

Experience and Expertise of Designing SDN MPI_Bcast

**Khureltulga Dashdavaa, Susumu Date, Hiroaki Yamanaka,
Eiji Kawai, Yasuhiro Watashiba, Kohei Ichikawa,
Hirotake Abe, Shinji Shimojo**

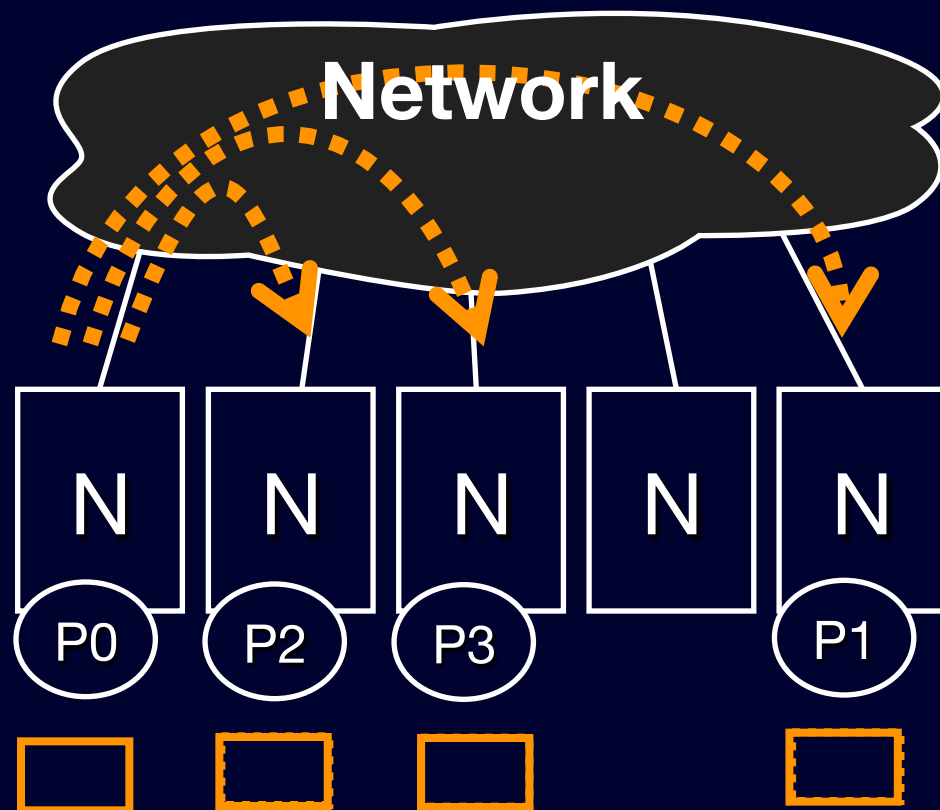
Outline

- **Background and Problem**
- **Software-Defined Network**
- **Research goal**
- **Proposed method: SDN_MPI_Bcast**
- **Evaluation**
- **Conclusion**

Background and Problem

- MPI is playing great role in parallel programming.
- However, MPI implementation is not optimized for common network hardware, such as Gigabit Ethernet.

Example: MPI Bcast (basic collective communication)



