Cloud Computing

蔡明宏 陳約廷

A View of Cloud Computing

Communications of the ACM Volume 53 Issue 4, April 2010

Author	Affiliation
Michael Armbrust, Armando Fox, Rean Griffith, Anthony D. Joseph, Randy Katz, Andy Konwinski, Gunho Lee, David Patterson, Ariel Rabkin, Ion Stoica, Matei Zaharia	UC Berkeley Reliable Adaptive Distributed Systems Laboratory (RAD Lab)

History-Based Harvesting of Spare Cycles and Storage in Large-Scale Datacenters

12th USENIX Symposium on Operating Systems Design and Implementation (OSDI ’16), November 2016

Author	Affiliation
Yunqi Zhang	University of Michigan & Microsoft Research
George Prekas	EPFL and Microsoft Research
Giovanni Matteo Fumarola	Microsoft
Marcus Fontoura	Microsoft
Íñigo Goiri	Microsoft
Ricardo Bianchini	Microsoft Research

History-Based Harvesting of Spare Cycles and Storage in Large-Scale Datacenters (2013)

Dynamic Resource Allocation Using Virtual Machines for Cloud Computing Environment

IEEE Transactions on Parallel and Distributed Systems Volume 24 Issue 6, June 2013

Author	Affiliation
Zhen Xiao	Peking University, Beijing
Weijia Song	Peking University, Beijing
Qi Chen	Peking University, Beijing

Open Source Cloud Technologies

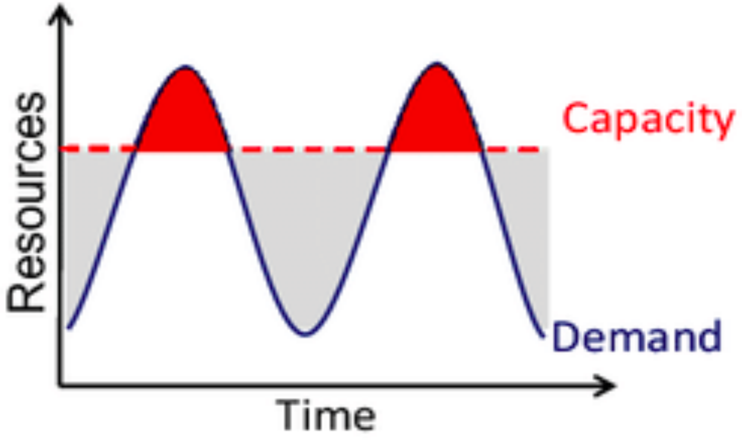
Third ACM Symposium on Cloud Computing (SoCC '12), October 2012

Author	Affiliation
Salman A. Baset	IBM T. J. Watson Research Center

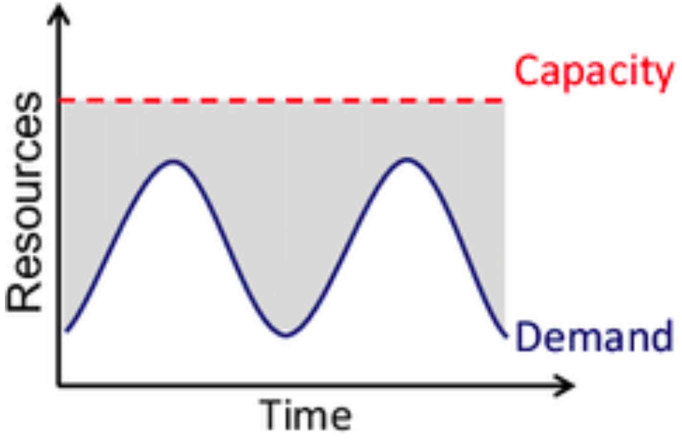
Why Cloud Computing? (Berkeley)

Cloud administrative	Client
• Economy of scale (profitable) • Brand effect Providers: Google, Microsoft, Amazon	• Pay as you go • Rewards conservative choices ➡️ Surge computing ➡️ Reduce UpFront cost

Cloud Management Difficulties



Under-Provisioning



Over-Provisioning

Utility is Key!

Static Resource Allocation

- Main Goal: Efficient Task Scheduling

To achieve:

- Better Server Utilization

History-Based Harvesting of Spare Cycles and Storage in Large-Scale Datacenters (2013)

History-Based Harvesting of Spare Cycles and Storage in Large-Scale Datacenters

12th USENIX Symposium on Operating Systems Design and Implementation (OSDI ’16), November 2016

Author	Affiliation
Yunqi Zhang	University of Michigan & Microsoft Research
George Prekas	EPFL and Microsoft Research
Giovanni Matteo Fumarola	Microsoft
Marcus Fontoura	Microsoft
Íñigo Goiri	Microsoft
Ricardo Bianchini	Microsoft Research

History-Based Harvesting of Spare Cycles and Storage in Large-Scale Datacenters (2013)

History-Based Harvesting of Spare Cycles & Storage in Large-Scale Datacenters

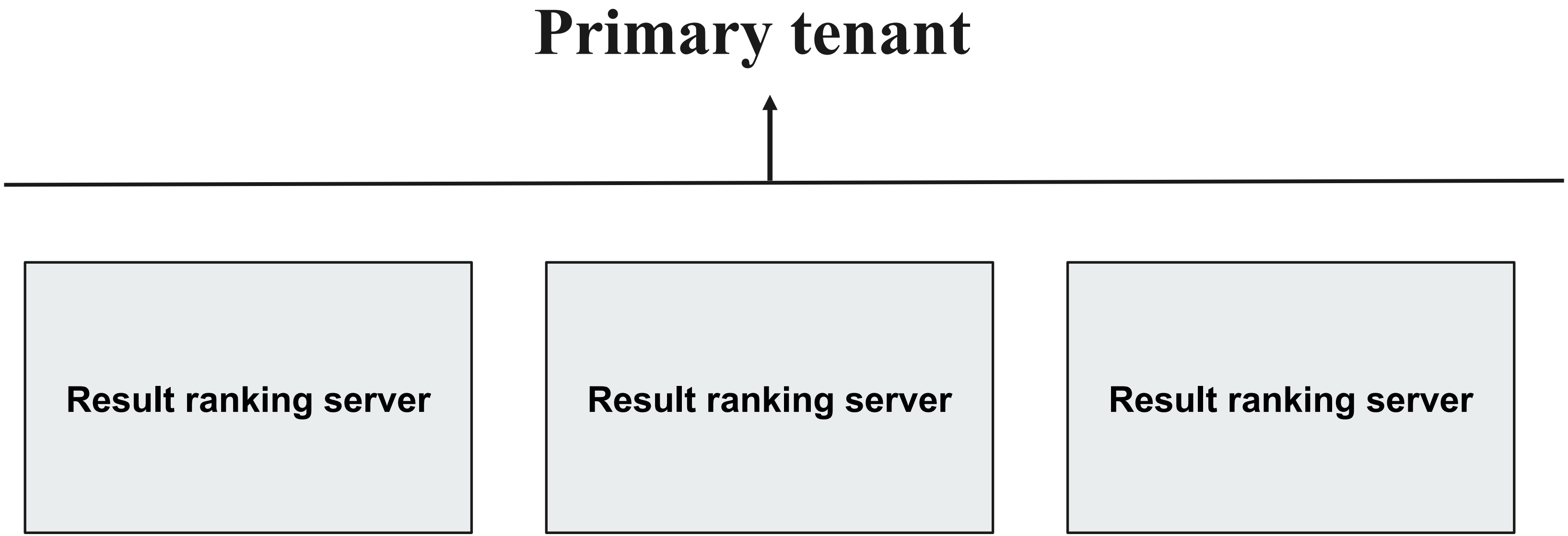
- Smart task scheduling
- Smart data placement

Smart Task Scheduling

- Motivation/Main contribution
- History data observation
- Algorithm
- Experienment

