Disaggregated Micro Data Center: Resource Allocation Considering Impact of Network on Performance

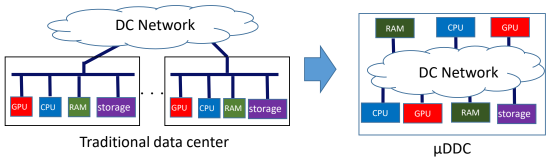
Akishige Ikoma, Yuichi Ohsita, Masayuki Murata
Graduate School of Information Science and Technology, Osaka University

09 JANUARY 2023 IEEE Consumer Communications & Networking Conference (CCNC) 1

1

Background

- Edge services using micro data centers
  - Smaller latency than the cloud
    - Effective for time-sensitive services
  - More limited resources compared with large data centers
- ➢Disaggregated Micro Data Center (µDDC)
  - µDDC is constructed of resources connected by a network
  - Achieve efficient resource utilization
    - Optimization per resource
  - Flexible scaling without server constraints

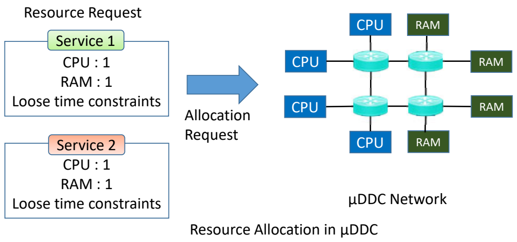


2

2

Problem of Resource Allocation in µDDC

- Resource request arrives before application start
- Allocate execution computational resources, memory resources, and routes between resources

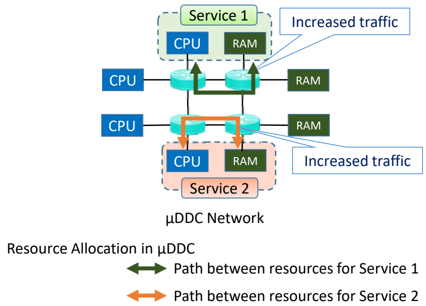


3

3

Problem of Resource Allocation in µDDC

- Resource request arrives before application start
- Allocate execution computational resources, memory resources, and routes between resources

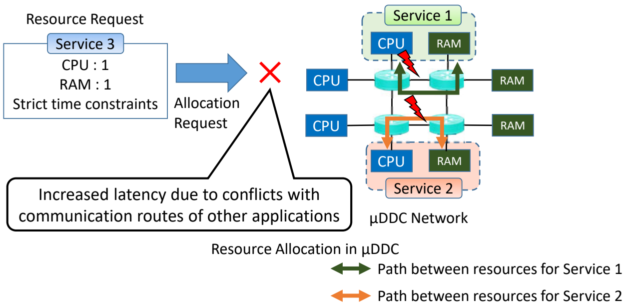


4

4

Problem of Resource Allocation in µDDC

- Resource request arrives before application start
- Allocate execution computational resources, memory resources, and routes between resources

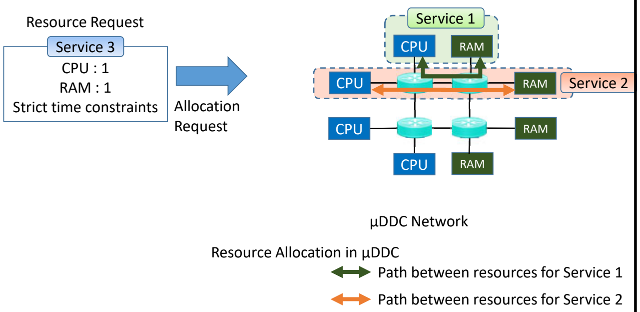


5

5

Problem of Resource Allocation in µDDC

- Resource request arrives before application start
- Allocate execution computational resources, memory resources, and routes between resources

