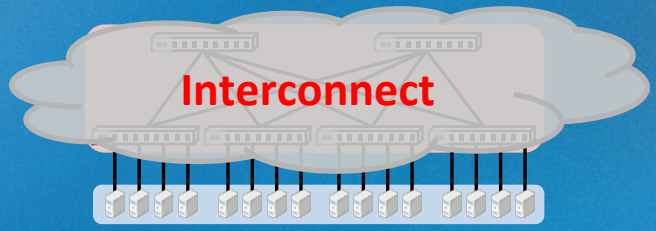
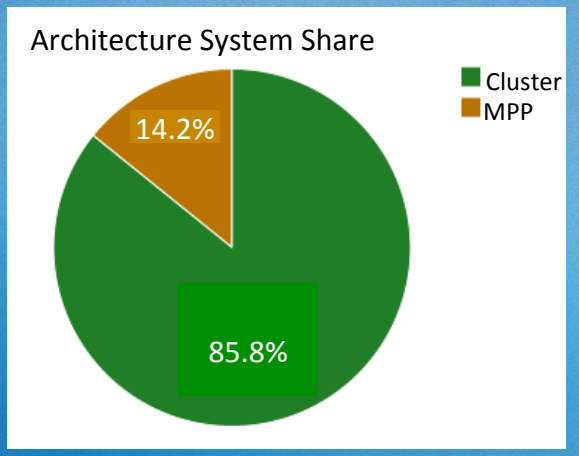
Background

- Current high-performance computing environment
  - The dominant trend is cluster system.



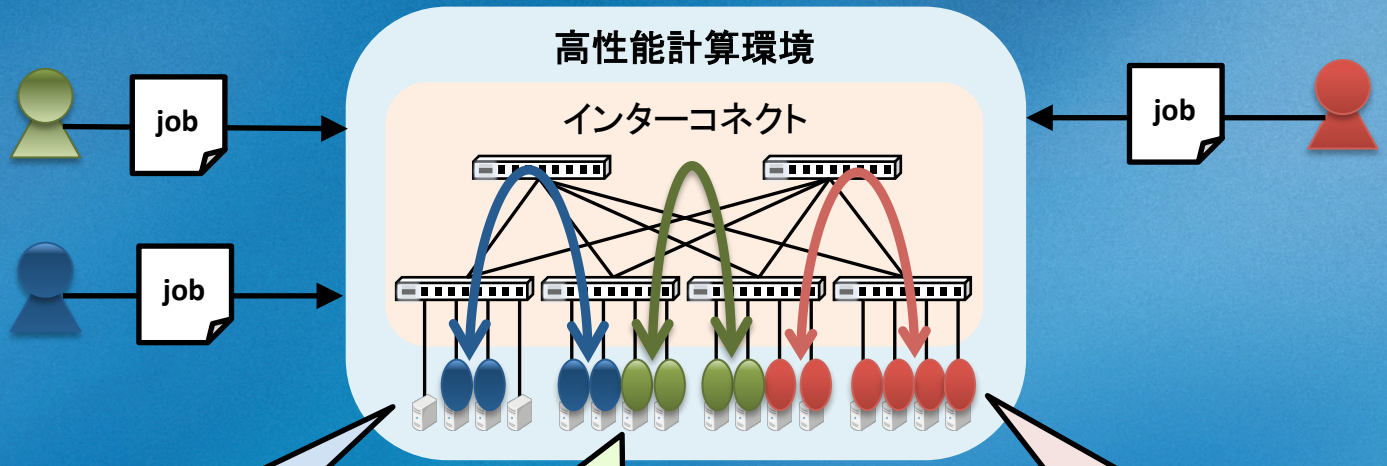


Statistics of TOP500 Supercomputer Sites (November 2014)

- Performance improvement
  - Increase of computing nodes
    => large-scale and distributed

- Structure of interconnect
  - Interconnect also becomes large-scale and complex for achieving high communication performance and fault tolerant.

Management on HPC environment



Efficient execution of parallel and distributed computation for gaining high computational performance.