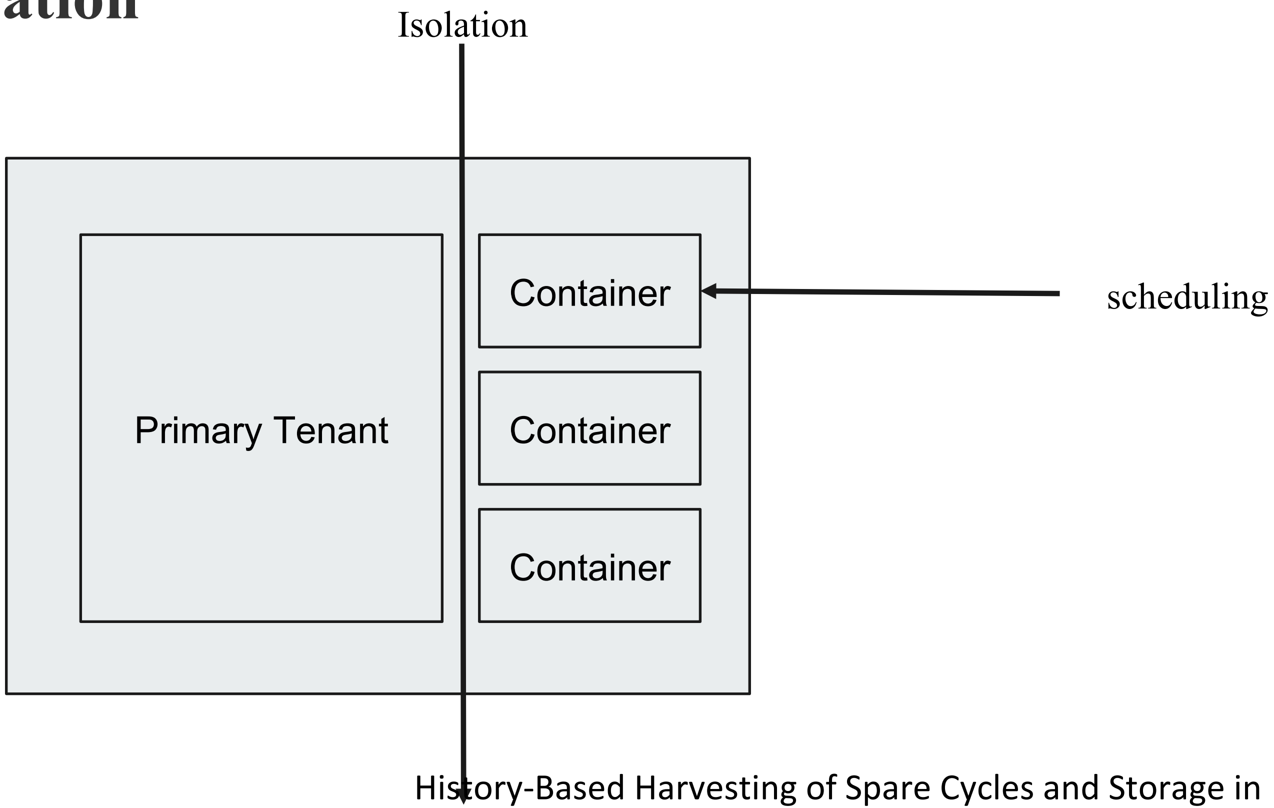
History-Based Harvesting of Spare Cycles and Storage in Large-Scale Datacenters (2013)

Define Primary Tenant



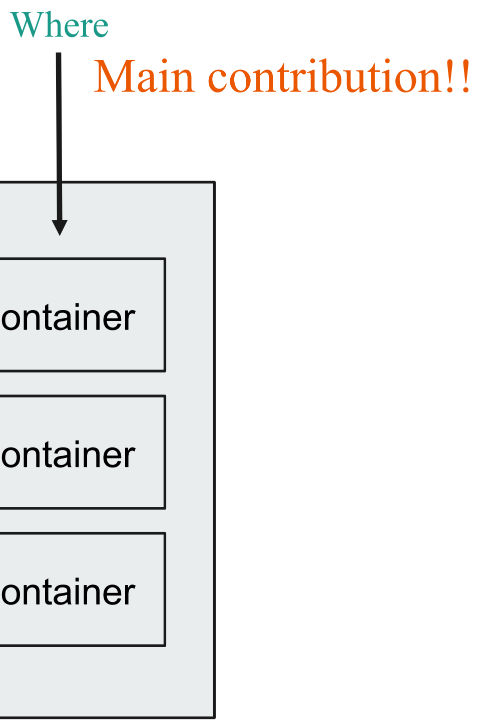
History-Based Harvesting of Spare Cycles and Storage in Large-Scale Datacenters (2013)

Service Co-Location



History-Based Harvesting of Spare Cycles and Storage in Large-Scale Datacenters (2013)

Service Co-Location



History-Based Harvesting of Spare Cycles and Storage in Large-Scale Datacenters (2013)

Smart Task Scheduling

- Motivation / Main contribution
- History data observation
- Algorithm
- Experienment

History-Based Harvesting of Spare Cycles and Storage in Large-Scale Datacenters (2013)

Tenant Characterization

Periodic

- User facing application

Constant

- Web crawling
- Batch Data analytics

Unpredictable

- Testing

History-Based Harvesting of Spare Cycles and Storage in Large-Scale Datacenters (2013)

Why historical is good reference?

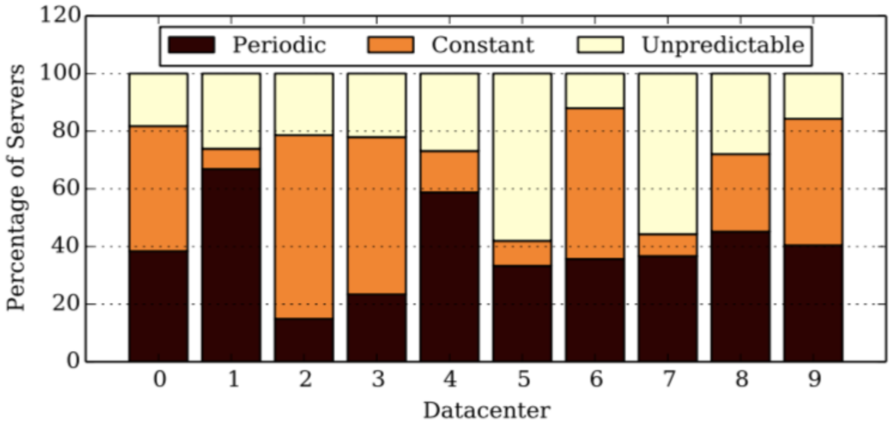


Figure 3: Percentages of servers per class.

Constant and Period Tenant occupies majority of server (75% average)

History-Based Harvesting of Spare Cycles and Storage in Large-Scale Datacenters (2013)

Smart Task Scheduling

- Motivation / Main contribution
- History data observation
- Algorithm
- Experienment

History-Based Harvesting of Spare Cycles and Storage in Large-Scale Datacenters (2013)

Objective

Priority / Type	1st	2nd	3rd
Long job	Constant	Periodic	Unpredictable

For Long job, want to prioritize Long Contant jobs

History-Based Harvesting of Spare Cycles and Storage in Large-Scale Datacenters (2013)

Objective

Type \ Priority	1st	2nd	3rd
Long job	Constant	Periodic	Unpredictable
Short job	Unpredictable	Periodic	Constant

History-Based Harvesting of Spare Cycles and Storage in Large-Scale Datacenters (2013)

Task scheduling

1. Weights classes for resource assignment
2. Resource requirement by DAG

History-Based Harvesting of Spare Cycles and Storage in Large-Scale Datacenters (2013)

Weighted setting

Priority / Type	1st	2nd	3rd
Long job	constant	periodic	unpredictable
Short job	unpredictable	periodic	constant
Medium job	periodic	constant	unpredictable

History-Based Harvesting of Spare Cycles and Storage in Large-Scale Datacenters (2013)

Job resource requirement

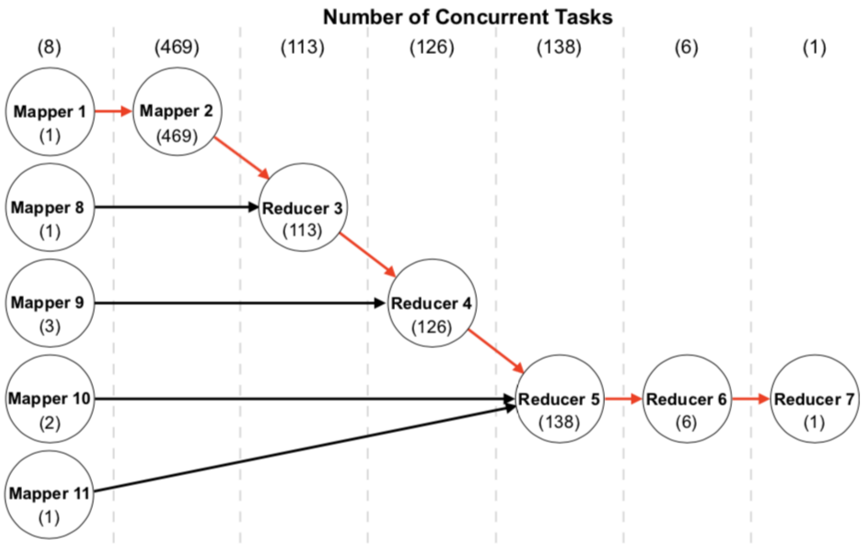


Figure 7: Example job execution DAG.