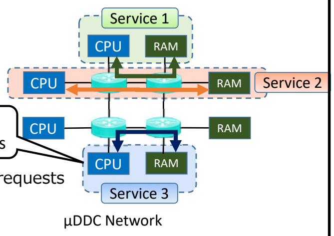


6

6

Problem of Resource Allocation in μDDC

- Resource request arrives before application start
- Allocate execution computational resources, memory resources, and routes between resources



Low-latency communications because no conflicts with other services

➢ Can accommodate strict constraint requests

Resource Allocation in μDDC

μDDC requires efficient allocation of network resources

7

7

Approach

- Research Objectives
  - Accommodate more applications by efficient use of network resources in μDDC
- Approach
  - We propose efficient resource allocation method
    - Efficient use of resources through link sharing for flexible routing
    - Considering future requests by preserving important resources

  - Model the impact of network on the application performance
  - Define resource allocation cost based on resource importance

8

8

Impact of network on performance

$$T^e = \frac{\sigma^c}{F} + \left\{\frac{V}{B} + T^l\right\} N$$

- $\frac{\sigma^c}{F}$: processing time in computational resource — Divide the number of clocks to run by the clock frequency
- $\frac{V}{B}$: transmission delay — Divide read data by bandwidth
- T^l: propagation delay +switching delay
- N: Number of accesses to memory

- Processing delay and the buffering delay for packet collision avoidance
- Derive buffering time by M/D/C queueing model
  - Derived from the arrival rate of packets passing over the link

9

9

Resource Allocation Method

Allocate execution resources for the service following the steps below

1. Model Impact of network on performance
   - Derive time to execute running services
2. Cost Assign to resources
   - Assign cost based on importance for future resource requests
3. Allocate resources and path between them
   - Allocate resources with the lowest cost
   - Allocate resources to finish the process within the acceptable time

Ensure resources are available for the future request while meeting service performance requirements

10

10

Resource Allocation Cost

- Avoid allocating resources likely needed for future resource allocations
  - These resources and links cost high
- Computational resource : $W_c^c = (|C_{Node_c}^s|) \cdot K_c$
  - High performance and many available resources
- Memory resource : $W_m^m = |M_{Node_m}^s|$
  - many available resources
- Link: $W_e^e = \begin{cases} \sum_{c\in C^s, m\in M^s} \left(\frac{N_{c,m}^r(e)}{N_{c,m}^r}\right)\left(\frac{W_c^c \cdot W_m^m}{H_{c,m}}\right) & e \notin E^{alc} \\ \epsilon & e \in E^{alc} \end{cases}$
  - Likely to be the path between important resources

- C^s : allocatable computational resource
- M^s : allocatable memory resource
- E^{alc} : allocated network resource
- K_c : FLOPS of computational resource c
- $N_{c,m}^r$: *shortest paths between resources c,m*
- $N_{c,m}^r(e)$: *shortest pathsthrough link e between resources c,m*
- $H_{c,m}$: *Shortest hops between resources c,m*

11

11

Resource Allocation Problem

- Allocate resources solving resource allocation problem
- constraints

$$\forall i \in N^v, \quad \sum_{j\in N^s} \delta_{i,j}^N = 1$$
$$\forall x \in E^v, \forall s,t \in N^s, \quad \sum_{y\in R_{s,t}} \delta_{x,y}^E = \delta_{n_x^{vs},s}^N \cdot \delta_{n_x^{vd},t}^N$$
$$\forall c \in N^c, \quad |C_c^s| - \sum_{c'\in C^v} \delta_{c',c}^N \geq 0$$
$$\forall m \in N^m, \quad |M_m^s| - \sum_{m'\in M^v} \delta_{m',m}^N \geq 0$$
$$\forall t \in S, \quad T_t^e \leq T_t^a$$

- objective function

$$minimize \quad \sum_{c\in N^c}\sum_{c'\in C^v} \delta_{c',c}^N (W_{c_c^s}^c) + \sum_{m\in N^m}\sum_{m'\in M^v} \delta_{m',m}^N (W_{m_m^s}^m) + \sum_{i,j\in N^s}\sum_{y\in R_{i,j}} \left[1_{\sum_{x\in E^v}\delta_{x,y}^E > 0}\left(\sum_{e\in y} W_e^e\right)\right]$$

