

# CloudThemis: I/O-Efficient MapReduce in the Cloud



**THEMIS**



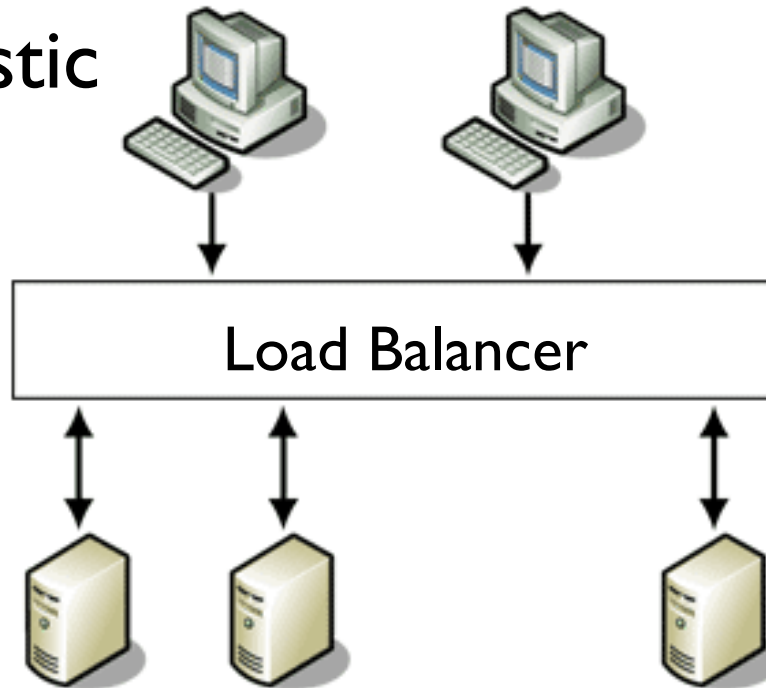
**TRITONSORT**

CNS Fall Review 2014 Progress Report  
Mike Conley, Amin Vahdat, George Porter



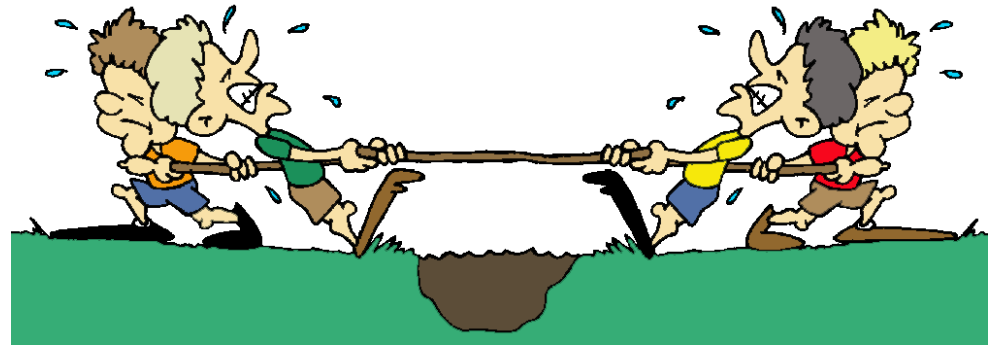
# The Cloud is Awesome

- Cheap, on-demand access
  - Low CAPEX and OPEX
- Access to new technologies
- Elastic



# The Cloud is Terrible

- Shared resources
  - Unpredictable performance
- I/O variance
  - Storage
  - Network
- Tail latency problem
  - Tightly coupled parallel apps suffer
  - e.g. MapReduce stragglers



# Central Question

- Is the following feasible?
  - Massive computation
  - Public cloud
  - High performance
  - Low cost

# Choosing a Cluster Configuration

- Option 1) Try every combination
  - Expensive, not very enlightening
- Option 2) Predict the best configuration
  - Benchmark to understand key bottlenecks
  - Pick best configuration for workload

# Choosing a Cluster Configuration

- Option 1) Try every combination
  - Expensive, not very enlightening
- Option 2) Predict the best configuration
  - Benchmark to understand key bottlenecks
  - Pick best configuration for workload
  - Amazon EC2
    - Large local storage

