A Reference Model for Virtual Machine Launching Overhead

Hao Wu, Shangping Ren, Gabriele Garzoglio, Steven Timm, Gerard Bernabeu, Hyun Woo Kim, Keith Chadwick, Seo-Young Noh

Presenter: Hao Wu

2014/5/28

Work supported by the U.S. Department of Energy under contract No. DE-AC02-07CH11359
And by joint CRADA FRA 2013-0001/ KISTI C13013 between KISTI and Fermilab
And supported in part by NSF under grant number CAREER 0746643 and CNS 1018731

Fermilab KiSTi IIT

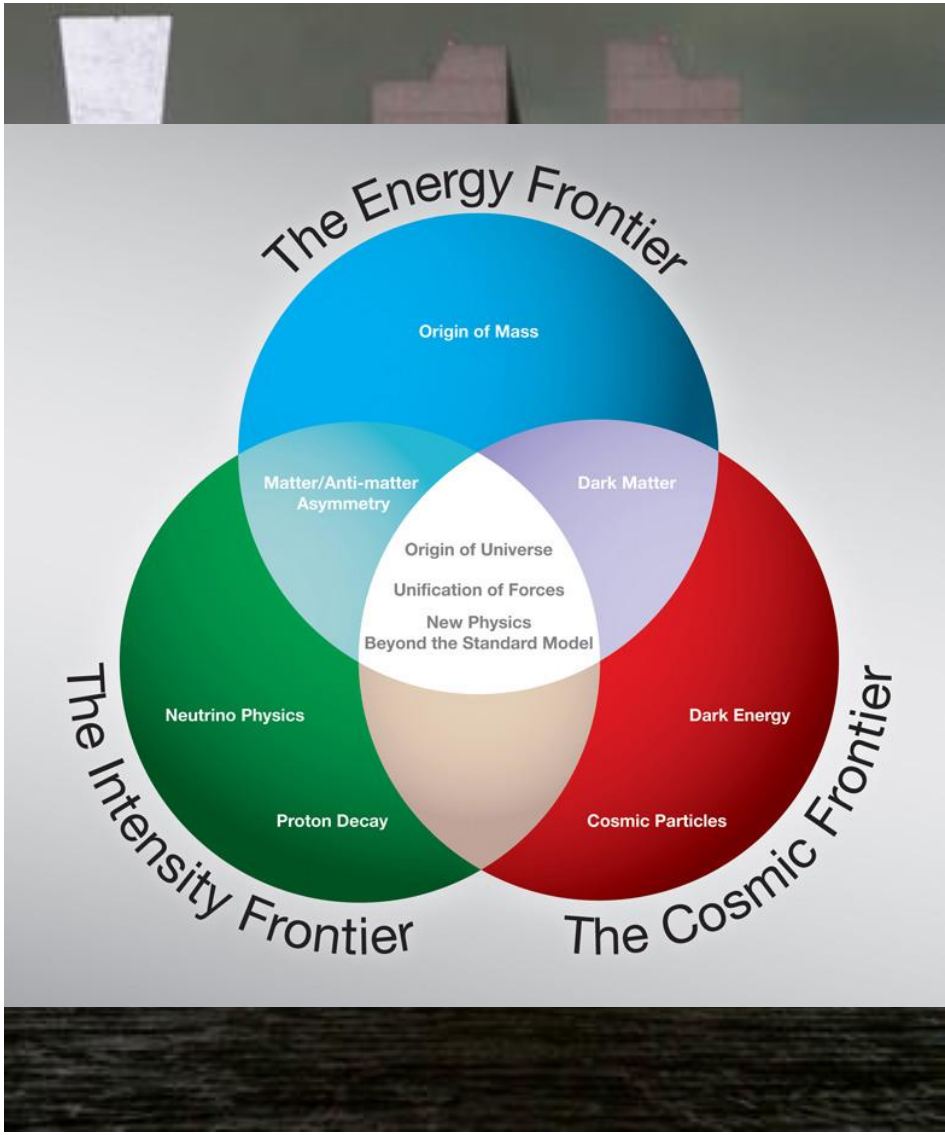
Outline

- **Introduction**
  - FermiCloud
  - Problem Statement
- **Observation**
  - Baseline
  - CPU Utilization Impact
  - IO Utilization Impact
  - Network Utilization Impact
  - Image Repository Impact
- **VM Launching Overhead Model**
- **Evaluation**
- **Conclusion**

Introduction: Fermilab

Fermi National Accelerator Laboratory:

- Leading United States particle physics laboratory
- Data center for:
  - Energy Frontier (CDF, D0, CMS)
  - Cosmic Frontier (DES, LSST, Darkside, etc.)
  - Intensity Frontier (NOvA, LBNE, Mu2e, MINOS, etc.)



Introduction: FermiCloud

- FermiCloud Project was established in 2009 with the goal of developing and establishing Scientific Cloud capabilities for the Fermilab Scientific Program.
- FermiCloud Infrastructure as a Service (IaaS) running since 2010.
- FermiCloud Project now focusing on On-demand Services for Scientific Users (PaaS, SaaS).

Introduction: Cloud Bursting

- Cloud bursting tool: GlideinWMS, vcluster
- Automatically allocate resources on collaborative private cloud (gCloud, etc.), community clouds (Rackspace, etc.), and public clouds (Amazon EC2, etc.)
- 3W challenges:
  1. What to burst?
  2. When to burst?
  3. Where to burst?

Introduction: Problem Statement

- VM launching overhead in literature:
  - Constant
  - Negligible
- FermiCloud daily operation:
  - Launch time has large variation
  - Consume CPU utilization and IO utilization

