Paper Review of Collaborative Mobile Edge Computing

Wei Wang

BUPT Ph.d candidate & UC Davis visiting student
Email: weiw@bupt.edu.cn, waywang@ucdavis.edu

UCDAVIS

Group Meeting, May. 19, 2017

Contents

- MEC vs. C-RAN
- Use-cases of collaborative MEC
- Heuristics for VM placement
- Summary

MEC vs. C-RAN

	MEC	C-RAN
Location	Co-located with base stations or aggregation points.	Centralized, remote data centers.
Deployment planning	Minimal planning with possible ad hoc deployments.	Sophisticated.
Hardware	Small, heterogeneous nodes with moderate computing resources.	Highly capable computing servers.
Fronthaul requirements	Fronthaul network bandwidth requirements grow with the total amount of data that need to be sent to the core network after being filtered/processed by MEC servers.	Fronthaul network bandwidth requirements grow with the total aggregated amount of data generated by all users.
Scalability	High	Average, mostly due to expensive fronthaul deployment.
Application delay	Support time-critical applications that require latencies less than tens of milliseconds.	Support applications that can tolerate round-trip delays on the order of a few seconds or longer.
Location awareness	Yes	N/A
Real-time mobility	Yes	N/A

Table 1. Comparison of features: MEC vs. C-RAN.

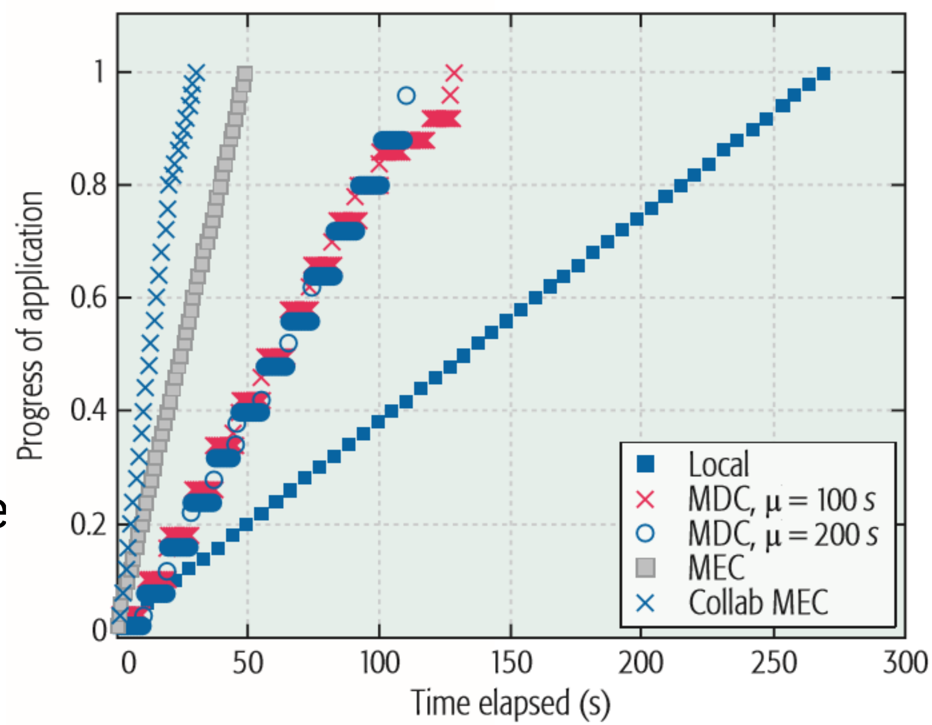
UCDAVIS
UNIVERSITY OF CALIFORNIA

Case-I Mobile Edge Orchestration

- Collaborative distributed computing framework, which is a extension from ETSI definition on MEC. 3 layers
- 1) End user, which implies both mobile and static end-user devices such as smartphones, sensors, and actuators
- 2) Edge nodes, which are the MEC servers co-located with the BSs
- 3) Cloud node, which is the traditional cloud-computing server in a remote data center
- 2 types of collaboration
- 1) horizontal collaboration at end-user layer and MEC layer
- 2) vertical collaboration between end users, edge nodes, and cloud nodes