Most "crowdy" moment as Job Resource Requirement

Smart Task Scheduling

- Motivation / Main contribution
- History data observation
- Algorithm
- Experienment

History-Based Harvesting of Spare Cycles and Storage in Large-Scale Datacenters (2013)

Primary Tenant Latency Comparation

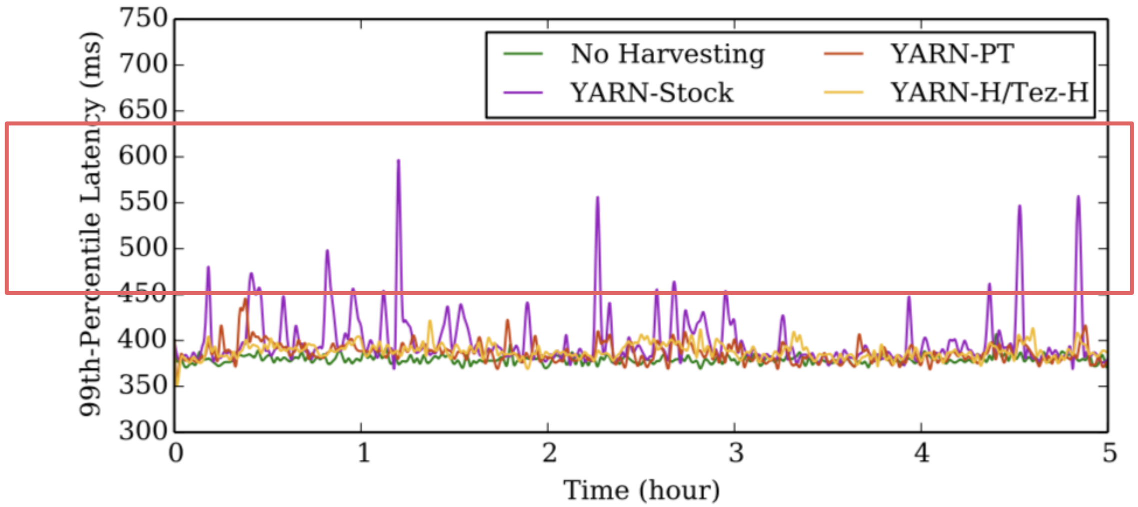
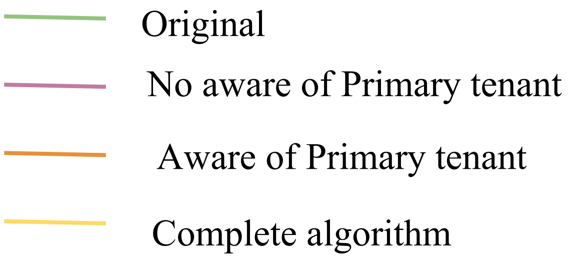
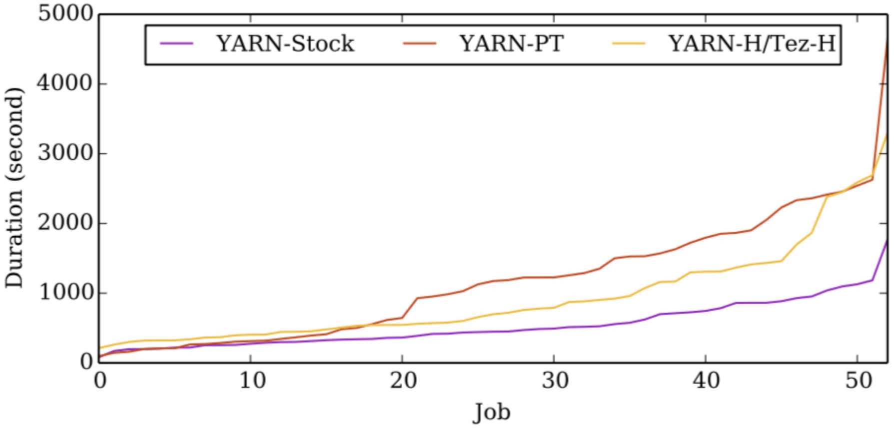


Figure 10: Primary tenant's tail latency in the real testbed for versions of YARN and Tez.



History-Based Harvesting of Spare Cycles and Storage in Large-Scale Datacenters (2013)

Secondary Tenant effect against Primary Tenant



Aware of Primary tenant

Complete algorithm

Not Aware of Primary tenant

Figure 11: Secondary tenants' run times in the real testbed for versions of YARN and Tez.

History-Based Harvesting of Spare Cycles and Storage in Large-Scale Datacenters (2013)

History-Based Harvesting of Spare Cycles & Storage in Large-Scale Datacenters

- Smart Task Scheduling
- Smart Data Placement

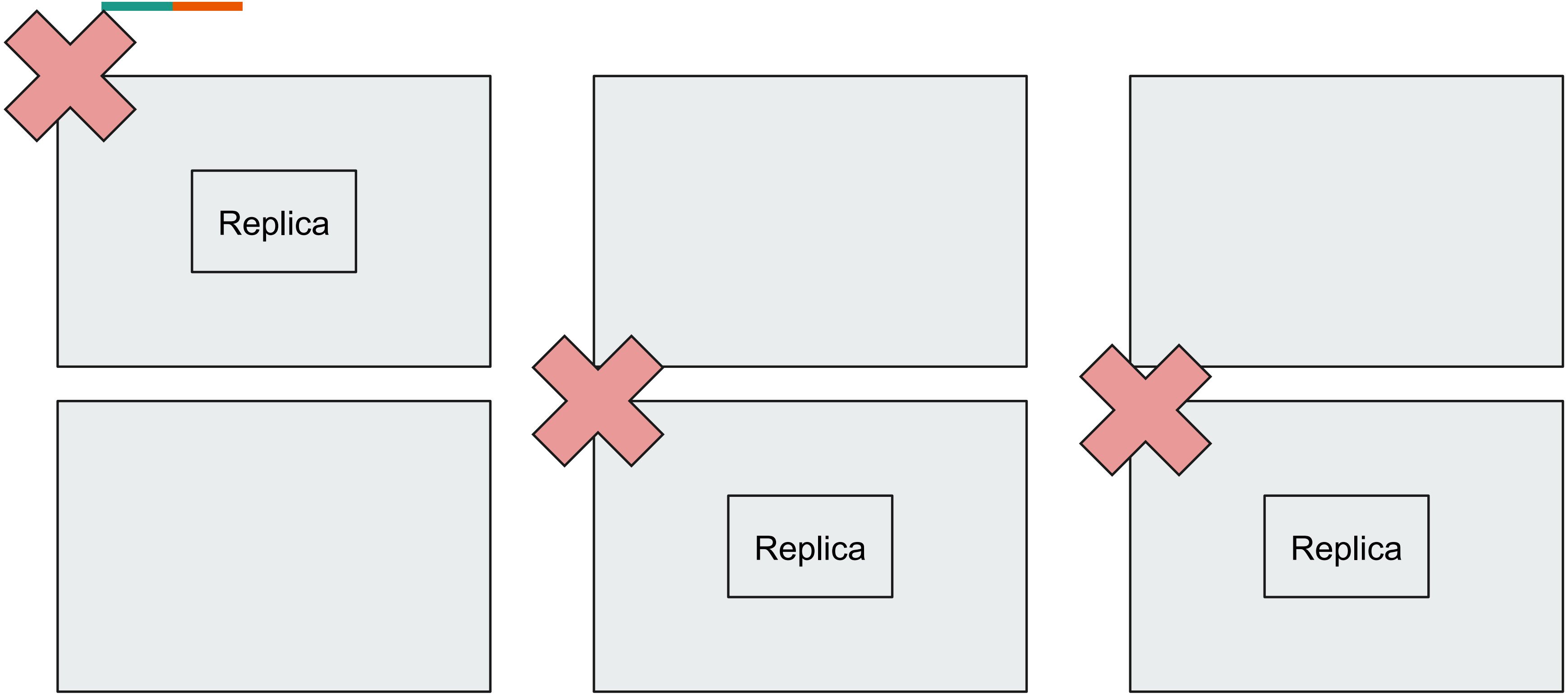
History-Based Harvesting of Spare Cycles and Storage in Large-Scale Datacenters (2013)

