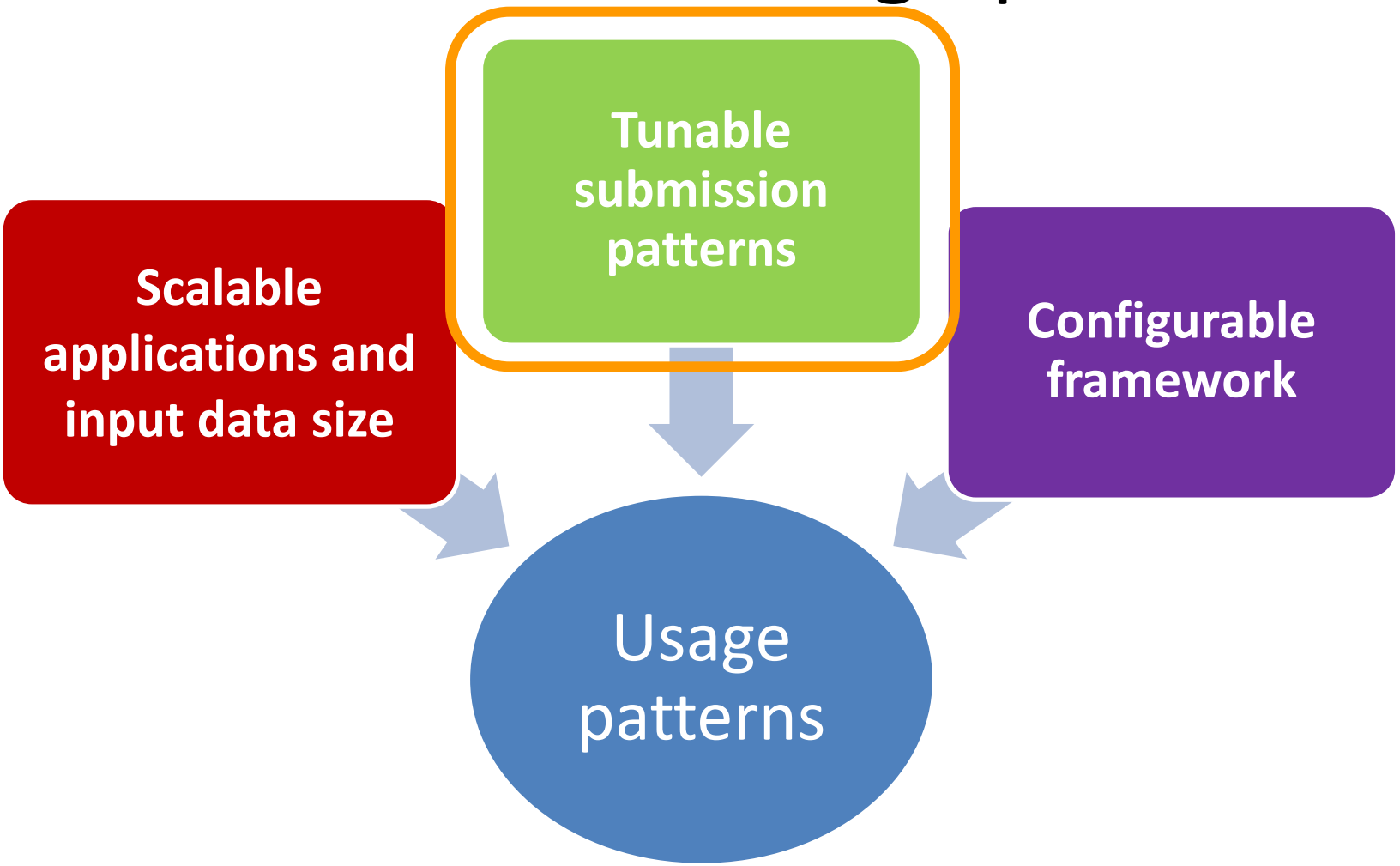
WebCrawling DataMining MachineLearning ImageProcessing	Naive Bayes & SVM HMM & IBCF & FPG	35%
TextIndexing LogProcessing	Basic Operations	31%
Reporting DataStorage	Hive	34%