N : Computing node

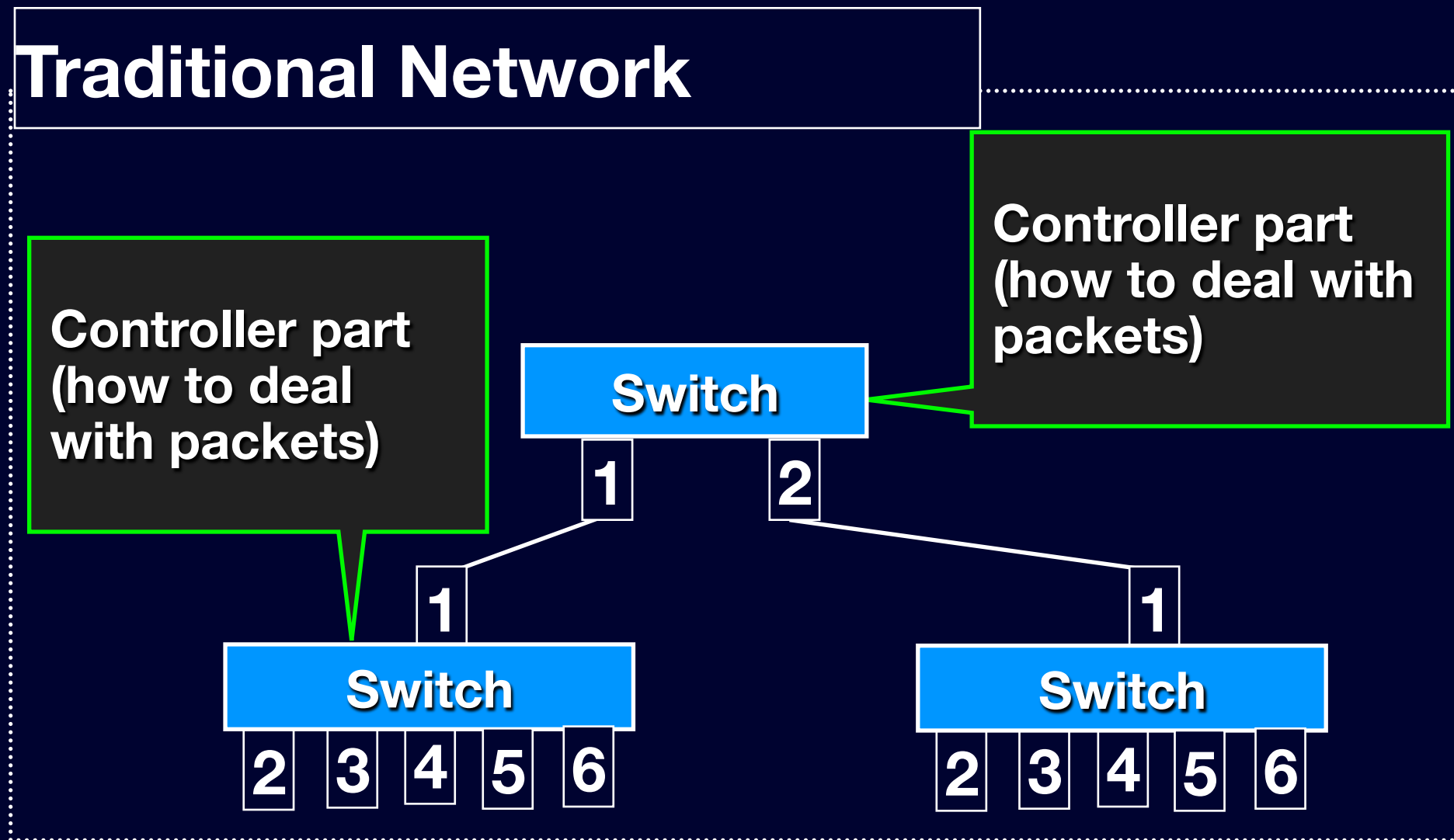
P0, P1, P2, P3 : Running process

Main reason is that network has “**static**” resource, and cannot be controlled “**dynamically**”.

➡ Collective communication uses multiple unicast-based communication.