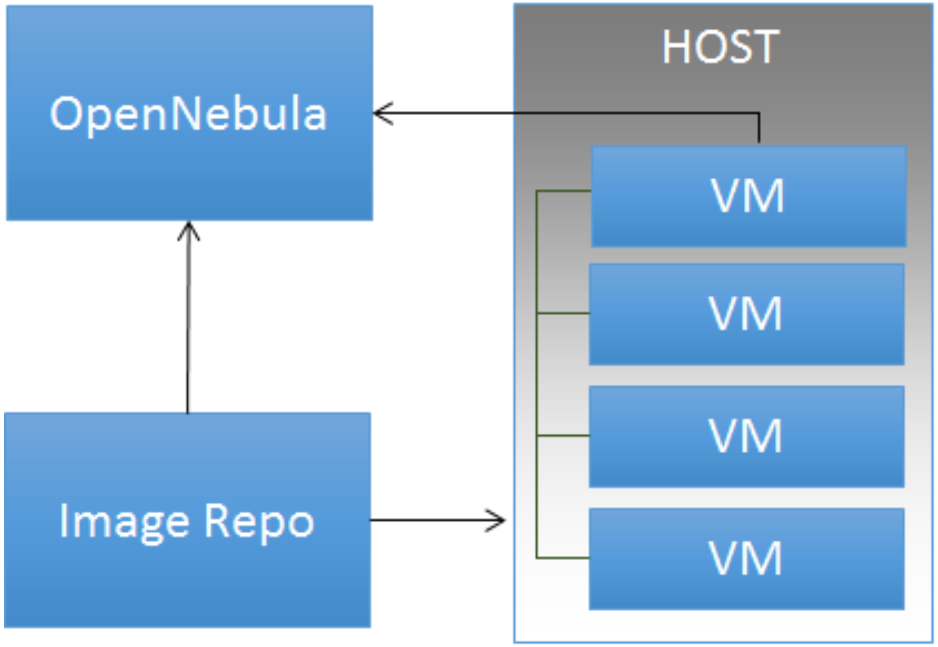
Example Scenarios

- Scenario one:
  - Only one host available in private cloud
  - Several VMs are being launched on the host
  - A new VM is launched
- Scenario two:
  - A task needs 5 more minutes to complete
  - A new VM needed for a new task
  - VM launch overhead is 6 minutes
- Without a reference model, VM launching process can lead resource and cost waste!!

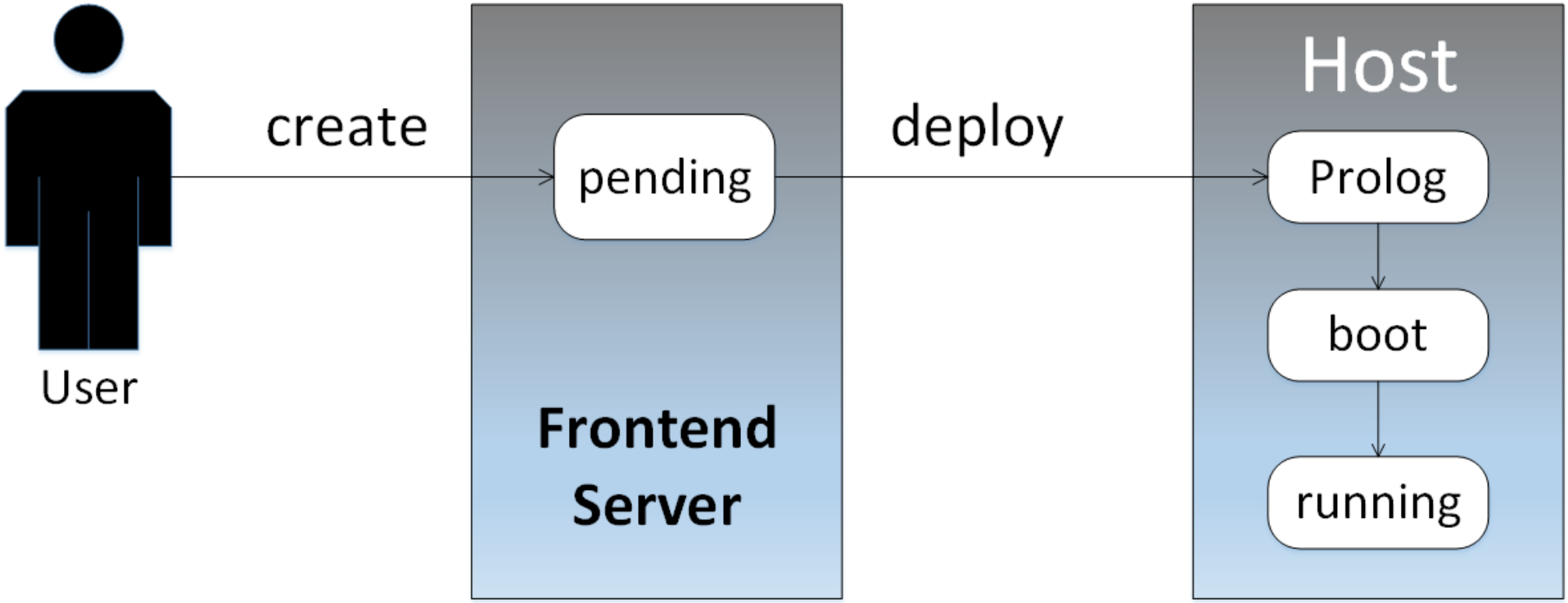
System Architecture

System Configuration:

- Front-end: 16-core Intel(R)Xeon(R) CPU E5640 @ 2.67GHz, 48GB memory.
- 15 VM hosts: 8-core Intel(R) Xeon(R) CPU X5355 @ 2.66GHz and 16GB memory
- Cloud Platform: OpenNebula



