Osaka University

Dynamic Placement of Virtual Network Functions based on Model Predictive Control

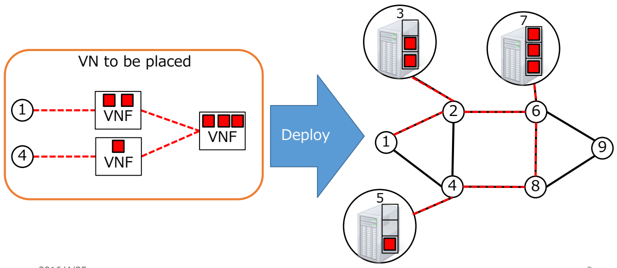
Kota Kawashima, Tatsuya Otoshi, Yuichi Ohsita, Masayuki Murata

Graduate School of Information Science and Technology, Osaka University

Osaka University

NFV (Network Functions Virtualization)

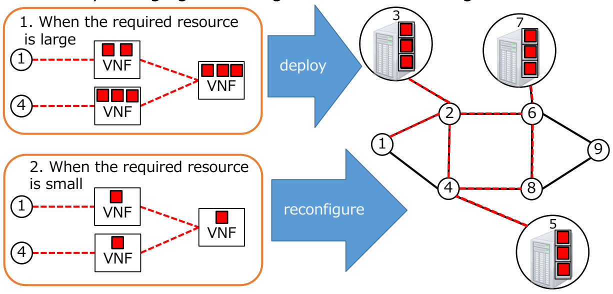
- The virtual network functions (VNFs) are hosted by ordinary server computers
- By placing the VNFs to the suitable server, the network services are provided efficiently



Osaka University

Dynamical placement of the VNFs

- Reconfigure the location of the VNFs according to the change of the required resources
  - By migrating the VNFs
  - By changing the configuration of the routing



Osaka University

The problem of existing methods

- The cost of the migration of the VNFs
  - The migration consumes network resources.
- The existing method considers only the currently required resources
  - The migration is not performed unless the necessity of the migration is detected

A large number of migrations may be required
-> Network instability

Our Method
- Detect the necessity of the migration from the predicted demands
  - Start migration in advance
  - Avoid a large number of migrations at the each time slot

Osaka University

Objective and Approach

- Objective
  - Establishment of a method which places the VNFs so as to follow the traffic variation
    - Start migration in advance of the change of the required resources
      - By considering the predicted future demands
    - Allocate sufficient resources to the VNFs without migrating a large number of VNFs at the same time
- Approach

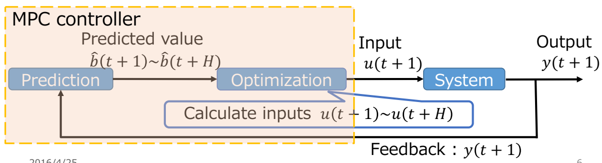
Applying MPC [1] to dynamic placement of VNFs
- Decide the placement based on the predicted value
- Robust control to prediction errors

[1] S. J. Qin and T. A. Badgwell, "A survey of industrial model predictive control technology," *Control Engineering Practice*, vol. 11, pp. 733-764, July 2003.

Osaka University

Model Predictive Control (MPC) [1]

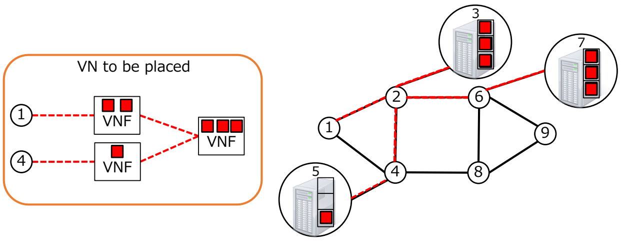
- Overview
  - Inputs setting to a system to make the output close to desired one
- Correction of prediction error by feedback
  - Controller implements only the calculated inputs for the next time slot
  - Controller observes the output and corrects the prediction
  - Controller recalculates the inputs with the corrected prediction