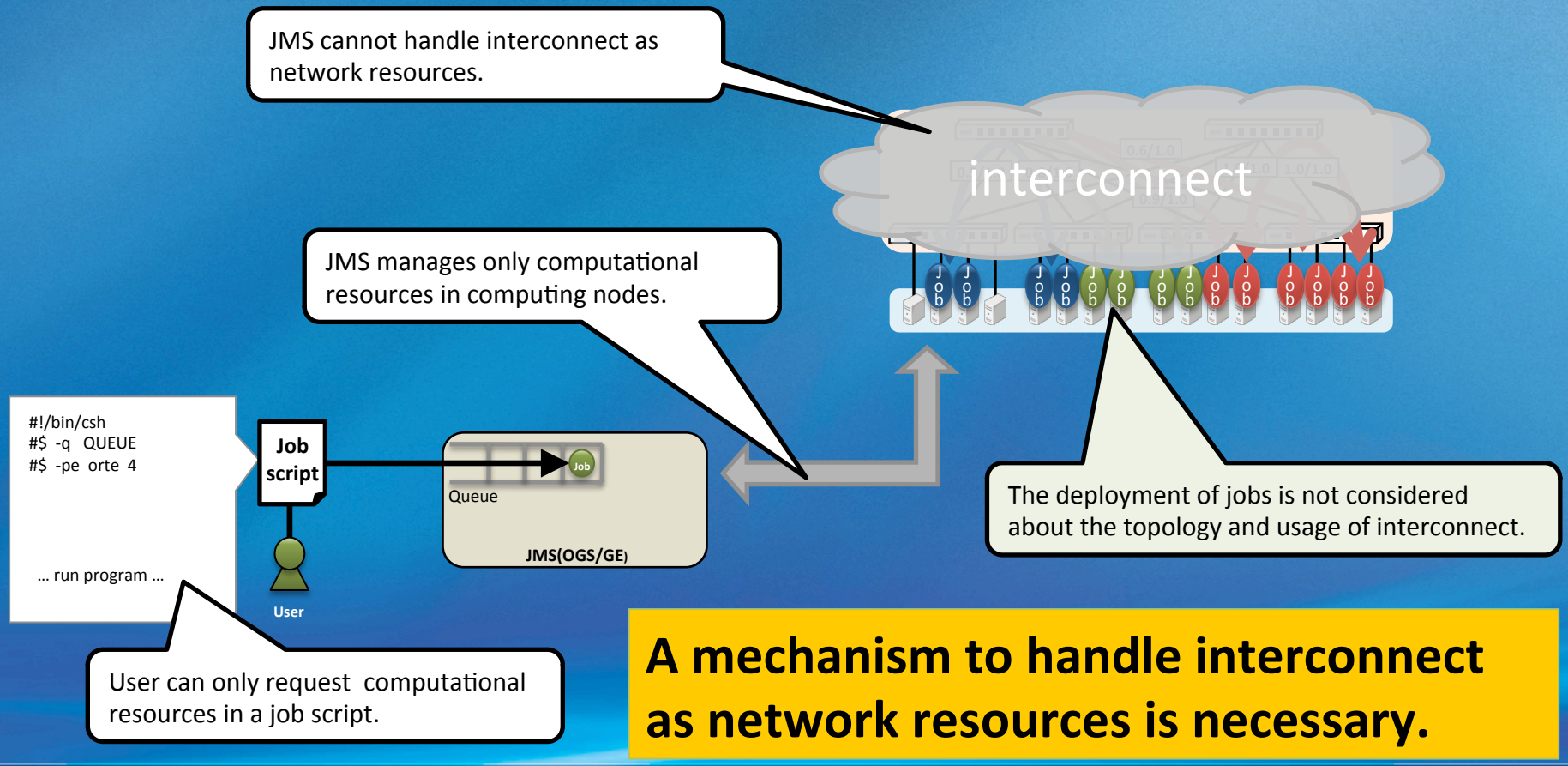
Efficient system operation by running multiple jobs concurrently.

Management and handling of various resources and computational request.

Efficient and flexible resource management on cluster system is essential for gaining high computational performance.

Traditional Resource Management

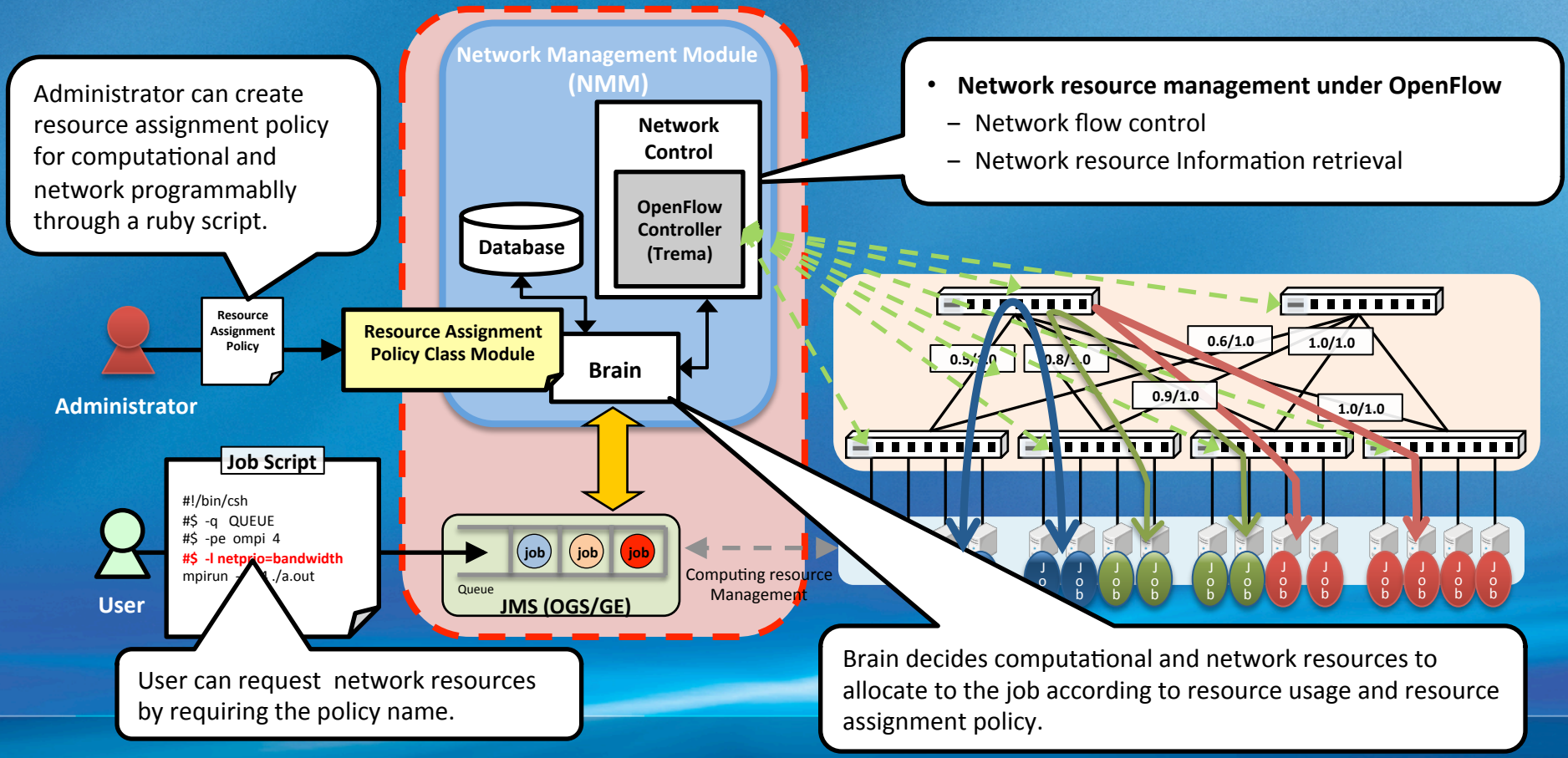
- Job Management System (JMS)
  - Traditional Resource Management System on HPC cluster systems
  - Managing and allocating only computational resources.



