Characterizing and Orchestrating VM Reservation in Geo-distributed Clouds to Improve the Resource Efficiency

Jiuchen Shi, Kaihua Fu, Quan Chen, Changpeng Yang, Pengfei Huang, Mosong Zhou, Jieru Zhao, Chen Chen, Minyi Guo

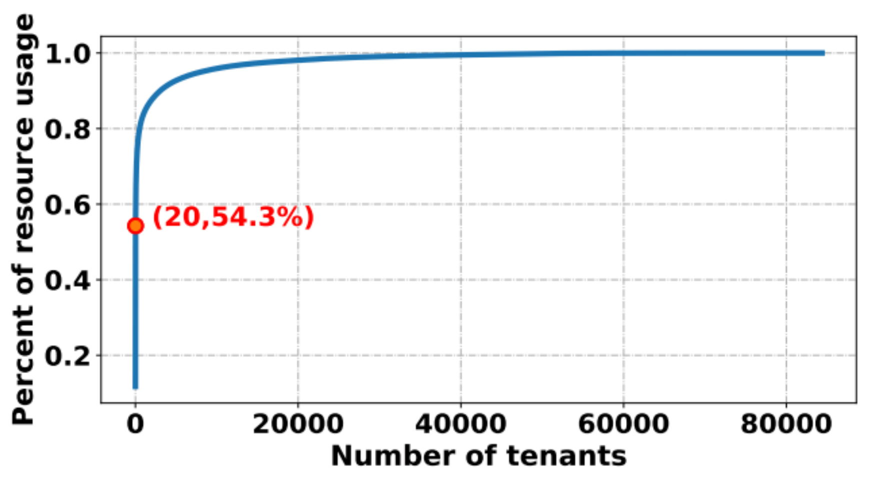
Shanghai Jiao Tong University, Huawei Cloud

Content

Introduction and Background

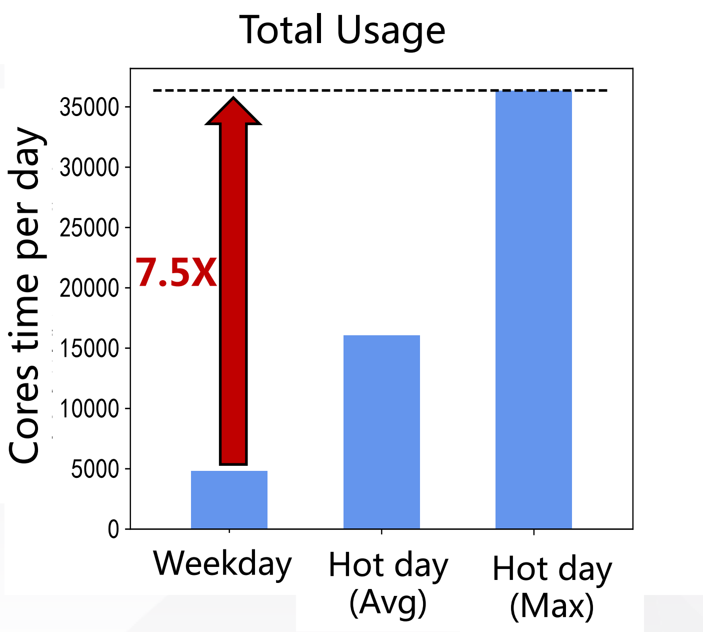
Large Tenants and Resource Reservation

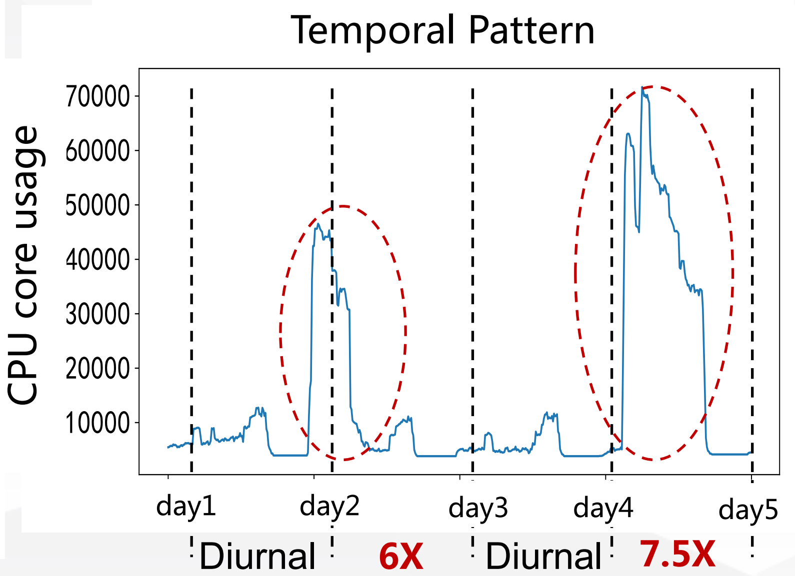
4 month data, about **85k** tenants
top **20** dominates **54.5%** of resource
Obvious **head effect**



Aggregated resource usage

Bursty Resource usage can be **7.5X**
Unplanned load burst
Service of **Reserved VMs** (AWS, Azure, Google, Huawei)