Osaka University

Objective of Our VNF Placement

- Minimize the number of active physical node at each time
- The cost of migrations should also be considered
  - Migration causes performance degradation



Osaka University

Placement of VNFs based on MPC (MPC-VNF-P)

- Formalization

$$minimize\ \frac{(1-w)}{H\cdot|N^P|}\sum_{0<t\le H}\sum_{n\in N^P} M_n^{Node}(t) + \frac{w}{2|N^{VNF}|}\bar{M}$$

Decrease the number of active physical nodes

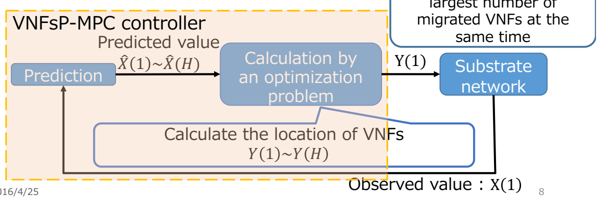
Decrease the cost of migration

w: weight for migration
H: predictive horizon
N^P: set of physical nodes
$M_n^{Node}(t)$: binary variable, that indicates the deployment of VNFs
N^{VNF}: set of VNF nodes
$\bar{M}$: variable to decrease the cost of migration
$M_{n^{vnf},n}^{Node}(t)$: binary variable, which indicates the placement of a VNF

$$St.\ 0<t\le H,\ \sum_{n\in N^P}\sum_{n^{vnf}\in N^{VNF}} |M_{n^{vnf},n}^{Node}(t) - M_{n^{vnf},n}^{Node}(t-1)| \le \bar{M}$$

Cost is defined by the largest number of migrated VNFs at the same time

- Procedure



Osaka University

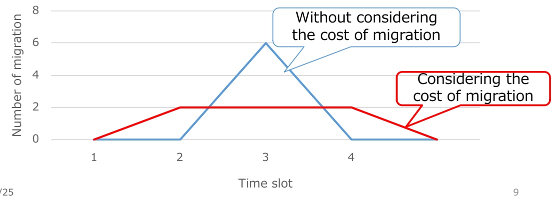
Decreasing migration by introducing $\bar{M}$

- Before introducing $\bar{M}$
  - Minimize the number of migration in predictive horizon

$$\sum_{0<t\le H}\sum_{n\in N^P}\sum_{n^{vnf}\in N^{VNF}} |M_{n^{vnf},n}^{Node}(t) - M_{n^{vnf},n}^{Node}(t-1)|$$

A large number of migration may occur

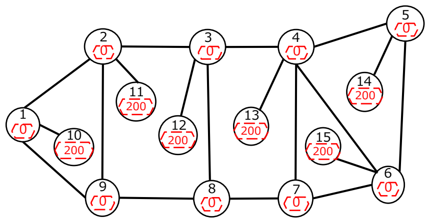
- Effect of introducing $\bar{M}$



Osaka University

Evaluation: Physical network environment

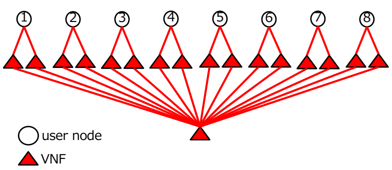
- The topology is based on the Internet2 topology
  - Six nodes are connected to the servers
  - Only the servers have the resources to host the VNFs
    - Each server has the resources whose capacity is 200
  - The bandwidth of each link has a sufficiently large value

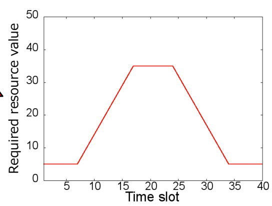


Osaka University

Evaluation: Virtual network environment

- The virtual network includes 8 user nodes and 17 VNFs
  - Two kinds of the VNFs
    - One handles the traffic near user and are connected to user nodes
    - The other is connected to all of the VNFs
- We generate the time variations of the required resources.