Smart Data Placement

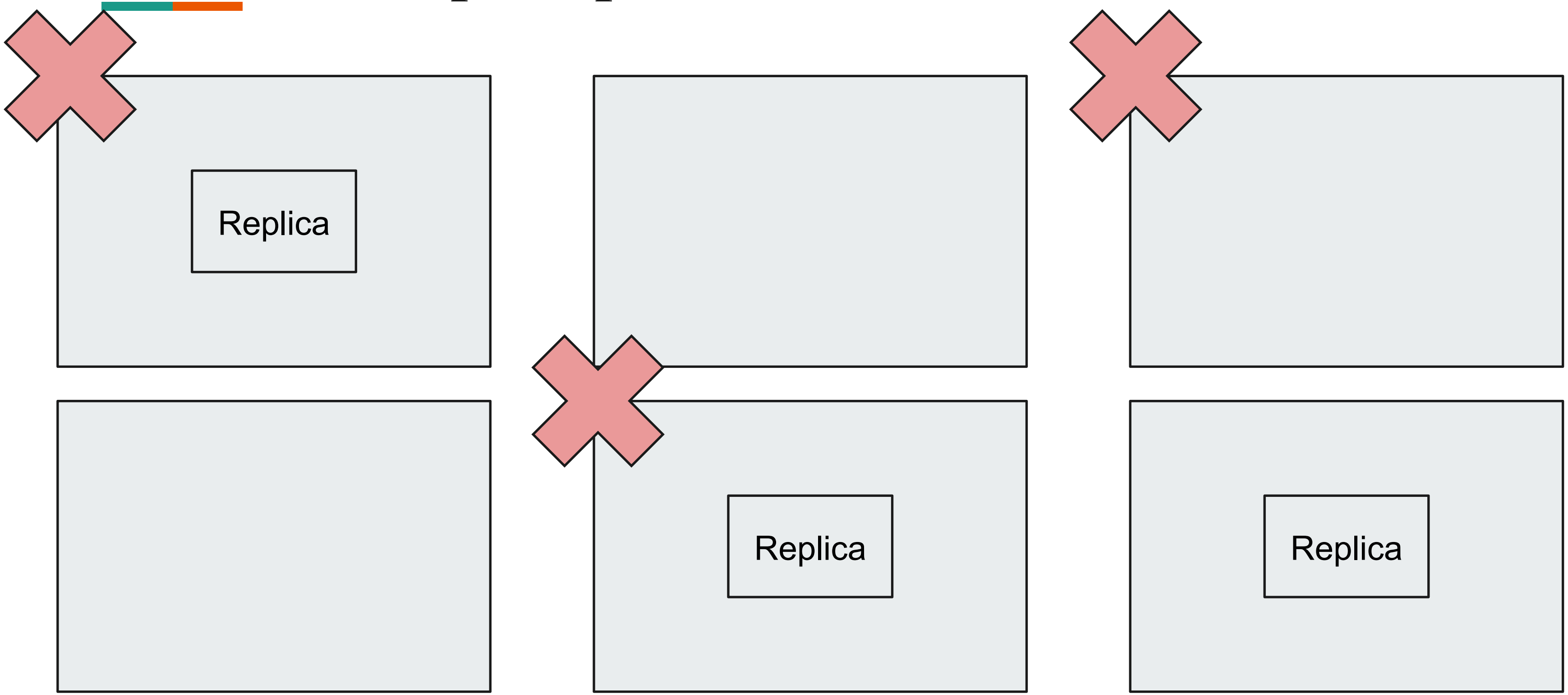
- Motivation / Main contribution
- History data observation
- Algorithm
- Experienement

History-Based Harvesting of Spare Cycles and Storage in Large-Scale Datacenters (2013)

Re-Image - Durability/Busy server - unavailable data



Ideal Data replica placement



Greedy Solution

Re-Image the disks the least / have lowest CPU utilizations.

Consistant Performance!!!

History-Based Harvesting of Spare Cycles and Storage in Large-Scale Datacenters (2013)

Smart Data Placement

- Motivation / Main contribution
- History data observation
- Algorithm
- Experienenment

History-Based Harvesting of Spare Cycles and Storage in Large-Scale Datacenters (2013)

ReImages frequency CDF

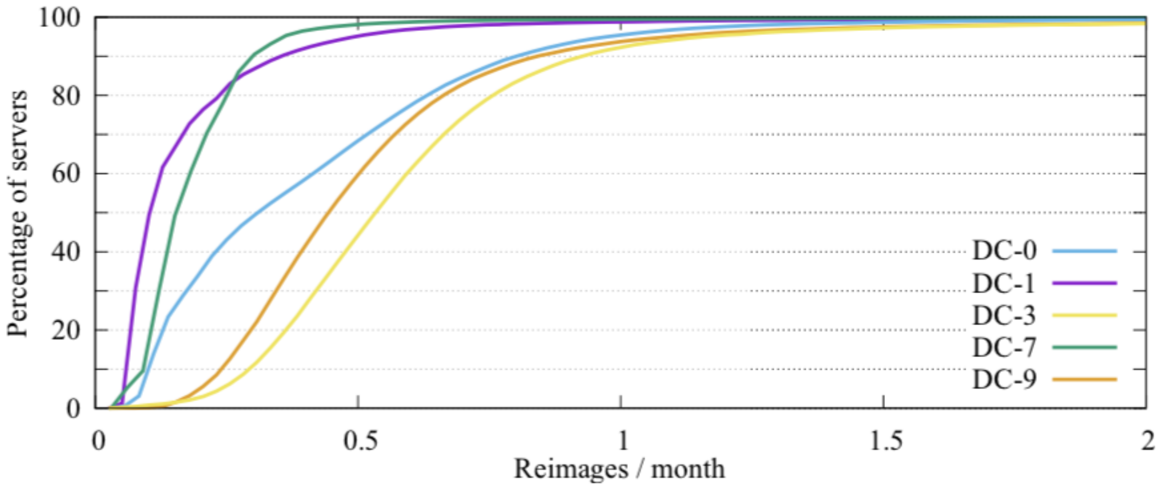


Figure 4: Per-server number of reimages in three years.

Diversity in reimage frequency

History-Based Harvesting of Spare Cycles and Storage in Large-Scale Datacenters (2013)