Case-I Mobile Edge Orchestration

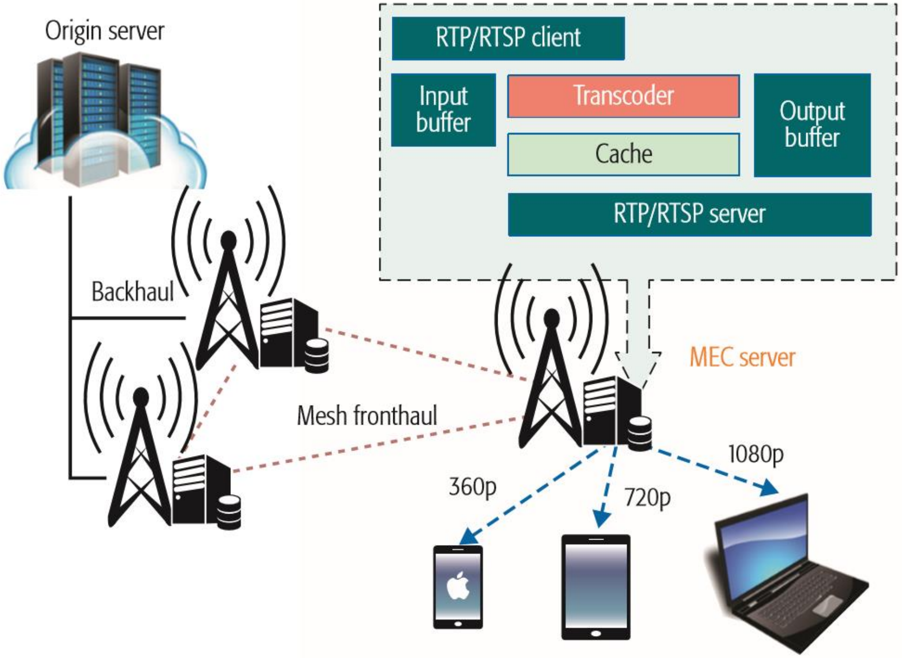
- Collaboration examples:
- Executing the application locally on the mobile device (Local)
- Distributing tasks to proximal mobile devices forming a mobile device cloud (MDC)
- Offloading the tasks to a single MEC server (MEC)
- To two collaborating MEC servers (collab MEC)



T. X. Tran, A. Hajisami, P. Pandey and D. Pompili, "Collaborative Mobile Edge Computing in 5G Networks: New Paradigms, Scenarios, and Challenges," in *IEEE Communications Magazine*, vol. 55, no. 4, pp. 54-61, April 2017.

Case-II collaborative video caching and processing

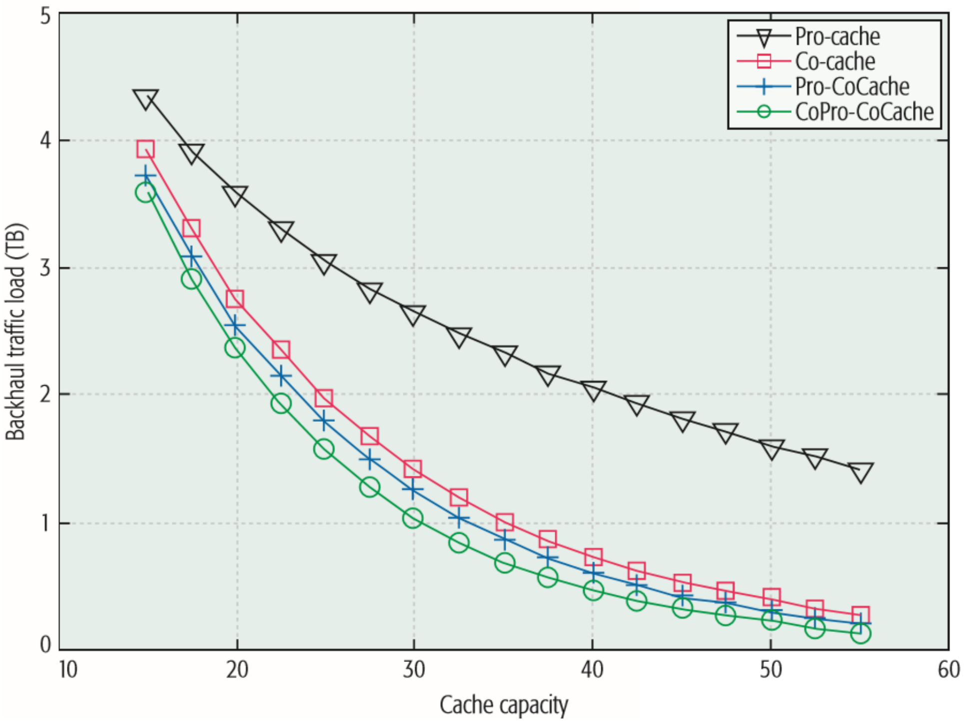
- Collaborative caching: a video request can be served using not only the local BS's cache, but also the cached copy at neighboring BSs via the backhaul links.
- Collaborative processing: MEC servers collaborate with each other to not only provide the requested video but also transcode it to an appropriate variant. Each variant is a bit rate version (e.g., 720p, 1080p) of the video, and a higher bit rate version can be transcoded into lower bit rate ones.