Data Set Sizes Demonstration

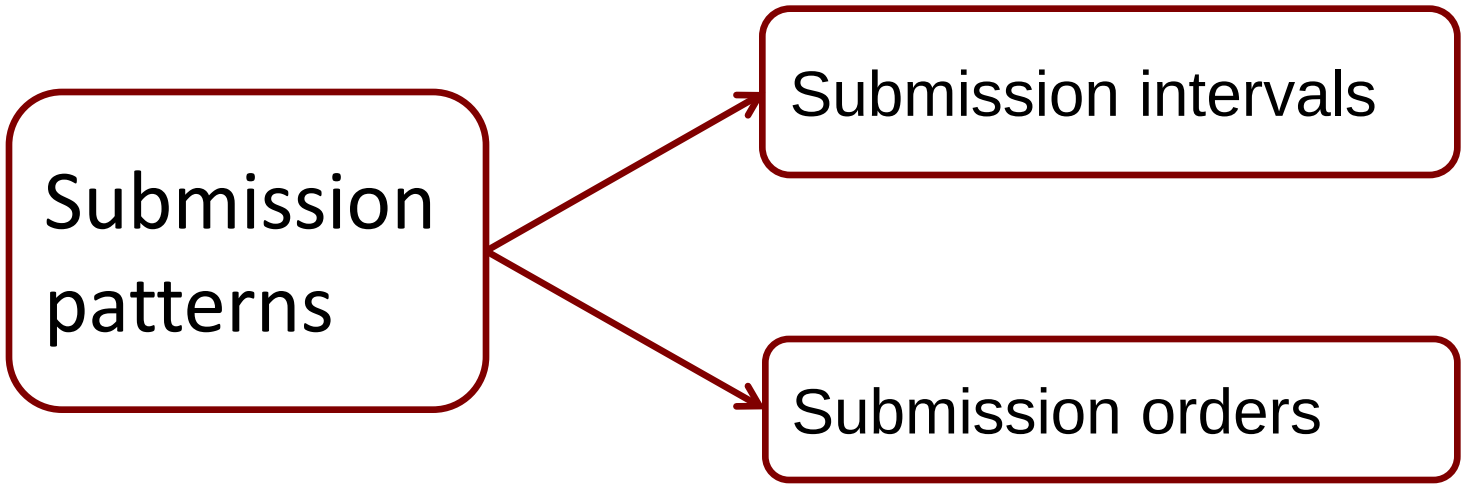
Map Number	Percentage	Size
<10	40.57%	128MB~1.25GB
10~500	39.33%	1.25GB~62.5GB
500~2000	12.03%	63.5GB~250GB
>2000	8.07%	250GB~

Workload Characterization on a Production Hadoop Cluster: A Case Study on Taobao

Workloads with usage patterns



Submission Patterns



Submission Intervals

Form the Facebook report, distribution of inter-arrival times was roughly exponential with a mean of 14 seconds.

Ddelay scheduling: A simple technique for achieving locality and fairness in cluster scheduling. In Proceeding In Proceedings of the 5th European conference on Computer systems.