# Evaluation Tool and Workload

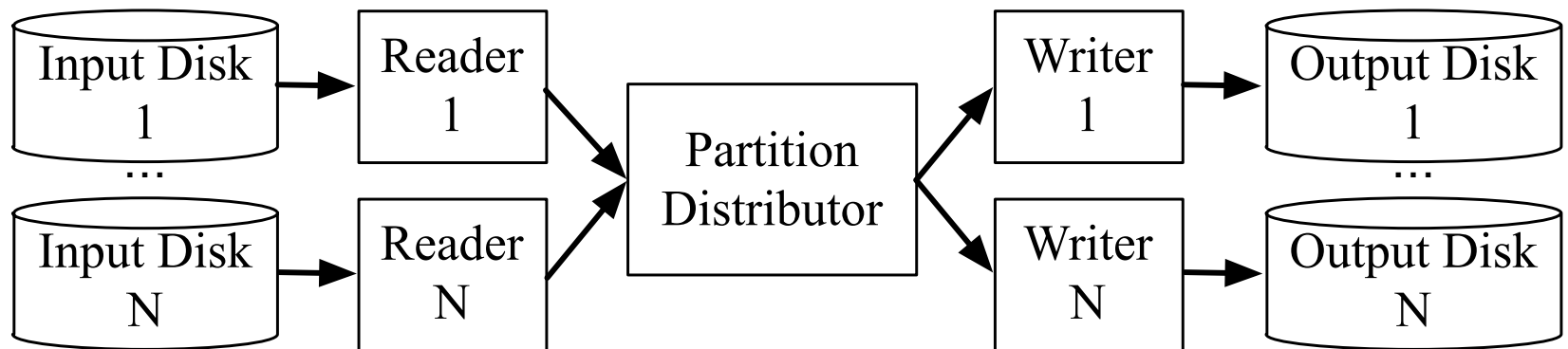
- Themis MapReduce
  - Restrict scope to out-of-memory workloads
  - For in-memory, systems like Spark are great
- Evaluation workload is 100TB sort
  - But first let's benchmark...





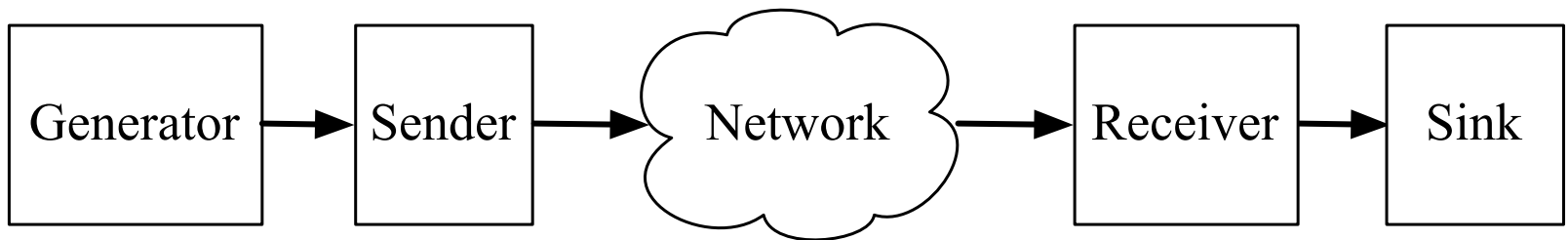
# DiskBench

- Mimic storage I/O in Themis
  - Without the MapReduce logic
  - More realistic benchmark than **dd**
  - As much code reuse as possible