UCDAVIS
UNIVERSITY OF CALIFORNIA

Case-II collaborative video caching and processing

- Benefits:
- The content origin servers need not generate all variants of the same video.
- Users with various capabilities and network conditions will receive videos that are suited for their capabilities, as content adaptation is more appropriately done at network edge.
- Collaboration among the MEC servers enhances cache hit ratio and balance processing load in the network.

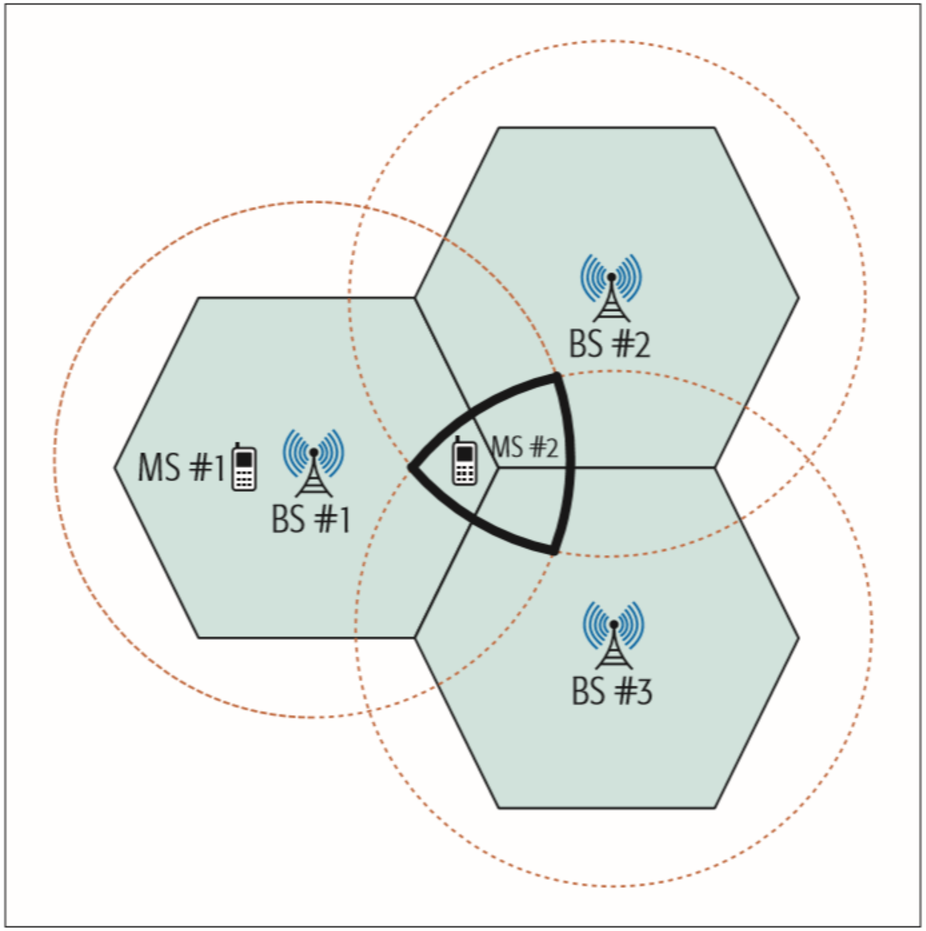


Case-III Two-layer interference cancellation

- Two-layer interference cancellation strategy for an uplink MEC-assisted RAN.
- based on the channel quality indicator (CQI) of each user, our solution identifies “where” to process its uplink signal so as to reduce complexity, delay, and bandwidth usage.
- In a MEC-assisted RAN, we have access to the computational processing at the BSs, and the signal demodulation of the cell center MSs can be done in local BSs (layer 1). This means that the system performance for cell center MSs relies on a simple single transmitter and receiver.
- Since the SINRs of cell edge MSs are often low, their signals should be transmitted to the BPU (layer 2) for further processing. In this case, the BPU has access to all the celledge MSs from different cells and is able to improve their SINRs via coordinated processing.

Case-III Two-layer interference cancellation

- MS #2 is a cell edge MS and is located in the interference region of BSs #2 and #3, there may be an intense interference from MS #2 to BSs #2 and #3; thus, coordinated interference cancellation at the upper layer is needed to cancel this interference, and the BS should transmit the raw data to the upper layer for further processing.



