DOCLITE: DOCKER CONTAINER-BASED LIGHTWEIGHT BENCHMARKING ON THE CLOUD

Supervisors:
Dr. Adam Barker
Dr. Blesson Varghese

Summer School 2015

Lawan Thamsuhang Subba

Structure of Presentation

- Introduction
- Problem Statement
- Related Work
- Solution
- Evaluation
- Limitations
- Conclusion

Introduction

- **What is Cloud Computing?**
  - Rent computing resources
  - Pay as you go Model
  - Elasticity
  - Competitive pricing
  - Quality of Service guarantees

- **Who are using it?**

Problem Statement

- Which service is right for me?
- Hundreds of different Cloud providers
  - Major providers are Amazon, Google and Microsoft
- Every provider offers virtual machines with different computing capacity and associated costs
  - Some virtual machines provided by Amazon include m2.xlarge (2 vCPUs, 17.1 GB Memory) and cr1.8xlarge (32 vCPUs, 244 GB Memory)
- Existing research shows differing performance
  - Hardware infrastructure
  - Operating system
  - Virtualization technology

Related Work

- VMmark
  - Scalable tile based benchmarks
  - Each tile represents workload
    - Mail Server, Web Server, File Server
  - Performance is an aggregate of all workloads

- CloudCmp
  - Compare cloud services along metrics like elastic computing, persistent storage and networking services
  - Used Benchmarking tools to capture performance attributes
  - Compared Amazon AWS, Google App Engine, Rackspace Cloudservers and Microsoft Azure

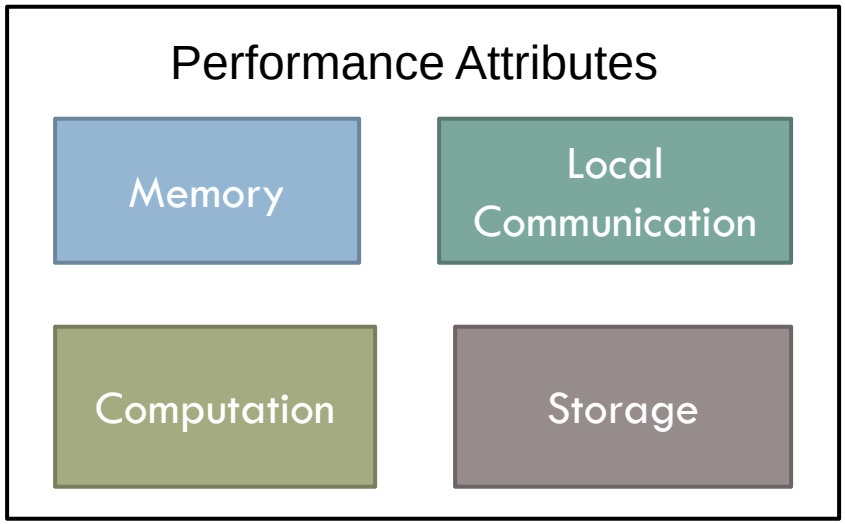
- CloudStone
  - Use Web applications on Cloud services to evaluate performance
  - Web application Olio and workload generation tools were to gauge performance under different environments

Heavyweight Benchmarks

- Developed at the Big data lab at the University of St Andrews
- None of the existing work considered Users computing requirements
- Main properties of Cloud service are performance, security and economics
  - Performance attributes can be measured in terms of
    - Memory
    - Local communication
    - Computation
    - Storage
- Performance attributes were captured using a benchmarking tool for 12 different virtual machines provided by Amazon
  - LMbench

Contd..

