See Spot Run: Using Spot Instances for MapReduce Workflows

Navraj Chohan[1]
Claris Castillo[2]
Mike Spreitzer[2]
Malgorzata Steinder[2]
Asser Tantawi[2]
Chandra Krintz[1]

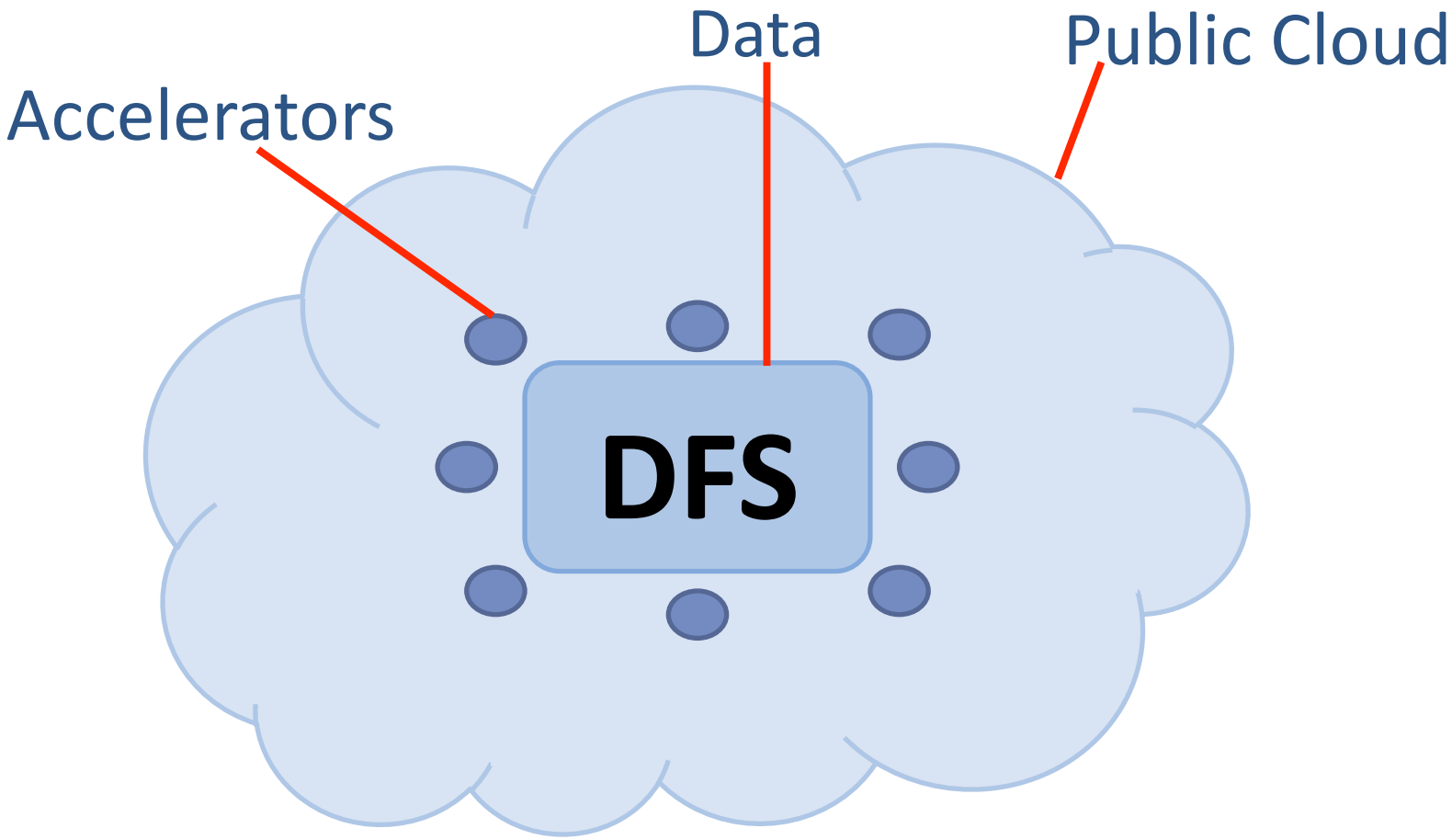
UC Santa Barbara [1]
IBM Research[2]