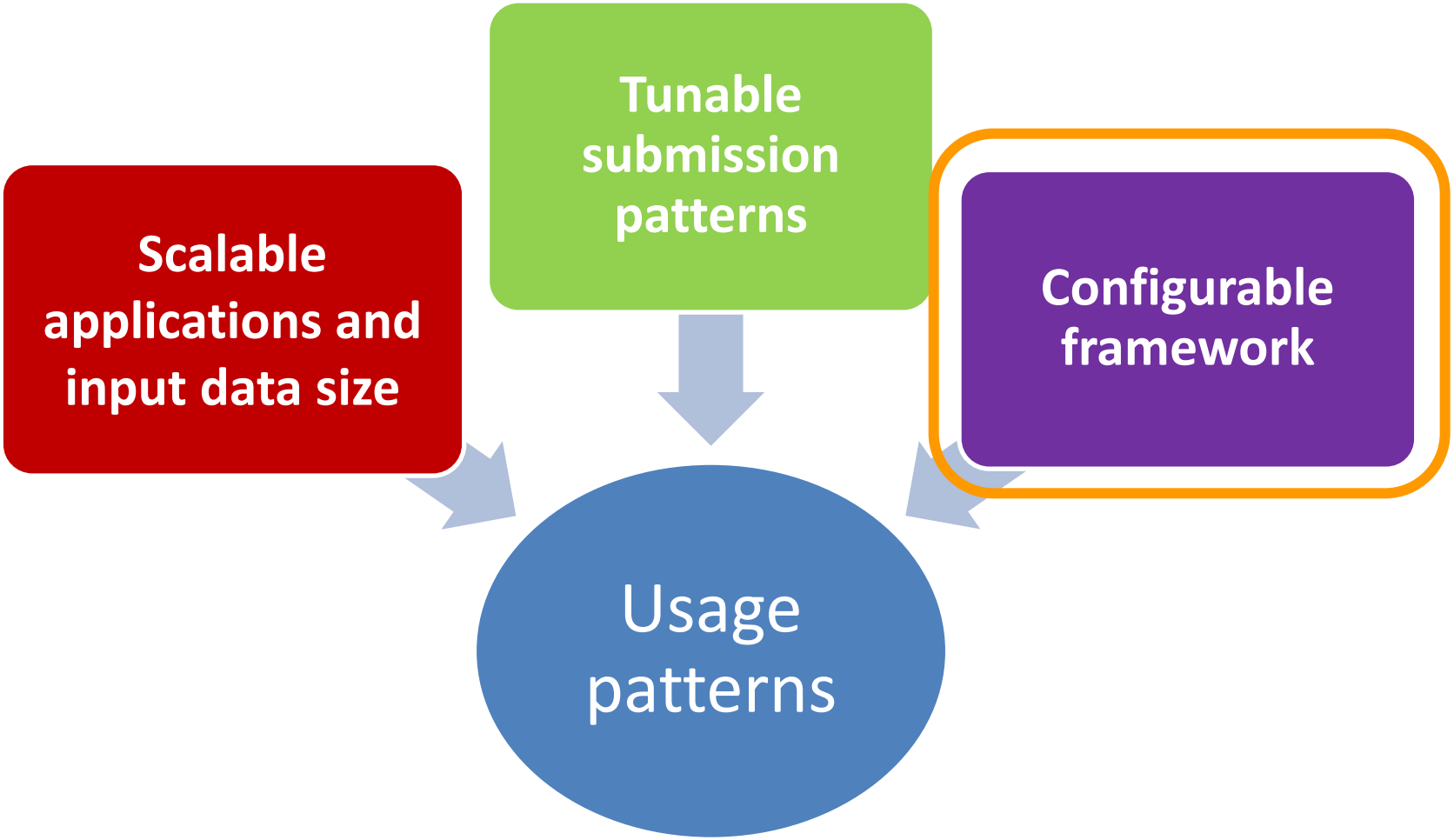
Probability density function

$$f(x;\lambda) = \begin{cases} \lambda e^{-\lambda x} & , \ x \geq 0, \\ 0 & , \ x < 0. \end{cases}$$

Submission Orders

- For the workloads with different resource sizes and different catelogs
  - Submitting jobs randomly
  - Submitting jobs with batch model

Workloads with usage patterns



Hadoop Configurations

Dimensions	Explanation
Map/Reduce Number	affect system utilization
Scheduling Policy	Hadoop chooses which job to run according to this policy
Main Parameters	mapred.tasktracker.map.tasks.maximum mapred.tasktracker.reduce.tasks.maxmum mapred.child.java.opts dfs.block.size

Hadoop Settings

Parameter	Value
Mapred.tasktracker.tasks.reduce.maximum	usually, this value is equal to the core number of current node
dis.block.size	default value is 64M, you can change it to ensure there won't be too much map number for most workloads
Map (adjust through the block size)	10~100 per node, and it's would be better if the execution time was more than 1min

Scheduling Policy

- Common schedule algorithms
  - First input first out
  - Fair-share scheduler
  - Capacity scheduler
- Fair-share scheduling can do a good job

Workload Characterization on a Production Hadoop Cluster: A Case Study on Taobao

CloudRank-D methodology: Our metrics

- Focus
  - From user perspective
  - Easy to compare and understand
- Metrics
  - Data processed per second or joule
- How to get it?