SDN-enhanced JMS

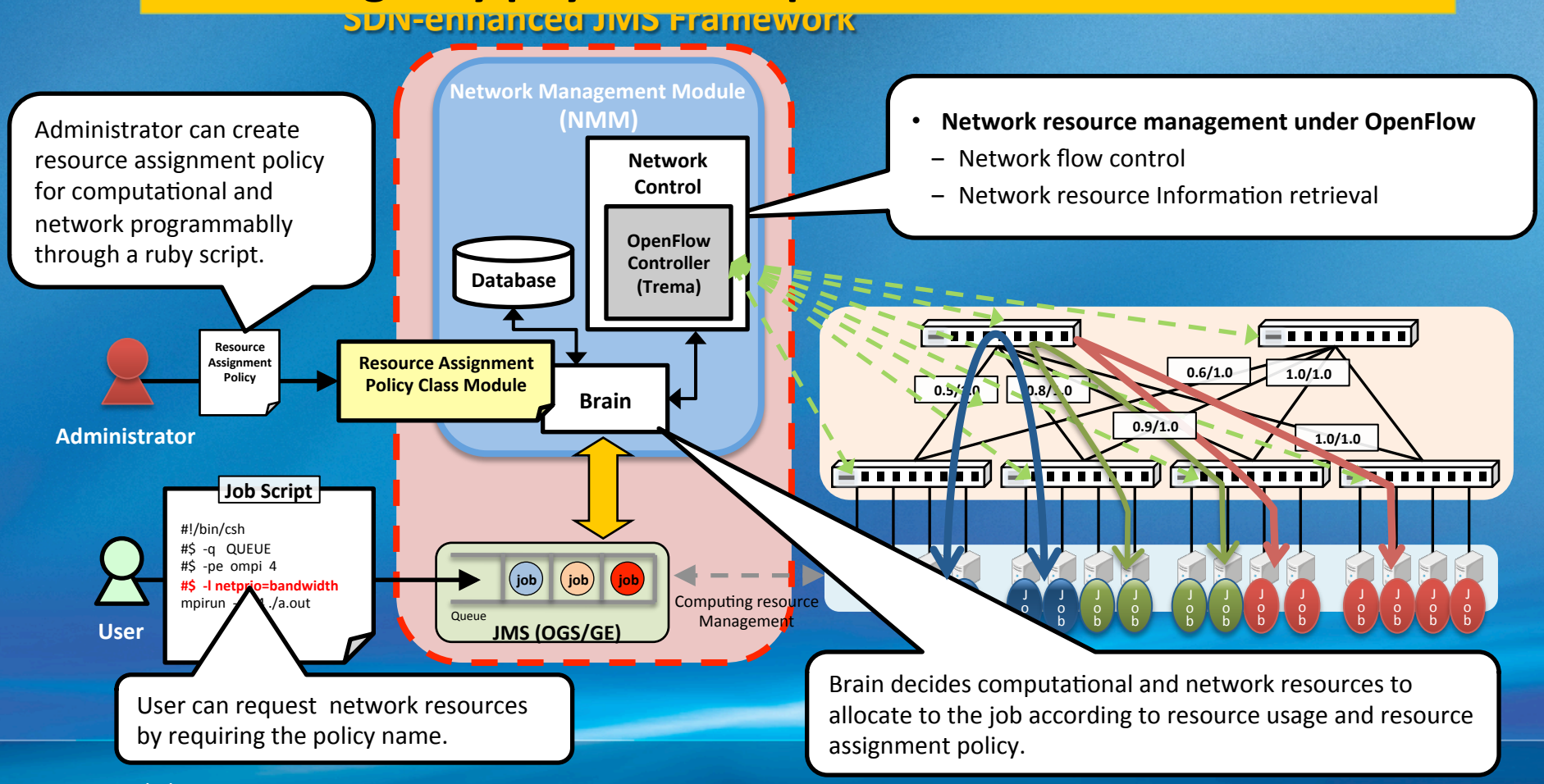
- **Leveraging Software Defined Networking / OpenFlow**
  - Retrieving topology and usage of interconnect
  - Managing communication paths assigned to each job as Flow Entries