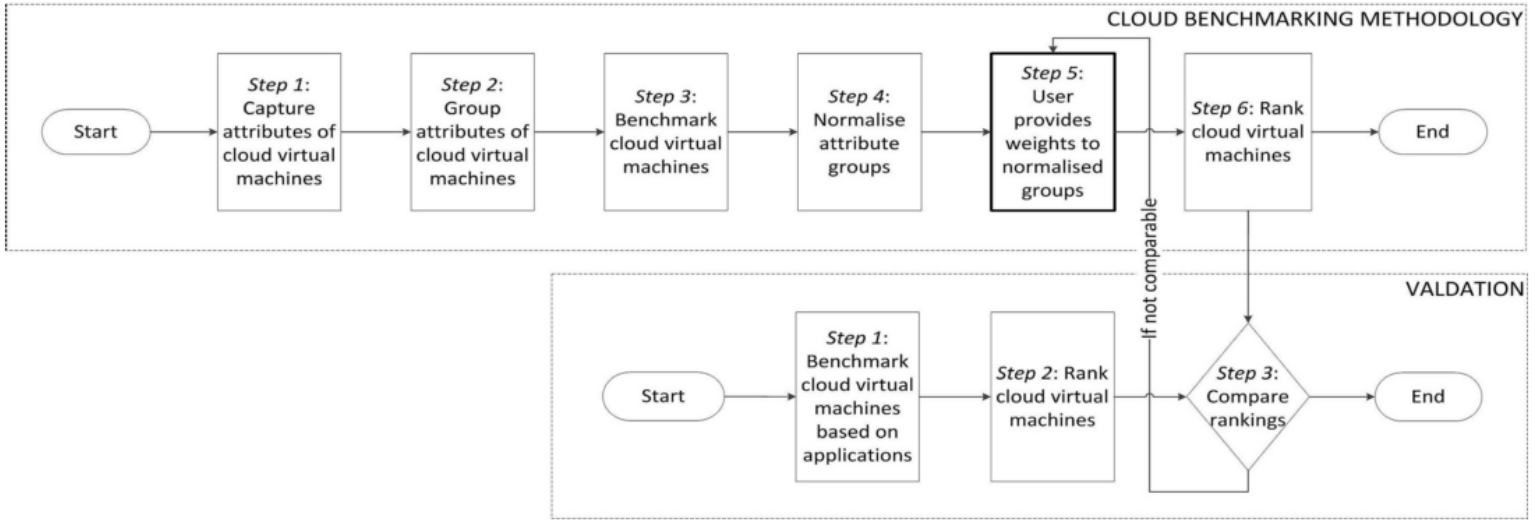
- Customers can express their computing needs by providing weights [0-5] for the four attribute groups



{memory, local communication, computation, storage}
{2, 2, 5, 0} → Customer A
{1, 1, 1, 5} → Customer B
{5, 2, 2, 2} → Customer C

- Virtual Machines are ranked by how likely they are to meet customers computing requirements.
  - 1) Virtual Machine A
  - 2) Virtual Machine B
  - 3) Virtual Machine C

Contd..



Benchmark ranks

Empirical ranks

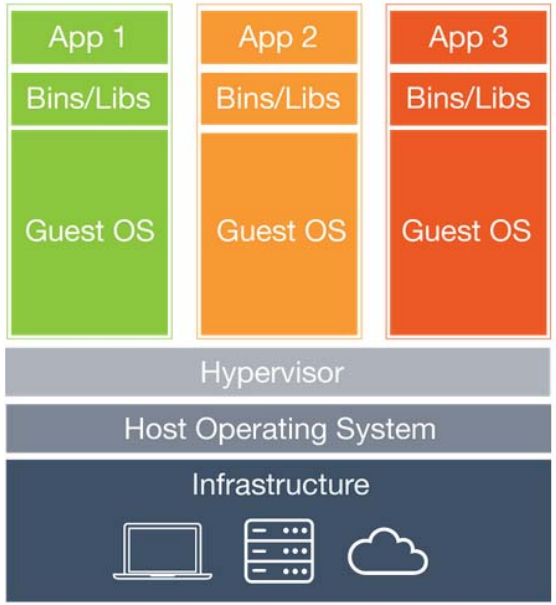
- Two computationally intensive real world applications
  - Aggregate risk analysis
  - Molecular dynamic simulation
- Validation by performing correlation between Benchmark ranks and Empirical ranks
- Heavyweight benchmarking was able to identify virtual machines that will meet users requirements with good measure of accuracy.
- However Heavyweight benchmarking proved time consuming and costly