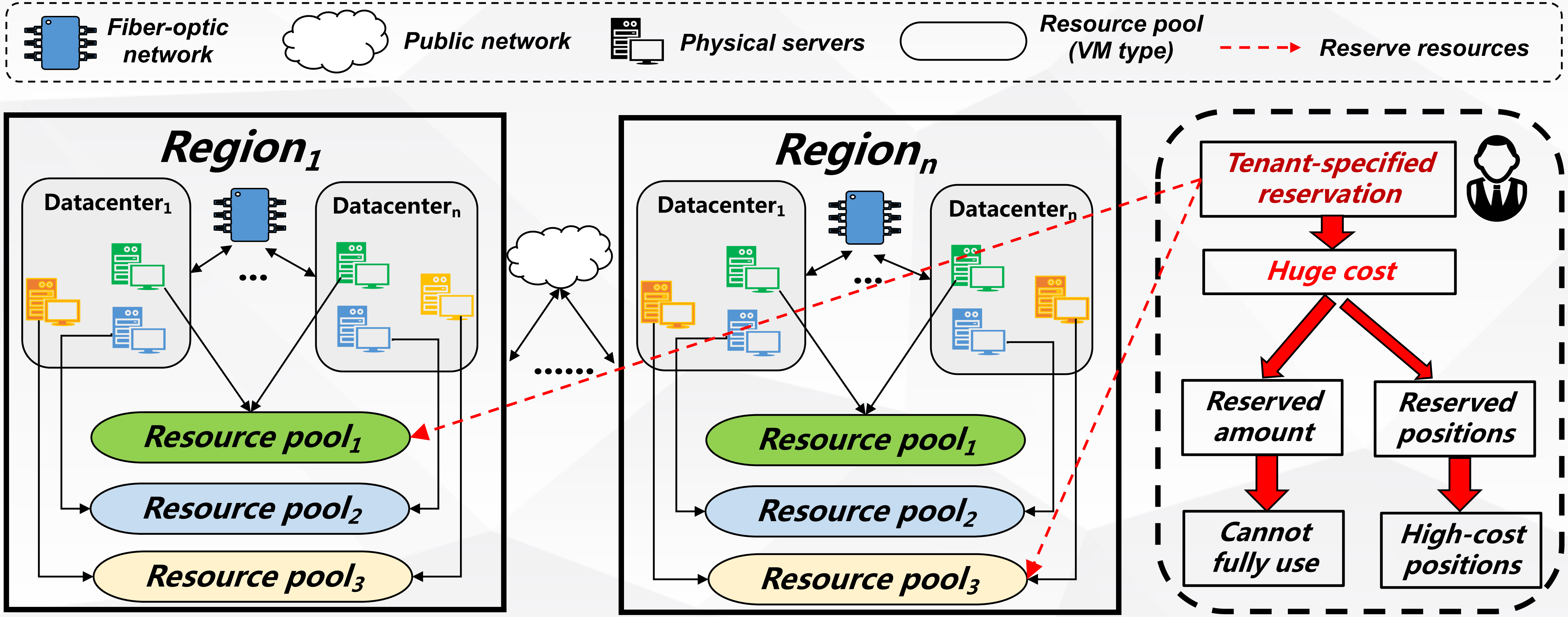




An example top tenant in Huawei Cloud

Geo-distributed Clouds and Reservation

Cloud Regions, Datacenter, Resource Pools, VM types, Tenant-specified reservation



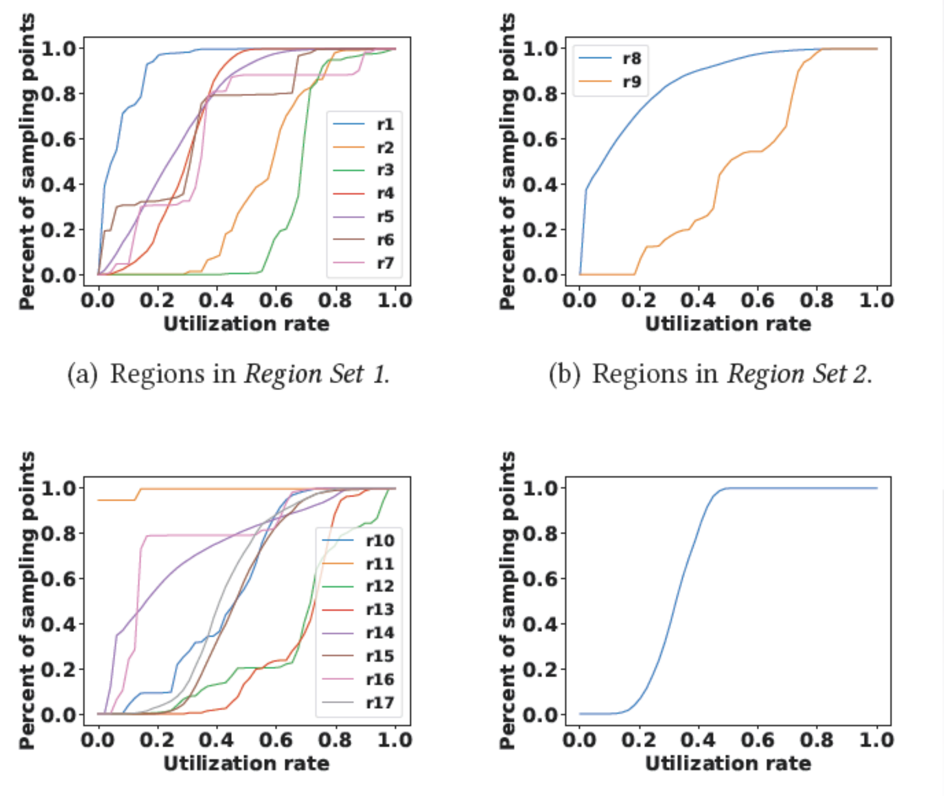
- **Geo-distributed Trace:** 2021, 4 month, top 20 tenants, 17 regions, sample every 10 mins

Resource Reservation Status

Utilization of Reserved Resources

Tenant-specified resource reservation

Utilization rate mainly distributed in 20%-50%, regions differ from each other



(a) Regions in *Region Set 1*.