Emergence of Software-Defined Network

Software-Defined Network (SDN) :
a new concept of network technology

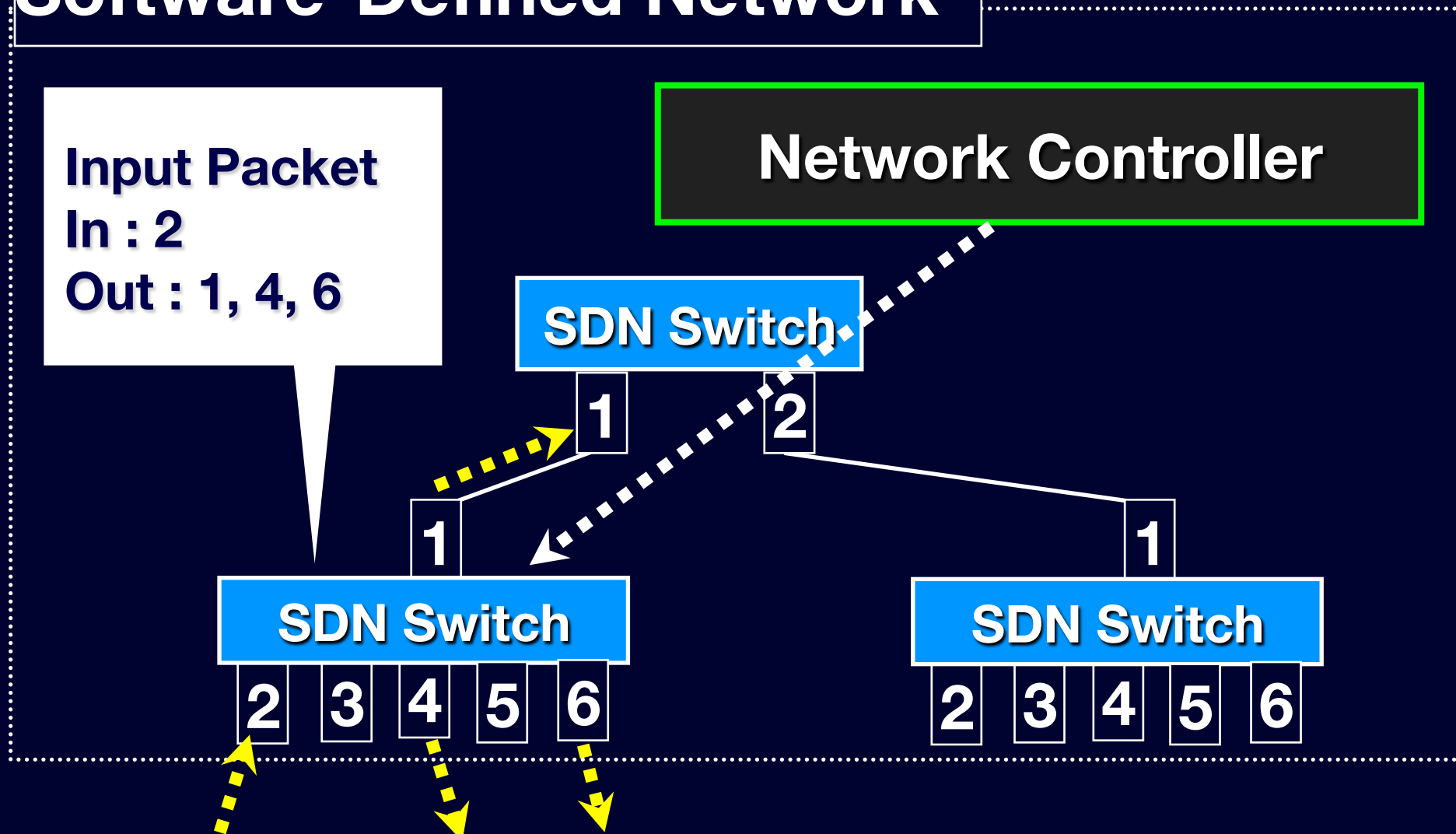


Emergence of Software-Defined Network

Software-Defined Network (SDN) :
a new concept of network technology

- In SDN, controller parts decoupled from devices (switches) and centralized to one place.