Solution

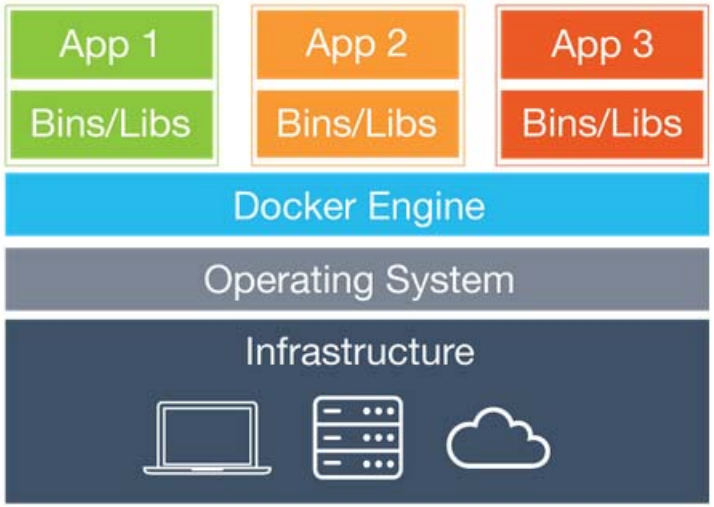
Docker Containers

- Docker is a open source virtualization technology. Which supports creation of containers which encapsulate all resources necessary for an application to run properly.

- Every container is running under the same kernel, but they have their own file system, processors, memory and devices