  DPS=Total ***data input size***/Total run time

  DPJ=Total ***data input size***/Total energy consumption

Contents

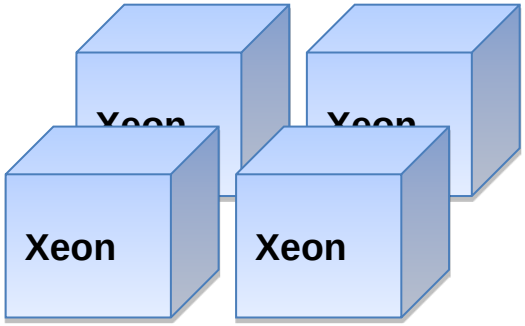
- Background & Motivation
- Introduction of CloudRank-D
- Use cases

How to use?

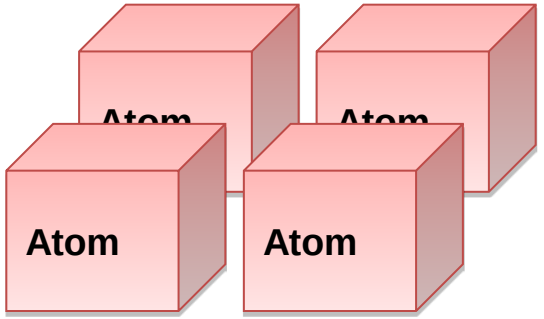
CloudRank-D

Use Case 1: Comparing Two Hardware Platforms

Cluster 1



Cluster 2



CPU Type		Intel CPU Core	
Intel ®Atom D510		2cores@1.66GHz	
TLB	L1 I/D Cache	L2 Cache	Memory
256 entries	24K	512K	2G

CPU Type		Intel CPU Core	
Intel ®Xeon		4cores@1.6GHz	
TLB	L1 I/D Cache	L2 Cache	Memory
256 entries	32K	4096K	4G

Two clusters comprise 128 nodes respectively.

Procedures

Step 1
Prepared hardware platform

Step 2
Customize workloads

Step 3
Run workloads

Step 4
Get results and optimize systems

Base Information

- Evaluating two private cloud systems
- Using all workloads we provide
- Deploying uniform software platform
- Adopting same configuration