Overview

- Data Analytic Cloud
- Instance Options
- MapReduce
- Spot Instances
- Evaluation

UCSB

Data Analytic Cloud



UCSB

Cloud Options

- Different VM Sizes
- Pricing Options
  - On-demand
  - Leased
  - Spot Instances

VM Resources

Instance Type	EC2 Compute Units	Memory (GB)	Storage (GB)	On-Demand Price (per hr)
m1.small	1	1.7	160	$0.095
c1.medium	5	1.7	350	$0.19
m1.large	4	7.5	850	$0.380
m2.xlarge	6.5	17.1	420	$0.570
m1.xlarge	8	15	1690	$0.760
c1.xlarge	20	7	1690	$0.760
m2.2xlarge	13	34.2	850	$1.340
m2.4xlarge	26	68.4	1690	$2.68

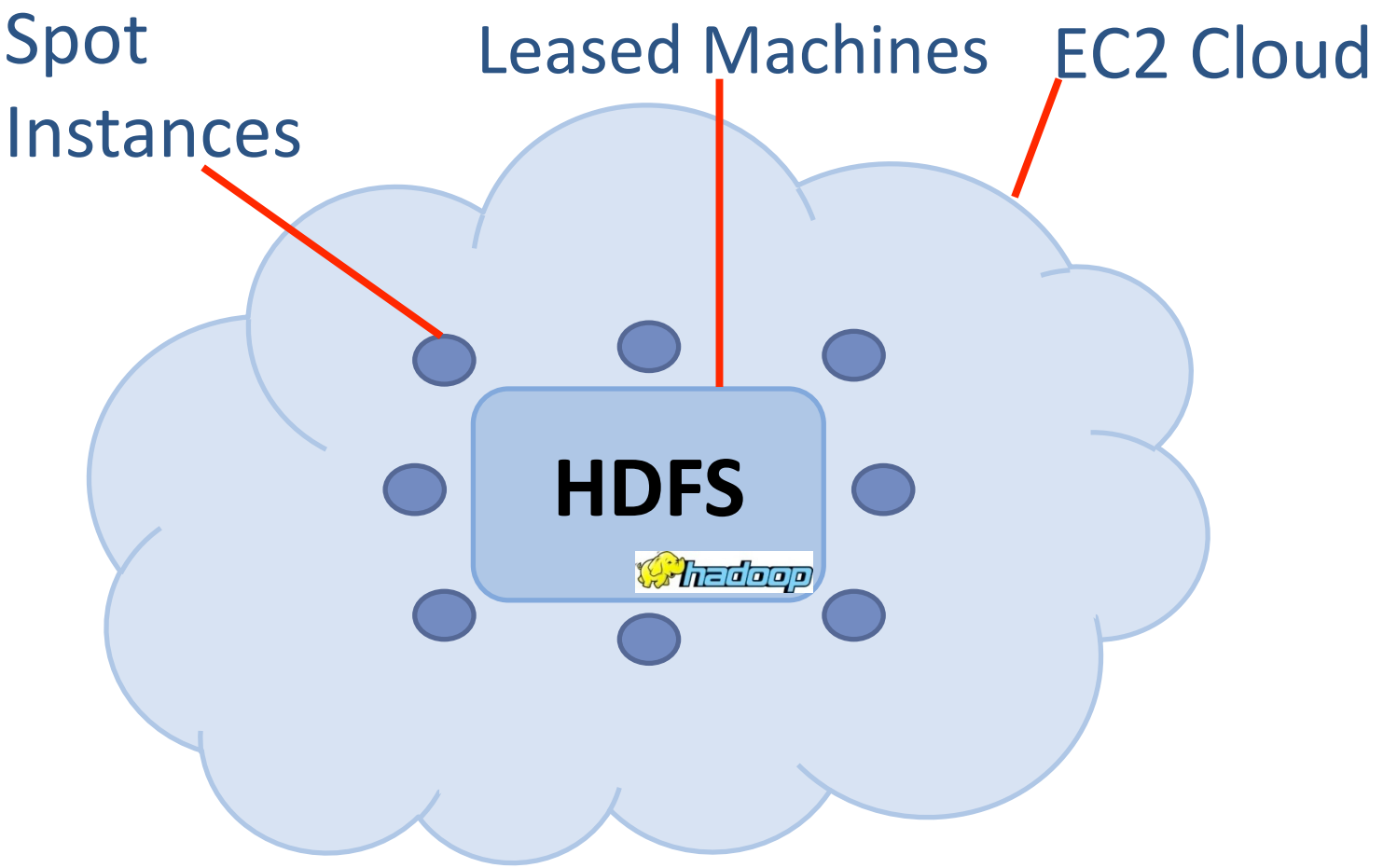
Pricing from http://aws.amazon.com/ec2/

Pricing Options

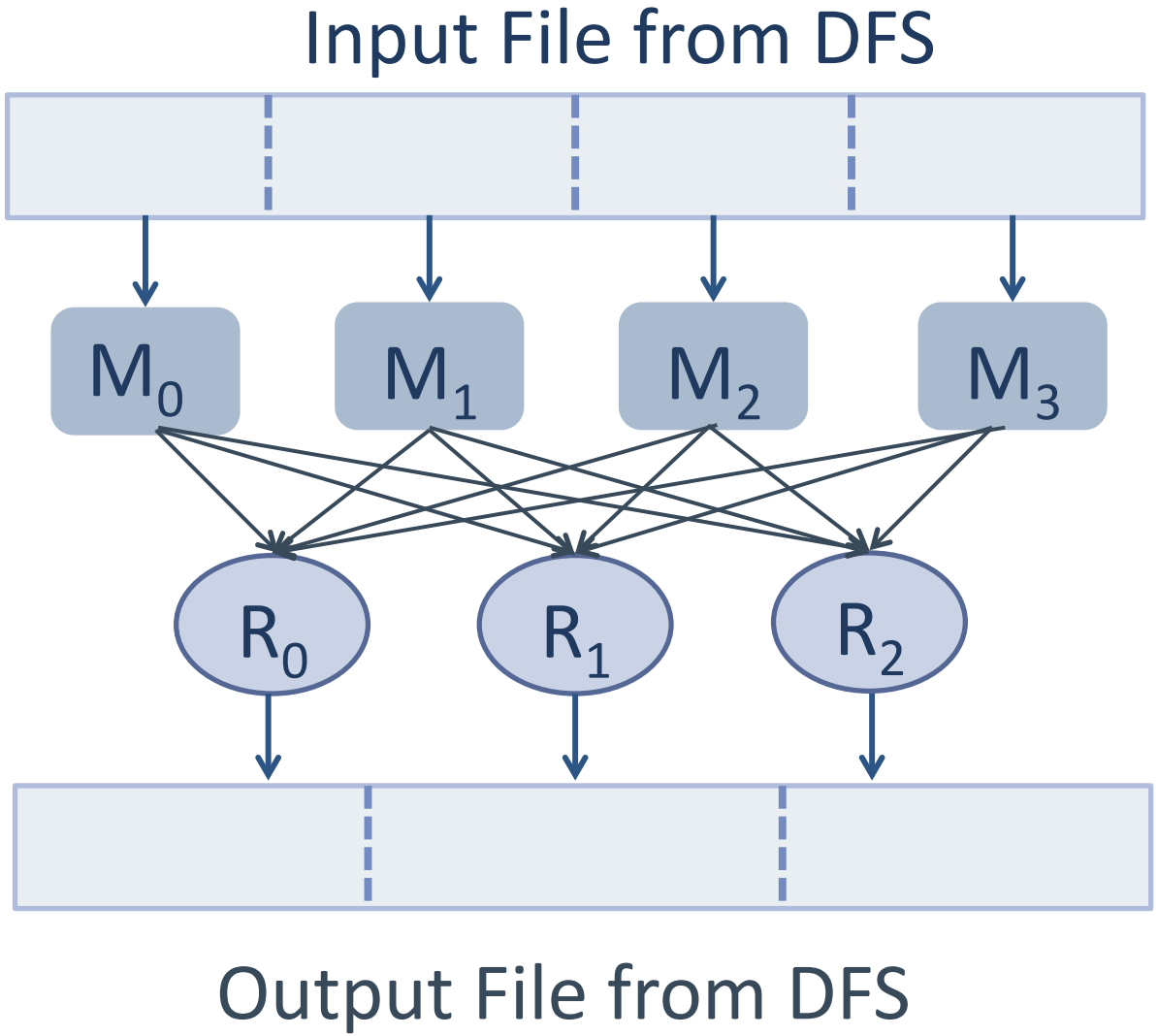
Instance Type	On-Demand Price (per hr)	Reserved-1 Year Price (per hr)	Reserved-3Year Price (per hr)	Spot Instance Average Price (per hr)
m1.small	$0.095	$0.056	$0.043	$0.0399
c1.medium	$0.19	$0.112	$0.087	$0.0798
m1.large	$0.380	$0.224	$0.173	$0.167
m2.xlarge	$0.570	$0.321	$0.246	$0.240
m1.xlarge	$0.760	$0.448	$0.347	$0.320
c1.xlarge	$0.760	$0.448	$0.347	$0.323
m2.2xlarge	$1.340	$0.784	$0.606	$0.559
m2.4xlarge	$2.68	$1.56	$1.21	$1.12