OpenNebula VM Launching Process



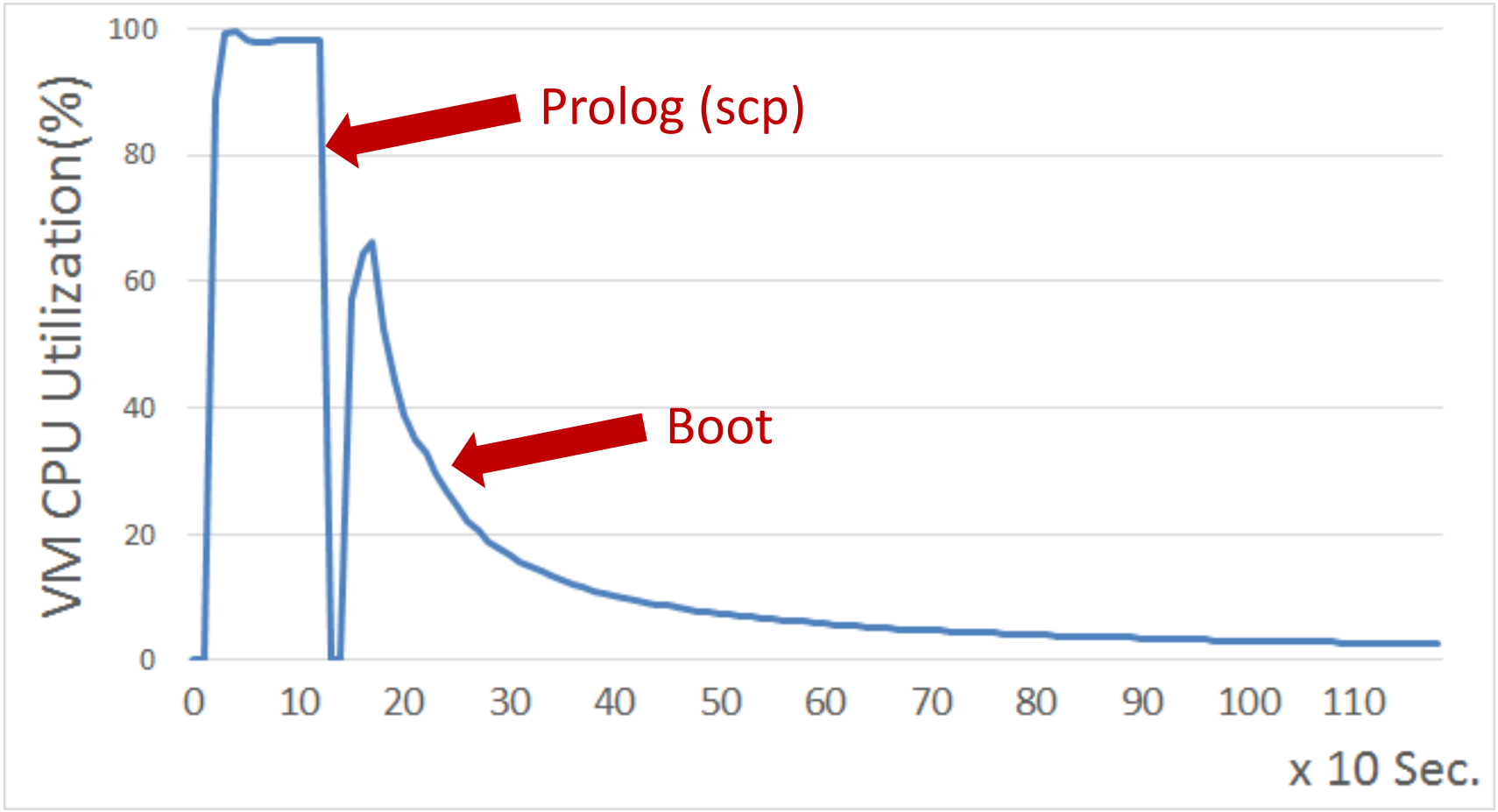
Methodology

- System information collection:
  - Noninvasive programs, i.e., Iostat, sar
- VM launch time:
  - Retrieve information from VM system log
  - Start time of SSHD service
- System CPU Utilization control:
  - cgroup

Terminology

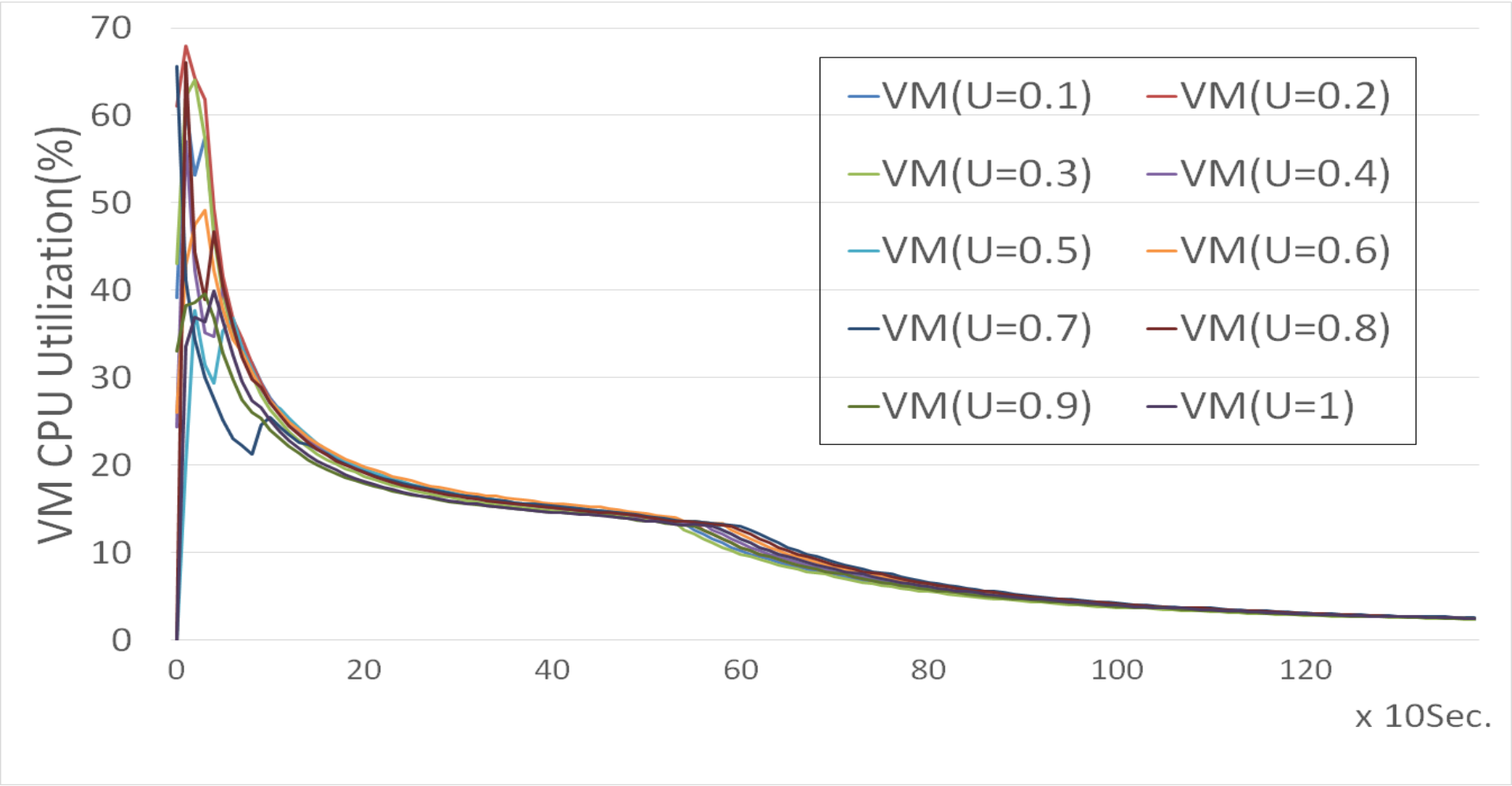
- System CPU Utilization
  - Total CPU utilizations consumed by all the processes on one host
- System IO Utilization
  - Total IO utilizations consumed by all the processes on one host
- VM CPU Utilization:
  - CPU utilization consumed by a virtual machine on a single core
- Prologue (Prolog)
  - The process of copying an image from image repository to the host machine