- Derived by metaheuristic method for NP-hard
  - We use Ant Colony Optimization (ACO)

12

12

Resource Allocation Problem

- Allocate resources solving resource allocation problem
- constraints

$$\forall i \in N^v,\quad \sum_{j \in N^s} \delta^N_{i,j} = 1$$
$$\forall x \in E^v, \forall s,t \in N^s,\quad \sum_{y \in R_{s,t}} \delta^E_{x,y} = \delta^N_{n^{vs}_x,s} \cdot \delta^N_{n^{vd}_x,t}$$
$$\forall c \in N^c,\quad |C^s_c| - \sum_{c' \in C^v} \delta^N_{c',c} \geq 0$$
$$\forall m \in N^m,\quad |M^s_m| - \sum_{m' \in M^v} \delta^N_{m',m} \geq 0$$

Resource mapping constraints
- Request resources and μDDC resources are one-to-one
- No more than one service can be allocated to a resource

- objective function

$$\text{minimize}\quad \sum_{c \in N^c} \sum_{c' \in C^v} \delta^N_{c',c}(W^c_{c^s}) + \sum_{m \in N^m} \sum_{m' \in M^v} \delta^N_{m',m}(W^m_{m^s_m}) + \sum_{i,j \in N^s} \sum_{y \in R_{i,j}} \left[1_{\sum_{x \in E^v} \delta^E_{x,y} > 0} \left(\sum_{e \in y} W^e_e\right)\right]$$

- Derived by metaheuristic method for NP-hard
  - We use Ant Colony Optimization (ACO)

13

13

Resource Allocation Problem

- Allocate resources solving resource allocation problem
- constraints

$$\forall i \in N^v,\quad \sum_{j \in N^s} \delta^N_{i,j} = 1$$
$$\forall x \in E^v, \forall s,t \in N^s,\quad \sum_{y \in R_{s,t}} \delta^E_{x,y} = \delta^N_{n^{vs}_x,s} \cdot \delta^N_{n^{vd}_x,t}$$
$$\forall c \in N^c,\quad |C^s_c| - \sum_{c' \in C^v} \delta^N_{c',c} \geq 0$$
$$\forall m \in N^m,\quad |M^s_m| - \sum_{m' \in M^v} \delta^N_{m',m} \geq 0$$
$$\forall t \in S,\quad T^e_t \leq T^a_t$$

- objective function

Time constraints
- Finish the process within the acceptable time

$$\text{minimize}\quad \sum_{c \in N^c} \sum_{c' \in C^v} \delta^N_{c',c}(W^c_{c^s}) + \sum_{m \in N^m} \sum_{m' \in M^v} \delta^N_{m',m}(W^m_{m^s_m}) + \sum_{i,j \in N^s} \sum_{y \in R_{i,j}} \left[1_{\sum_{x \in E^v} \delta^E_{x,y} > 0} \left(\sum_{e \in y} W^e_e\right)\right]$$

- Derived by metaheuristic method for NP-hard
  - We use Ant Colony Optimization (ACO)

14

14

Resource Allocation Problem

- Allocate resources solving resource allocation problem
- constraints

$$\forall i \in N^v,\quad \sum_{j \in N^s} \delta^N_{i,j} = 1$$
$$\forall x \in E^v, \forall s,t \in N^s,\quad \sum_{y \in R_{s,t}} \delta^E_{x,y} = \delta^N_{n^{vs}_x,s} \cdot \delta^N_{n^{vd}_x,t}$$
$$\forall c \in N^c,\quad |C^s_c| - \sum_{c' \in C^v} \delta^N_{c',c} \geq 0$$
$$\forall m \in N^m,\quad |M^s_m| - \sum_{m' \in M^v} \delta^N_{m',m} \geq 0$$
$$\forall t \in S,\quad T^e_t \leq T^a_t$$