Pricing from http://aws.amazon.com/ec2/

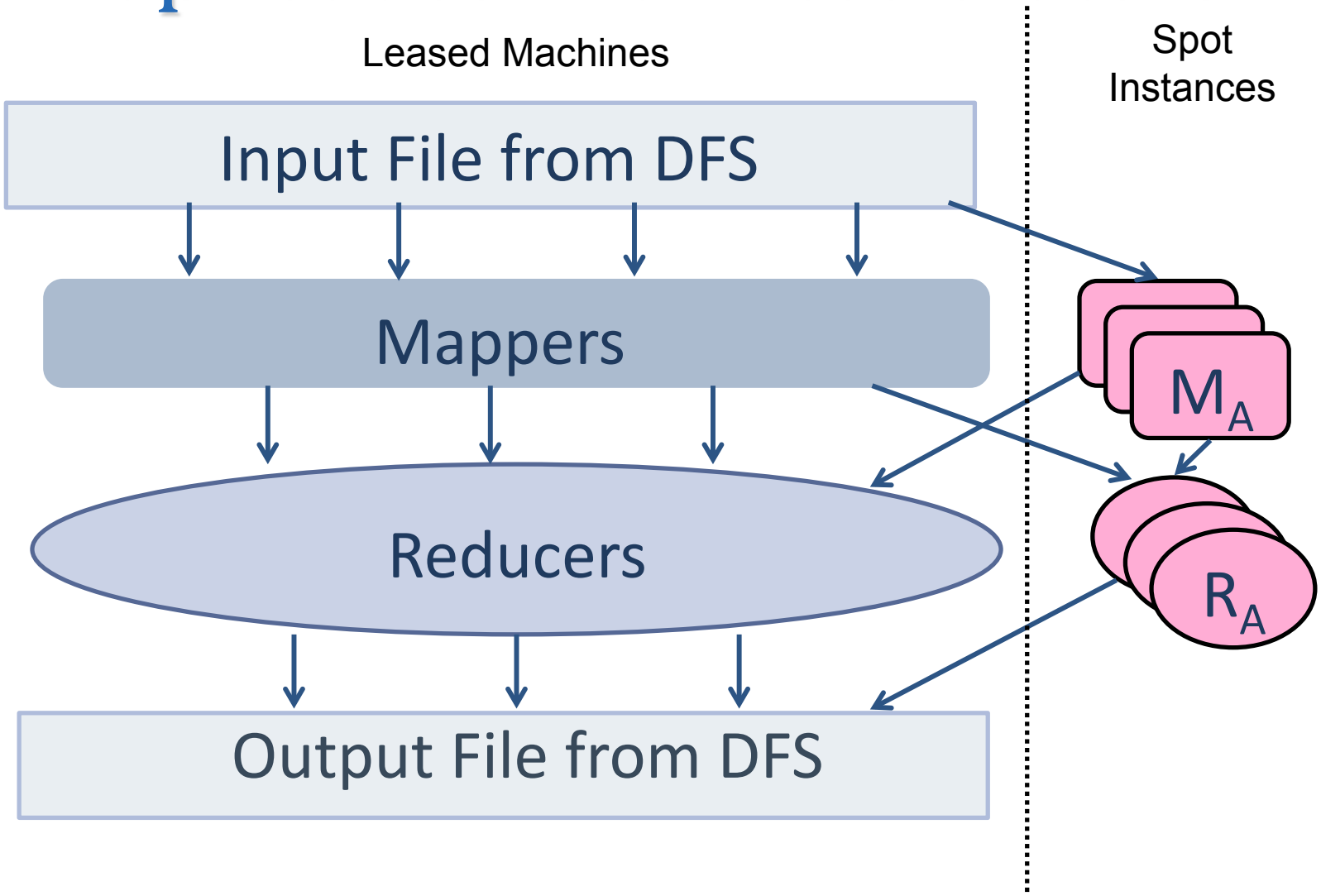
Building Blocks



MapReduce



MapReduce with Accelerators



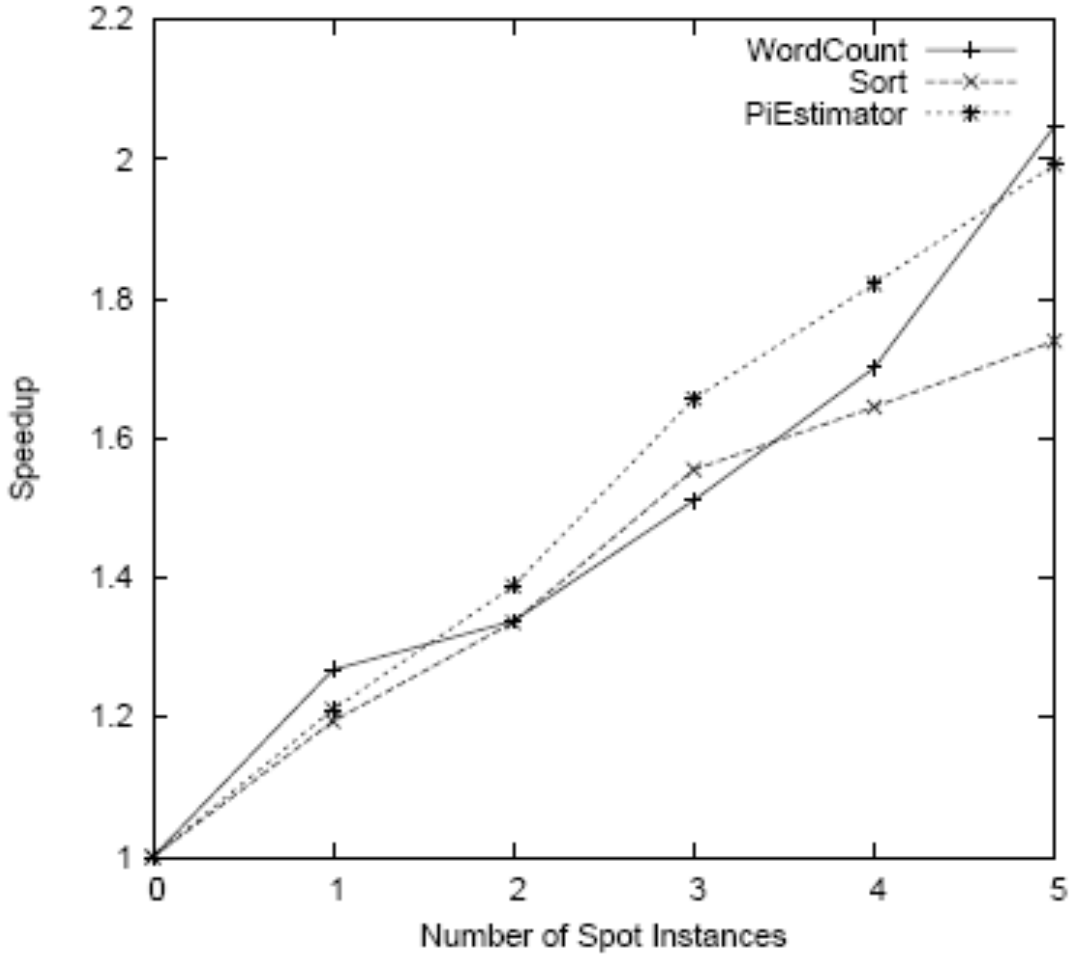
How Spot Instances Work

- Make a max bid on a spot instance