SDN-enhanced JMS Framework



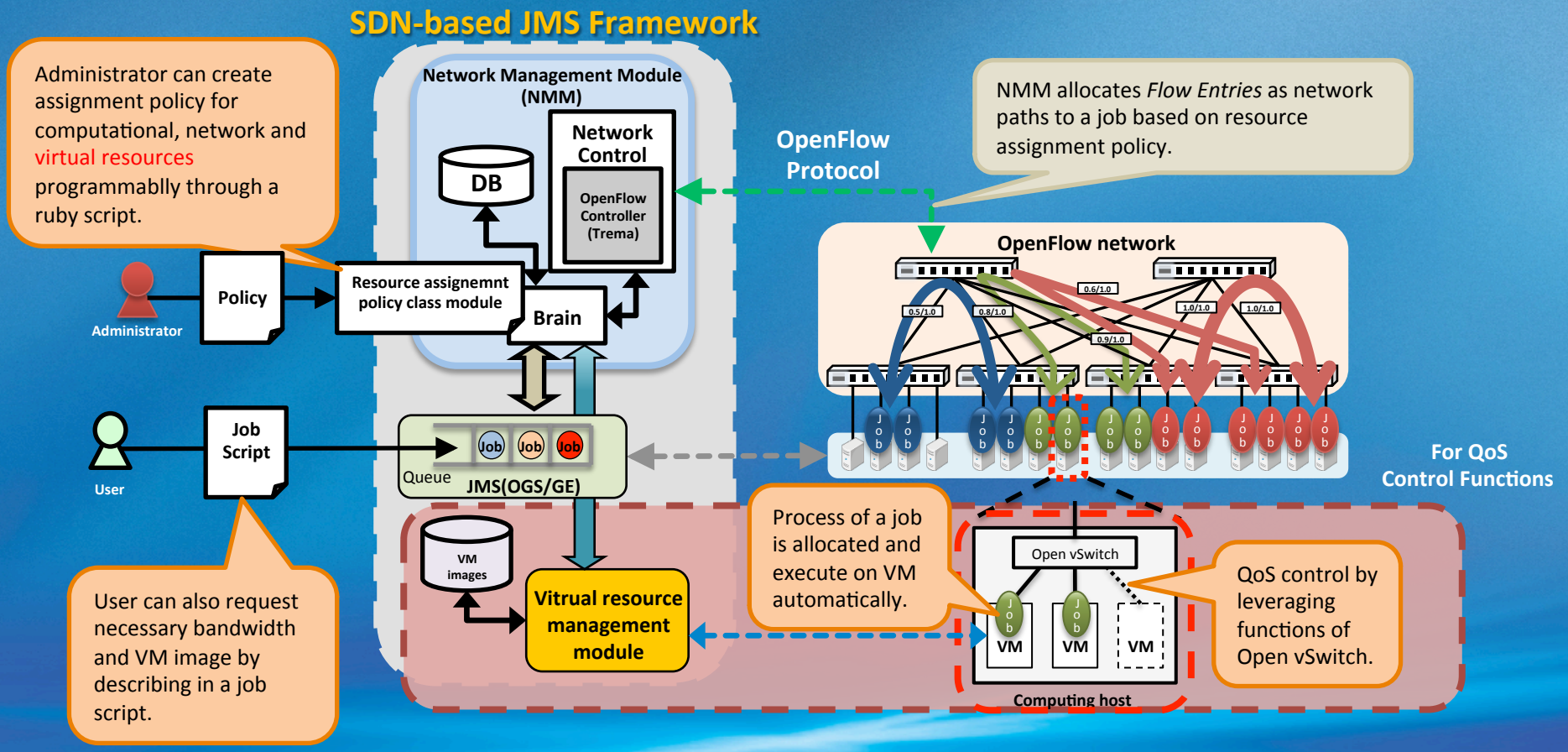
SDN-enhanced JMS

- **- Available bandwidth of communication paths allocated to jobs is not guaranteed**
- **- Handling only physical computational resources**