Change of ReImages frequency Group CDF

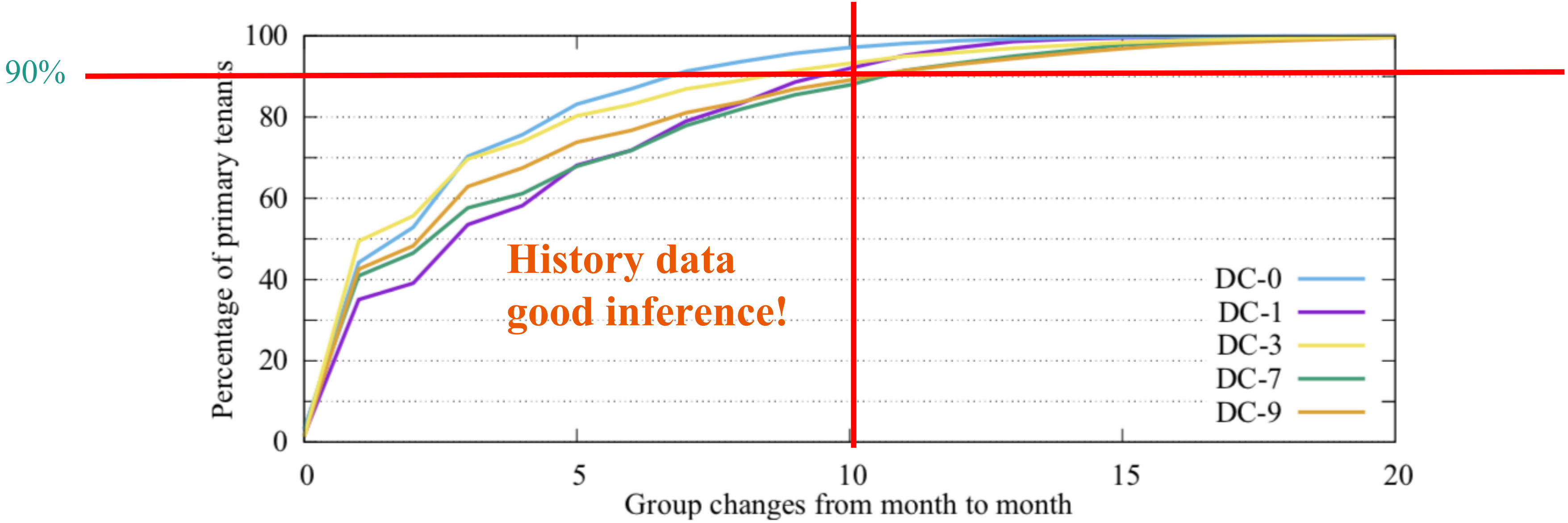


Figure 6: Number of times a primary tenant changed reimage frequency groups in three years.

History-Based Harvesting of Spare Cycles and Storage in Large-Scale Datacenters (2013)

- **Diversity in reimage frequency**
- **History data can be good inference!**

History-Based Harvesting of Spare Cycles and Storage in Large-Scale Datacenters (2013)

Smart Data Placement

- Motivation / Main contribution
- History data observation
- Algorithm
- Experienement

History-Based Harvesting of Spare Cycles and Storage in Large-Scale Datacenters (2013)

Take advantage of diversity!

utilization / Reimage	High	Medium	Low
Frequent			
Intermediate			
Infrequent			

smart data placement

Reimage \ utilization	High	Medium	Low
Frequent	☺		
Intermediate			
Infrequent			

Take advantage of diversity!

Reimage \ utilization	High	Medium	Low
Frequent	☺		
Intermediate			☺
Infrequent		☺	

Classification Example

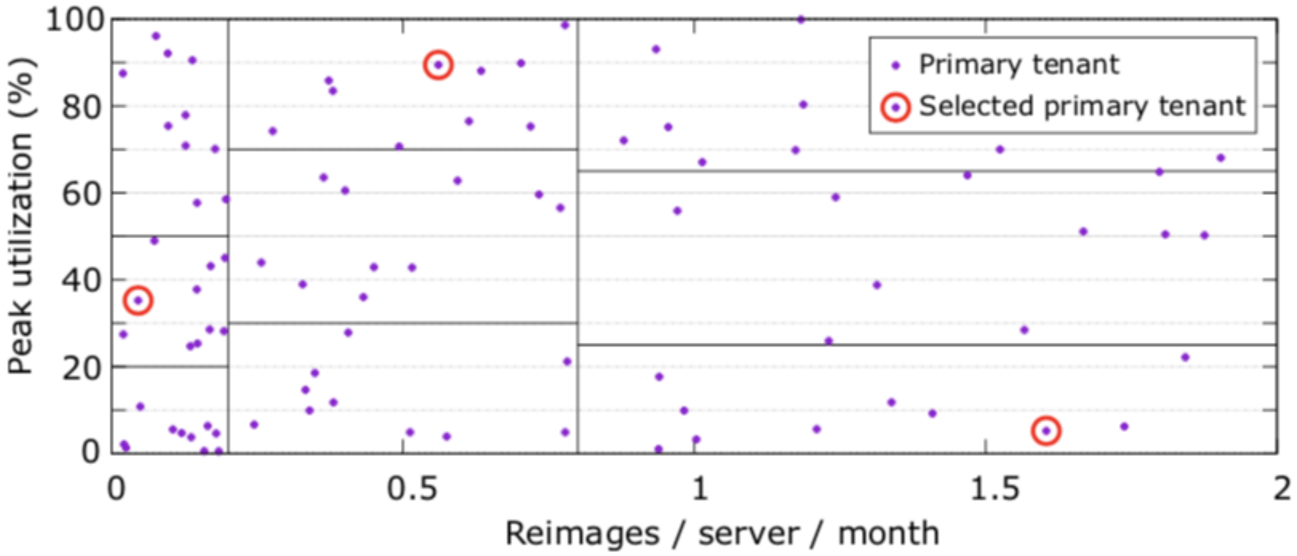


Figure 8: Two-dimensional clustering scheme.

Smart Data Placement

- Motivation / Main contribution
- History data observation
- Algorithm
- Experienment

History-Based Harvesting of Spare Cycles and Storage in Large-Scale Datacenters (2013)

Primary Tenant Latency Comparison

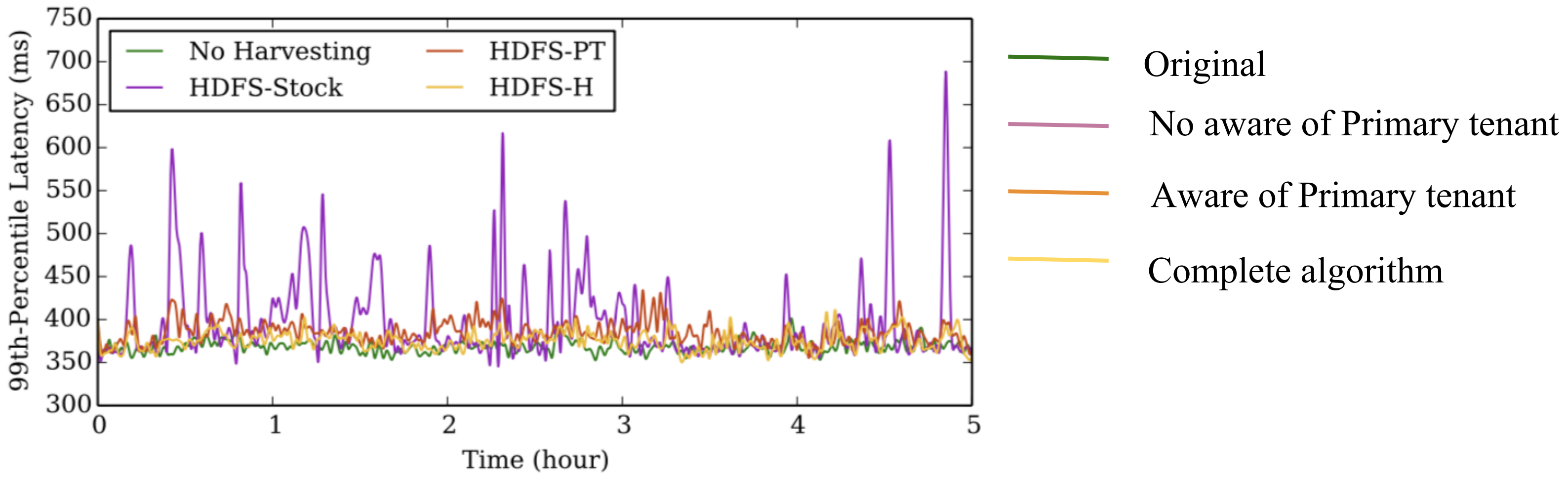


Figure 12: Primary tenant's tail latency in the real testbed for versions of HDFS.