UC DAVIS
UNIVERSITY OF CALIFORNIA

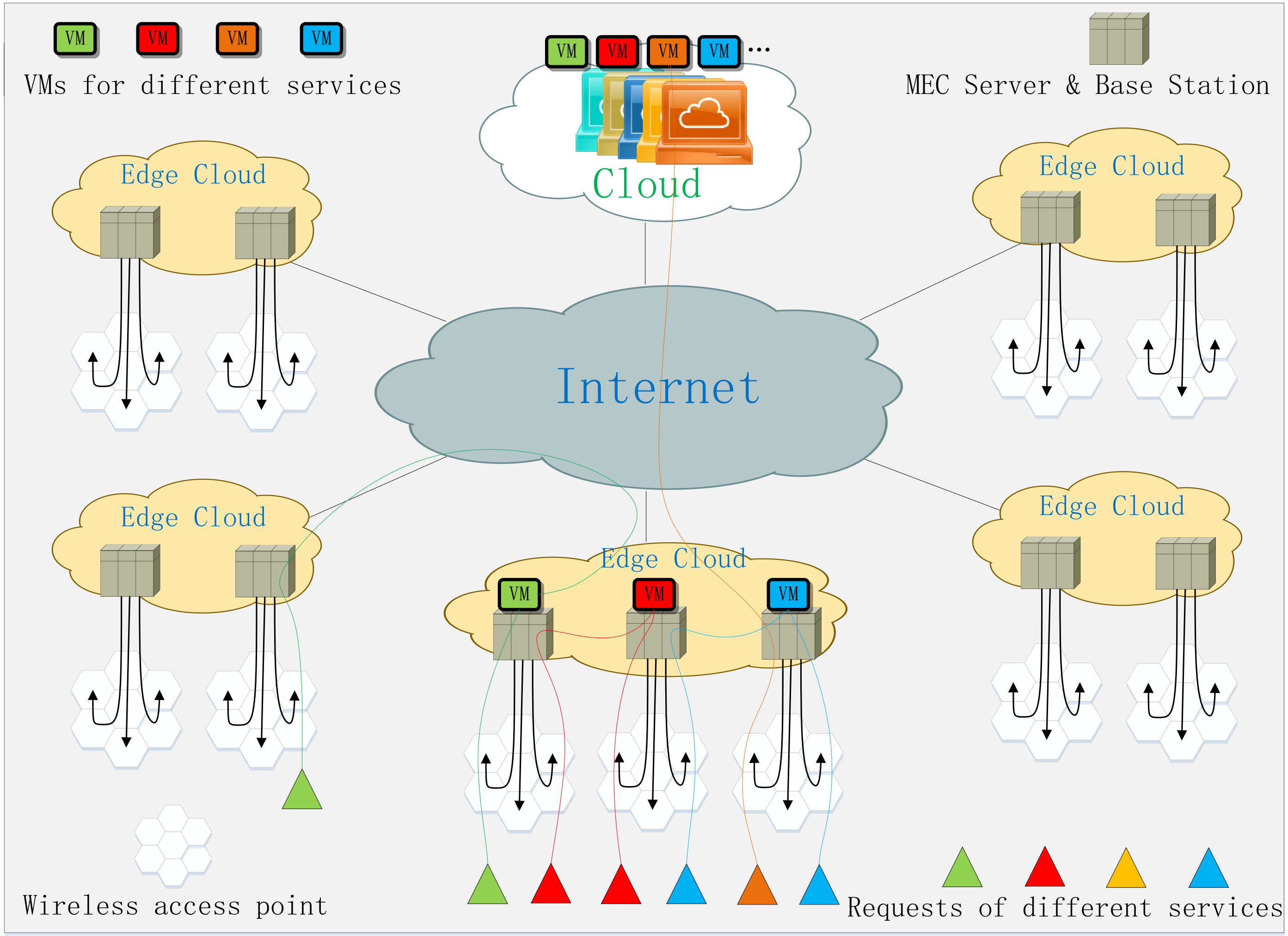
Recall of VM placement and workload assignment

- O

(1)