Software-Defined Network



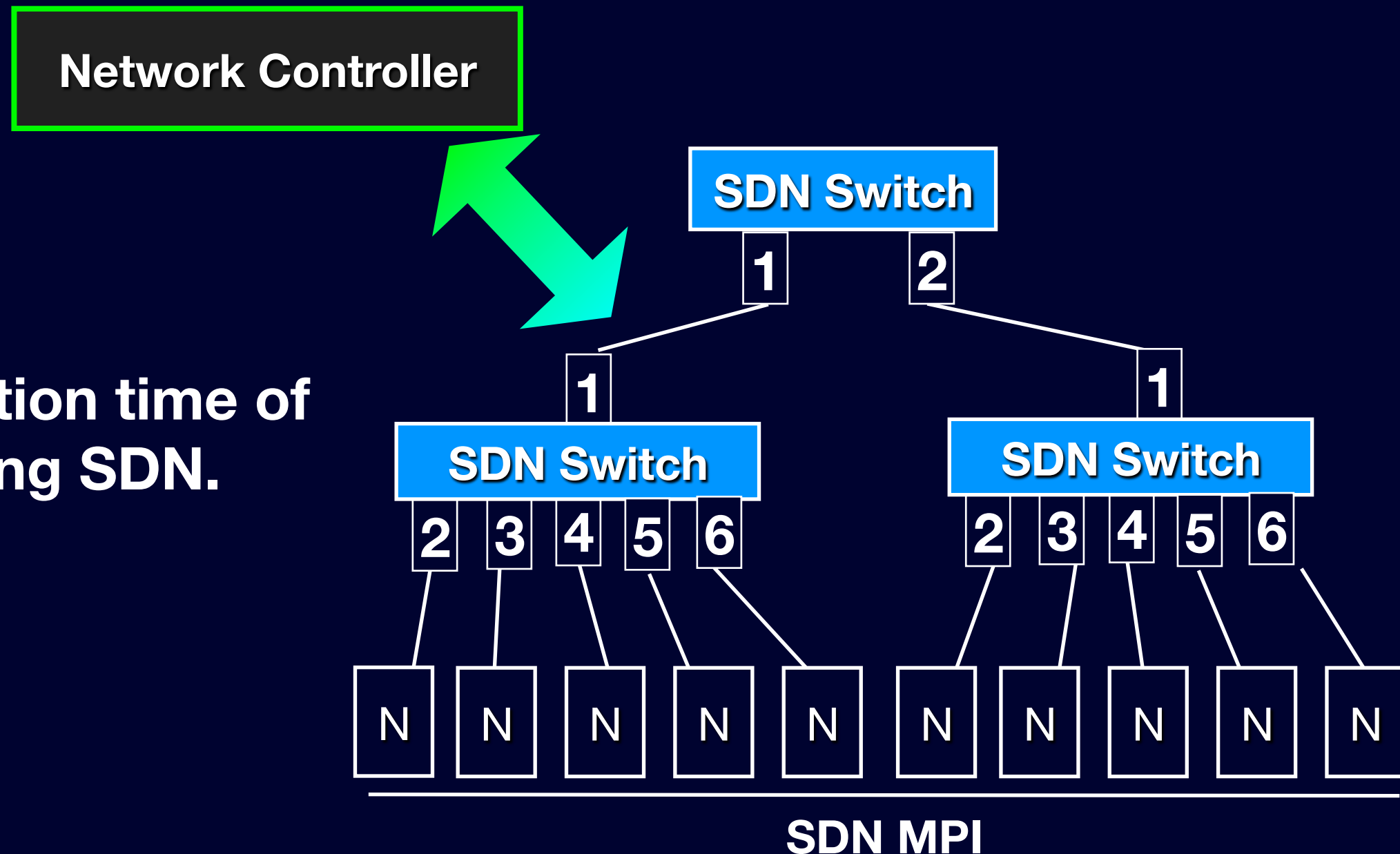
Example:

- Controls out going port or ports of incoming packet on switches.

➔ Network controller can control network “**dynamically**”.