Baseline VM Launching Overhead



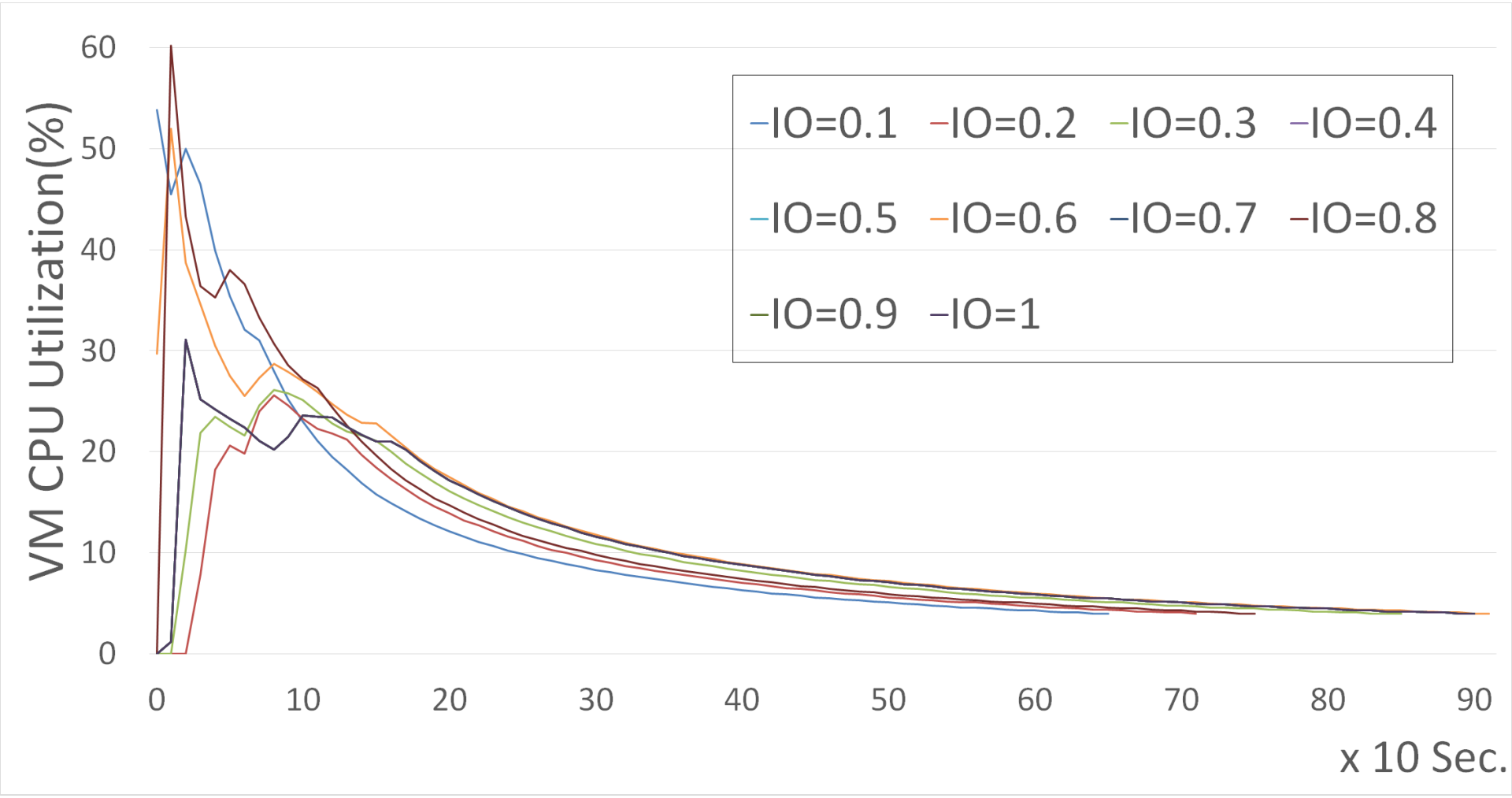
VM configuration: one virtual core, 2GB memory, 16GB raw image

System CPU Utilization Impact



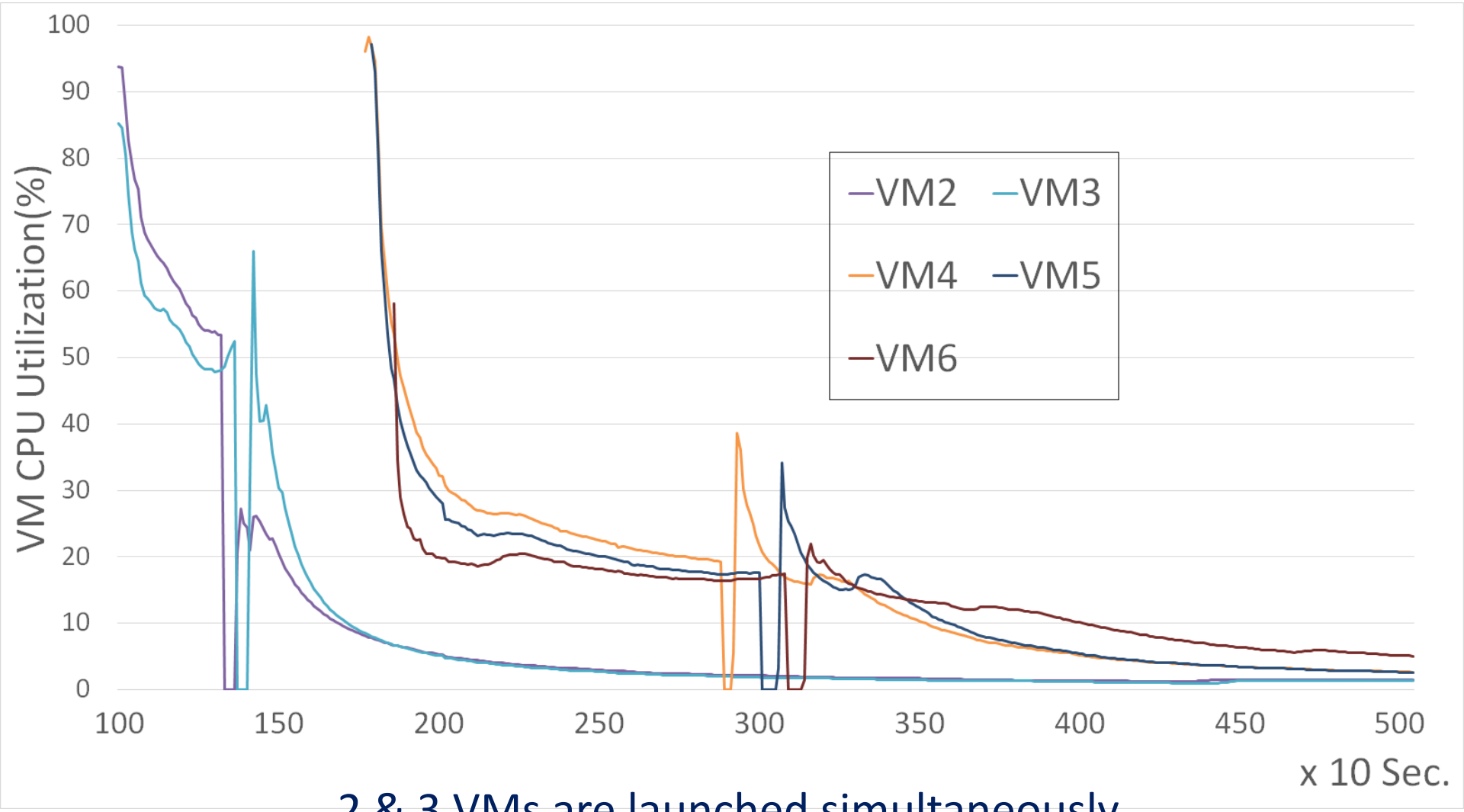
VMs are launched under different system CPU utilization