- Spot instance is available if
  - Max bid > market price
- Not available if
  - Max bid ≤ market price
- Always pay market price
- Pay for full hour if terminated by user
- Free partial hour if terminated by Amazon

MapReduce with Spot Instances

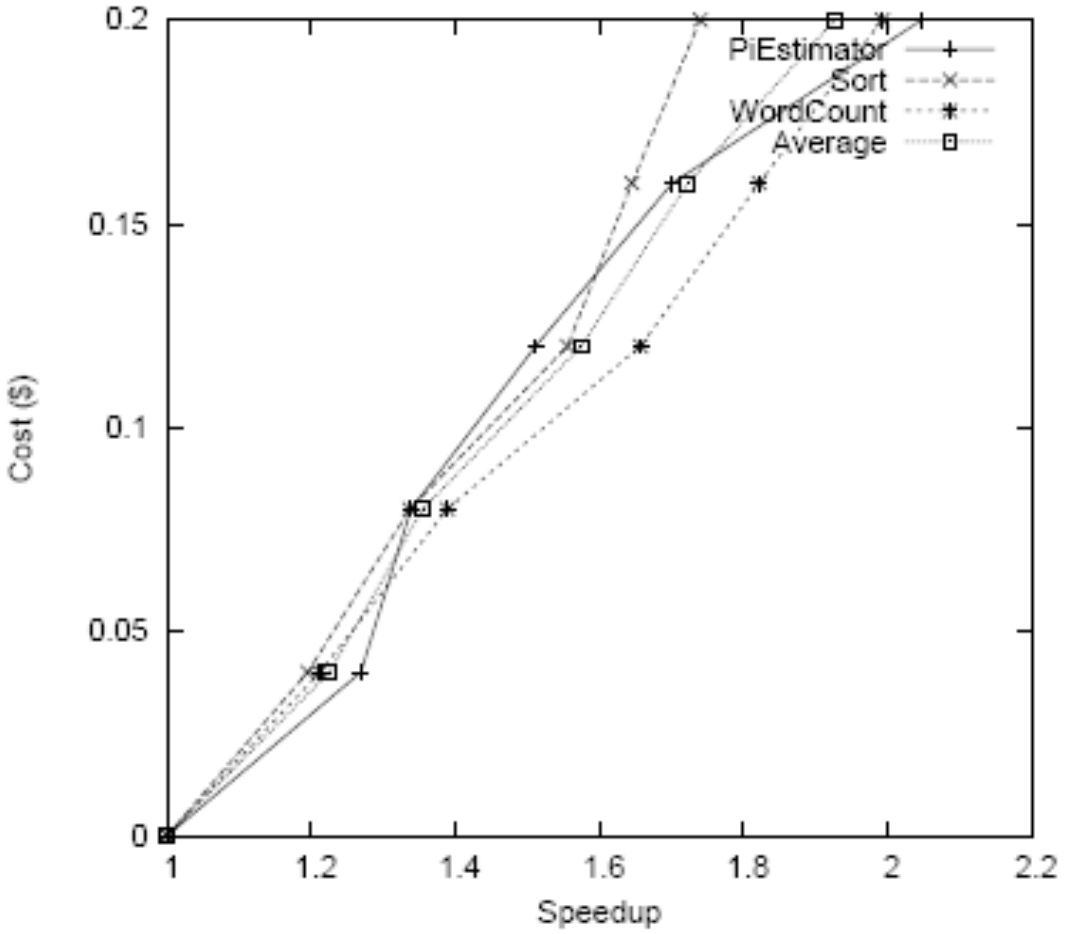
- MR paradigm
  - Embarrassingly parallel jobs
  - Fault tolerant
  - Transient workers
  - Workers pull data
- Spot Instances
  - Provide transient and (relatively) inexpensive resources