History-Based Harvesting of Spare Cycles and Storage in Large-Scale Datacenters (2013)

Lost block Comparison

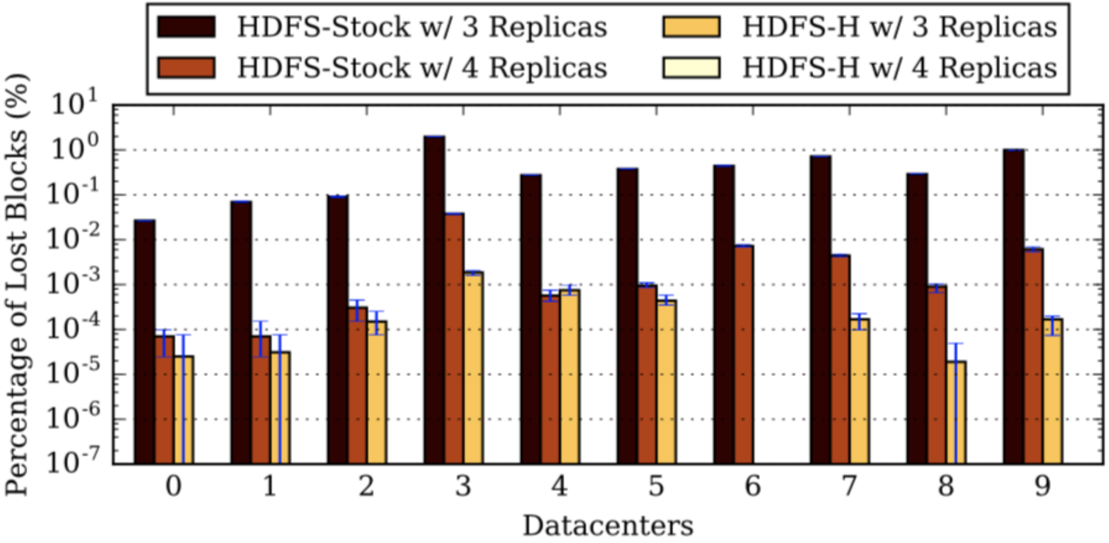
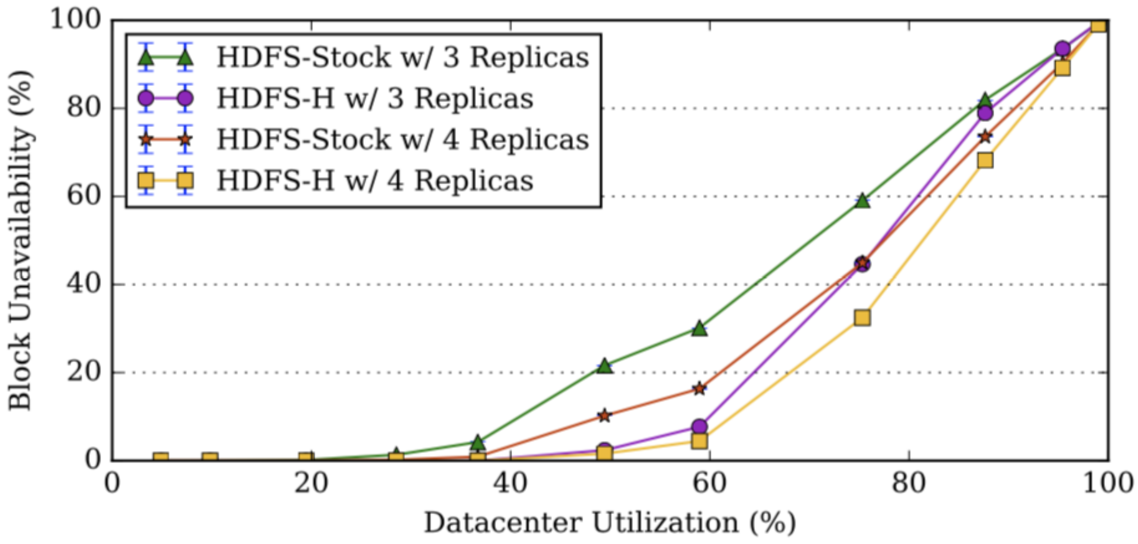


Figure 15: Lost blocks for two replication levels.

Original w/ 3 replicas
Original w/ 4 replicas
Algorithm w/3 replicas
Algorithm w/4 replicas

History-Based Harvesting of Spare Cycles and Storage in Large-Scale Datacenters (2013)

Failed Access Comparison



Original w/ 3 replicas

Original w/ 4 replicas

Algorithm w/3 replicas

Algorithm w/4 replicas

Figure 16: Failed accesses under linear scaling.

Main contribution Summary

- Interesting View of Historical Data
- Primary / Secondary Tenant Perspective
- Data Placement Algorithm
- Historical Data as Good Measure of Future

History-Based Harvesting of Spare Cycles and Storage in Large-Scale Datacenters (2013)

Dynamic Resource Allocation Using Virtual Machines for Cloud Computing Environment

IEEE Transactions on Parallel and Distributed Systems Volume 24 Issue 6, June 2013

Author	Affiliation
Zhen Xiao	Peking University, Beijing
Weijia Song	Peking University, Beijing
Qi Chen	Peking University, Beijing

Dynamic Resource Allocation

- Main Goal: Effective Load Prediction

A good Framework to achieve:

- Prevent Overload (Under-Provisioning)
- Green Computing when Possible (Over-Provisioning)

Dynamic Resource Allocation Using Virtual Machines for Cloud Computing Environment (2012)

Cloud Manager Framework - Skewness

Skewness (a measure of balance)

→ Minimize skewness

→ Balance resource utilization

$$skewness(p) = \sqrt{\Sigma_{i=1}^{n}(r_i/\bar{r} - 1)}$$

- $\bar{r}$: average utilization of all resources
- r_i: utilization of resource i

Dynamic Resource Allocation Using Virtual Machines for Cloud Computing Environment (2012)

Cloud Manager Framework - Migration

Trade-Off: Load Balancing ←→ Green Computing

Migration:

- 🔥 Hot Migration for HotSpot (too High utilization)
- ☃️ Cold Migrartion for ColdSpot (too Low utilization)

Dynamic Resource Allocation Using Virtual Machines for Cloud Computing Environment (2012)

Cloud Manager Framework

✅ Skewnesss

✅ Hotspot, Coldspot

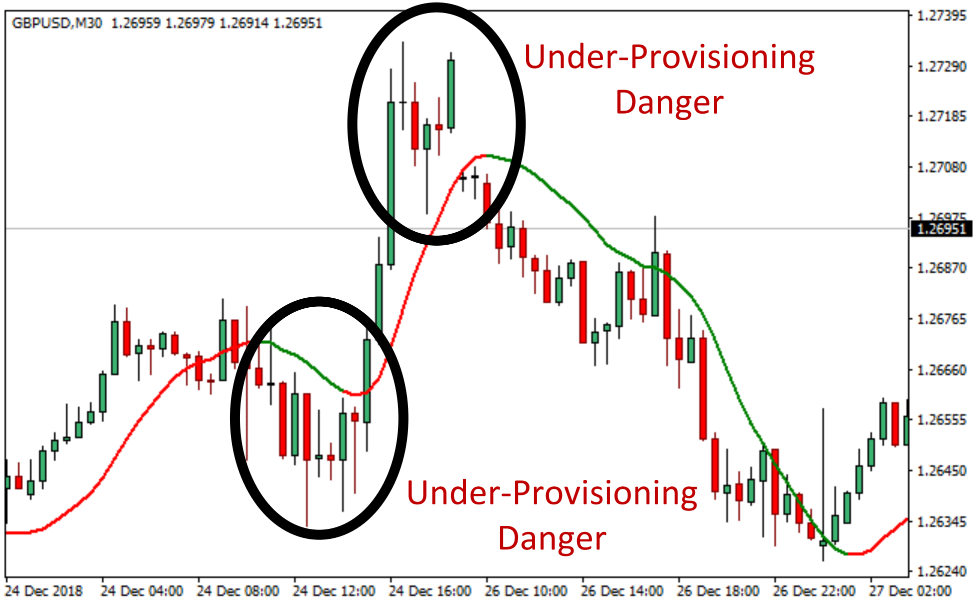
➡️ Load Prediction

Dynamic Resource Allocation Using Virtual Machines for Cloud Computing Environment (2012)

Load Prediction - Original

No High Error!

