

Toward efficient and flexible resources provisioning on SDN-enhanced Job Management System Framework

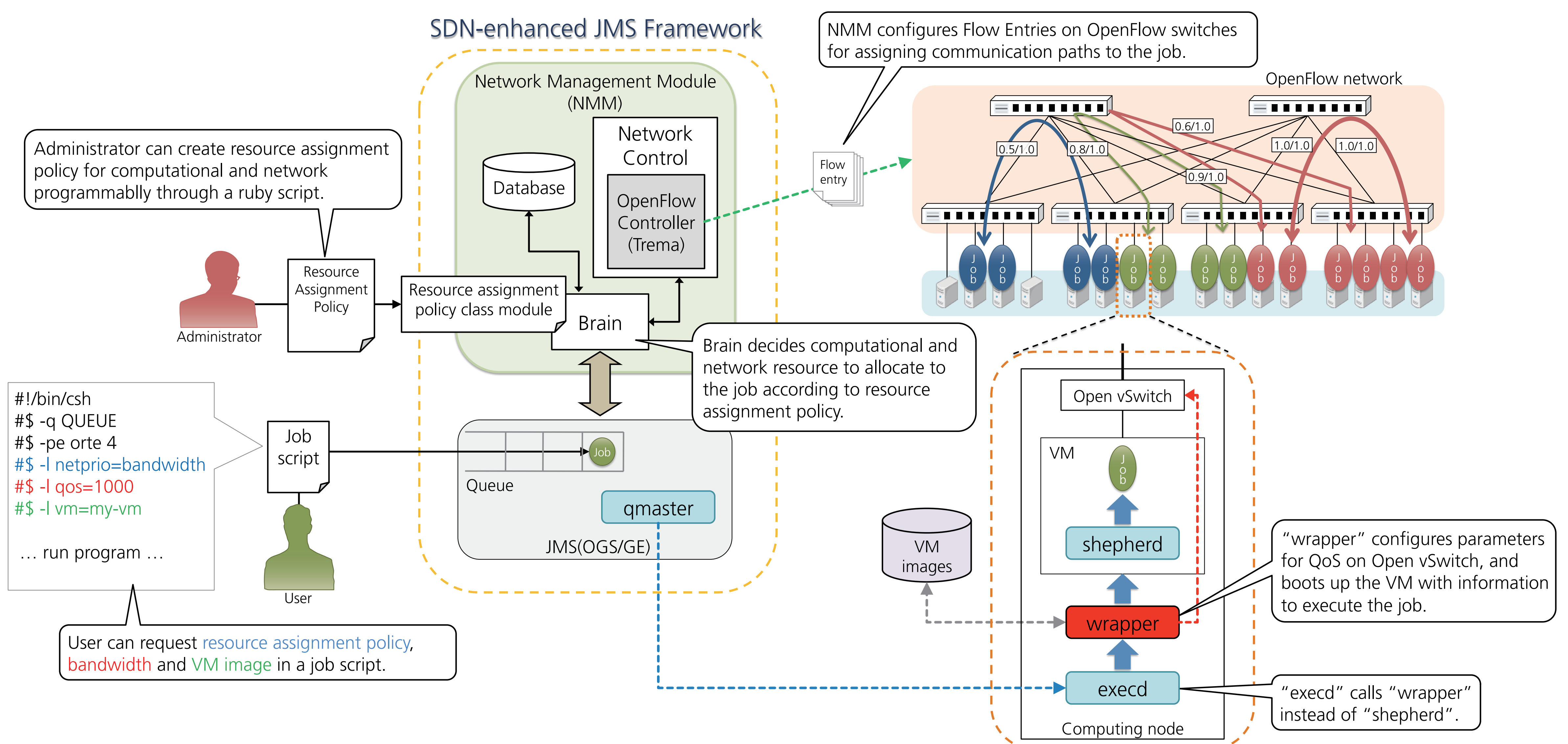
Cybermedia Center, Osaka University, Japan

Motivation and Objectives

Nowadays, usage patterns of users' computations on high-performance cluster system have been diversified for large-scale simulations and analyses in the various science fields. Since an HPC cluster system needs to accommodate multiple jobs concurrently, efficient and flexible resource management is essential for providing high performance computing capabilities for multiple users. However, most Job Management Systems (JMSs) available today, which are deployed on HPC cluster system for computational workload distribution and balancing purposes, determine resource allocation to each jobs only based on computational resources such as CPU and memory. In this research, we realize a novel JMS framework for handling various kinds of resources such as network resources and virtualized computational resources.