- objective function

Allocate resources and routes to minimize costs

$$\text{minimize}\quad \sum_{c \in N^c} \sum_{c' \in C^v} \delta^N_{c',c}(W^c_{c^s}) + \sum_{m \in N^m} \sum_{m' \in M^v} \delta^N_{m',m}(W^m_{m^s_m}) + \sum_{i,j \in N^s} \sum_{y \in R_{i,j}} \left[1_{\sum_{x \in E^v} \delta^E_{x,y} > 0} \left(\sum_{e \in y} W^e_e\right)\right]$$

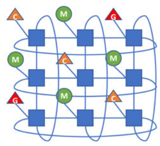
- Derived by metaheuristic method for NP-hard
  - We use Ant Colony Optimization (ACO)

15

15

Evaluation Environment

- Simulate allocation process in a 3×3 2D torus network



multi-core fiber（4 optical fiber cores）
Switch
Memory Pool (250 memory resources)
CPU Pool(28 computational resources)
GPU

- Evaluate the number of blocked resource requests
  - Blocked when request fails to satisfy the acceptable time
- 4 types requests arrive in 300 minutes
  - arrive at a rate of 20% per minute
  - Lifetime is 60 minutes
  - Request 4 only requires GPU

Acceptable time

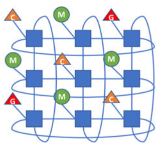
Request	time
Request 1	3000ms
Request 2	500ms
Request 3	150ms
Request 4	150 ms

16

16

Evaluation Environment

- Simulate allocation process in a 3×3 2D torus network



multi-core fiber（4 optical fiber cores）
Switch
Memory Pool (250 memory resources)
CPU Pool(28 computational resources)
GPU

- Evaluate the number of blocked resource requests
  - Blocked when request fails to satisfy the acceptable time
- 4 types requests arrive in 300 minutes
  - arrive at a rate of 20% per minute
  - Lifetime is 60 minutes
  - Request 4 only requires GPU

Acceptable time

Request	time
Request 1	3000ms
Request 2	500ms
Request 3	150ms
Request 4	150 ms

Strict performance requirements

17

17

Evaluation

- Change the arrival rate of each request

Arrival rate of each request in each requests sequence

	Request 1	Request 2	Request 3	Request 4
Request sequence 1	75%	5%	5%	15%
Request sequence 2	35%	15%	35%	15%
Request sequence 3	5%	5%	75%	15%

- Comparison of network resource allocation policies
  - RA-CNP: proposed method
  - SP: allocate the shortest path between resources
  - NP: allocate low latency path between resources
- Comparison of Link Usage Policies
  - RA-CNP w/o S: each link is used by one application only
  - RA-CNP w/o A: prohibit aggregating multiple optical fiber cores

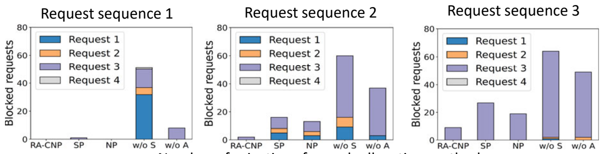
18

18

Result

- In all cases, RA-CNP accommodated the most requests
  - RA-CNP blocked in environments with a high arrival rate of requests with a high amount of required resources
  - More than twice fewer rejections than comparative methods
- Reduce the blockings by sharing and aggregating link

- RA-CNP can accommodate more applications without affecting the required performance



Number of rejections for each allocation method

Conclusion

- We proposed RA-CNP that achieved an efficient usage of network in μDDC
  - We modeled the impact of network on performance
  - Preserves important resources for future requests
  - Efficient use of network through link sharing and aggregation

- We demonstrated to accommodate more requests to satisfy the required performance

- Future work
  - Evaluation of our method on a larger scale μDDC
  - Investigation of network structure suitable for μDDC
    - Network topology and resource placement