Research goal

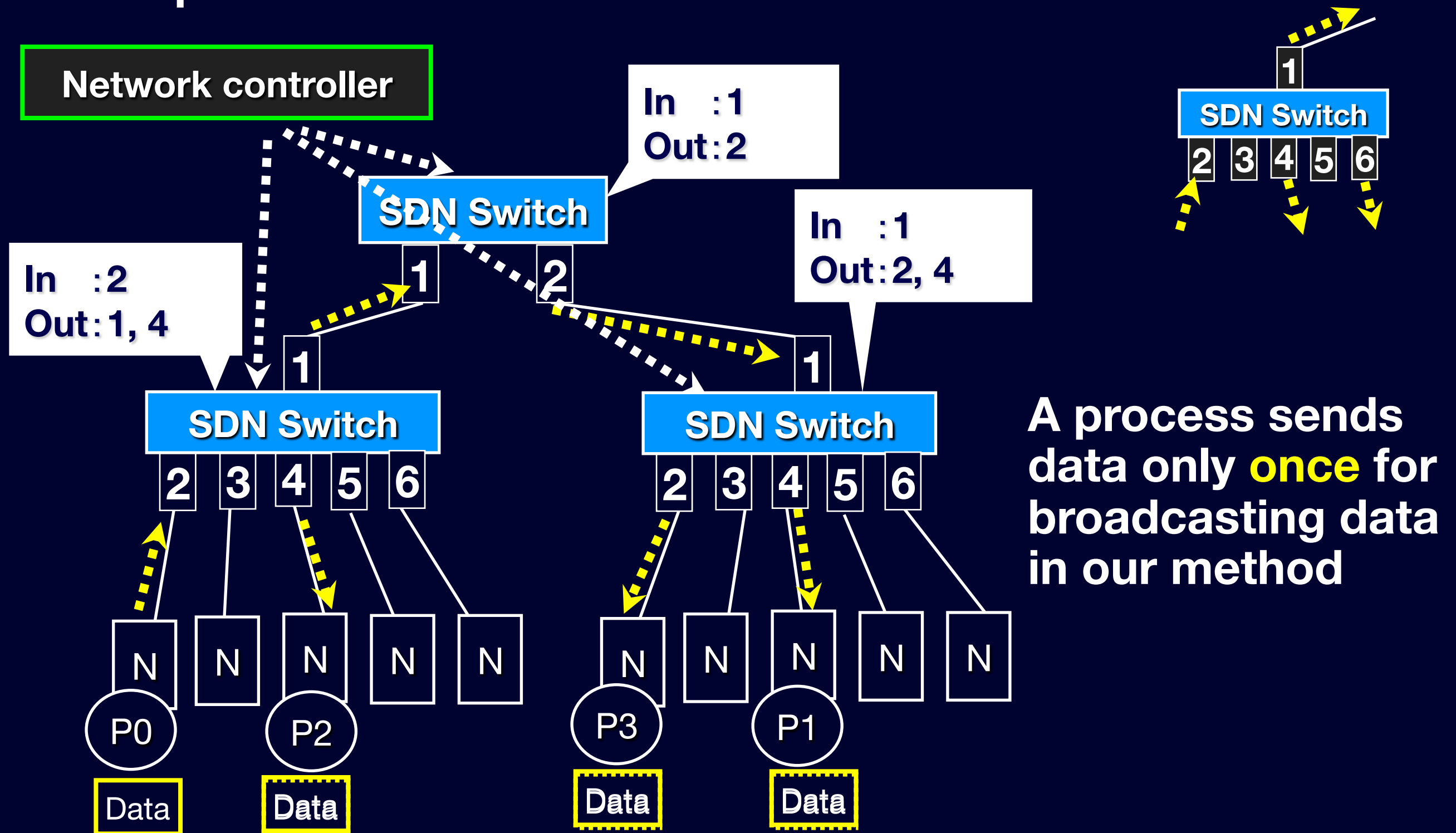
- Integrate “**dynamic**” controller ability of SDN into MPI in order to overturn the assumption that network is “**static**” resource.



- Reduce execution time of MPI_Bcast using SDN.

Proposed Method: SDN_MPI_Bcast

An SDN feature that Network controller can choose outgoing ports of incoming packet on switches is adapted to MPI_Bcast as a duplication action.