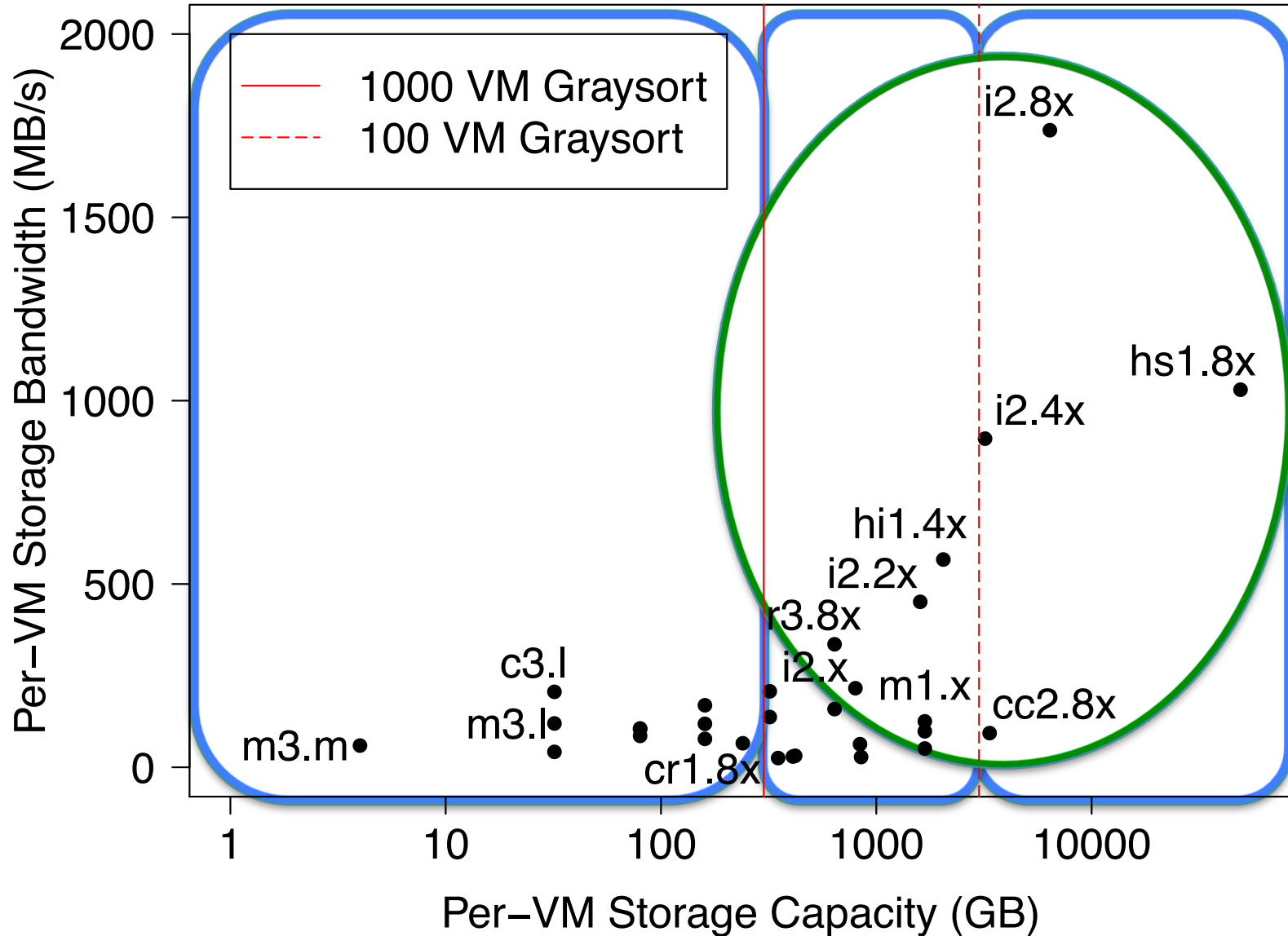


# NetBench

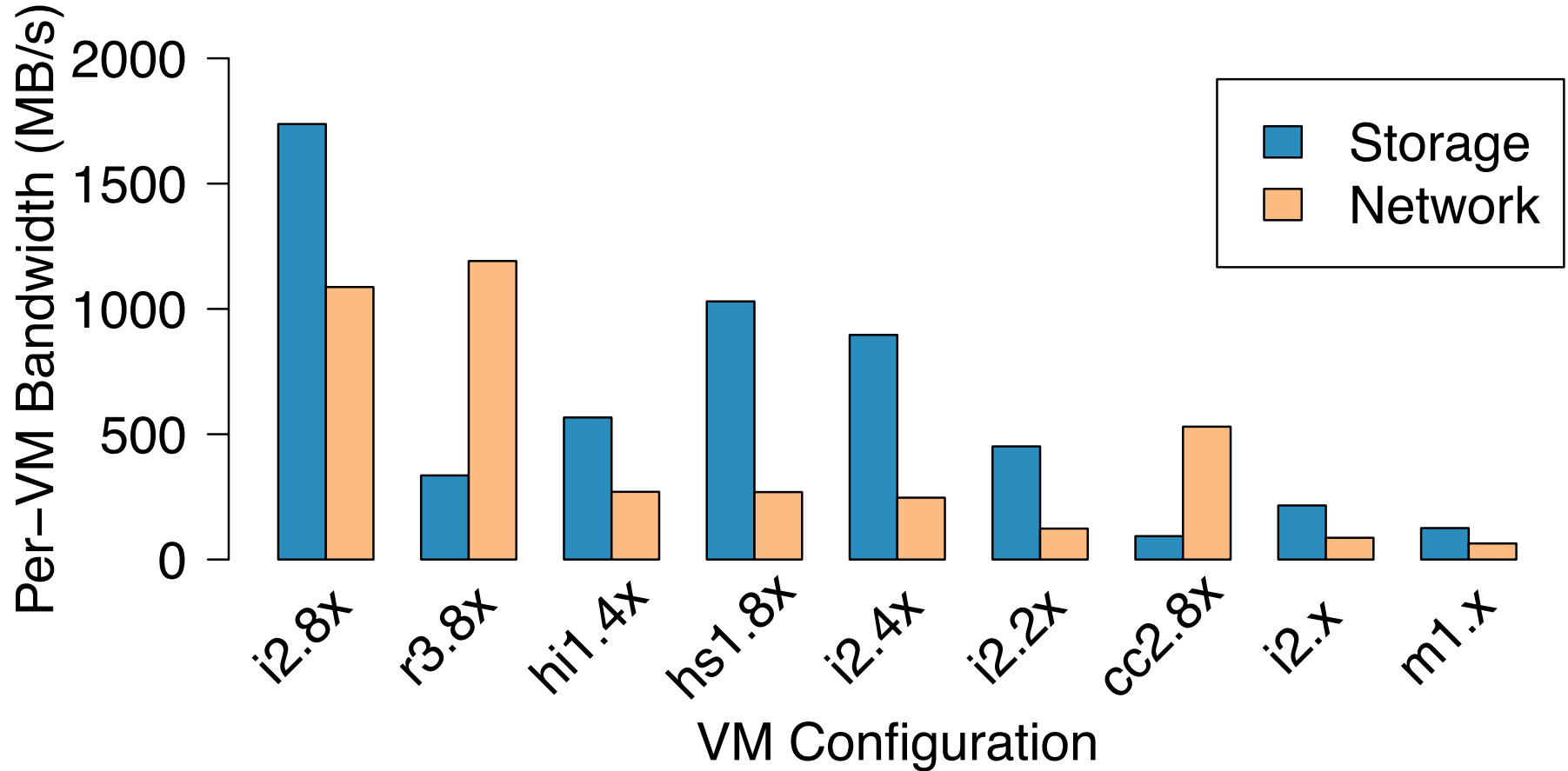
- Mimic network I/O in Themis
  - Simulate the shuffle phase of MapReduce
  - More realistic benchmark than **iperf**