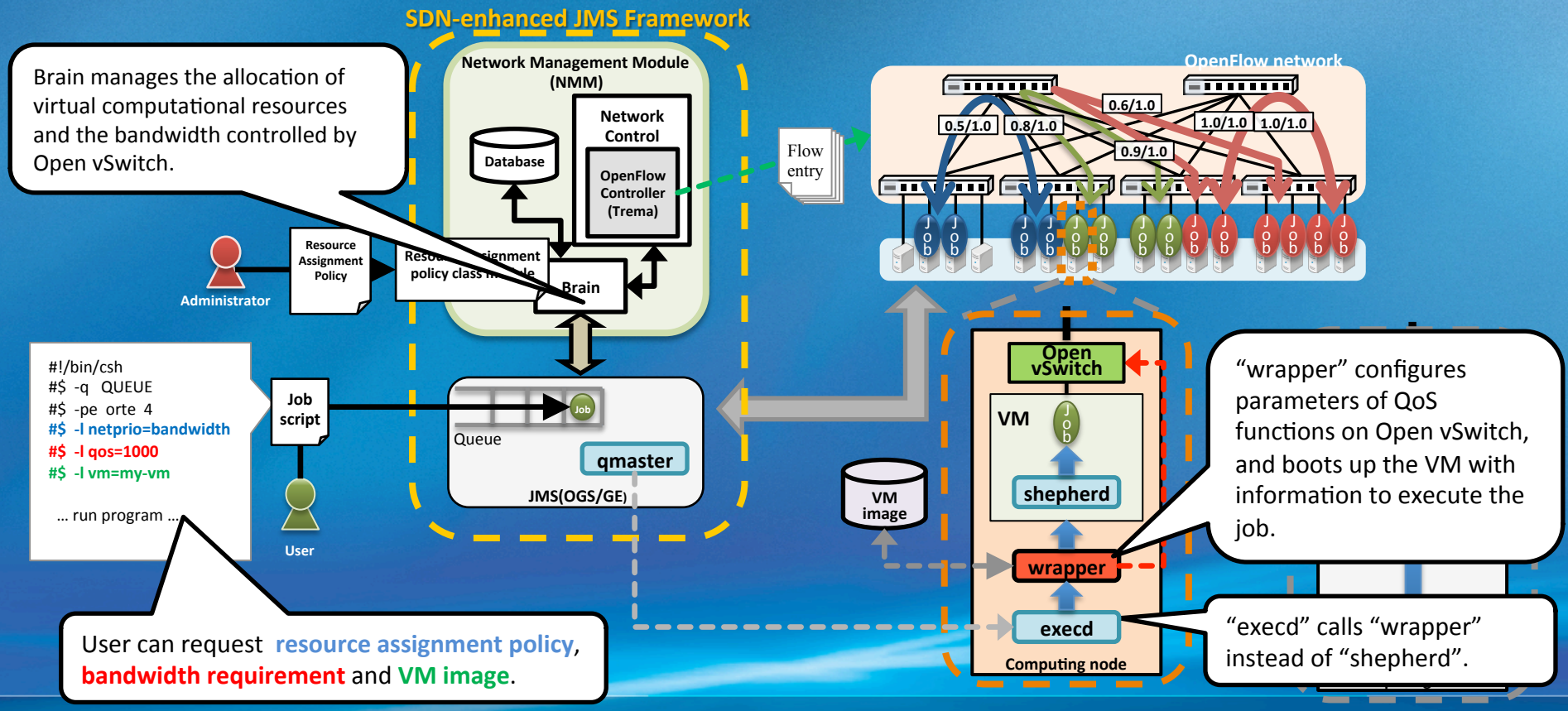
Proposed resource management

- **Allocation of virtual computational resources to a job**
- **QoS control by leveraging functions of Open vSwitch**



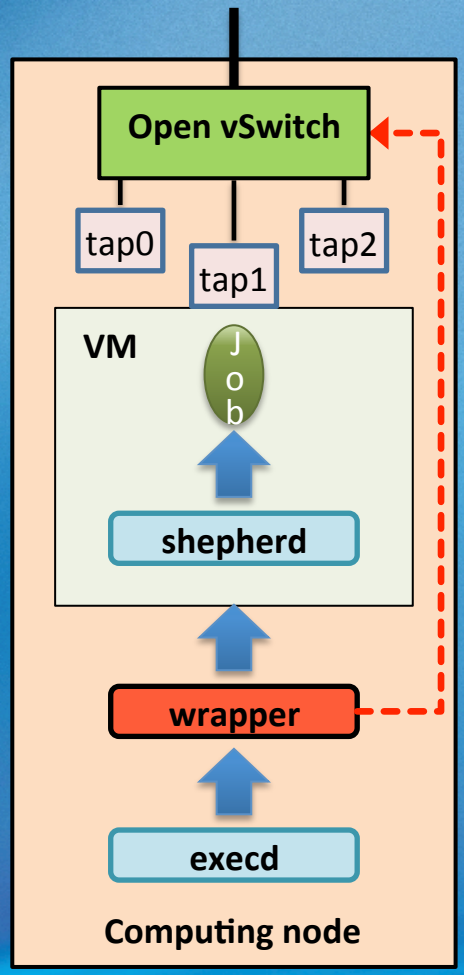
System architecture

- **Implementation of wrapper module**
  - **Boot up user-request VM and deploy process of job on the VM**
  - **Configure parameters of QoS functions on Open vSwitch**
- **QoS control by leveraging functions of Open vSwitch**



wrapper program

- OGS/GE has a functionality for running a wrapper program for shepherd process.
- Role of wrapper program
  - Retrieving information for virtual machine and job management.
  - Making CD image for guest OS on virtual machine
  - Starting up virtual machine and running shepherd
  - Configuring parameter of Open vSwitch
- Tap is managed as slot.