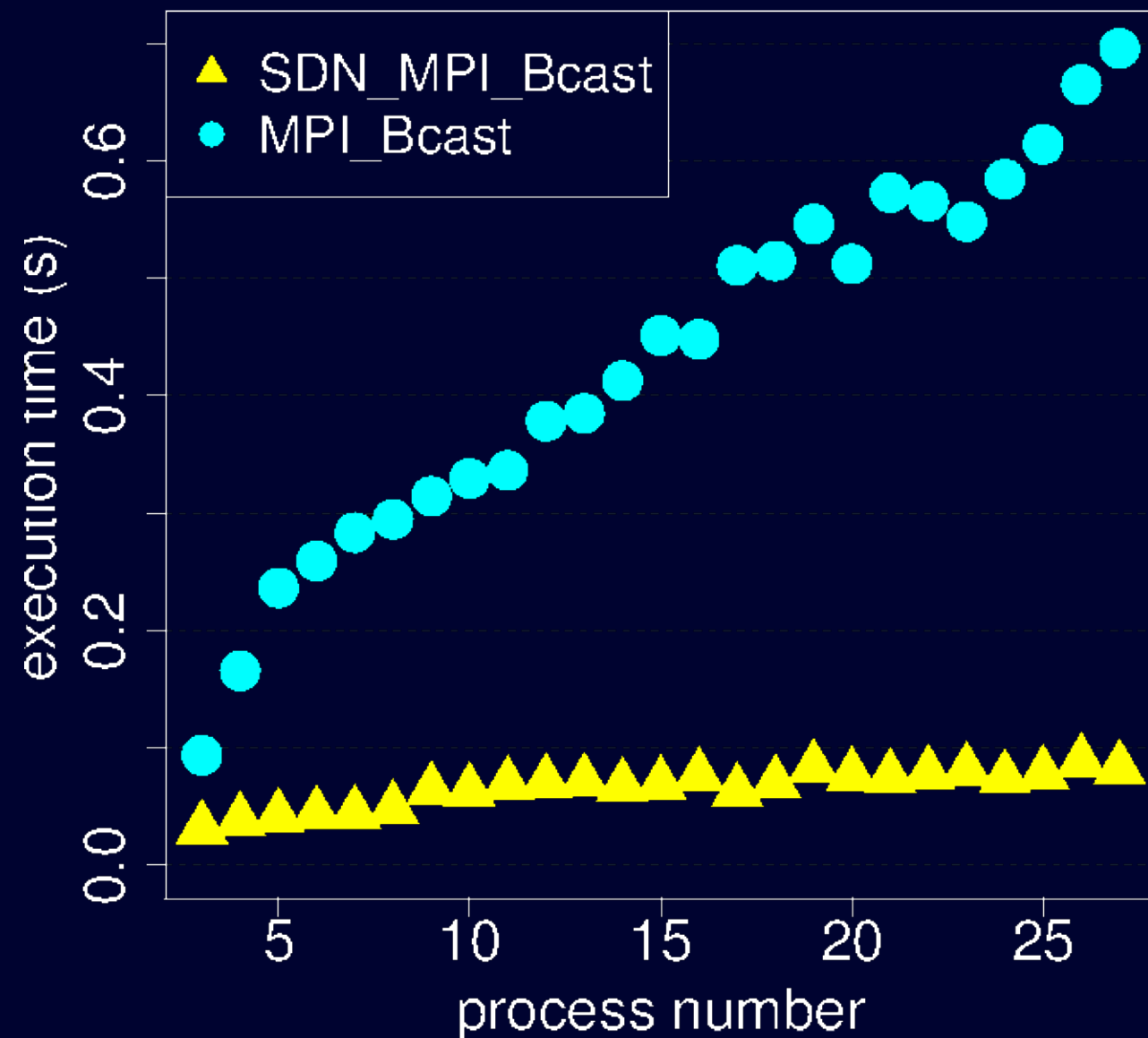


Experiment Result

- Execution time of SDN_MPI_Bcast and MPI_Bcast.