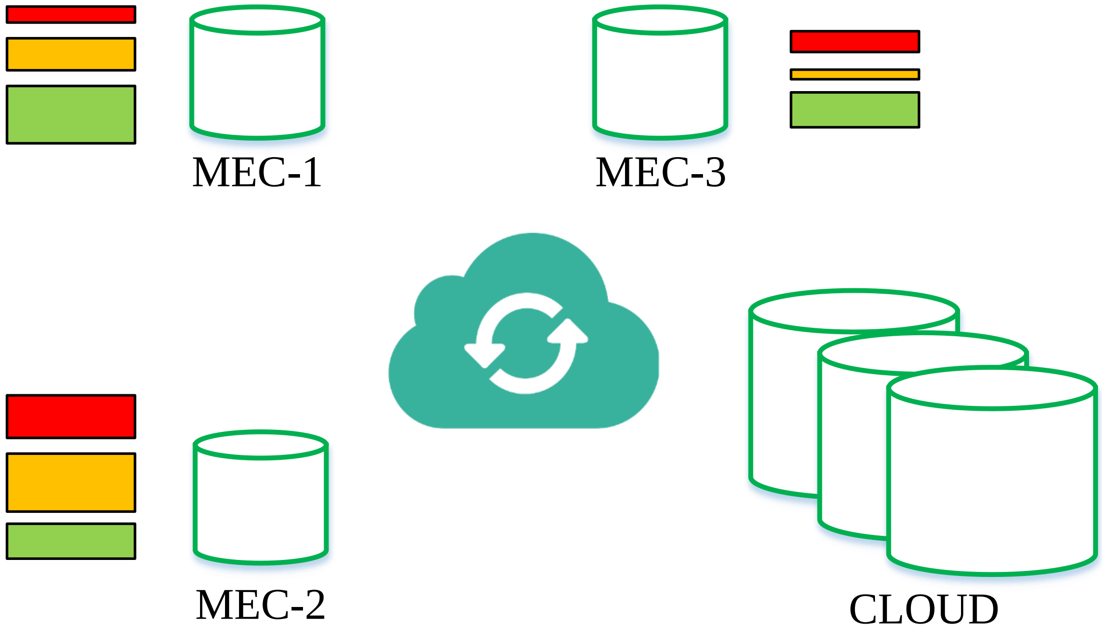
(2)

(3)



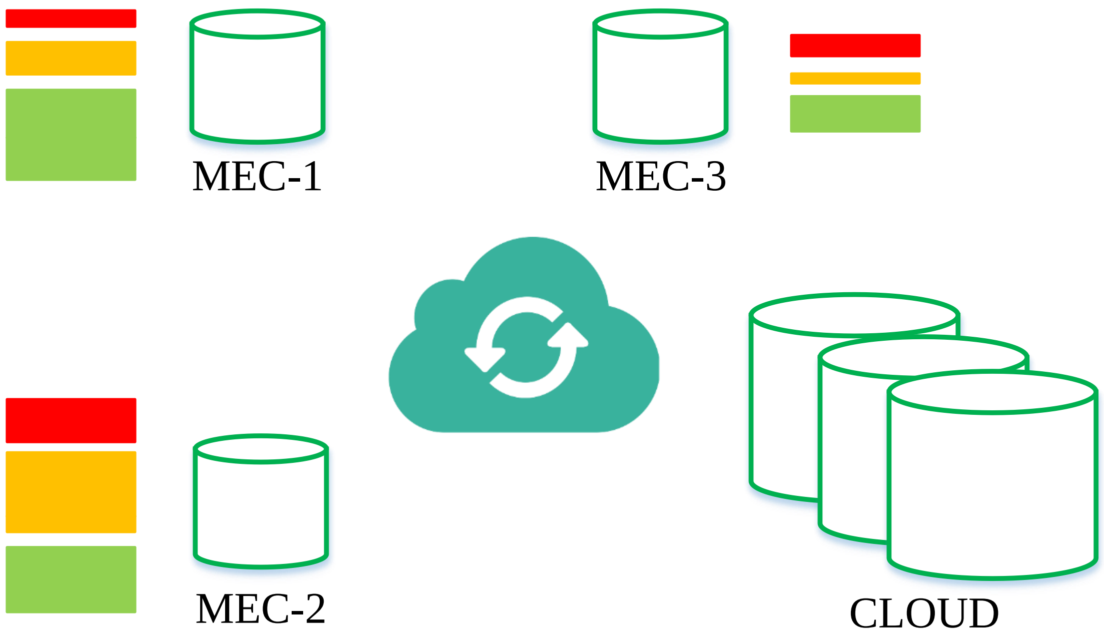
Heuristic for VM placement

- Given workloads of each APP from each Edge
- How to place VMs for each APP at each Edge?
- Initial placement + Flow exchange