Configure VM environment

- Basic configuration for virtual machine is set up by auto-running CD image generated by SDN-enhanced JMS.

- Contents of CD image
  - option value for booting up virtual machine
  - hostname
  - size of memory
  - network information (IP address, MAC addres, hosts)
  - environment variable for SDN-enhanced JMS

- Modify node list for parallel computation
  - rename hostname of computing node to hostname of virtual machine

Additional resource requirement

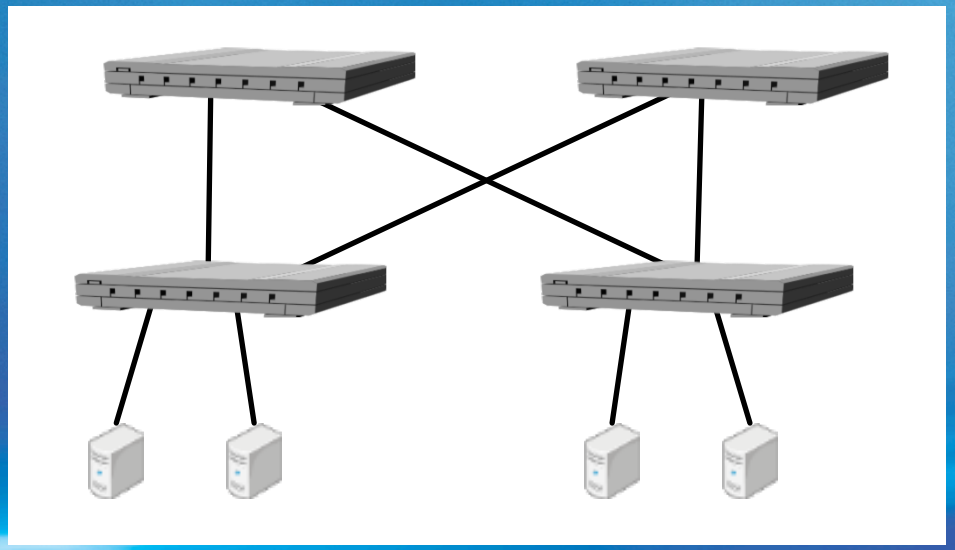
- All of additional resource requirement are provided as “-l” options of *qsub* command.

- Additional option
  - netprio : type of network resource allocation
    ex) bandwidth, hop, and so on
  - qosband : value of QoS rate limiting parameter in OpenvSwitch
  - vmimage : Path of VM image.
  - vmmemory: size of memory to allocate to VM
  - vmopt : boot option of VM

Demo environment

- Constitution of cluster system
  - The number of hosts: 28 nodes
  - Network
    - OpenFlow Switch : NEC UNIVERGE PF5240
    - Bandwidth: 1Gbps

Specifiation of computing host	
OS	CentOS 6.2
CPU	IntelXeon E5-2620(2.00GHz) x2
Memory	64GB
Network	on board Intel I350 GbE



Conclusion and Future work

- This presentation shows implementation and behavior of SDN-enhanced JMS with the functionalities of managing virtual computational resources and bandwidth control.

- Future work
  - Development of resource assignment policy for handling these new resources
  - Evaluation of effectiveness of these functionalities and new resource assignment policy