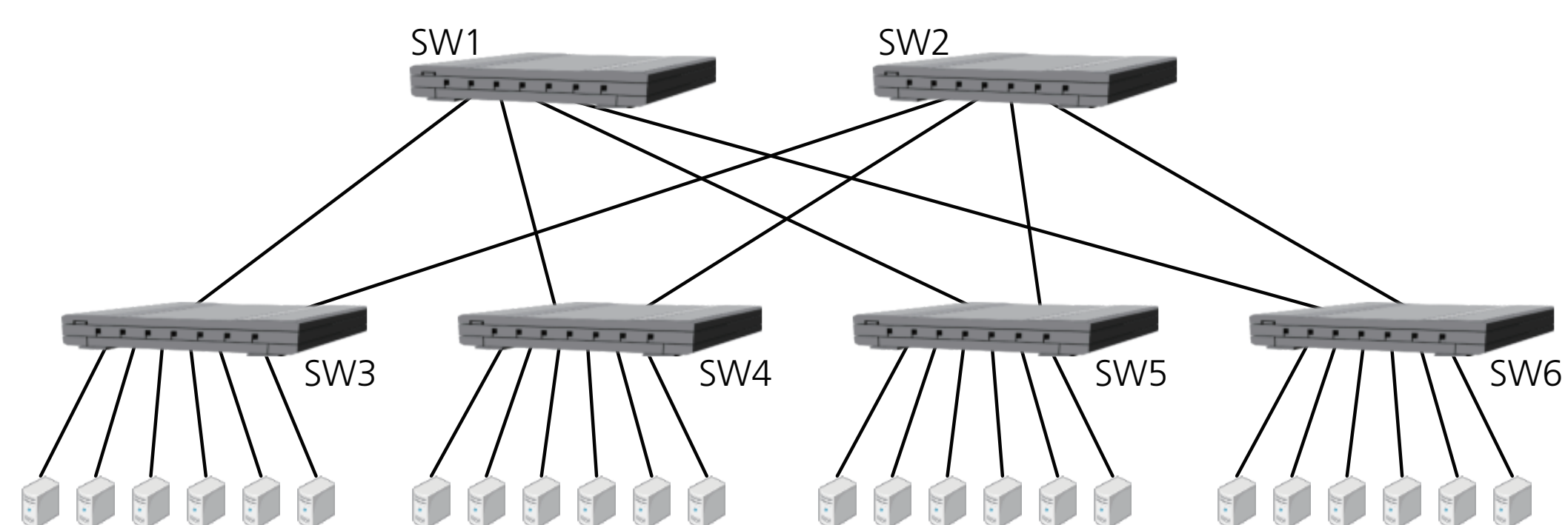
Proposed SDN-enhanced Job Management System Framework

We have been studying and developing a novel network-aware JMS integrated *Software-Defined Networking (SDN)* concept, which can dynamically control an entire network in a centralized manner, into a traditional JMS [1]. The mechanisms to manage the network resources are designed and implemented as *Network Management module (NMM)* leveraging OpenFlow, which is an implementation of the SDN concept. The SDN-enhanced JMS can allocate both computational and network resources to each job according to the resource usage on a cluster system and *Resource Assignment Policy* defined by administrator. Moreover, we have also been developing a mechanism for deploying job's processes to virtual machines (VMs) on computing nodes, and guaranteeing available bandwidth on communication paths allocated to a job by using QoS functions of Open vSwitches (OVSS) connected with VMs.