Hypervisor based Virtualization

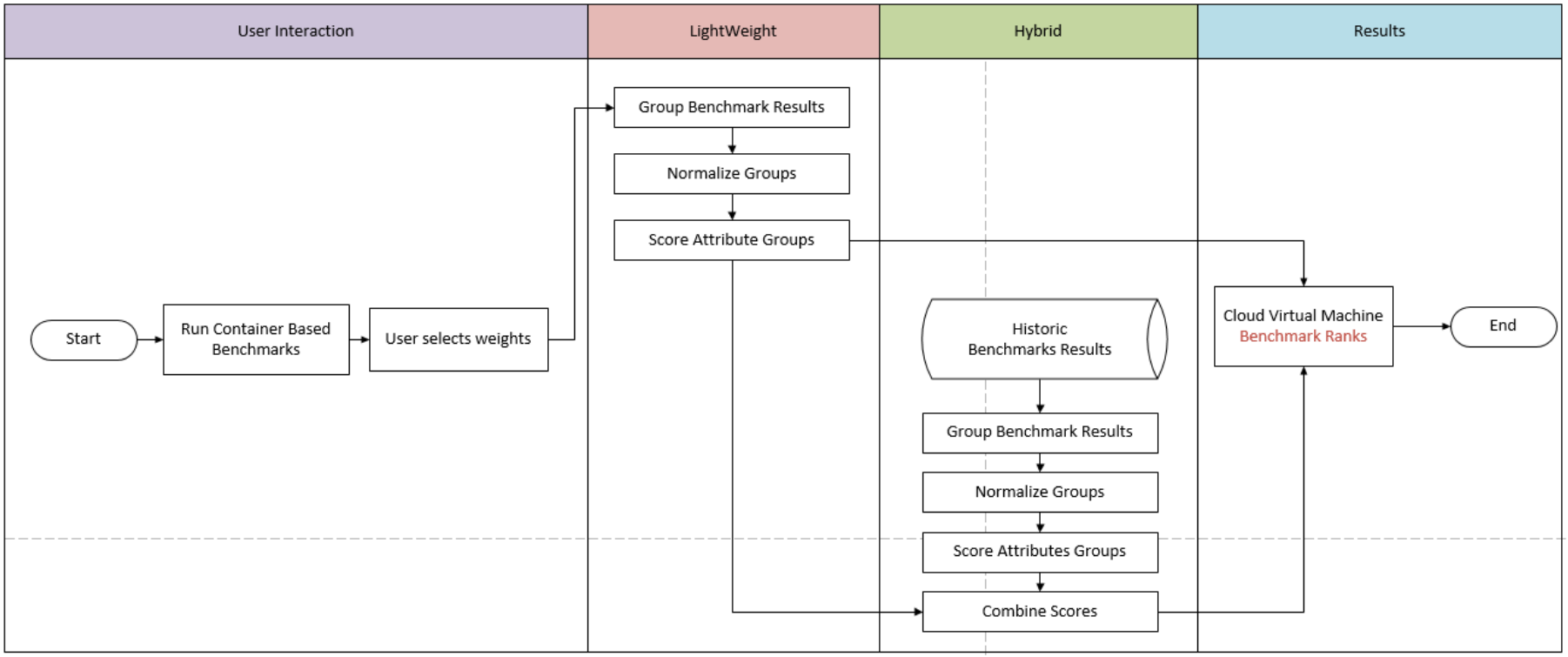


Container based Virtualization

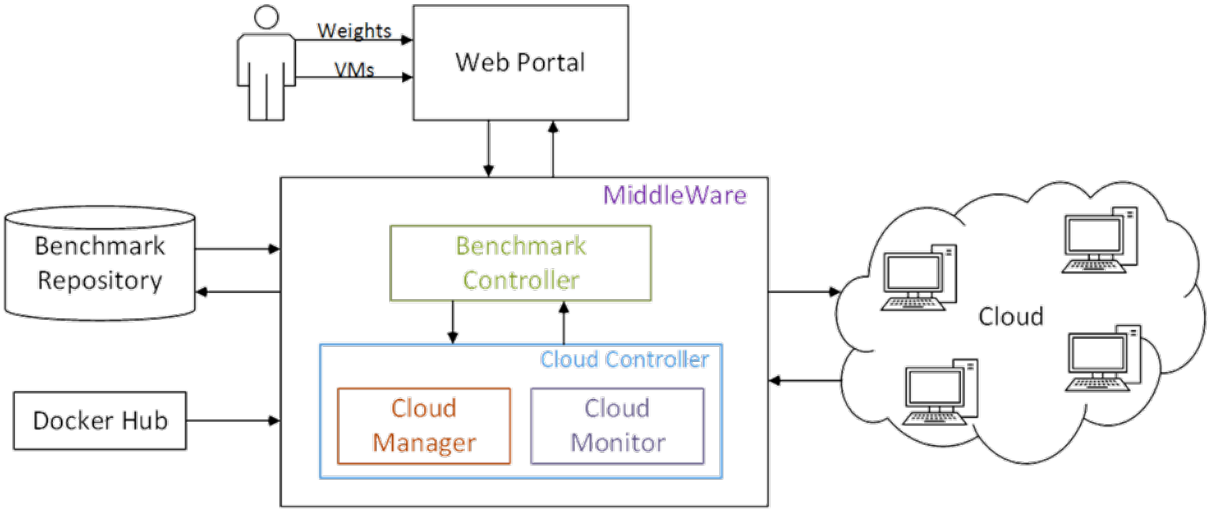
Lightweight and Hybrid Benchmarking

- Benchmark a subset of the system resources for the same Amazon VMs used by the Heavyweight approach
- Create Docker containers, allocate computing resources and run benchmarking tools inside them
  - Memory (100 Mb, 500 Mb, 1000 Mb)
  - CPU Cores (Single Core, All Cores)
  - Storage
- We have historic benchmark data from the Heavyweight approach
- Our benchmark ranks are evaluated using the Empirical ranks generated for the Heavyweight approach

Contd..


User Interaction
LightWeight
Hybrid
Results
Start
Run Container Based Benchmarks
User selects weights
Group Benchmark Results
Normalize Groups
Score Attribute Groups
Historic Benchmarks Results
Group Benchmark Results
Normalize Groups
Score Attributes Groups
Combine Scores
Cloud Virtual Machine Benchmark Ranks
End

Docker Container-based Lightweight (DoCLite) Application



- Web Portal
  - Bootstrap and HTML5
- MiddleWare
  - MVC.NET
- Benchmark Repository
  - FTP server for benchmark results
- Docker Hub
  - Can be used to Store custom benchmark containers

Evaluation

Lightweight Benchmarks

Amazon VM	Sequential Ranking				Parallel Ranking			
	Emp	100Mb	500Mb	1000Mb	Emp	100Mb	500Mb	1000Mb
m1.xlarge	10	10	10	10	8	10	10	10
m2.xlarge	5	4	4	5	10	8	8	8
m2.2xlarge	5	7	6	7	7	9	9	9
m2.4xlarge	5	6	7	6	5	6	6	6
m3.xlarge	3	3	3	3	9	7	7	7
m3.2xlarge	3	5	5	5	6	4	3	4
cr1.8xlarge	1	1	1	1	2	1	1	1
cc2.8xlarge	2	2	2	2	1	2	2	2
hi1.4xlarge	8	8	8	8	4	3	4	3
hs1.8xlarge	9	9	9	9	3	5	5	5
Correlation		94%	94%	94%		83%	80%	83%