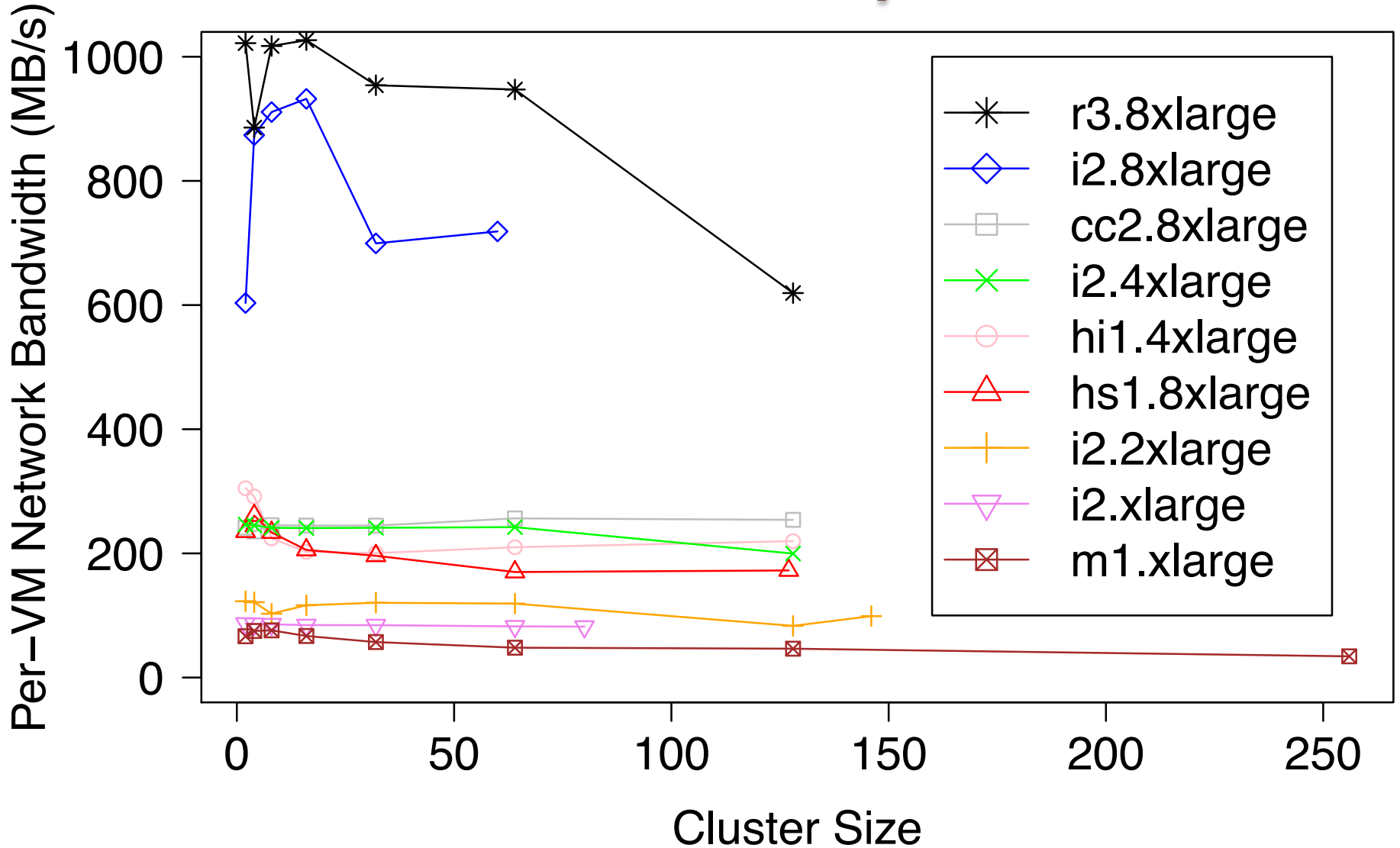
# Per-VM Storage Performance



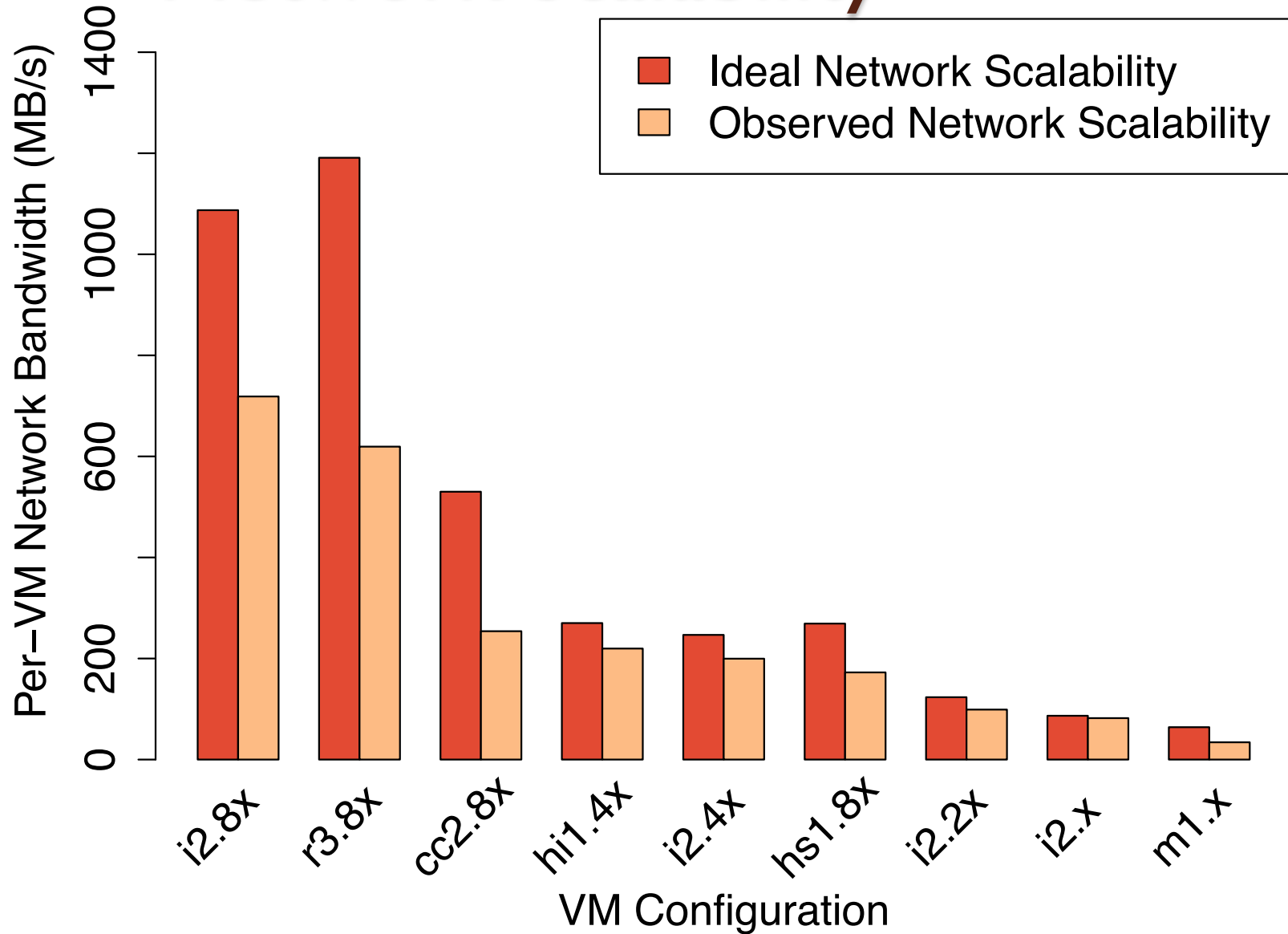
# Per-VM Network Performance