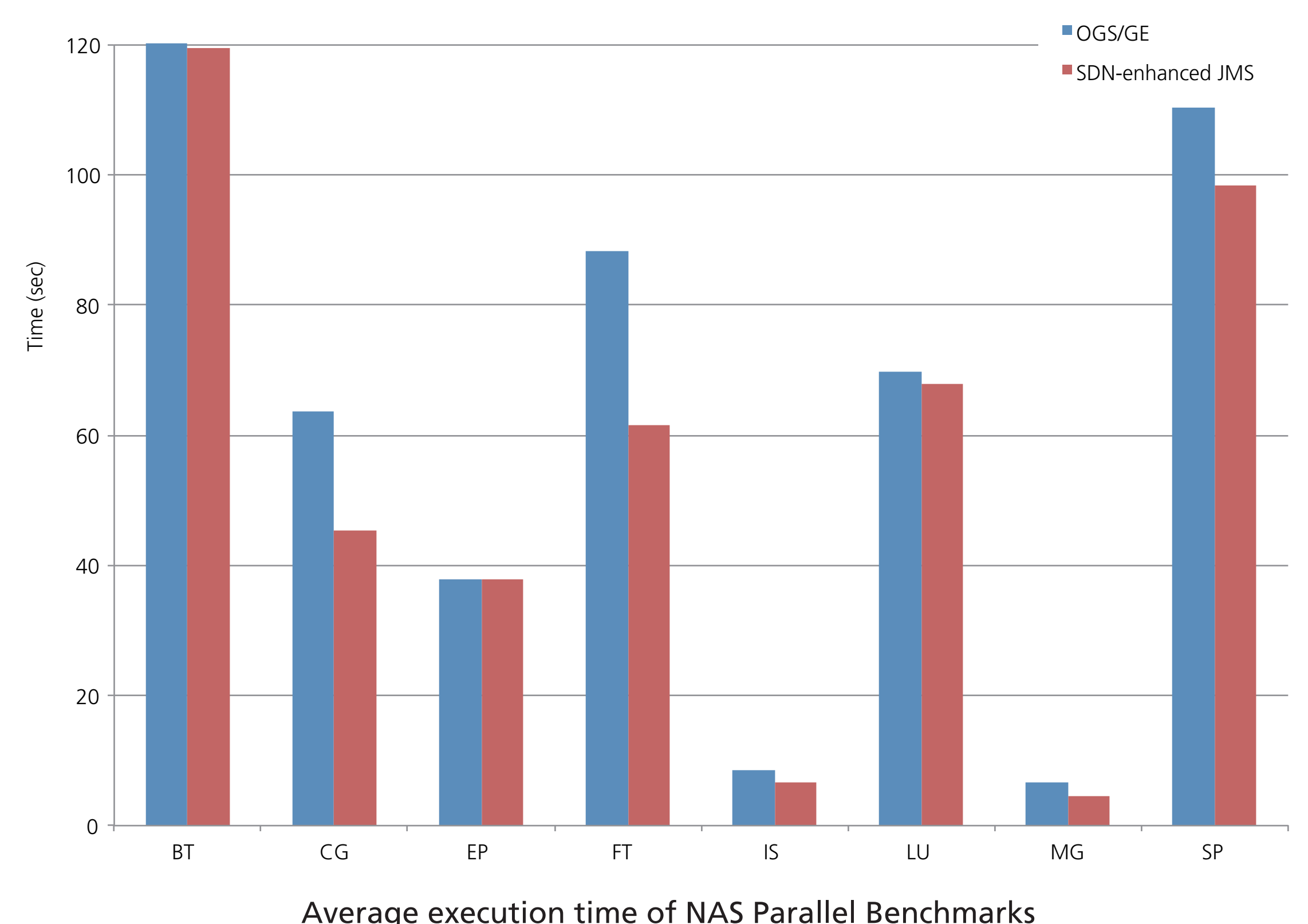


Evaluation

To assess the efficiency of explicitly allocating communication paths to each jobs by the SDN-enhanced JMS, we conducted an experiment to compare the average execution time of jobs on a fat-tree cluster system. Note that processes of jobs were allocated to real computational resources and a single computing node accommodated only one process. In the experiment, we submitted multiple jobs, each of which ran NAS Parallel Benchmarks with class B and generated 4 processes. As a result, our proposed SDN-enhanced JMS succeeded to reduce the average execution time of jobs.



2-tier fat-tree cluster system:
The number of computing nodes is 28 and every link is connected on a Gigabit Ethernet LAN.



Acknowledgments

This research was supported in part by the collaborative research of the National Institute of Information and Communications Technology (NICT) and Osaka University (Research on High Functional Network Platform Technology for Large-scale Distributed Computing).

[1] Y. Watashiba, Y. Kido, S. Date, H. Abe, K. Ichikawa, H. Yamanaka, E. Kawai, H. Takemura, "Prototyping and Evaluation of a Network-aware Job Management System on a Cluster System Leveraging OpenFlow", The 19th IEEE International Conference On Networks (ICON 2013), Dec. 2013.