Job Speedup



UCSB

Speedup Cost

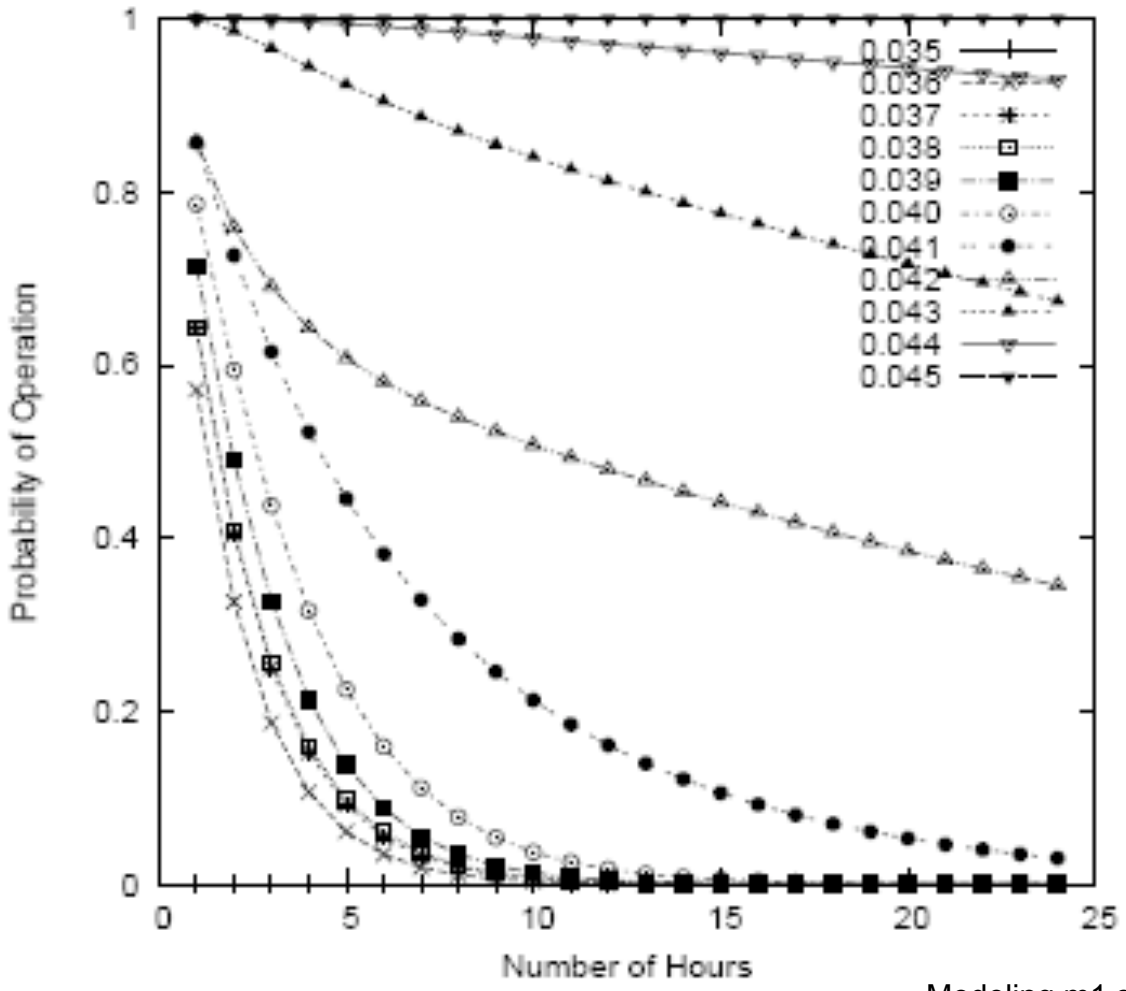


UCSB

Downside of Spot Instances

- Termination has a cost
- VM uptime probability is a function of the user's maximum bid price
- Work will have to be redone
  - Operational nodes must pick up the slack
  - This includes map output which has been already consumed by a reducer

VM Lifetime

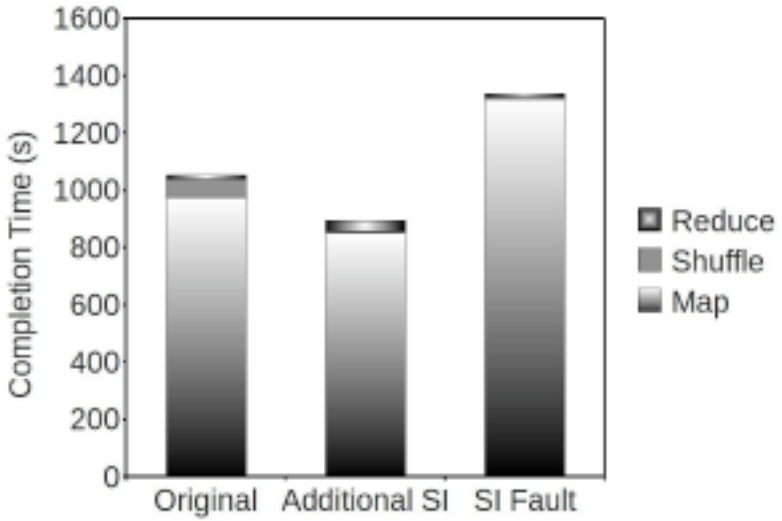


Modeling m1.small instance using data from cloudexchange.net

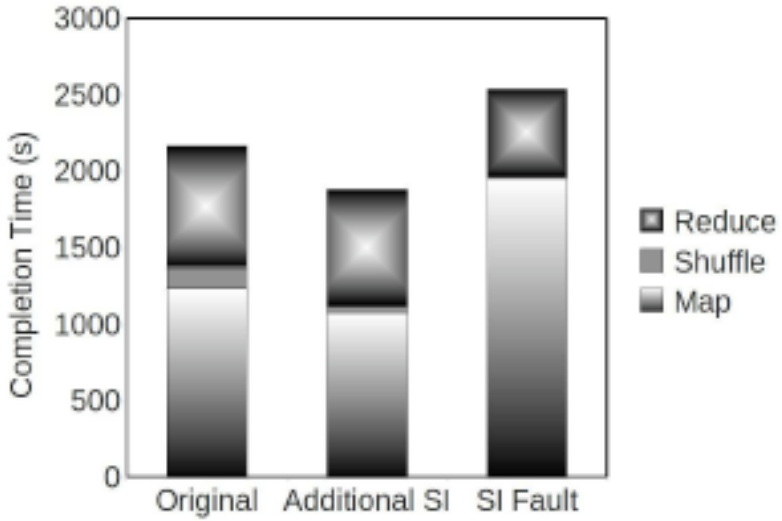
UCSB

MapReduce with Faults

Fault injected at half-way point of original job



WordCount



Sort

UCSB

Handling Faults Efficiently

- Have Hadoop track which map output has been consumed by a reducer to avoid re-execution
- Store intermediate data (map output) in HDFS*
- Lower fault detection time
  - Default: 10 minutes

*Steven Y. Ko et al. from HotOS09'

Summary

- Spot instances provide inexpensive resources for transient workloads
- MapReduce jobs speedup with more resources
- Spot instance termination hurts a job's time to completion

Questions?

UCSB