Original: Exponential (Weighted) Moving Average



- Slower trend
- Conservative trend

Dynamic Resource Allocation Using Virtual Machines for Cloud Computing Environment (2012)

Load Prediction - FUSD

Improved: Fast Up Slow Down Algorithm

- Aggressive estimation when rising trend
- Conservative estimation when descending trend

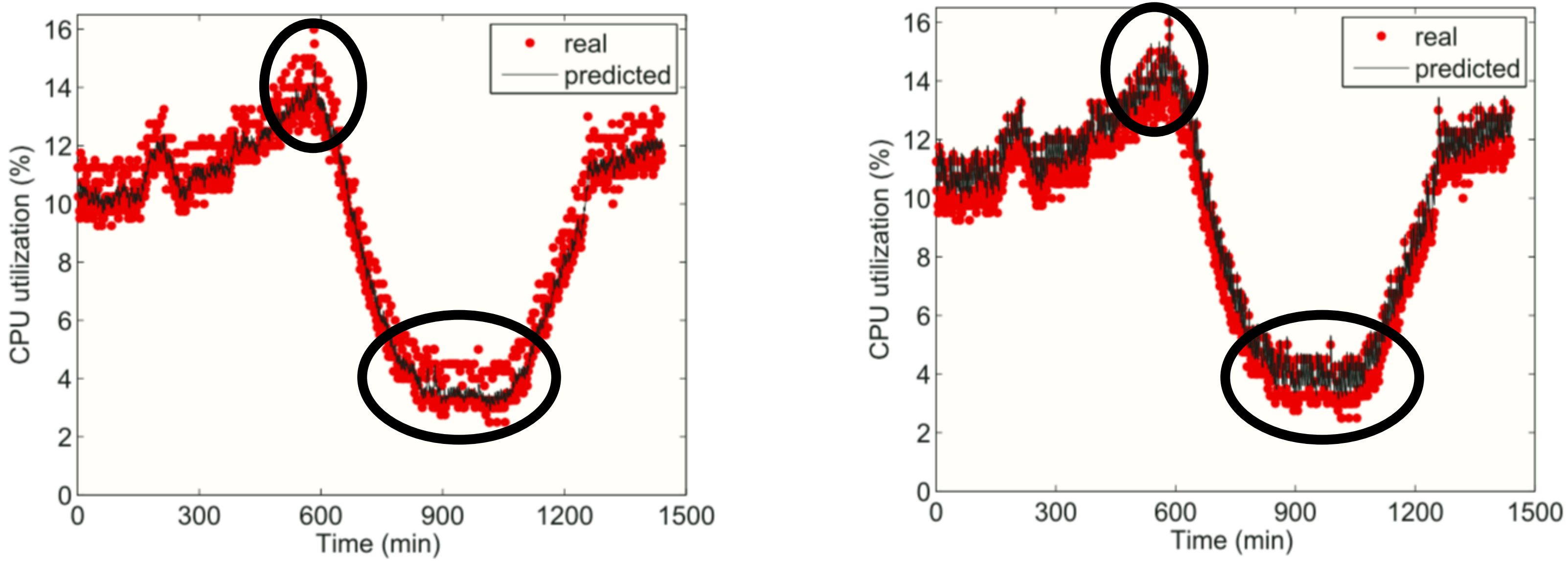
$$E(t) = -|\alpha| \times E(t-1) + (1+|\alpha|) \times O(t), -1 \leq \alpha \leq 0$$

- $\uparrow \alpha$ for increasing trend
- $\downarrow \alpha$ for descending trend

根據（上升/下降） trend 而有不同參數

Dynamic Resource Allocation Using Virtual Machines for Cloud Computing Environment (2012)

Load Prediction - FUSD

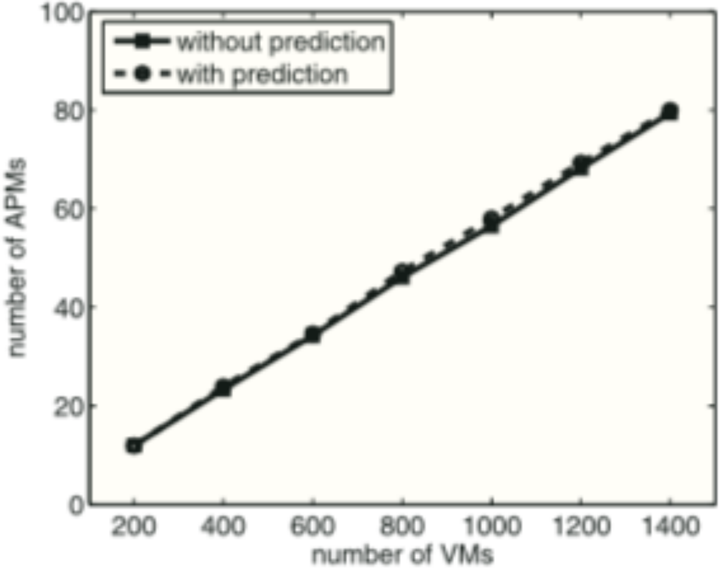


(a) EWMA: $\alpha = 0.7, W = 1$

(b) FUSD: $\uparrow \alpha = -0.2, \downarrow \alpha = 0.7, W = 1$

Dynamic Resource Allocation Using Virtual Machines for Cloud Computing Environment (2012)

Dynamic Resource Allocation - Results



(b) number of APMs

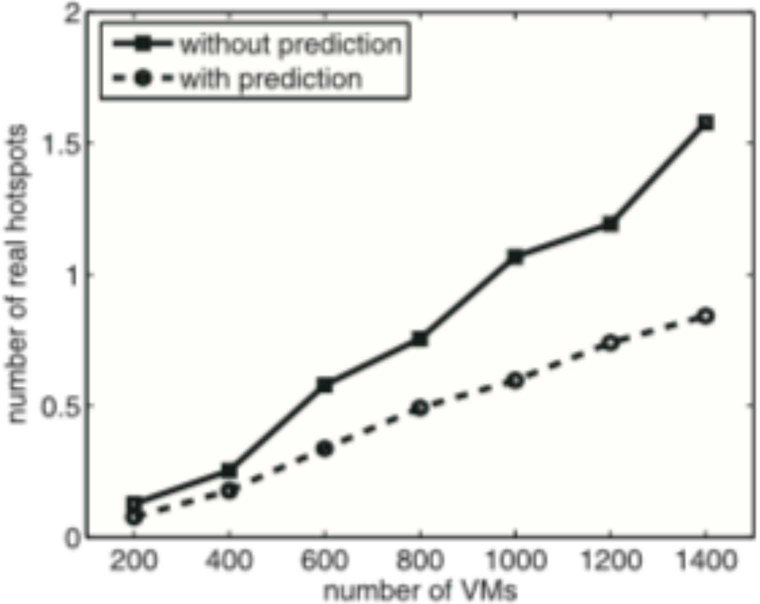
(APM = Active Physical Machine)

With vs. Without Prediction

Results are almost the same!

Dynamic Resource Allocation Using Virtual Machines for Cloud Computing Environment (2012)

Dynamic Resource Allocation - Results

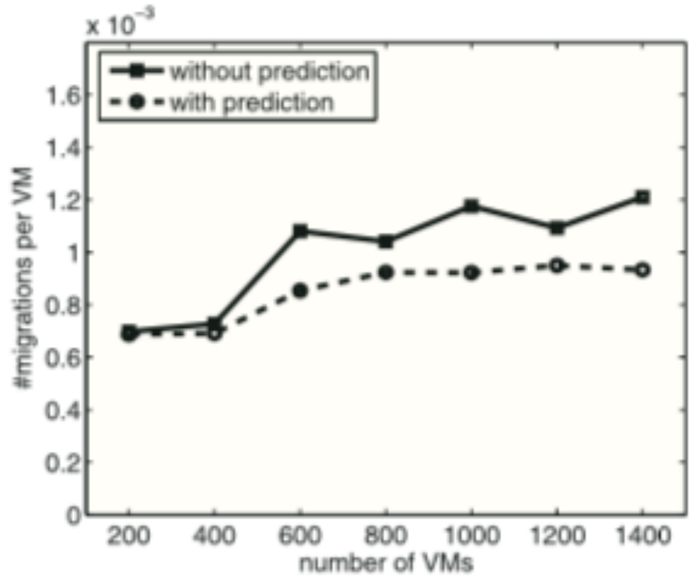


(a) number of hot spots

Notable decrease in hotspot

Dynamic Resource Allocation Using Virtual Machines for Cloud Computing Environment (2012)