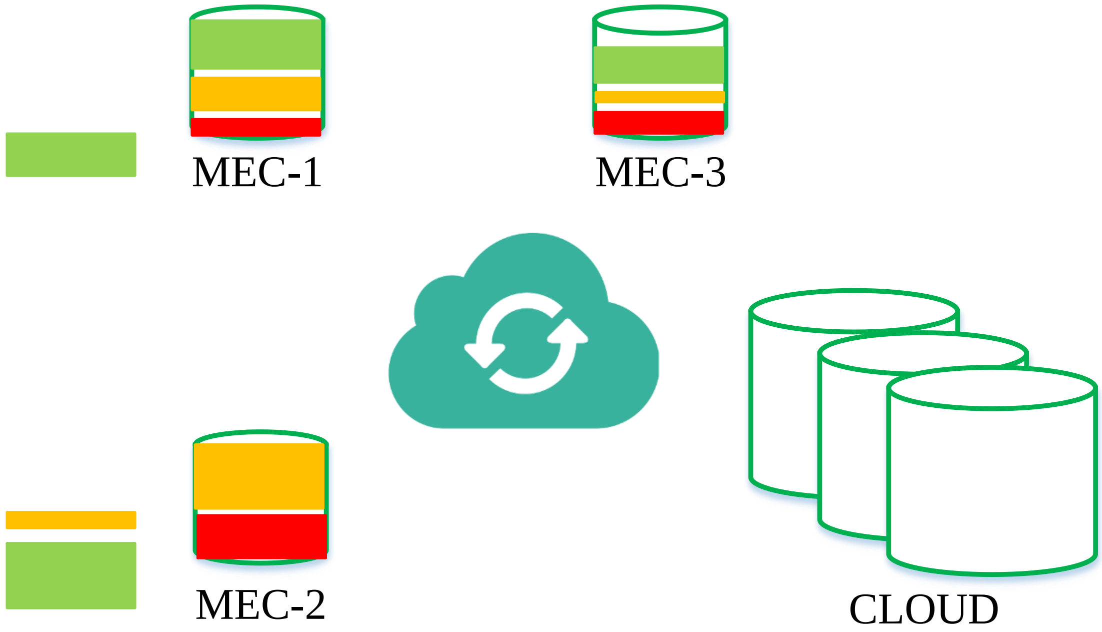
Two-phase initial placement

- Phase 1 – Local Placement



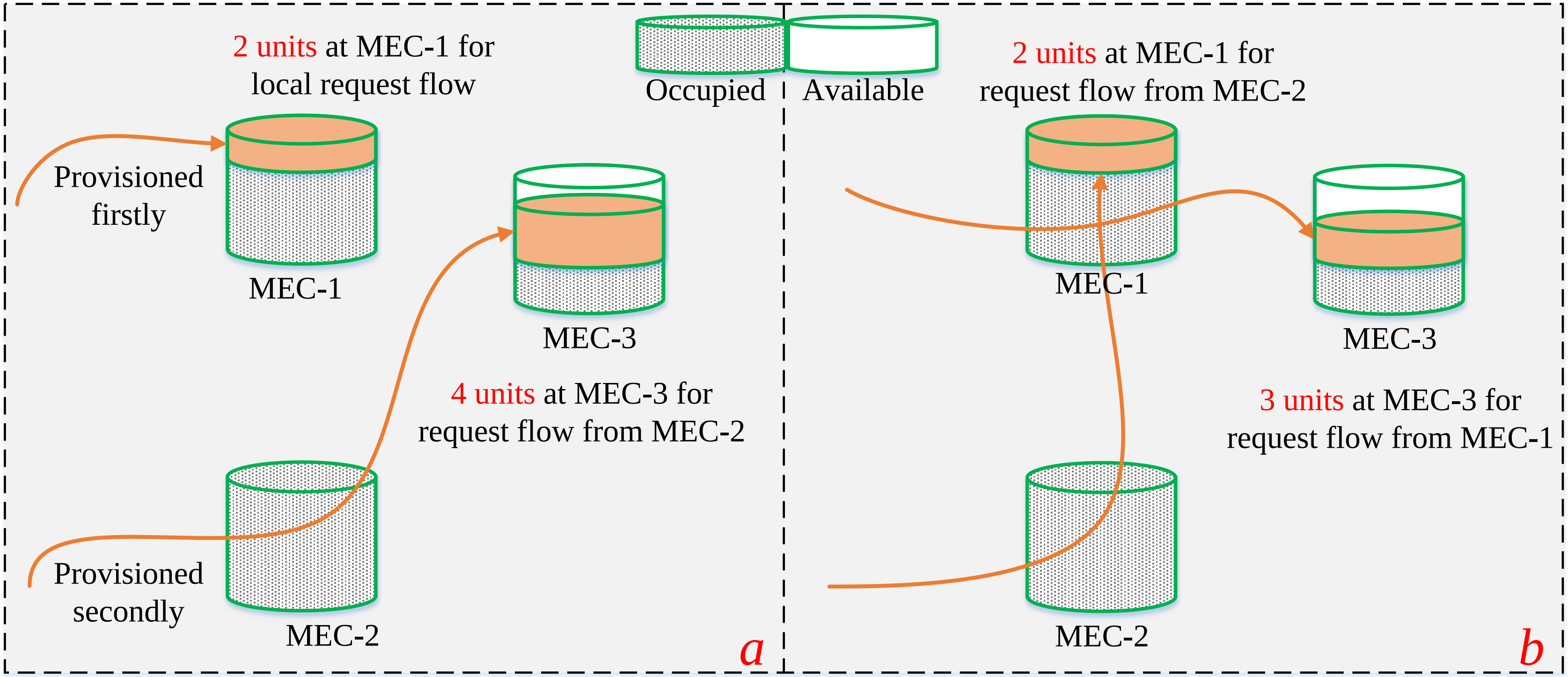
Two-phase initial placement

- Phase 2 – Remote Placement
- Principles: Low-latency flow first; Bigger flow first; Nearest destination first