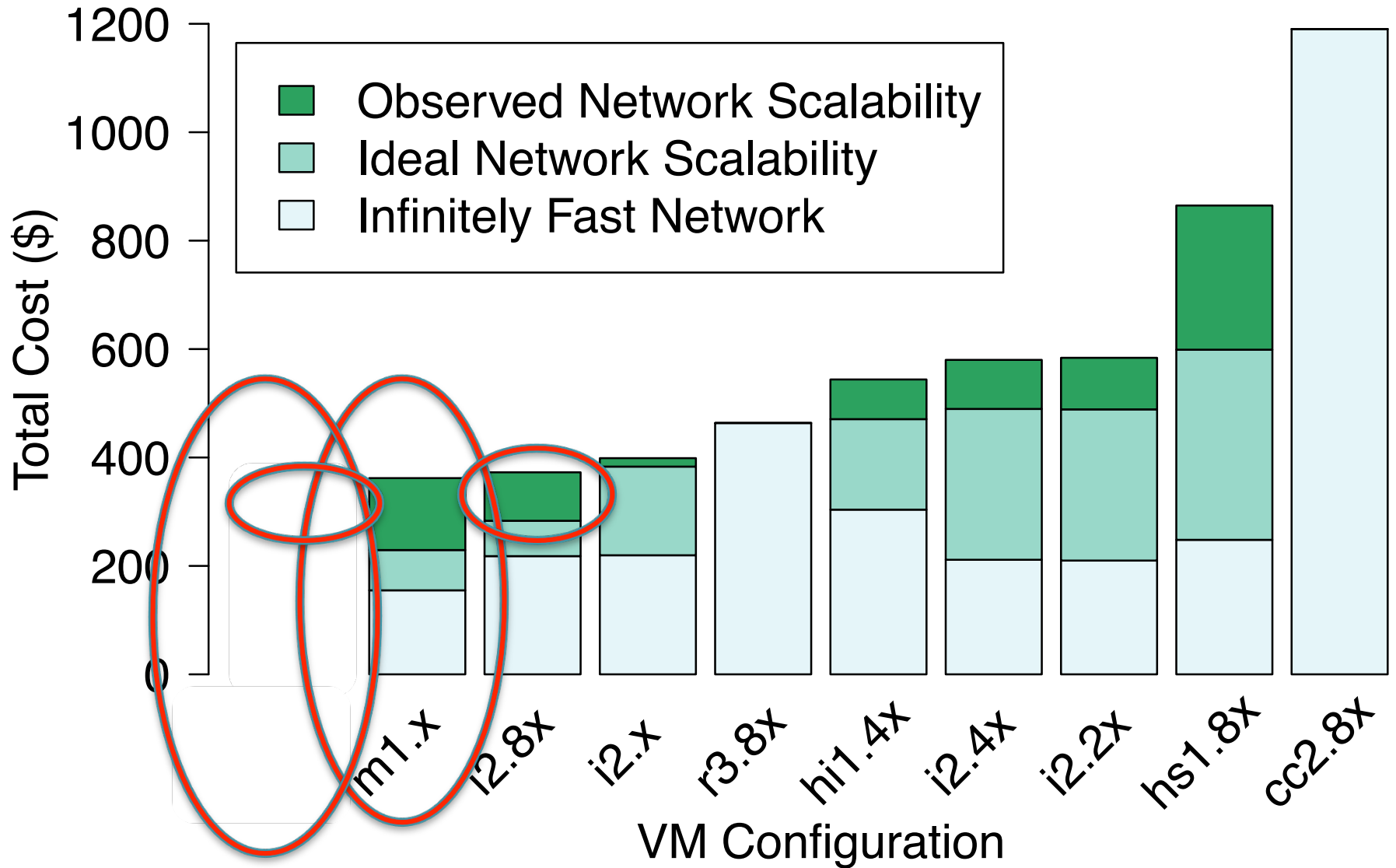
# Network Scalability



# Network Scalability



# Predicted Cost of 100TB Sort



# You Benchmarked EC2... So What?

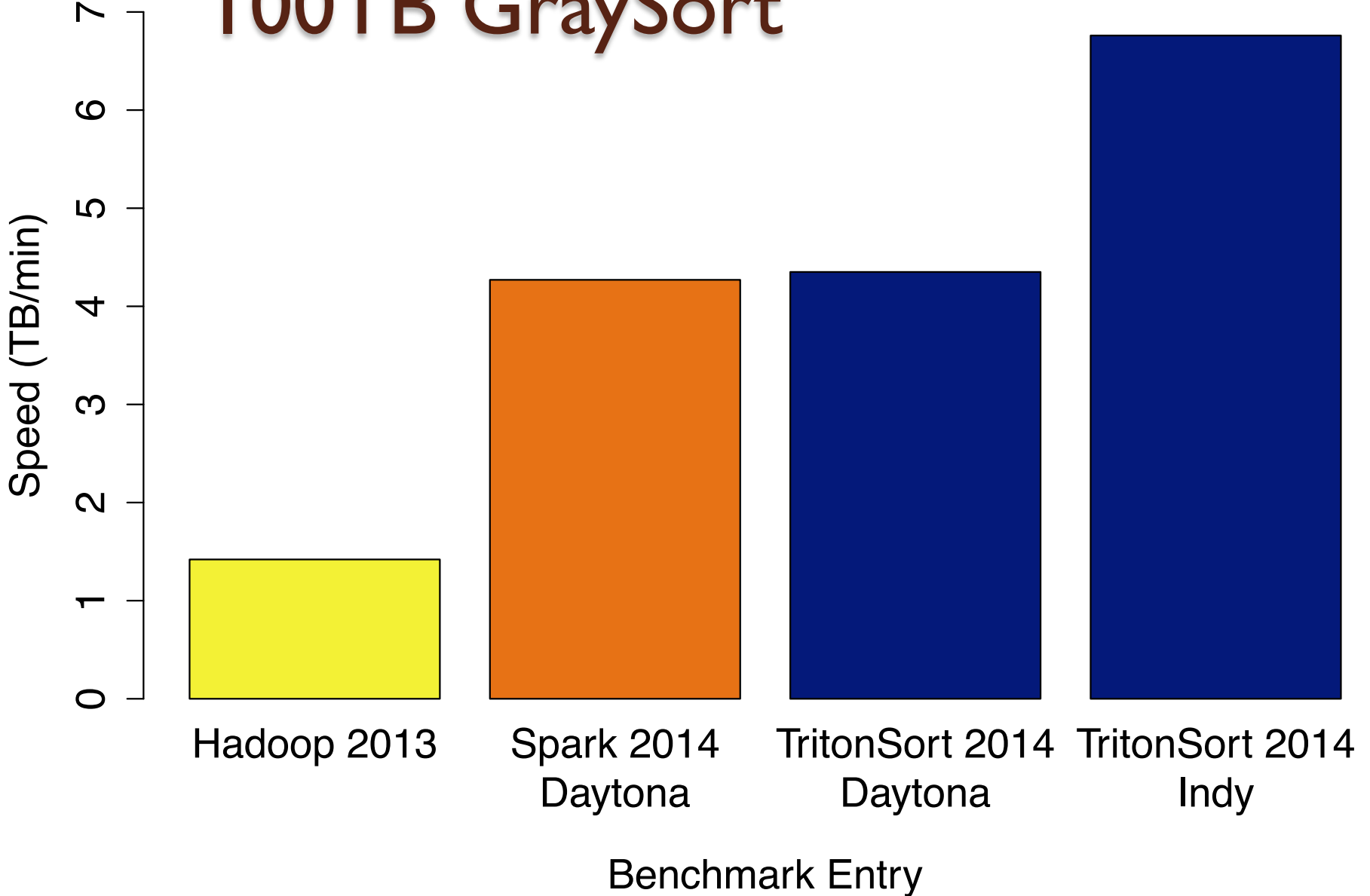
- Unofficial 2014 SortBenchmark Results
  - Indy/Daytona 100TB GraySort
  - Indy 60 second MinuteSort
  - Indy/Daytona 100TB CloudSort