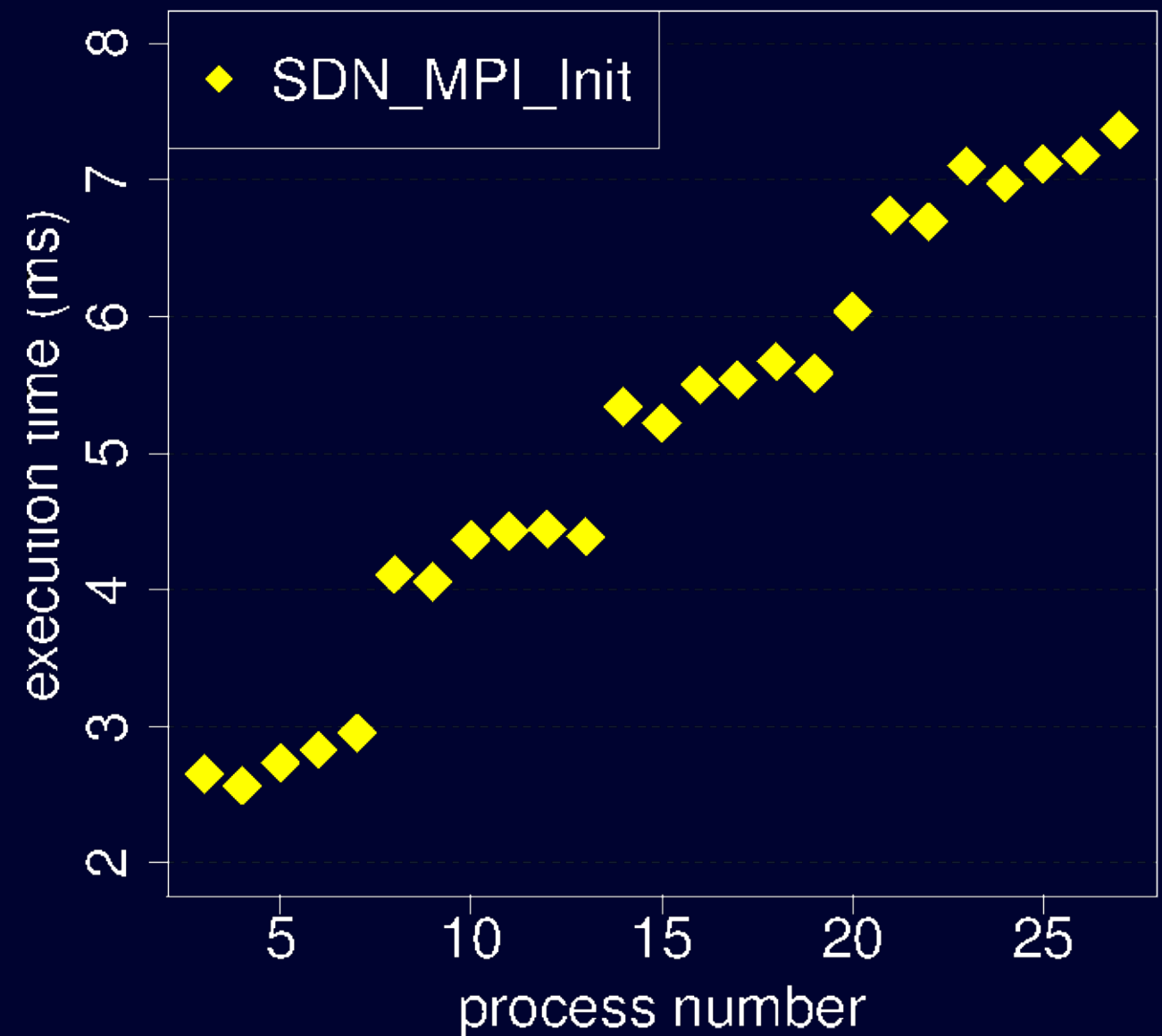
Change process numbers

- Data size: 15.6 KiB (constant)



- Execution time on initialization.
 - Install Duplication Rules.

Change process number



Conclusion

- **Introduced architecture of MPI_Bcast Leveraged Software-Defined Network.**
- **Our experiment result confirmed that SDN_MPI_Bcast is feasible.**
- **Future Works**
 - **Stability and reliability of prototype implementation.**
 - **Aim to SDN MPI.**