Osaka University

Compared method

- MinActiveNode
  - Minimize the number of active physical nodes without considering the cost of migration
- NoMPC
  - The predicted required resources only at the next time slot are used, considering the cost of migration
- MPC-VNF-P
  - Proposed method

	MinActiveNode	NoMPC	MPC-VNF-P
Control parameters	$H=1$ $w=0$	$H=1$ $w=0.03$	$H=3,5$ $w=0.03$

Other simulation environments

- Prediction method
  - Simple line fitting to past time series
- Metrics
  - Maximum resource utilization
    - The largest resource utilization, which is defined by

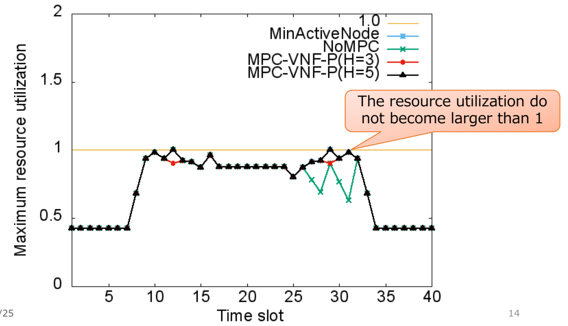
$$\max_{n^p \in N^p} \left(\frac{1}{u^p_{n^p}} \sum_{n^{vnf} \in N^{VNF}_{n^p}} u^v_{n^{vnf}} \right)$$

  - Number of active physical nodes
    - The number of physical nodes hosting at least one VNFs
  - Number of migrated VNFs
    - The number of VNFs which are migrated at each time slot

u^p_n: the resource capacity of physical node n^p
$N^{VNF}_{n^p}$: the set of virtual nodes hosted by the physical node n^p
$u^v_{n^{vnf}}$: the resource required by VNF node n^{vnf}

Maximum resource utilization

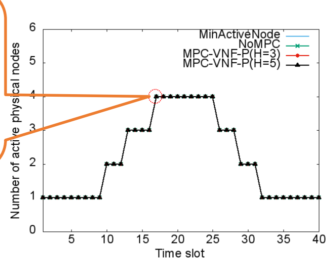
- All methods map the virtual network properly
  - VNFs are migrated before the lack of resources is caused by using the predicted values.



Number of active physical nodes

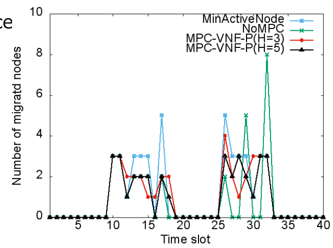
- All methods change the number of active physical nodes according to the time variation of the required resources
- MPC-VNF-P indicates the same performance compared with MinActiveNode

- The future required resources are predicted to increase, while the actual required resources stop increasing
- MPC-VNF-P avoids the increase of the number of active physical nodes
  - Correcting the prediction errors
  - Calculating the locations of VNFs again



Number of migrated VNFs

- MinActiveNode and NoMPC require a larger number of migrations
  - MinActiveNode does not consider the cost of migration
  - NoMPC does not consider the future required resources
- MPC-VNF-P avoids a large number of migrations at any time slot
  - Start migration in advance by using the predicted values



Summary and future work

- Summary
  - Proposition of MPC-VNF-P
    - We introduce the idea of placement of VNFs based on MPC
    - Our method starts migration in advance of traffic variation
      - By considering the predicted future demands
  - Evaluation of MPC-VNF-P
    - We show that MPC-VNF-P allocates sufficient resources without migrating a large number of VNFs at the same time
    - We show that our method handles the time variation of the demands
- Future work
  - The evaluation using the actual traffic traces
  - Establishing a distributed algorithm of the dynamic placement of the VNFs