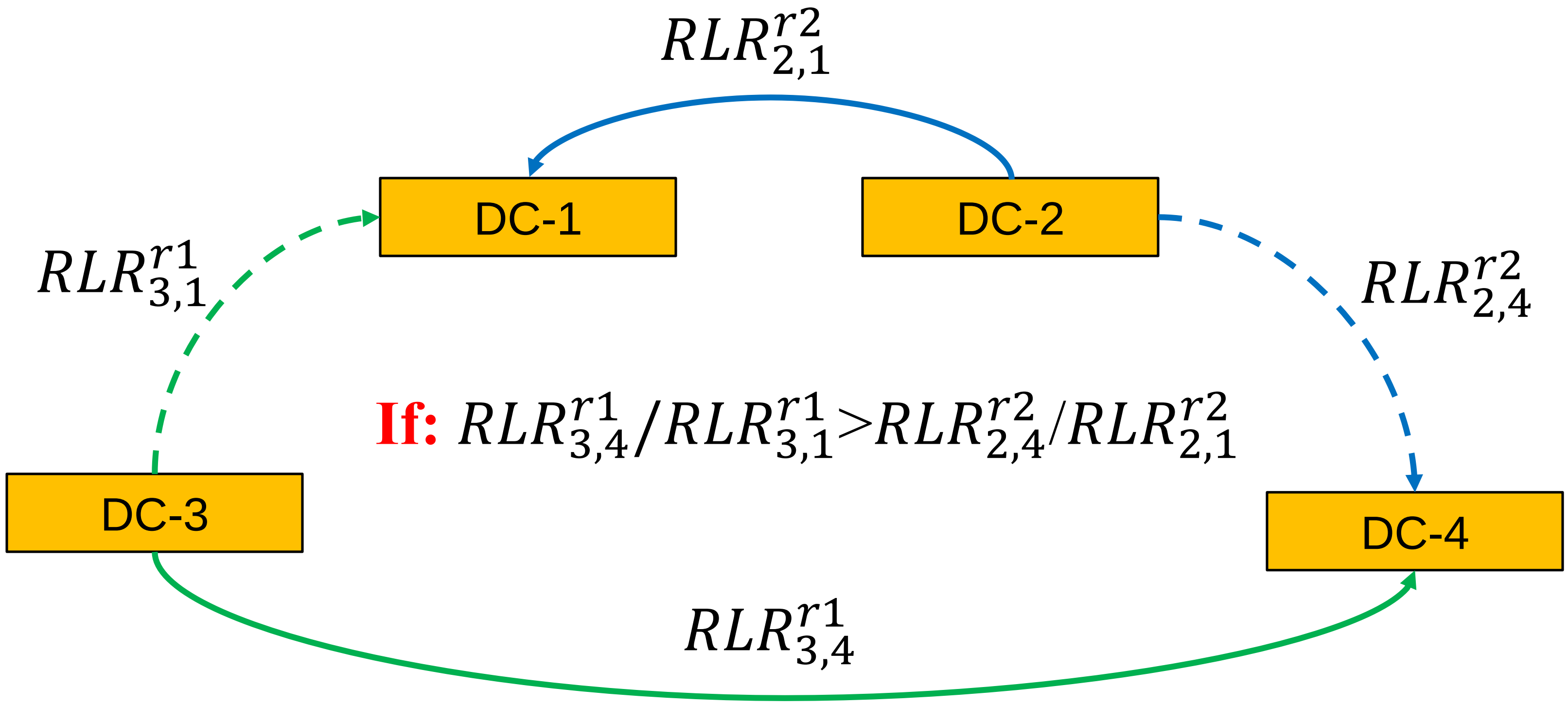
Flow exchange

- Basic rule for remote placement: Longer distance between src and dst causes higher inter-DC latency, and less time is left for queueing and processing. Thus, more VMs are required to process faster.
- Un-optimal condition:



Flow exchange (cont.)