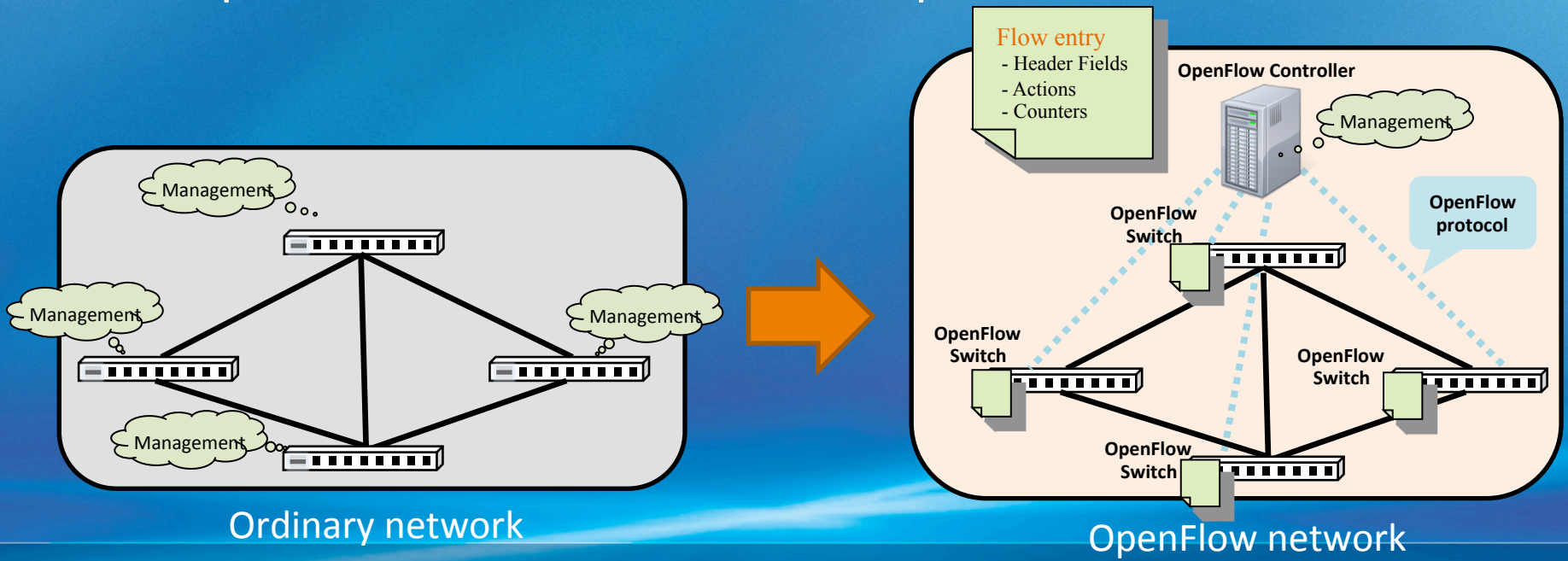
- Cluster system of Osaka University
  - SDN-enhanced JMS will be deployed.
  - This system will be connected with PRAGMA-ENT.

- Home page
  http://hpc-sdn.ime.cmc.osaka-u.ac.jp/sdnjms/
  - source code
  - doccument

Thank you for your attention !