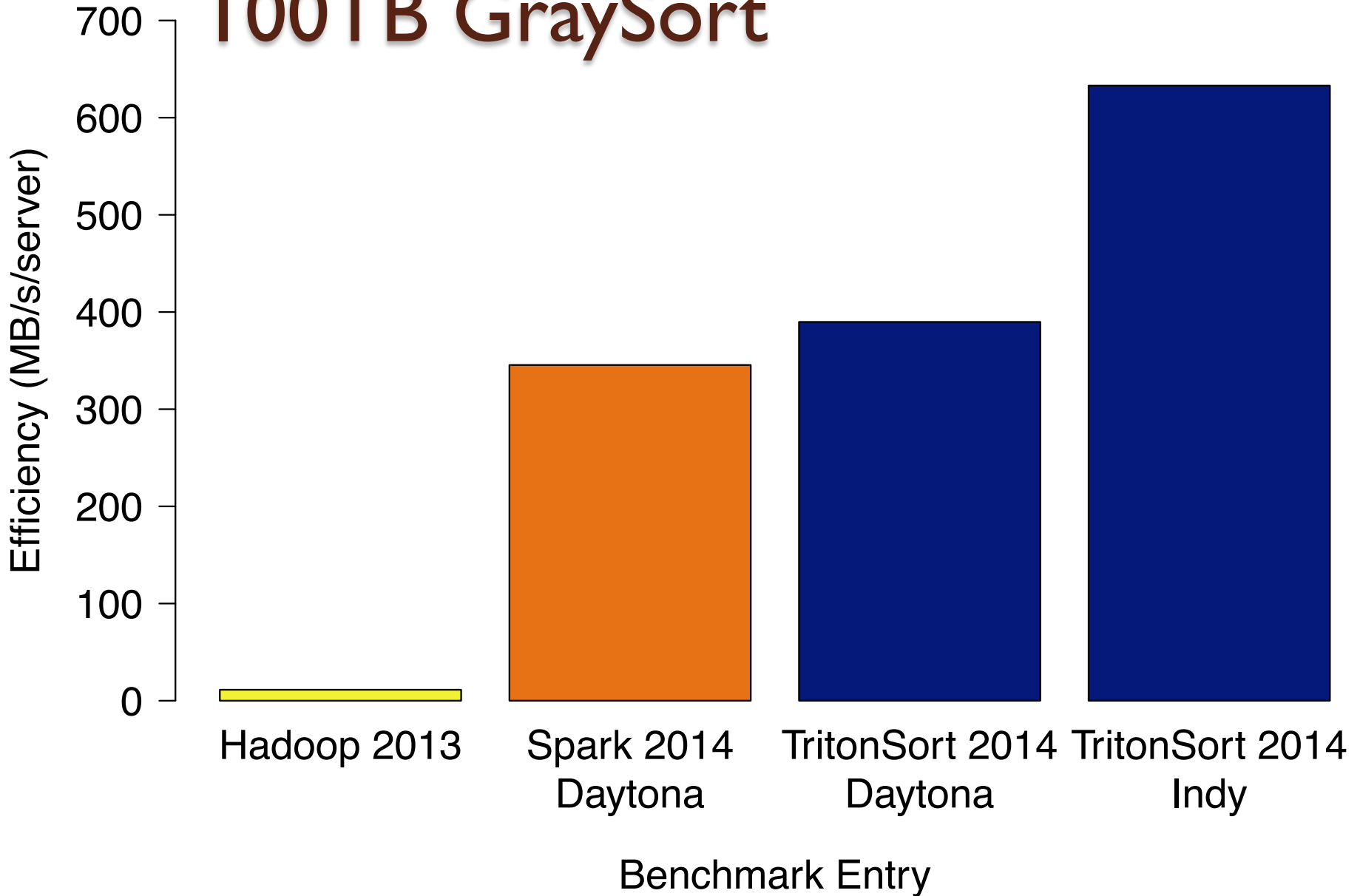
# You Benchmarked EC2... So What?