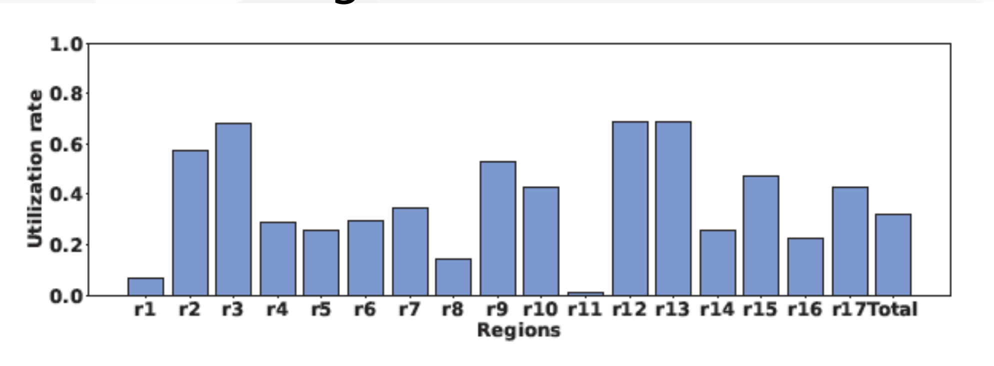
(b) Regions in *Region Set 2*.

(c) Regions in *Region Set 3*.

(d) Overall resource utilization rate.

Distribution of resource reservation rate

Total Avg. utilization rate = 32.3%



Average resource utilization rate

Note:

Region Set 1 : r1-r7

Region Set 2 : r8-r9

Region Set 3 : r10-r17

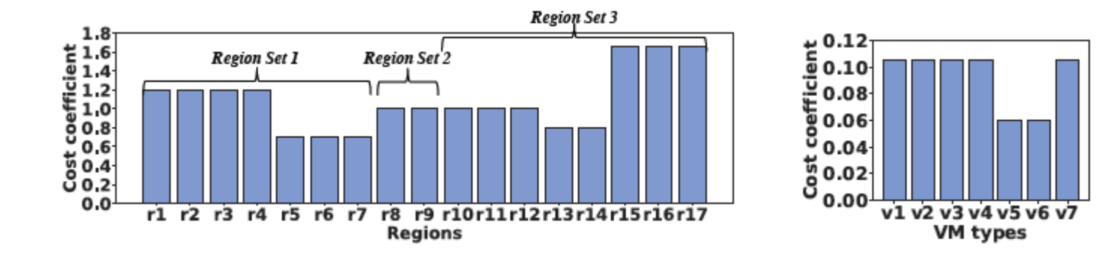
"Tenant-specified reservation" causes huge resource waste

Resource Usage Distribution

17 regions, 7 VM types, different cost coefficients

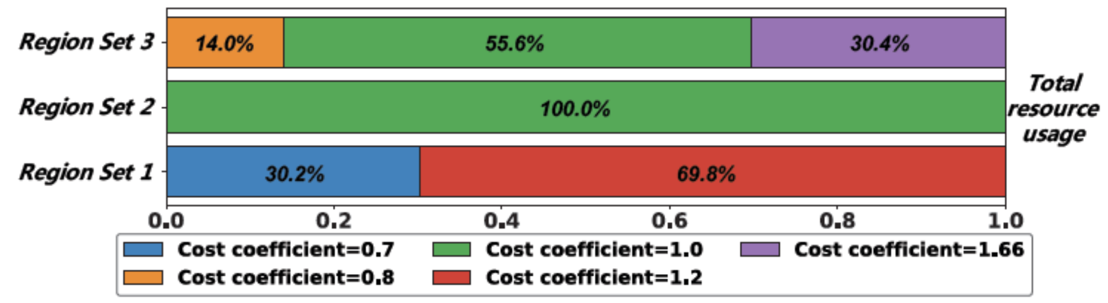
3 region set, regions in same *Region set* are Geographically Close, 7 VM types can exchange

Regions: Only 30.2% and 14.0% of resources distributed in low-cost.
VM types: Only 19.6% of resource distributed in low-cost.

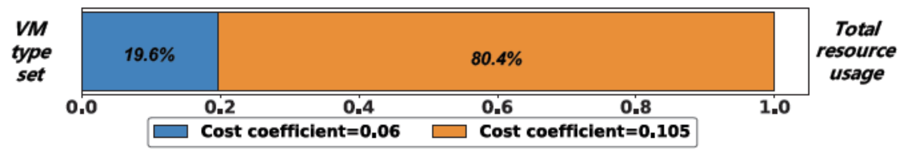


(a) Regions. (b) VM types.

Cost coefficients of regions and VM types



Resource usage distribution of regions



Resource usage distribution of VM types

"Tenant-specified reservation" not tends to low-cost region/VM type

Main Takeaways

What can we learn from current reservation status?

- Low utilization rate, **on-demand reservation** can reduce waste;
- Not aware of region/VM type costs, potential to lead **to low cost**;

- Define tenant "**Acceptable Spatial Ranges**" to ensure SLO for tenants
- Acceptable Spatial Ranges:

  reserved resource can **exchange** within some regions/VM types;

  decide **positions** and **resource amount** in corresponding ranges.

"On-demand + low-cost reservation" needs to explore "temporal and spatial patterns" of large tenants

SJTU

Temporal and Spatial Patterns

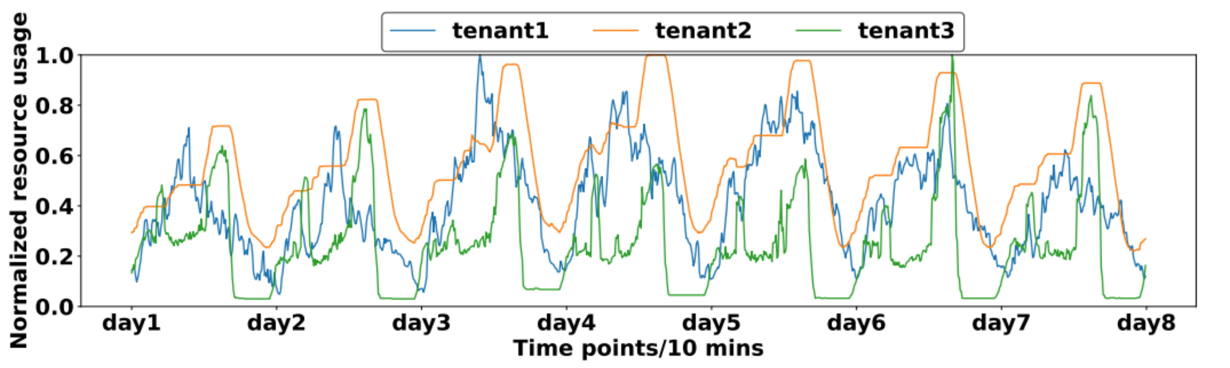
Temporal Usage Patterns

Diurnal usage pattern:

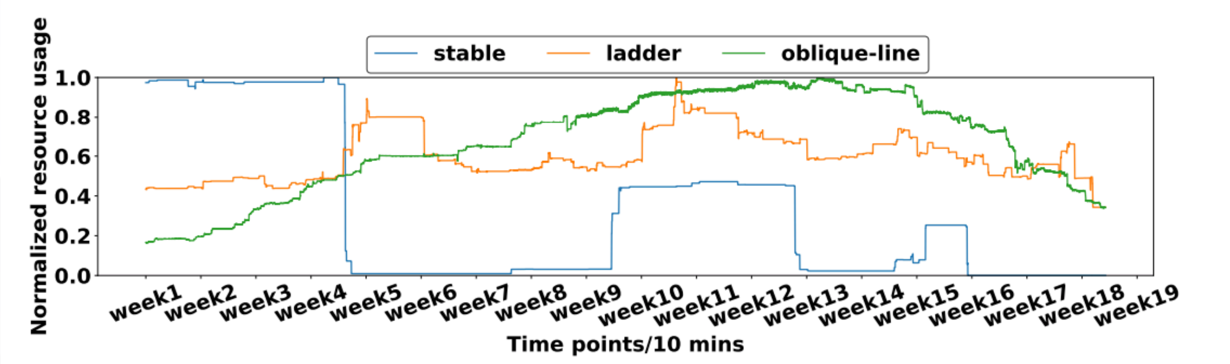
day and night mode in short and long term
Time series prediction methods: LSTM/ARIMA

Persistent usage pattern:

stable usage in pattern in long term, (1) stable;(2)ladder;(3)oblique-line; average values, linear regression analysis methods.



Diurnal pattern



Persistent pattern

Predictable patterns use corresponding prediction methods

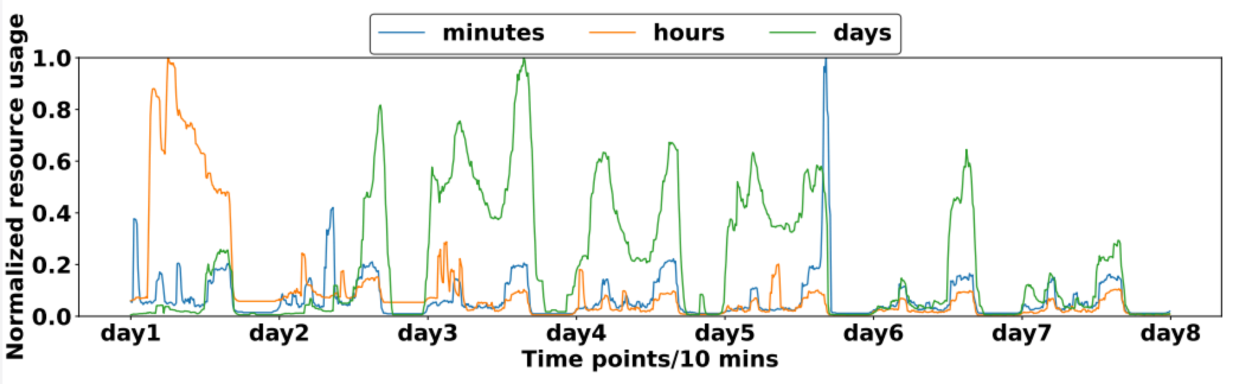
Temporal Usage Patterns

Bursty usage pattern:

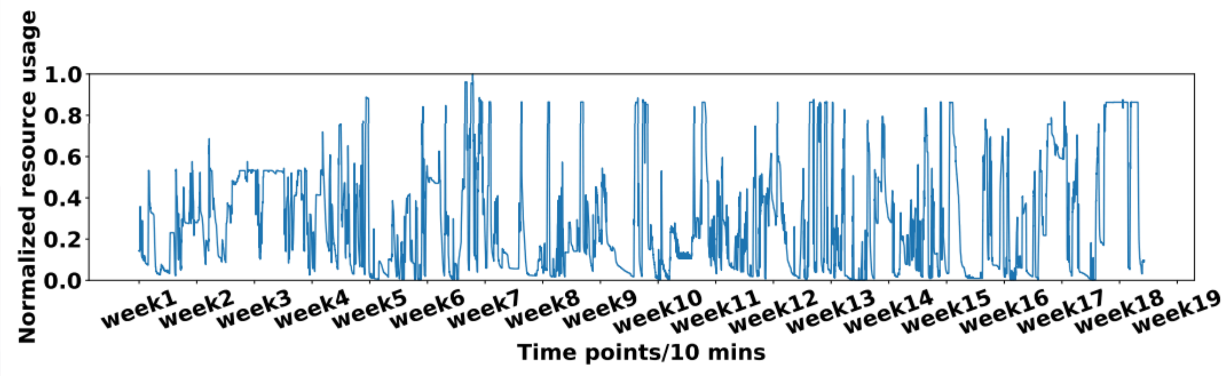
Diurnal pattern in normal time, existing bursty resource usage;
Bursty amount can be 3X-7X; Bursty duration can be minutes, hours, and days;
Online schedule and compensate for bursty time

Irregular usage pattern:

random in short and long term; Unpredictable, online schedule and compensate.