Case Study 1: W={5, 3, 5, 0}

Amazon VM	Sequential Ranking				Parallel Ranking			
	Emp	100Mb	500Mb	1000Mb	Emp	100Mb	500Mb	1000Mb
m1.xlarge	9	10	10	10	9	10	10	10
m2.xlarge	7	5	5	4	10	8	8	8
m2.2xlarge	6	7	6	7	7	9	9	9
m2.4xlarge	5	6	7	6	5	6	6	6
m3.xlarge	4	3	3	3	8	7	7	7
m3.2xlarge	3	4	4	5	6	4	4	4
cr1.8xlarge	1	1	1	1	1	1	1	1
cc2.8xlarge	2	2	2	2	2	2	2	2
hi1.4xlarge	8	8	8	8	3	3	3	3
hs1.8xlarge	10	9	9	9	4	5	5	5
Correlation		93%	92%	89%		90%	90%	90%

Case Study 2: W={4, 3, 5, 0}

- Benchmark tool was executed inside containers in Sequential (Single CPU core) and Parallel mode (All CPU Cores). For each mode containers were allocated memory of 100Mb, 500 Mb and 1000 Mb

- No significant difference for the 3 container resource allocation cases

Hybrid Benchmarks

Amazon VM	Sequential Ranking				Parallel Ranking			
	Emp	100Mb	500Mb	1000Mb	Emp	100Mb	500Mb	1000Mb
m1.xlarge	10	10	10	10	8	10	10	10
m2.xlarge	5	5	5	5	10	9	9	9
m2.2xlarge	5	7	7	7	7	8	8	8
m2.4xlarge	5	6	6	6	5	6	6	6
m3.xlarge	3	3	3	3	9	7	7	7
m3.2xlarge	3	4	4	4	6	4	4	4
cr1.8xlarge	1	1	1	1	2	1	1	1
cc2.8xlarge	2	2	2	2	1	2	2	2
hi1.4xlarge	8	8	8	8	4	3	3	3
hs1.8xlarge	9	9	9	9	3	5	5	5
Correlation		97%	97%	97%		86%	86%	86%