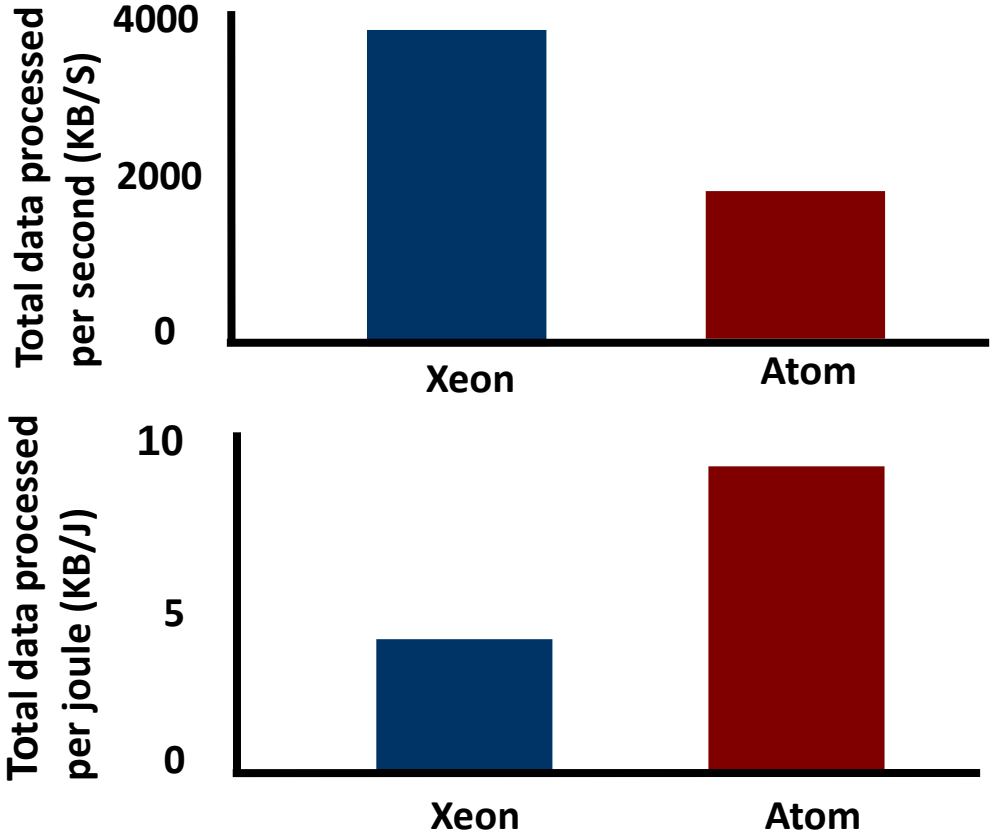
Software Configuration

software stack	Hadoop	version 0.20.2
	Hive	version 0.6.0
	Mahout	version 0.6
map/reduce slot	4 map slots and 2 reduce slots	
Hadoop system configuration	default	
Hadoop scheduling algorithm	fair schedule	

Run your workloads

Job Submission Patterns	You can submit the workloads according to the exponential distribution with a specified mean submission interval --- 14 seconds
	Submission order : Random

An example of result

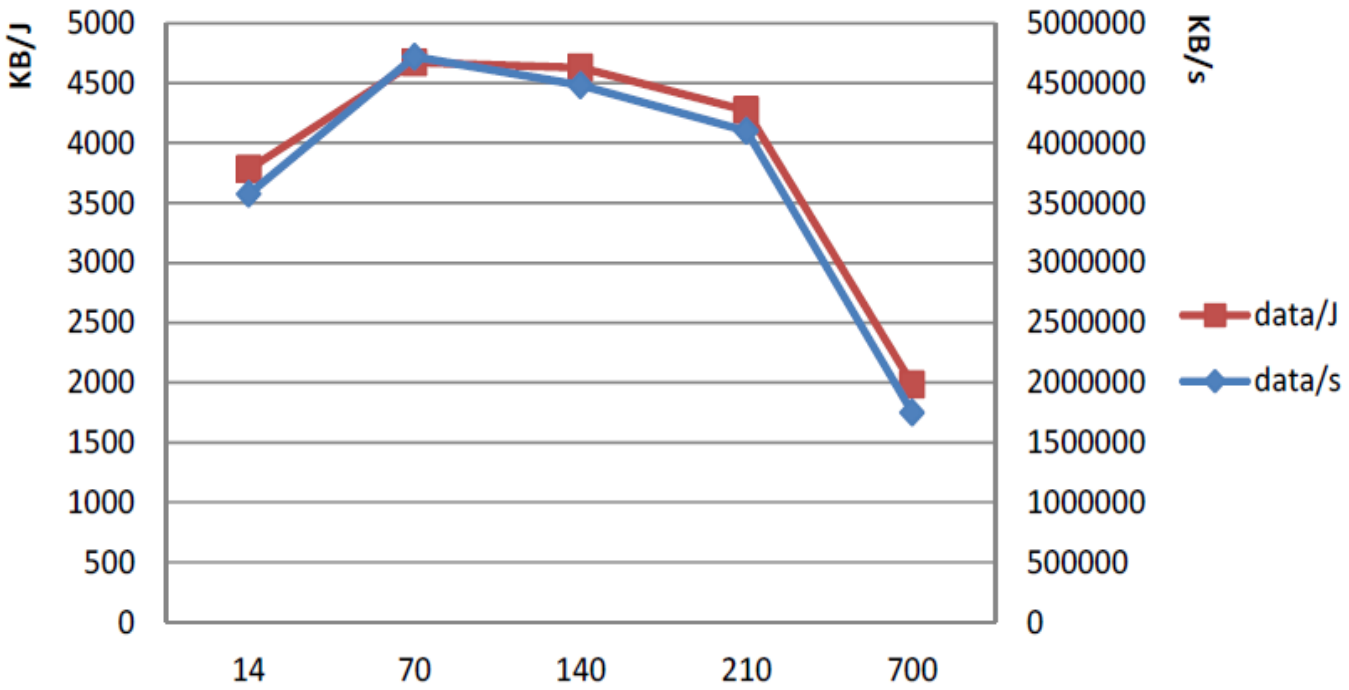


- Xeon
  - less time, more energy
- Atom
  - more time, less energy

The comparion between Xeon Atom on two metrics

Optimized (Cont')

- Tuning the interval



We can see that the best performance occurred when the interval value is 70.

Use Case 2: Scheduling Evaluation

I have designed a new Hadoop scheduling algorithm, but I don't have the workloads for test.