System IO Utilization Impact



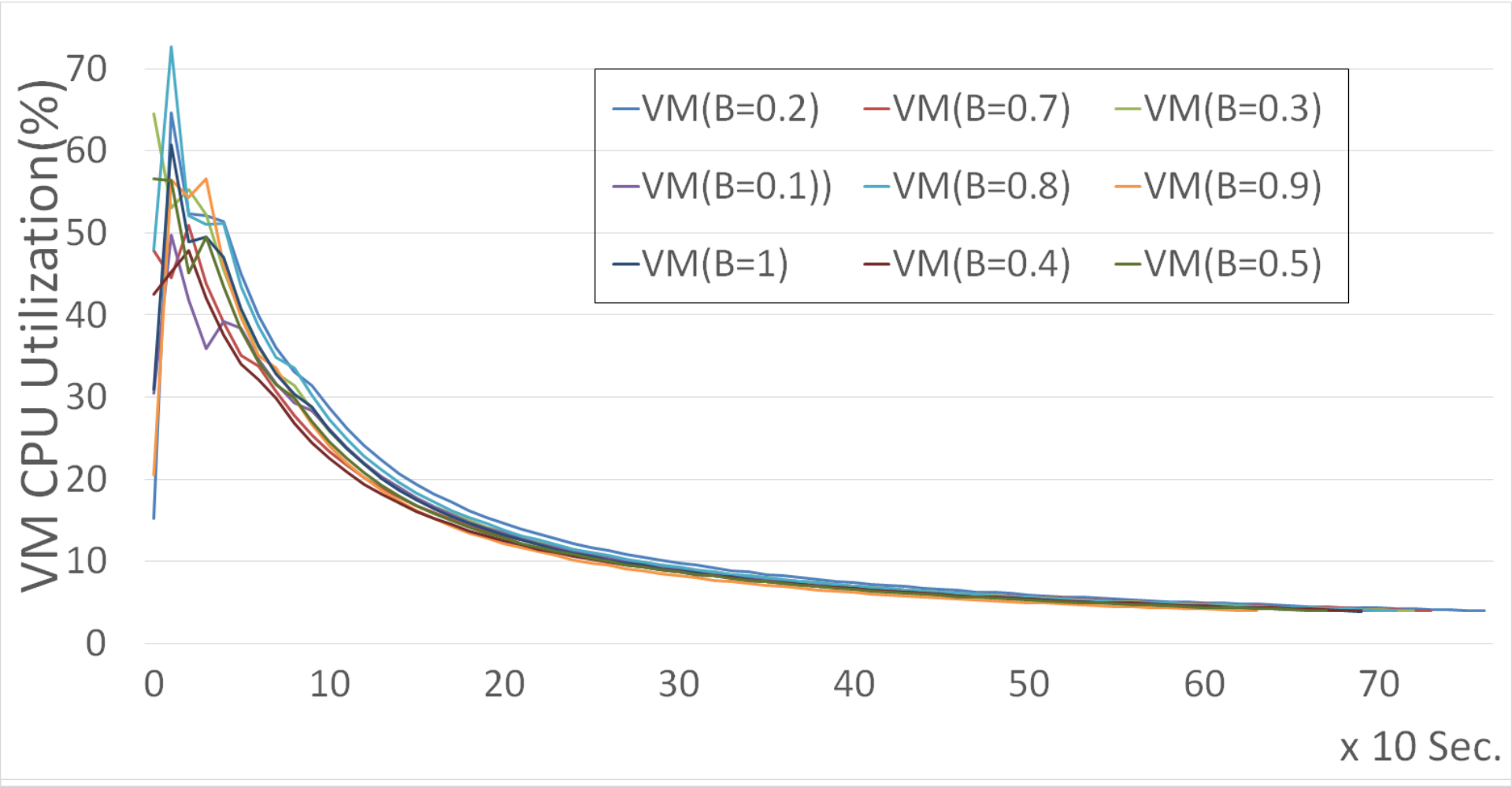
VMs are launched under different system IO utilization