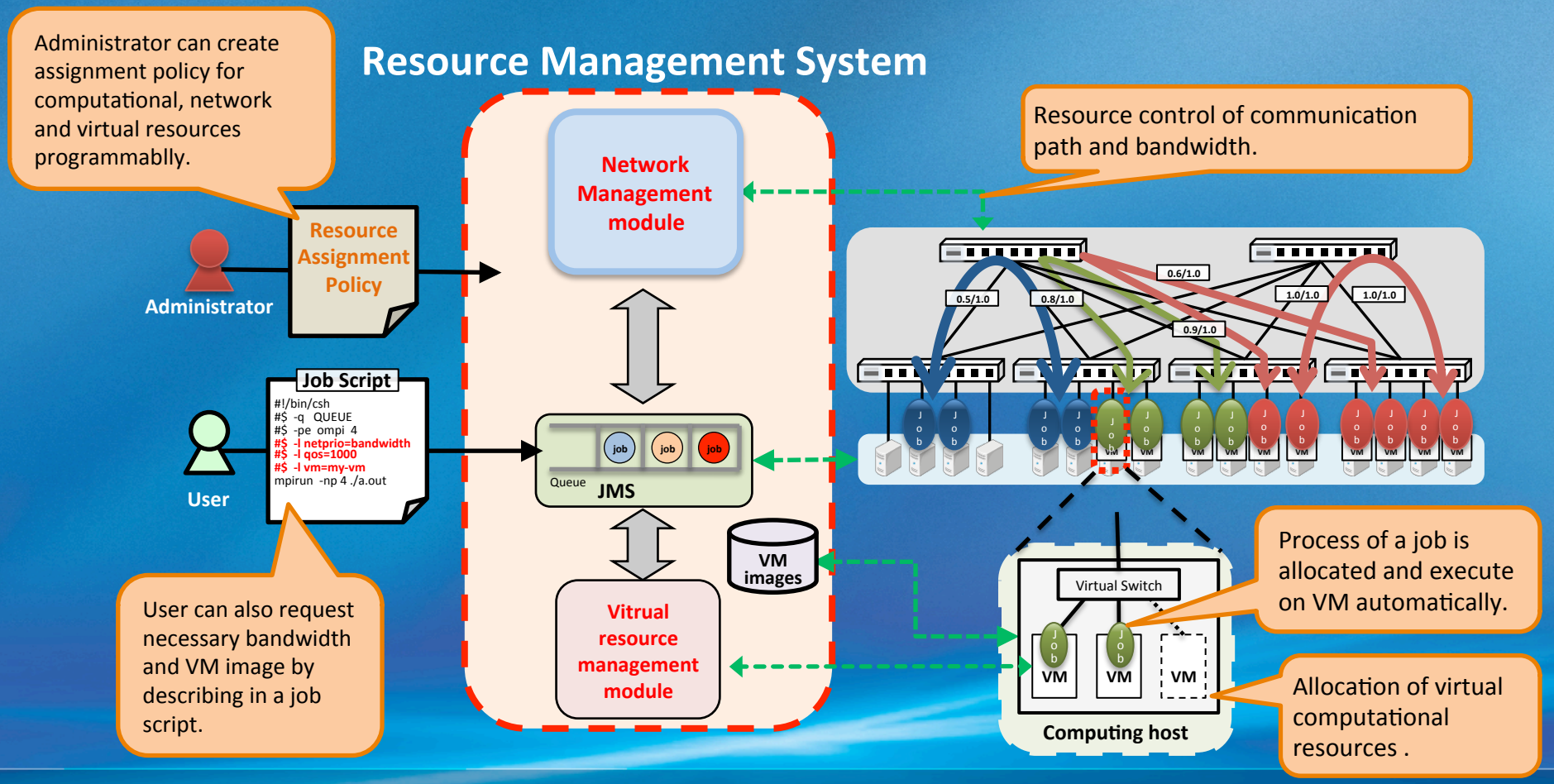
SDN / OpenFlow

- Software-Defined Networking（SDN）
  - Decomposing to control function and data transfer function
  - Aggregating network management into a software application
- SOpenFlow
  - Implementation of the SDN concept



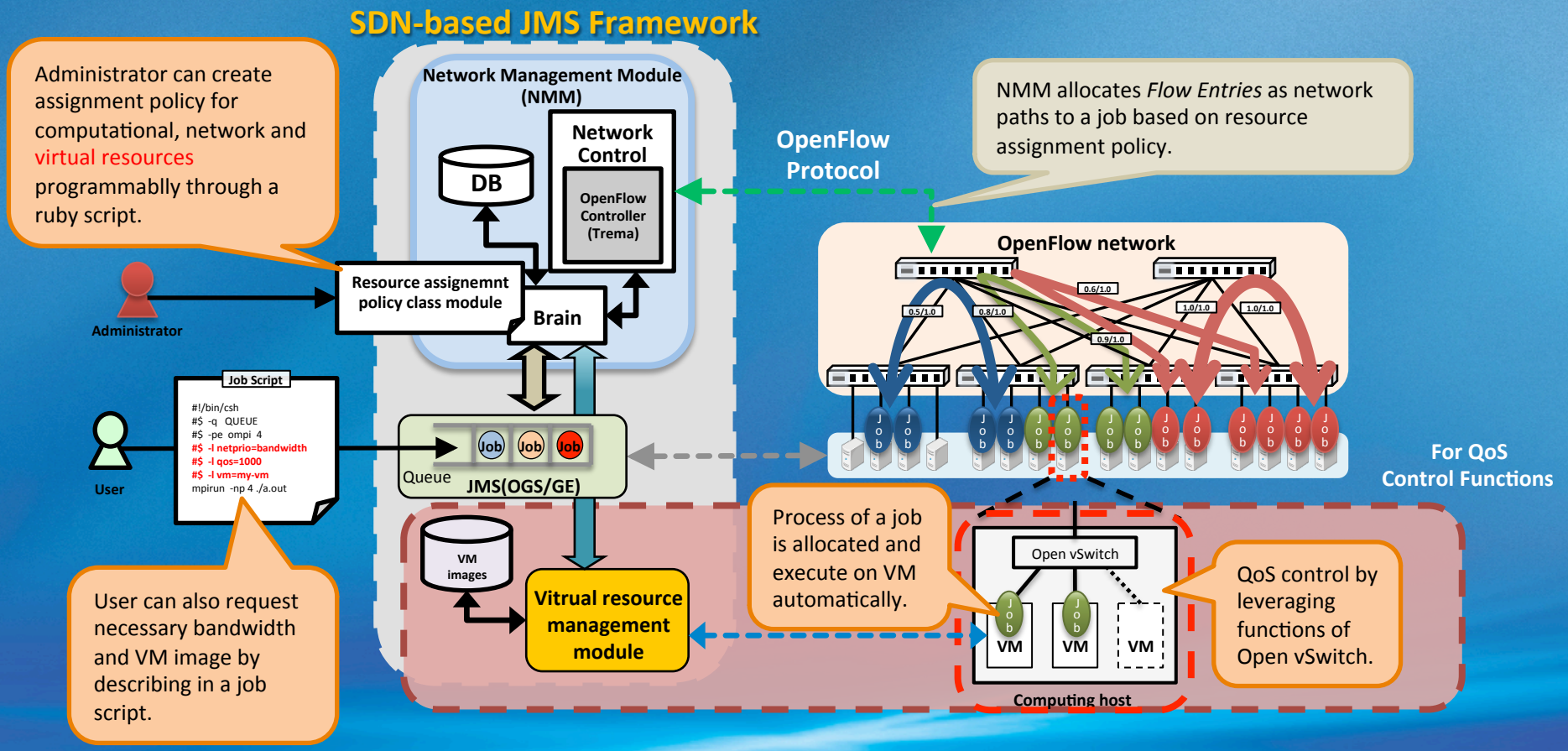
Research goal

- **Efficient and flexible resource management system**
  - Handling various computational and network resources.



Proposed resource management

- **Allocation of virtual computational resources to a job**
- **QoS control by leveraging functions of Open vSwitch**



Architecture

- Implementation of wrapper module
  - Boot up user-request VM and deploy process of job on the VM
  - Configure parameters of QoS functions on Open vSwitch
- QoS control by leveraging functions of Open vSwitch