Bursty tenants



Irregular pattern

Unpredictable patterns rely on online schedule and compensate

Spatial Usage Patterns

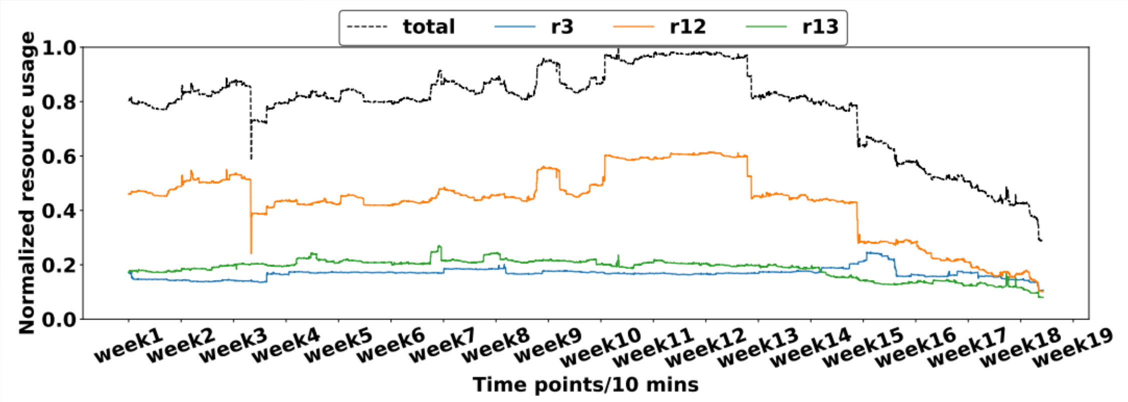
Using a single major region/VM type:

Focus on its temporal patterns for resource usage prediction.

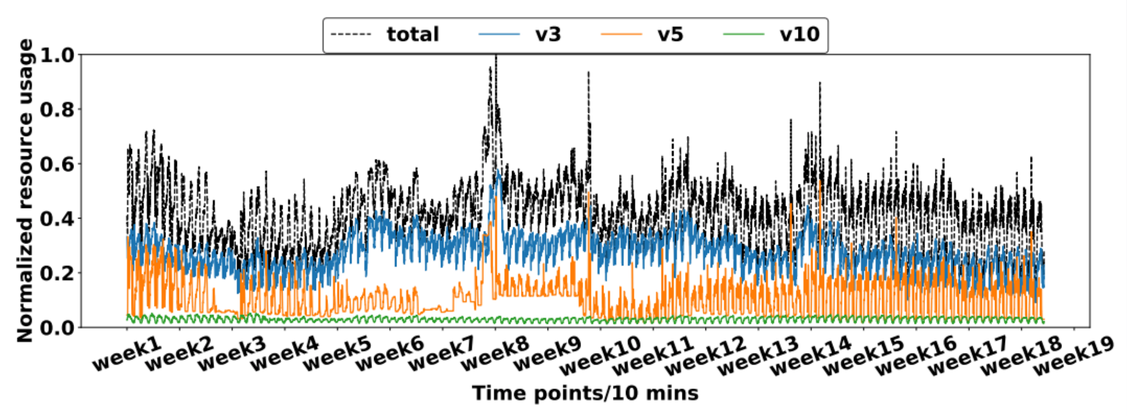
Using multiple regions/VM types with stable usage division:

Stable division and similar patterns between regions/VM types;

"predicted temporal usage × stable division" to predict usage on regions/VM types.



Multiple regions in stable mode



Multiple VM types in stable mode

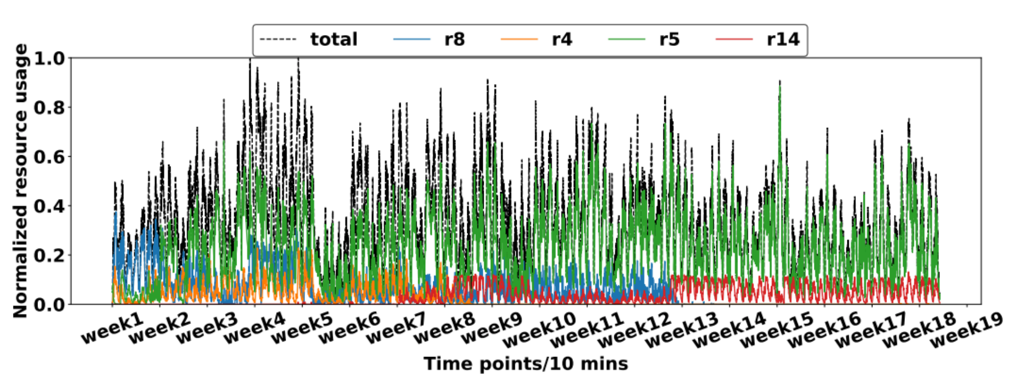
Mainly consider temporal patterns for prediction and reservation

Spatial Usage Patterns

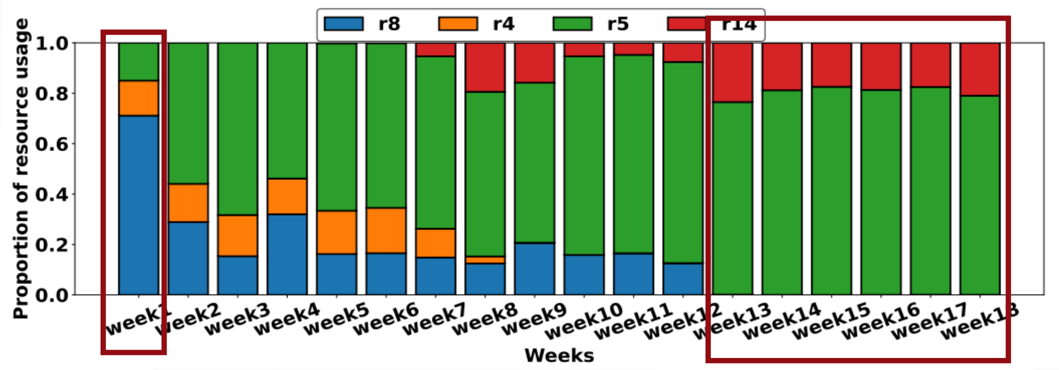
Using multiple regions/VM types with dynamic division:

Dynamic division and different patterns between regions/VM types;

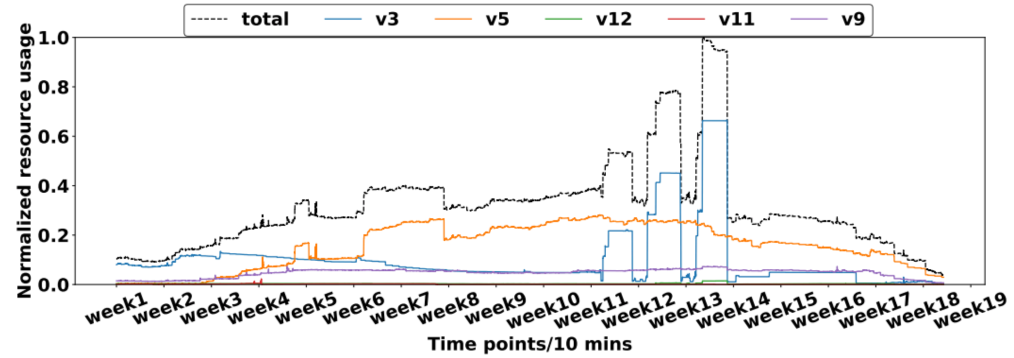
Combine tenants "Acceptable spatial ranges", different from temporal patterns.



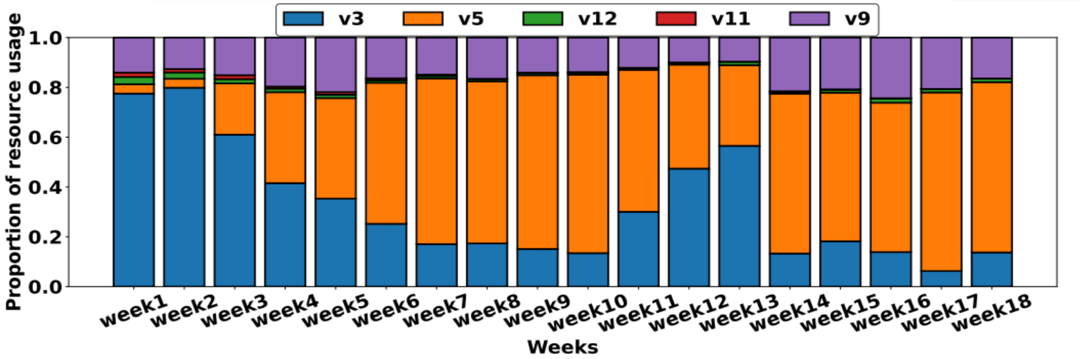
Multiple regions in dynamic mode



Breakdown of dynamic regions



Multiple VM types in dynamic mode



Breakdown of dynamic VM types

Prediction and reservation on tenants "Acceptable spatial ranges"

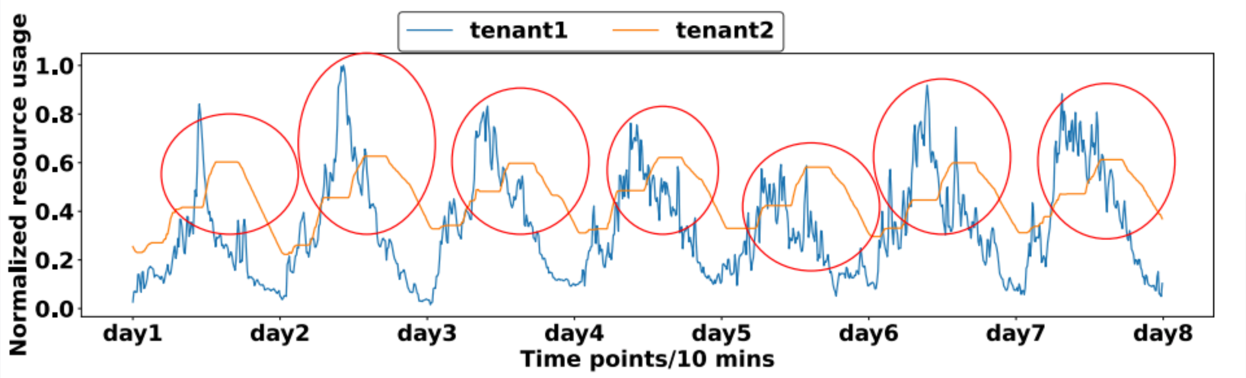
Potentials to Improve Resource Reservation Efficiency

Temporal potential:

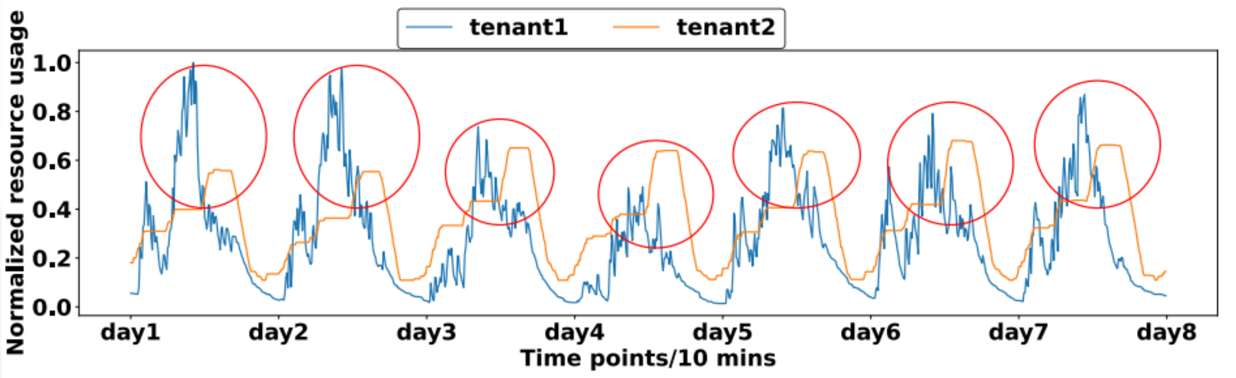
Diurnal tenants have complementary usage, temporal peak shaving reduces reservation;
The example can reduce 8.4% of resource reservation.

Spatial potential:

Tenants on different spatial (region/VM type) have complementary usage;
Further reduce resource reservation, the example can reduce 13.2%.



Temporal Potential of Peak Shaving
(tenant1 and tenant2 on region *r4* VM type *v5*)



Spatial Potential of Peak Shaving
(tenant1 on *r14* *v5* & tenant2 on *r10* *v1*)

Temporal + Spatial peak shaving can reduce resource reservation

Main Takeaways

What can we learn from temporal and spatial patterns?

- Predictable tenants needs corresponding prediction methods;
- Unpredictable tenants needs online schedule and compensate;
- Tenants have obvious spatial patterns:

  (1) different usage patterns on regions/VM types;

  (2) dynamic usage division on different regions/VM types.
- Predict/Reserve on tenants “Acceptable spatial ranges” .
- Temporal and Spatial potentials of peak shaving.

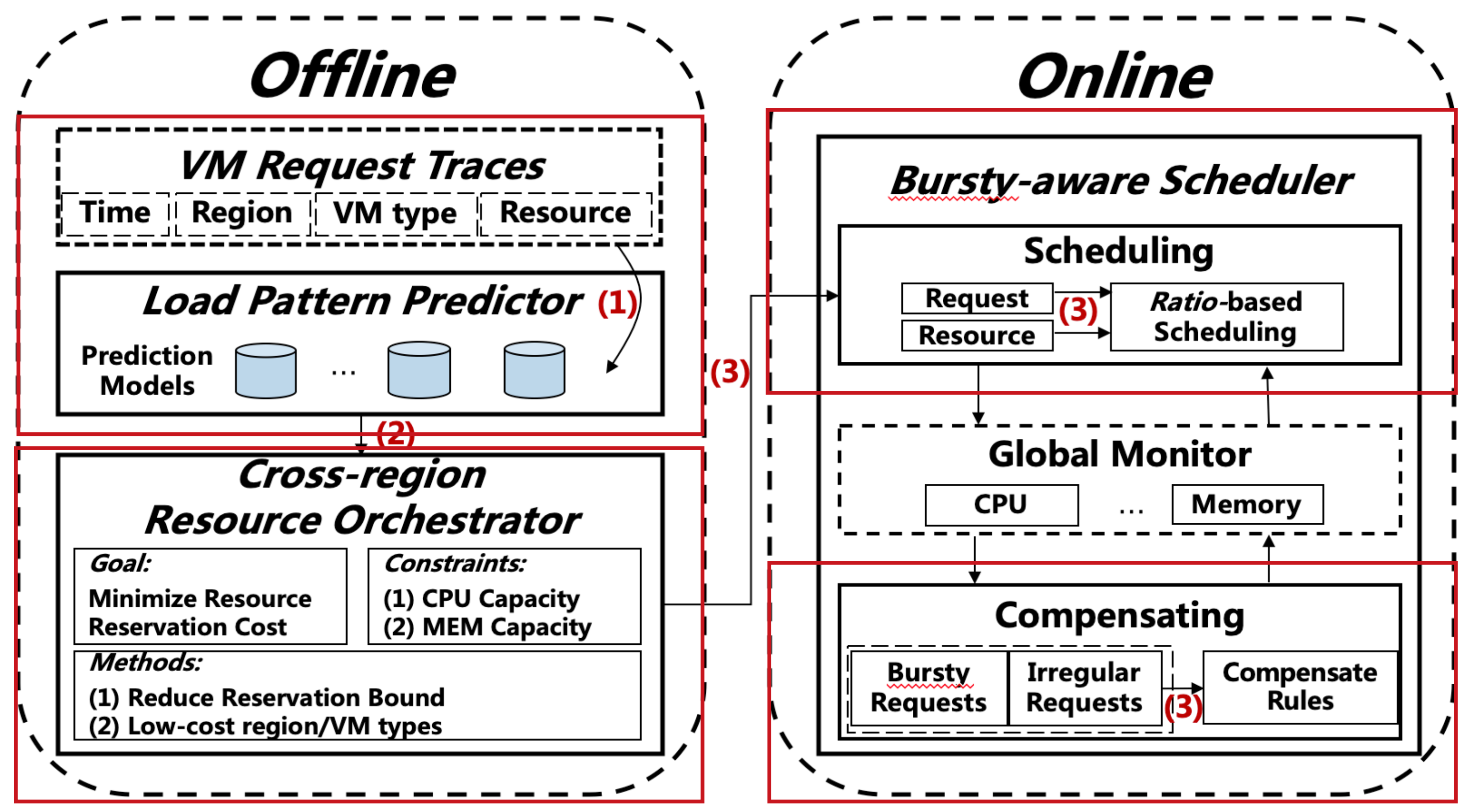
ROS: (1) Prediction; (2) Orchestration; (3) Schedule and Compensation