Simultaneous Launches



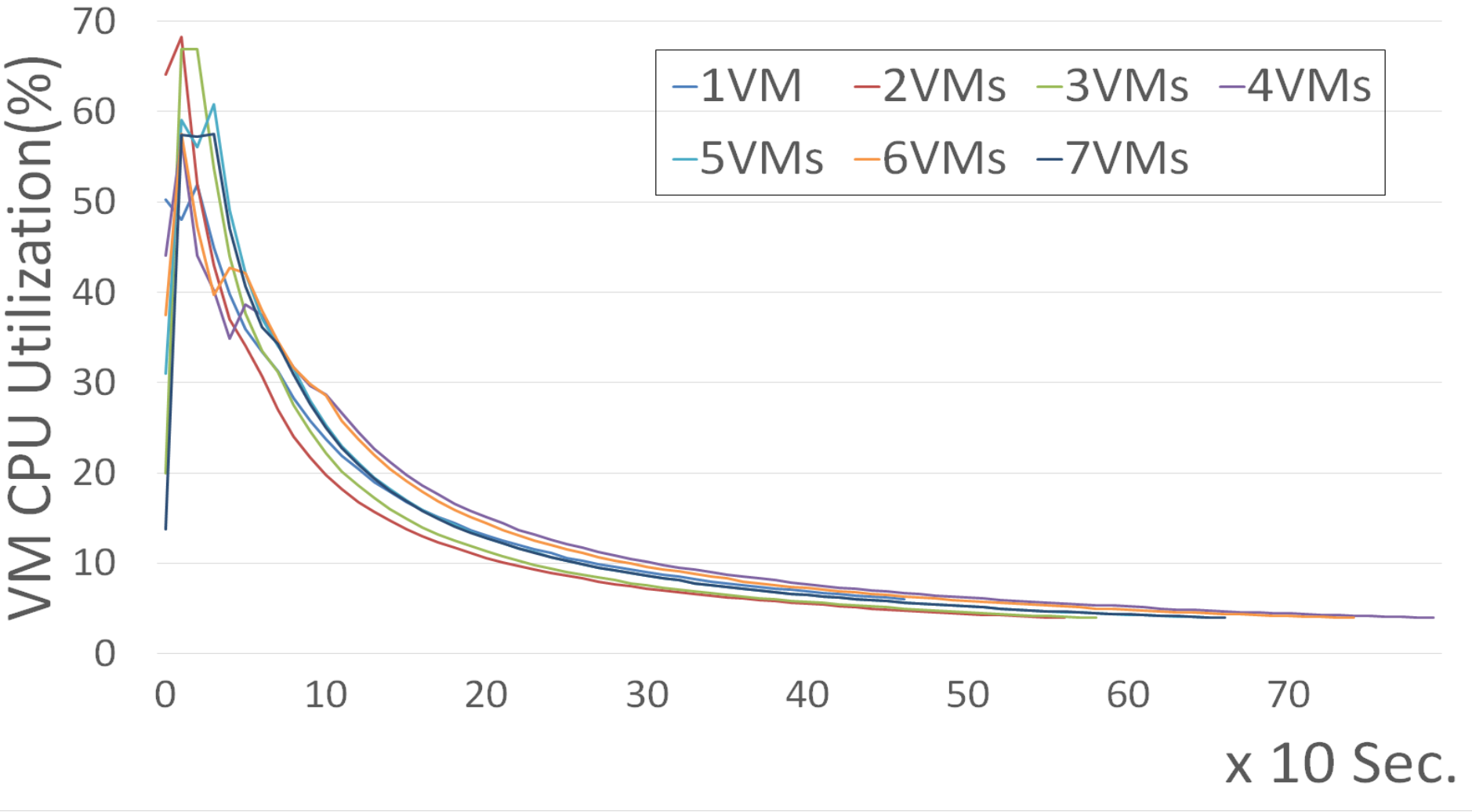
2 & 3 VMs are launched simultaneously

Network Traffic Impact



VMs are launched under different system network utilization

Image Repository Impact



VMs are launched on different hosts simultaneously

Summary of observation

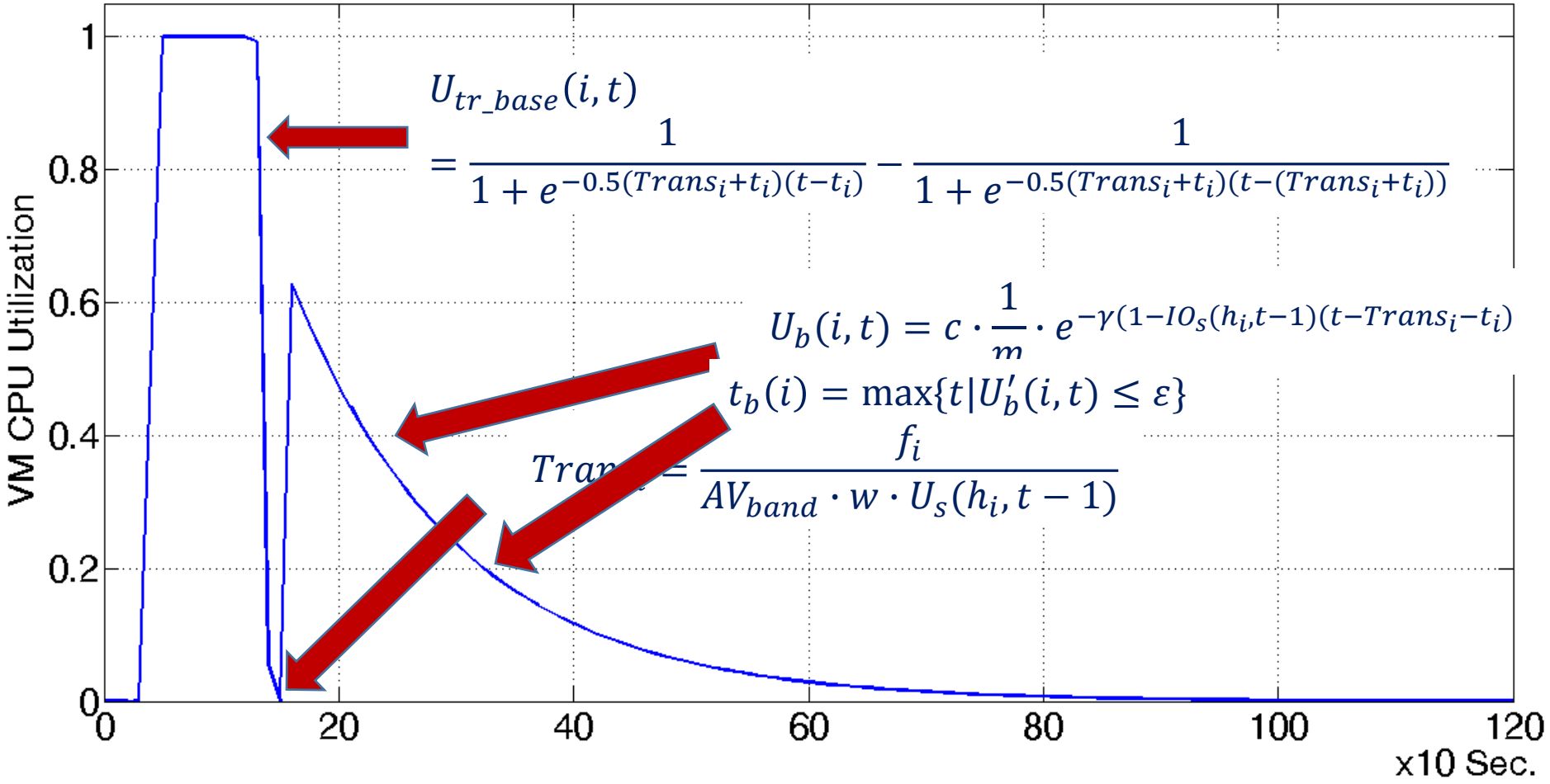
- VM launching overhead:
  - Prolog (image copying/transferring)
  - Booting overhead
- Prolog:
  - Significant variations
- Booting overhead:
  - Steady, i.e. has less variations