How to evaluate the scheduling ? And let people trust the evaluations results.

Using CloudRank-D

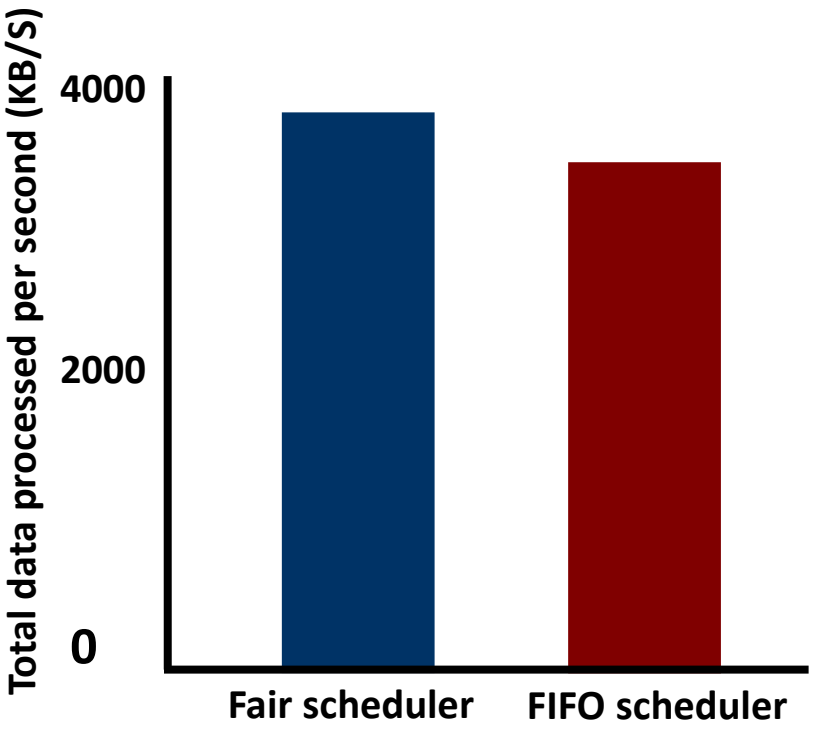
Step 1
Building foundation platform with different scheduling policy

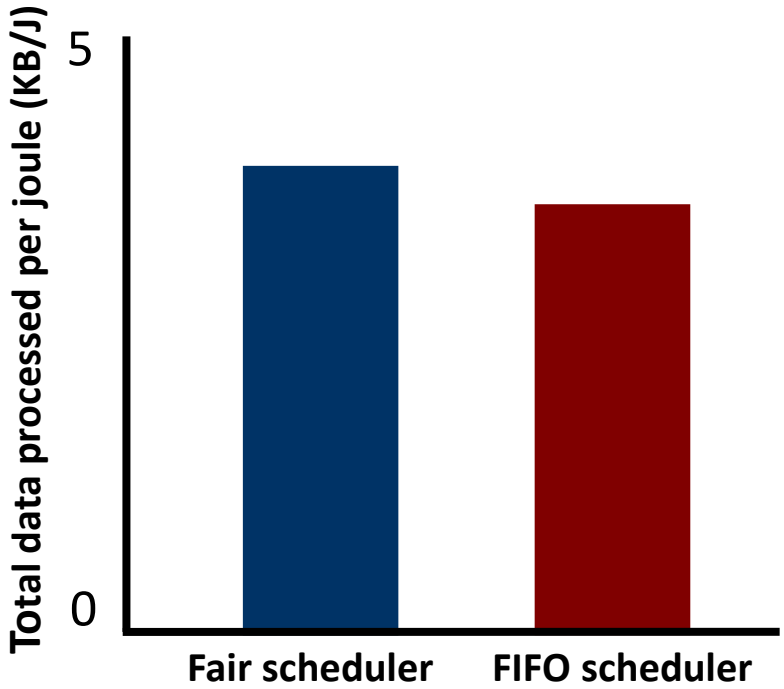
Step 2
Customizing workloads with productive scenarios

Step 3
Running workloads

Step 4
Getting the metrics under different scheduling policy

Our Result





We can see that fair scheduler works better than FIFO scheduler.

- Contact us
  - Websit: http://prof.ict.ac.cn/CloudRank/
  - Email: quanjing@ict.ac.cn

Thanks