ROS Methodology

System Overview

ROS: Resource orchestration and VM scheduling policy

- **Load Pattern Predictor:** identify patterns of tenants, and predict resource usage
- **Cross-region Resource Orchestrator:** Orchestrate resource reservation to optimize the total cost
- **Bursty-aware Scheduler:** Scheduling VMs, and compensating for bursty with cost-minimized rules



Cross-region Resource Orchestration

Optimization model for orchestration

Note:

resource type: one VM type on one region

- Input: Predicted resource usage of tenants
- Output: orchestrate matrix *Ratio*, *reserved lines* of regions
- Peak shaving between tenants and cost coefficients of resource types

Cost coefficients **Temporal and Spatial peak shaving**

$$\text{Minimize:}\quad C_{total} = \Sigma_{j=1}^{n} \boxed{C_j} \times \boxed{max(\Sigma_{i=1}^{m}(f_i(t) \times ratio_{ij})|_{t=0}^{24h}}) $$

Goal: Optimize cost and reservation

$$\text{s.t.}\quad \begin{cases} \forall i, & \Sigma_{j=1}^{n} ratio_{ij} = 1 & i = 1 \cdots m \\ \forall j, & max(\Sigma_{i=1}^{m} CPU_{ij}(t)|_{t=0}^{24h}) < Cap_{CPU}^{j} & j = 1 \cdots n \\ \forall j, & max(\Sigma_{i=1}^{m} MEM_{ij}(t)|_{t=0}^{24h}) < Cap_{MEM}^{j} & j = 1 \cdots n \end{cases}$$

① Satisfy SLO and resource demand

② Capacity of region/VM type

Evaluation of ROS

Evaluation Setup

Trace dataset:

- VM requests of large tenants during 2021.04-07

Spatial ranges:

17regions divide into 3 region sets based on geographical positions

- *Region Set 1* : 7 regions
- *Region Set 2* : 2 regions
- *Region Set 3* : 8 regions

Typical VM types :

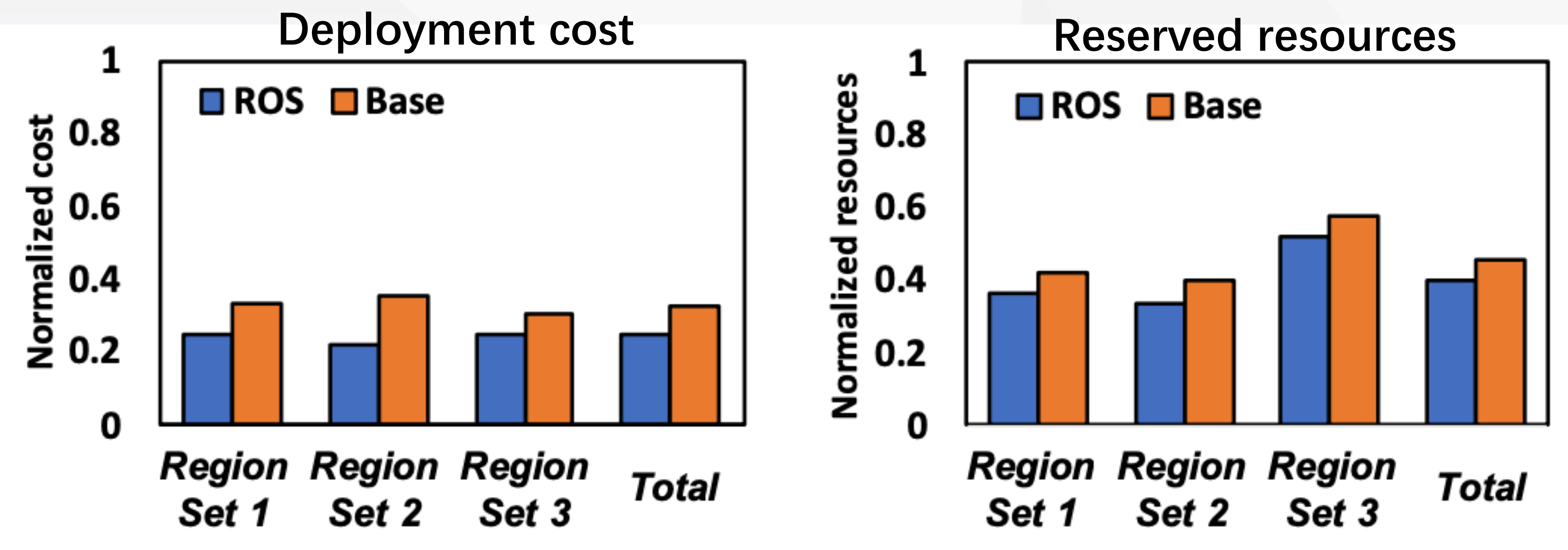
- v1, v2, v3, v4, v5, v6, v7
- Tenant workloads can be switched inside the 7 VM types

Baselines:

- Tenant-specified reservation
- Geo-distributed capacity planning (Narayanan et al.)

Reducing the Deployment Cost

- Compared with “**Tenant-specified” strategy**:
  ROS can reduce **75.4% of deployment cost** and **60.1% of resource reservation.**
- Compared with “**Geo-distributed capacity planning**”:
  ROS can reduce **24.7% of deployment cost** and **12.2% of resource reservation.**



Overall normalized cost and reserved resources

“Low-cost regions/VM types” and “Complementary patterns”

Lessons Learned

Lessons Learned

- Tenants do not understand their resource demand, prediction and low-cost region/VM type advices can reduce cost and improve utilization;
- It is better to consider the resource reservation from the cloud provider side, which can orchestrate multiple tenants together ;
- Tenants accept some low-cost regions, so cloud providers can build DCs on them, and improve utilization through adaptive orchestrator;
- Since tenants have tendency and tolerance to VM types, cloud providers can provide VM types accordingly to further improve resource efficiency

上海交通大學
SHANGHAI JIAO TONG UNIVERSITY

The end

饮 水 思 源　爱 国 荣 校