Notations

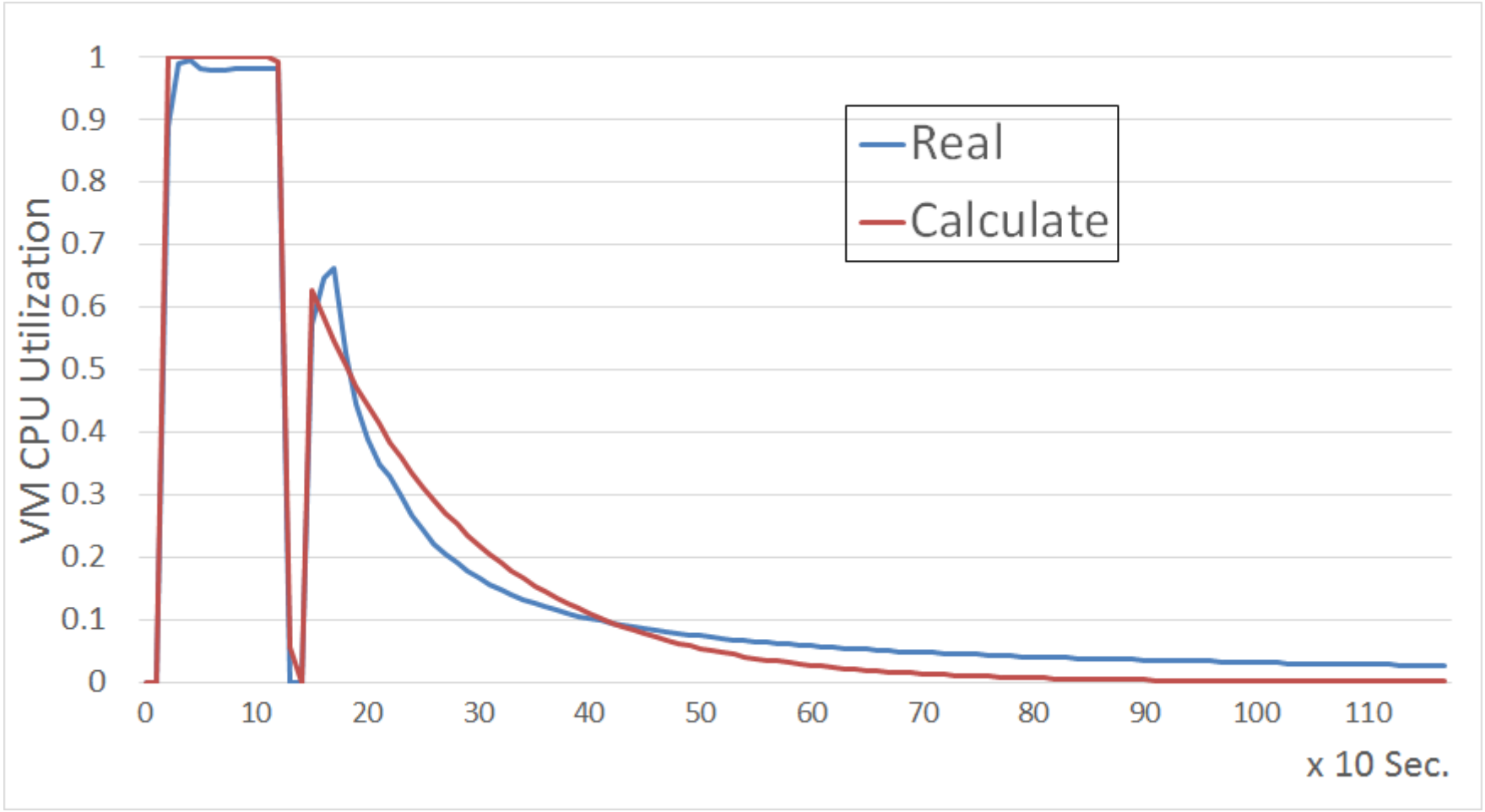
- VM: $v = (f, t, h)$
  - f : $the\ image\ size$
  - t : $is\ the\ virtual\ machine\ release\ time$
  - h: $is\ the\ host\ machine$

- VMs deployed on host $h_i = V_{h_i}\{v_1, v_2, \dots, v_n\}$
- Disk IO bandwidth: B_d
- Network IO bandwidth: B_n
- Available disk IO bandwidth:$av_d(h_i, t)$
- Available network bandwidth:$av_n(h_i, t)$
- Available image repository bandwidth:$av_i(t)$