- Definition: Remote/Local resource ratio: RLR
- $RLR^{r}_{s,d} = \frac{R^{r}_{d}}{R^{r}_{s}}$, required resource at d / s

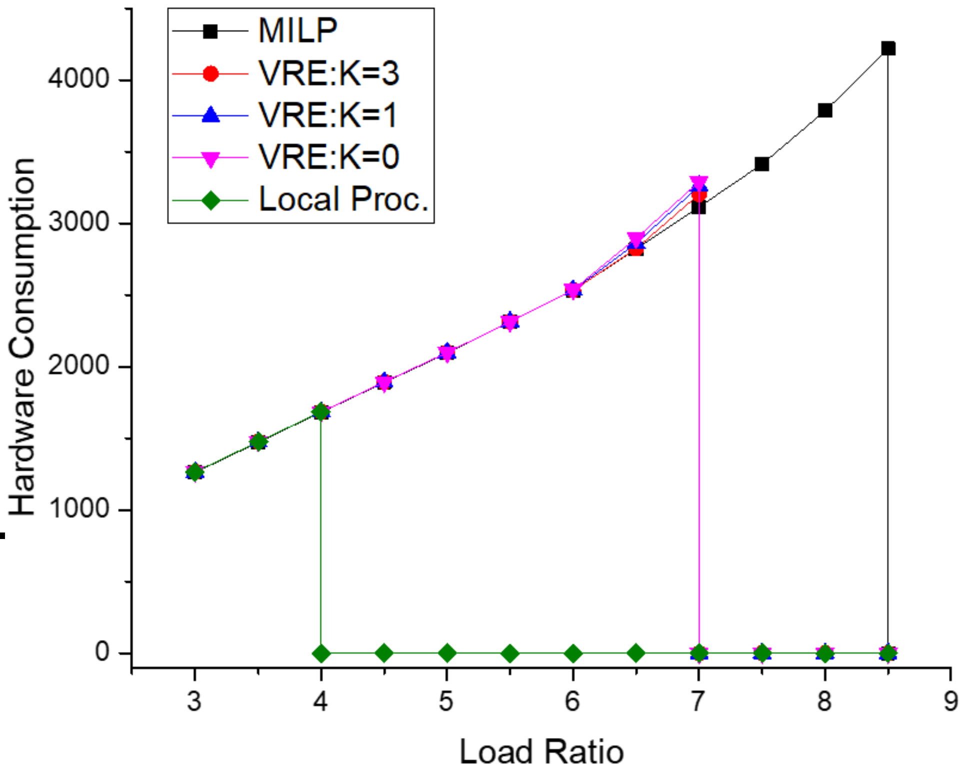


UCDAVIS
UNIVERSITY OF CALIFORNIA

Flow exchange (cont.)

Algorithm: VM and Request flow Exchange (VRE)

1. **for each** cross-DC big sub-flow $\mathrm{br}^a_{s,d}$, where s $\neq d$.
2. calculate its required resource $HW_{\mathrm{br}^a_{s,d}}$ at d.
3. **for each** DC node v ∈ V
4. estimate required resource $HW_{\mathrm{br}^a_{s,v}}$ if migrate $\mathrm{br}^a_{s,d}$ to v
5. **if** $HW_{\mathrm{br}^a_{s,v}} < HW_{\mathrm{br}^a_{s,d}}$, **then**
6. **for each** $\mathrm{br}^a_{o,v}$, which originates from o to v
7. **if** d is reachable for $\mathrm{br}^a_{o,v}$, **then**
8. calculate $HW_{\mathrm{br}^a_{o,v}}$ its required at v
9. estimate $HW_{\mathrm{br}^a_{o,d}}$ if migrate it to d
10. **if** $\frac{HW_{\mathrm{br}^a_{s,d}}}{HW_{\mathrm{br}^a_{s,v}}} > \frac{HW_{\mathrm{br}^a_{o,d}}}{HW_{\mathrm{br}^a_{o,v}}}$, **then**
11. migrate $\mathrm{br}^a_{o,v}$ to d and migrate $\mathrm{br}^a_{s,d}$ to v
12. **end if, end if, end if**
13. **end for, end for, end for**



Summary

- Collaborative MEC is a interesting topic, which can be studied in different contexts:
- 1) MEC for radio networks (mobility, Comp)
- 2) MEC for specific applications (location or context-aware, etc.)
- 3) MEC for Internet service provisioning.

- Collaboration can be considered from several aspects:
- 1) inter-DC(MEC server or cloud) collaboration
- 2) MEC as assistant for network devices or functions

- Low latency is the original goal for MEC, and it is also the strict constraint for MEC collaboration.

Thank you!

Wei Wang