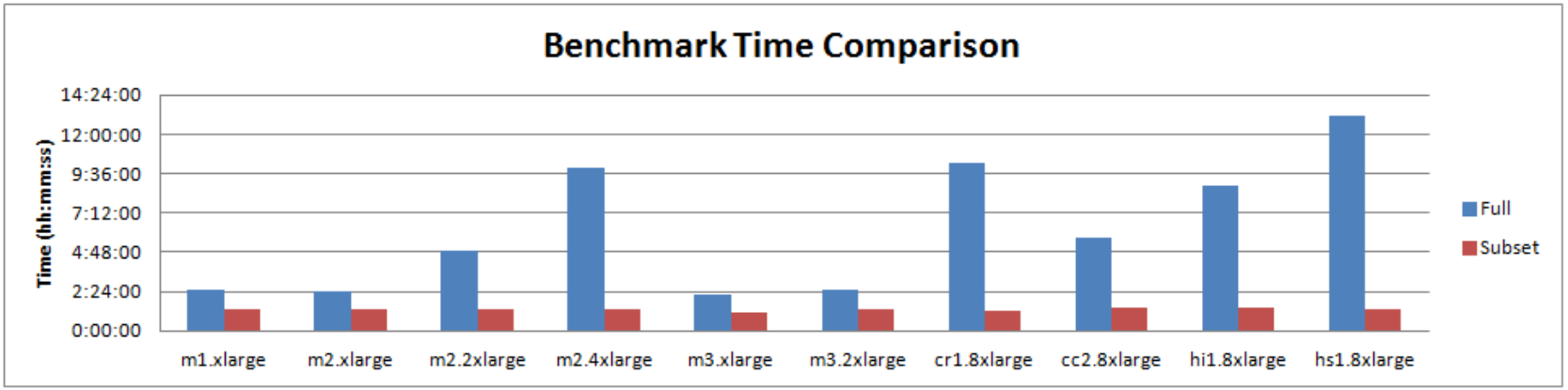
Case Study 1: W={5, 3, 5, 0}

Amazon VM	Sequential Ranking				Parallel Ranking			
	Emp	100Mb	500Mb	1000Mb	Emp	100Mb	500Mb	1000Mb
m1.xlarge	9	10	10	10	9	10	10	10
m2.xlarge	7	5	5	5	10	9	9	9
m2.2xlarge	6	7	7	7	7	8	8	8
m2.4xlarge	5	6	6	6	5	6	6	6
m3.xlarge	4	3	3	3	8	7	7	7
m3.2xlarge	3	4	4	4	6	4	4	4
cr1.8xlarge	1	1	1	1	1	1	1	1
cc2.8xlarge	2	2	2	2	2	2	2	2
hi1.4xlarge	8	8	8	8	3	3	3	3
hs1.8xlarge	10	9	9	9	4	5	5	5
Correlation		93%	93%	93%		93%	93%	93%

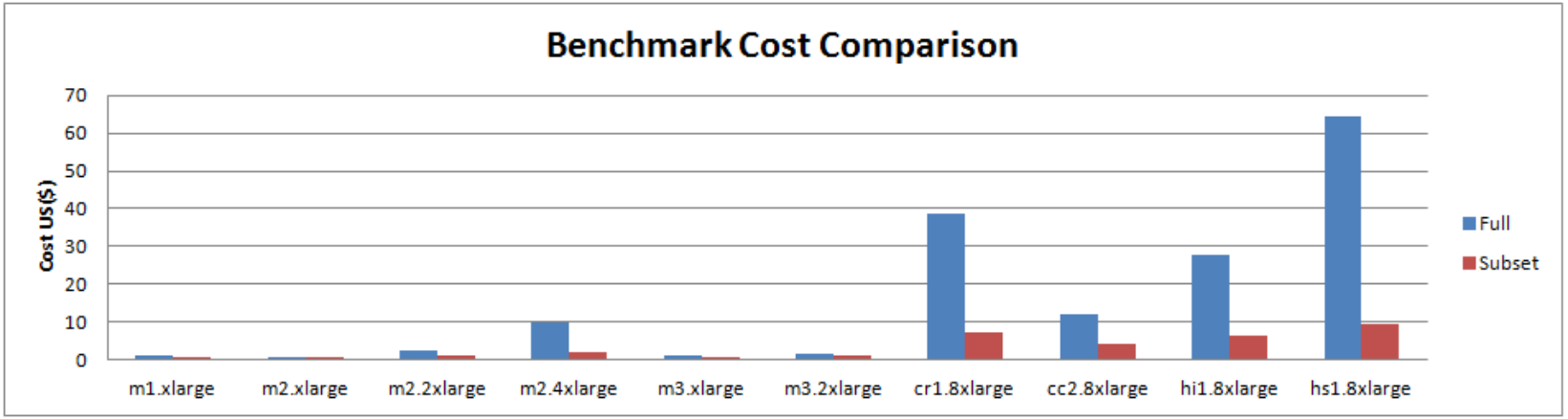
Case Study 2: W={4, 3, 5, 0}

- Slightly better correlation values for Hybrid benchmarks
- But Hybrid benchmarks rely on Historic Heavyweight benchmark results making it costly and time consuming

Lightweight vs Heavyweight Time and Cost Comparison



- Reduction in benchmark time from 63 hours to 14 hours (78%)



- Reduction in benchmark cost from 160 $ to 33 $ (79%)

Limitations

- Evaluation was performed for computationally intensive applications only
- Our application can help the customer find suitable virtual machines within Amazon only
- Compared to other works in the domain, we used only a single benchmarking tool (LMbench)
- Cloud providers are constantly updating their infrastructure and performance may change at a different date and time

Conclusion

- Lightweight benchmarks show promise in the evaluation of Cloud services
- The Lightweight approach reduced both benchmark time and cost
- The DoCLite tool can be downloaded from the Big data lab website of the University of St Andrews
- Publication
  - *Container-Based Cloud Benchmarking*
    - IEEE CloudCom 2015 (Under Review)
    - Blesson Varghese, Lawan Thamsuhang Subba, Long Thai and Adam Barker

Questions