- Unofficial 2014 SortBenchmark Results
  - Indy/Daytona 100TB GraySort
  - Indy 60 second MinuteSort
  - Indy/Daytona 100TB CloudSort

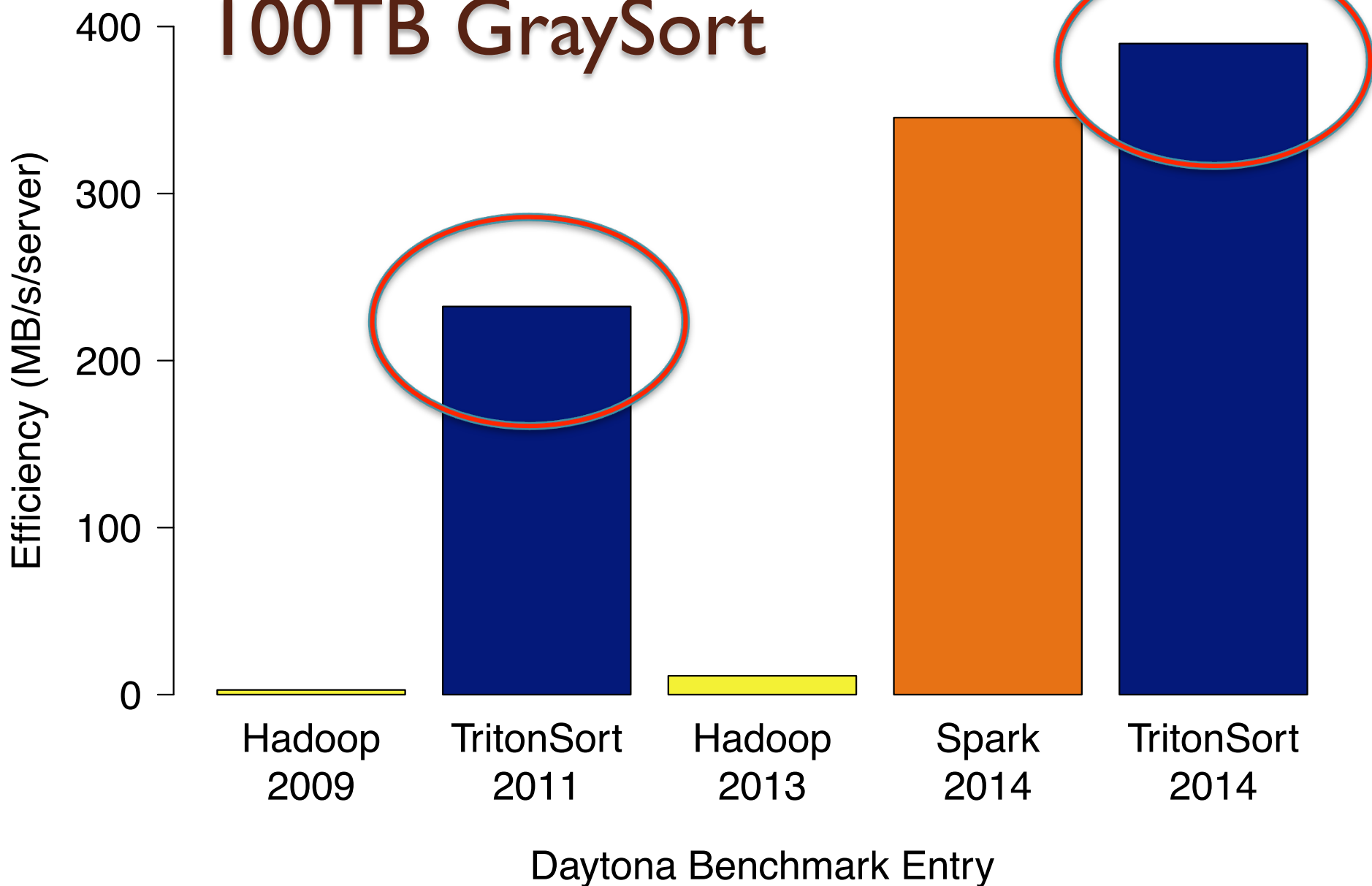
# 100TB GraySort



# 100TB GraySort



# 100TB GraySort



# Cost Analysis

- Indy GraySort
  - 178 i2.8xlarge VMs
  - 888 seconds
  - \$299.45
- Model predicts \$324.57
- Cost difference due to
  - Day-to-day and hour-to-hour variance
  - Different set of physical servers

# Central Question

- Is the following feasible?
  - Massive computation
  - Public cloud
  - High performance
  - Low cost

100TB Sort

Amazon EC2

> 4x Faster

< \$300



# Conclusion

- Investigate the Cloud as a viable environment for big data applications
- Benchmark EC2 to discover bottlenecks
- Run high performance 100TB sort
  - More than 4x faster than 2013 record
  - 11x fewer servers



<http://themis.sysnet.ucsd.edu/>

# Thank You!

- Cisco, NetApp, NSF, FusionIO, Amazon
- CNS (you!)
- Questions?