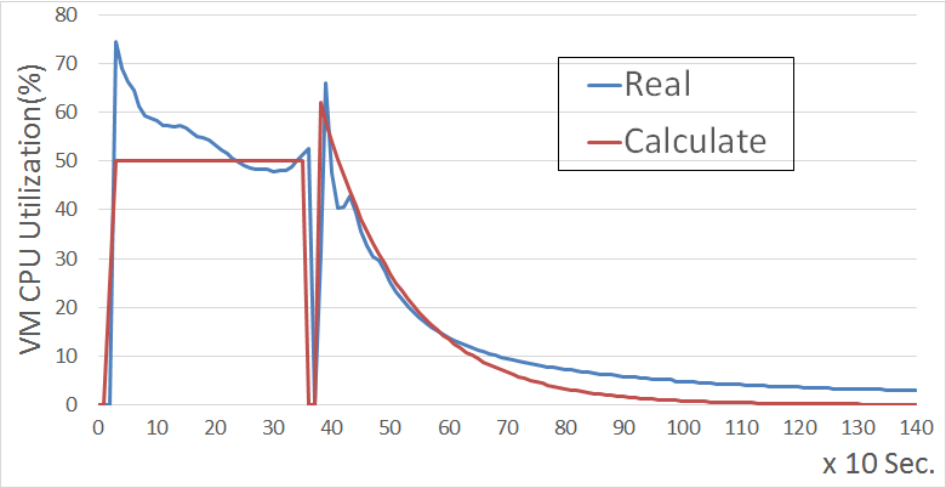
Modeling Overhead

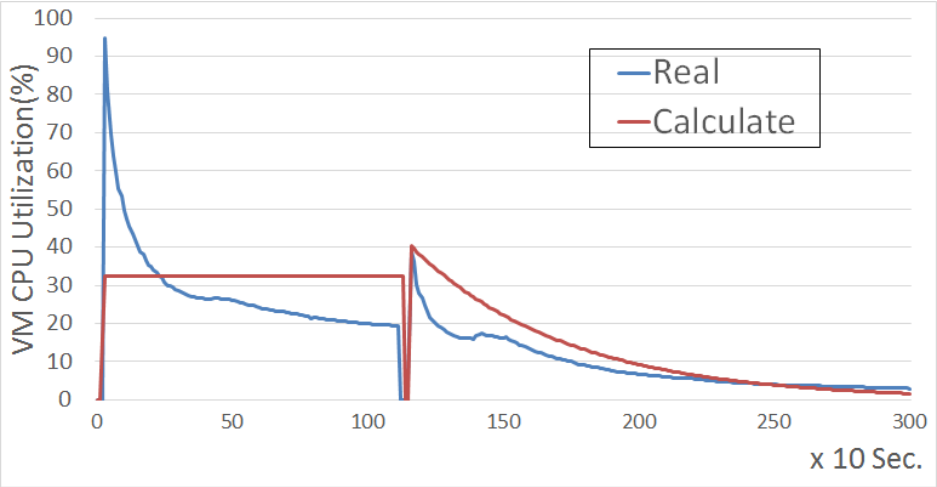


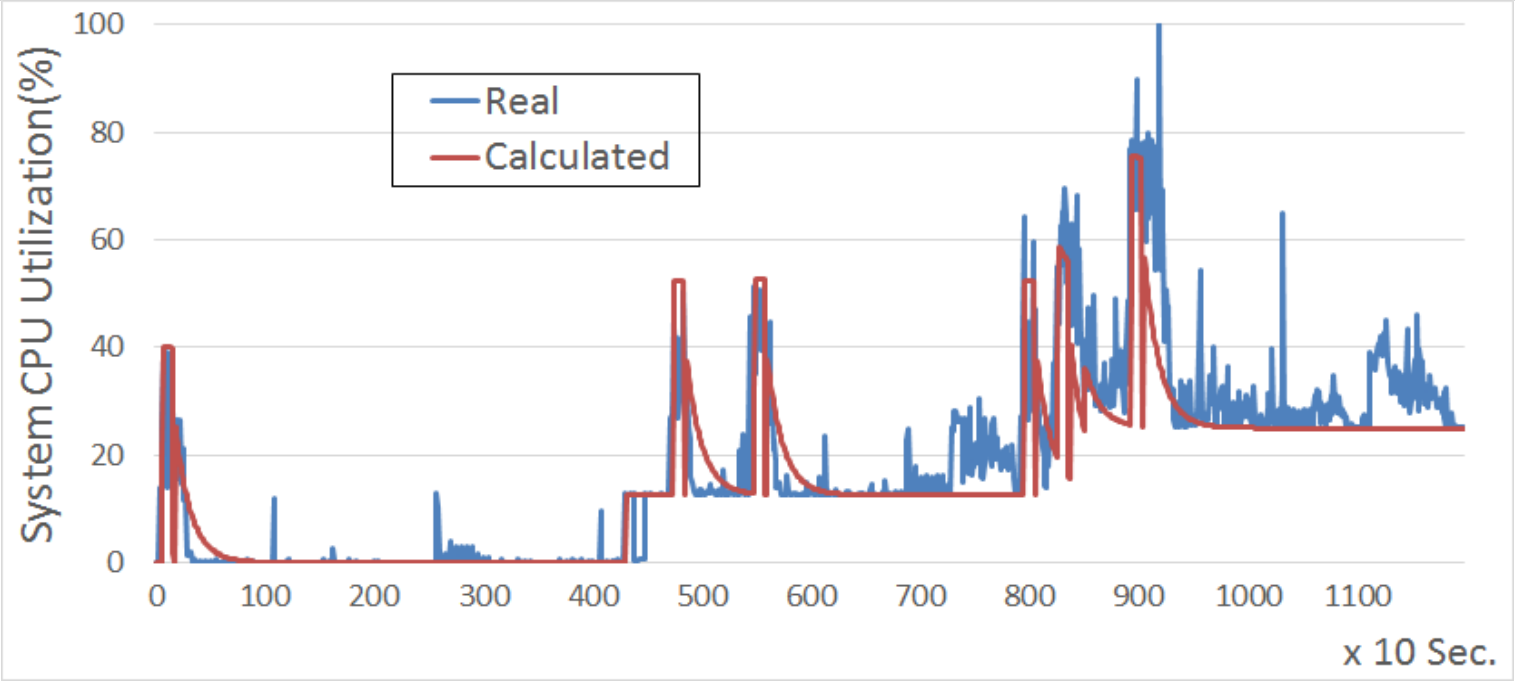
Evaluation: Baseline



Evaluation


VM CPU Utilization(%)
Real
Calculate
x 10 Sec.


VM CPU Utilization(%)
Real
Calculate
x 10 Sec.


System CPU Utilization(%)
Real
Calculated
x 10 Sec.

Evaluation: Overall Performance

	Utilization Difference	Time Difference
Baseline	0.03%	2.3%
2 Sim. Launches	4.46%	5.1%
3 Sim. Launches	6.28%	1.7%
Random. Launches	4.91%	6.9%
Overall	3.92%	4.0%

Conclusion

- Problem Addressed:
  - VM launching overhead has large variation
  - Without a VM launching overhead reference may lead to resource and cost waste
- Contribution:
  - Proposed a VM launching overhead reference model based on real FermiCloud operational data
  - CPU overhead
  - IO overhead
  - Timing overhead
- Evaluation indicates the proposed model can fairly predict VM launching overhead

Acknowledgement

U.S. DEPARTMENT OF ENERGY

Fermilab

NSF

KiSTi
www.kisti.re.kr
한국과학기술정보연구원
Korea Institute of Science and Technology Information

ILLINOIS INSTITUTE OF TECHNOLOGY

Thanks!
Questions?