Dynamic Resource Allocation - Results



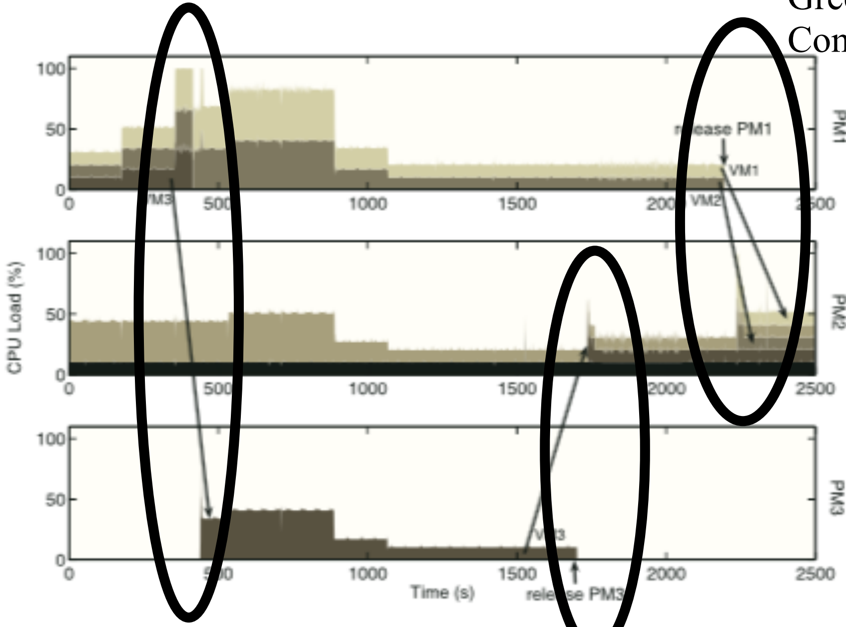
(c) number of migration

Notable decrease in migrations

Dynamic Resource Allocation Using Virtual Machines for Cloud Computing Environment (2012)

Dynamic Resource Allocation - Results

Green
Computing



Load
Balancing

Green
Computing

Demonstration on Migration

Dynamic Resource Allocation Using Virtual Machines for Cloud Computing Environment (2012)

Dynamic Resource Allocation

- Main Goal: Effective load prediction

- Contribution:

    - ✅ Cloud Manager Framework
    - ✅ Fast Up Slow Down Algorithm (FUSD)
    - ✅ Dynamic Green Computing

Dynamic Resource Allocation Using Virtual Machines for Cloud Computing Environment (2012)

Comparison

Different aspects of utilization:

- Static utility optimization via Primary/Secondary tenant
    - Better allocating with concept of Secondary Tenant
    - Good Initialization
- Dynamic utility optimization via Minimizing Skewness
    - Live Migration
    - Support Green Computing

Idea

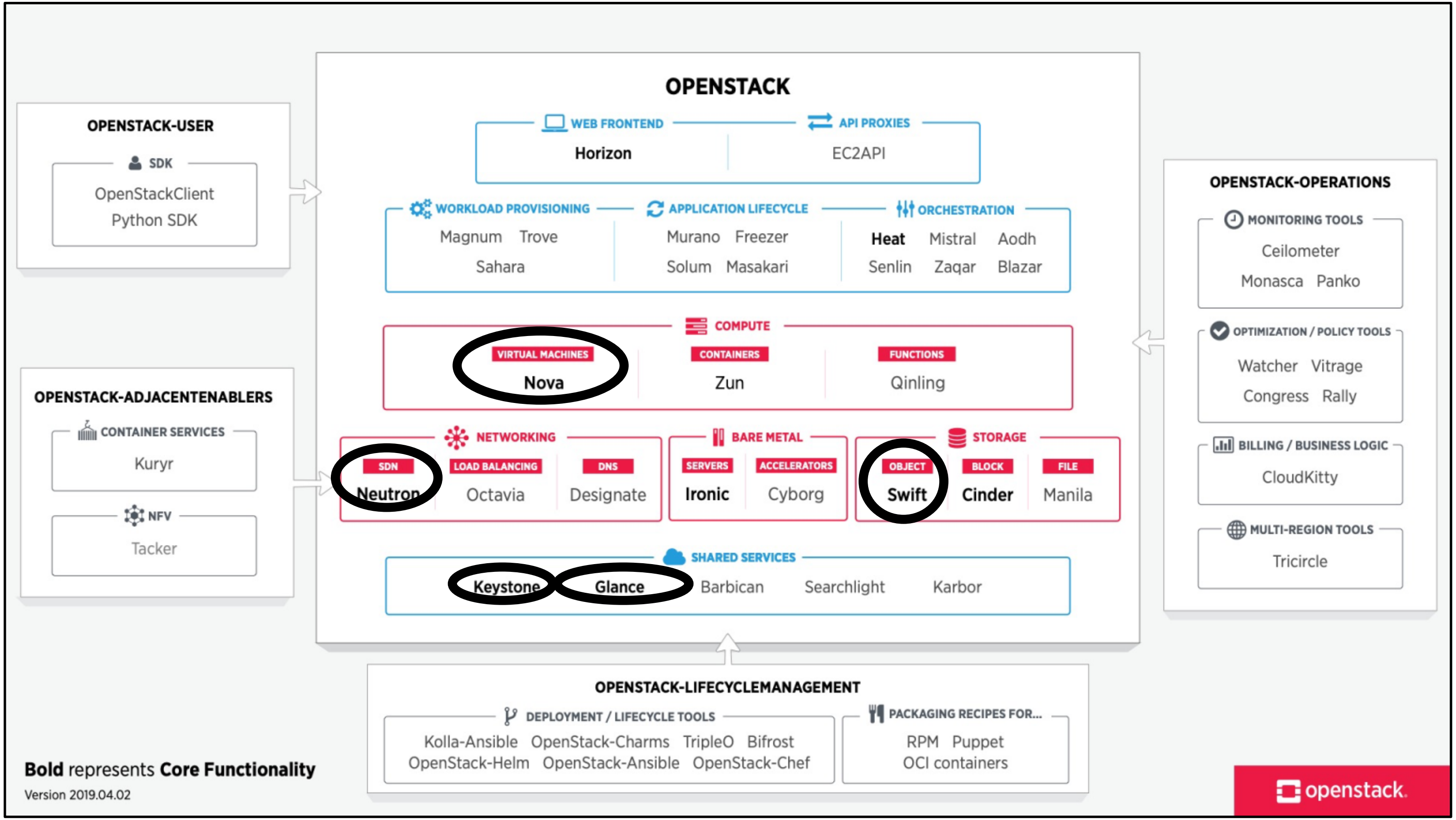
- Dynamic allocation: Better prediction via ML
- Static vs. Dynamic?
- Static + Dynamic

Cloud Framework - OpenStack

- Framework for Private Cloud
- Infrastructure as a Service (IaaS)
- Modulized component

openstack®

Open Source Cloud Technology (2012)


OPENSTACK
WEB FRONTEND
Horizon
API PROXIES
EC2API
WORKLOAD PROVISIONING
Magnum Trove
Sahara
APPLICATION LIFECYCLE
Murano Freezer
Solum Masakari
ORCHESTRATION
Heat Mistral Aodh
Senlin Zaqar Blazar
COMPUTE
VIRTUAL MACHINES
Nova
CONTAINERS
Zun
FUNCTIONS
Qinling
NETWORKING
SDN
Neutron
LOAD BALANCING
Octavia
DNS
Designate
BARE METAL
SERVERS
Ironic
ACCELERATORS
Cyborg
STORAGE
OBJECT
Swift
BLOCK
Cinder
FILE
Manila
SHARED SERVICES
Keystone Glance Barbican Searchlight Karbor
OPENSTACK-USER
SDK
OpenStackClient
Python SDK
OPENSTACK-ADJACENTENABLERS
CONTAINER SERVICES
Kuryr
NFV
Tacker
OPENSTACK-OPERATIONS
MONITORING TOOLS
Ceilometer
Monasca Panko
OPTIMIZATION / POLICY TOOLS
Watcher Vitrage
Congress Rally
BILLING / BUSINESS LOGIC
CloudKitty
MULTI-REGION TOOLS
Tricircle
OPENSTACK-LIFECYCLEMANAGEMENT
DEPLOYMENT / LIFECYCLE TOOLS
Kolla-Ansible OpenStack-Charms TripleO Bifrost
OpenStack-Helm OpenStack-Ansible OpenStack-Chef
PACKAGING RECIPES FOR...
RPM Puppet
OCI containers
Bold represents Core Functionality
Version 2019.04.02
openstack

OpenStack - Fast Deployment

- ➡️ Image Management (Glance)
- ➡️ Network Management (Neutron)
- ➡️ VM Management (Nova)
- ➡️ Storage Management (Swift)
- ➡️ Identity Authentication (KeyStone)

OpenStack - MAAS (Metal As A Service)

- Image Management
- Network Management
- Basic Configurations (user, bashrc, ssh…)

➡️ Fast Deployment

MAAS