An MPI concept with efficient control of network functionality based on SDN

Shimojo Lab., Osaka University
Khureltulga Dashdavaa, Munkhdorj Baatarsuren,
Keichi Takahashi, Yoshiyuki Kido, and Shinji Shimojo

Message Passing Interface

- Message Passing Interface (MPI) is important programming model since widely spread of cluster system in HPC.

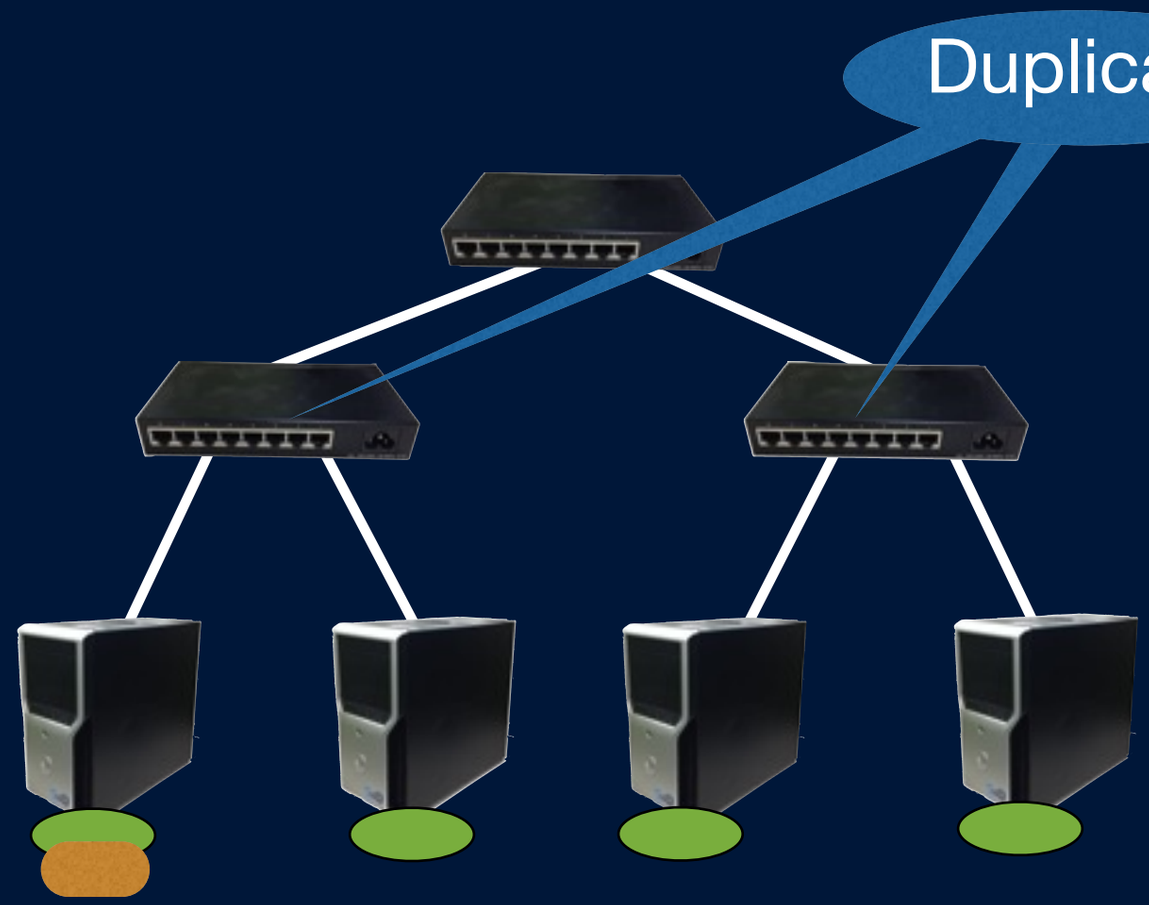
Cluster is parallel and distributed HPC architecture that uses a lot of low performance computers.

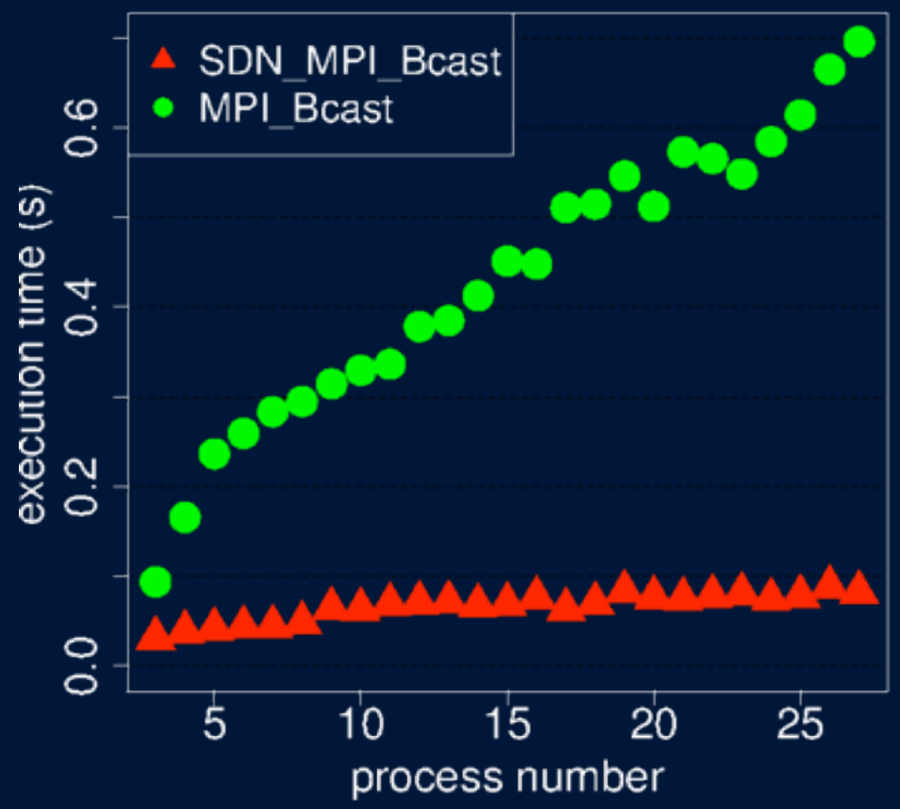
- top500 HPC's 86.8% is cluster.

- However, MPI collective communication's overhead becomes problem as scale of cluster grows.

Previous works

- SDN-MPI-Bcast achieves to boost MPI-Bcast up leveraging Software Defined Networking (SDN).
  - sets multicast tree path from source to others.

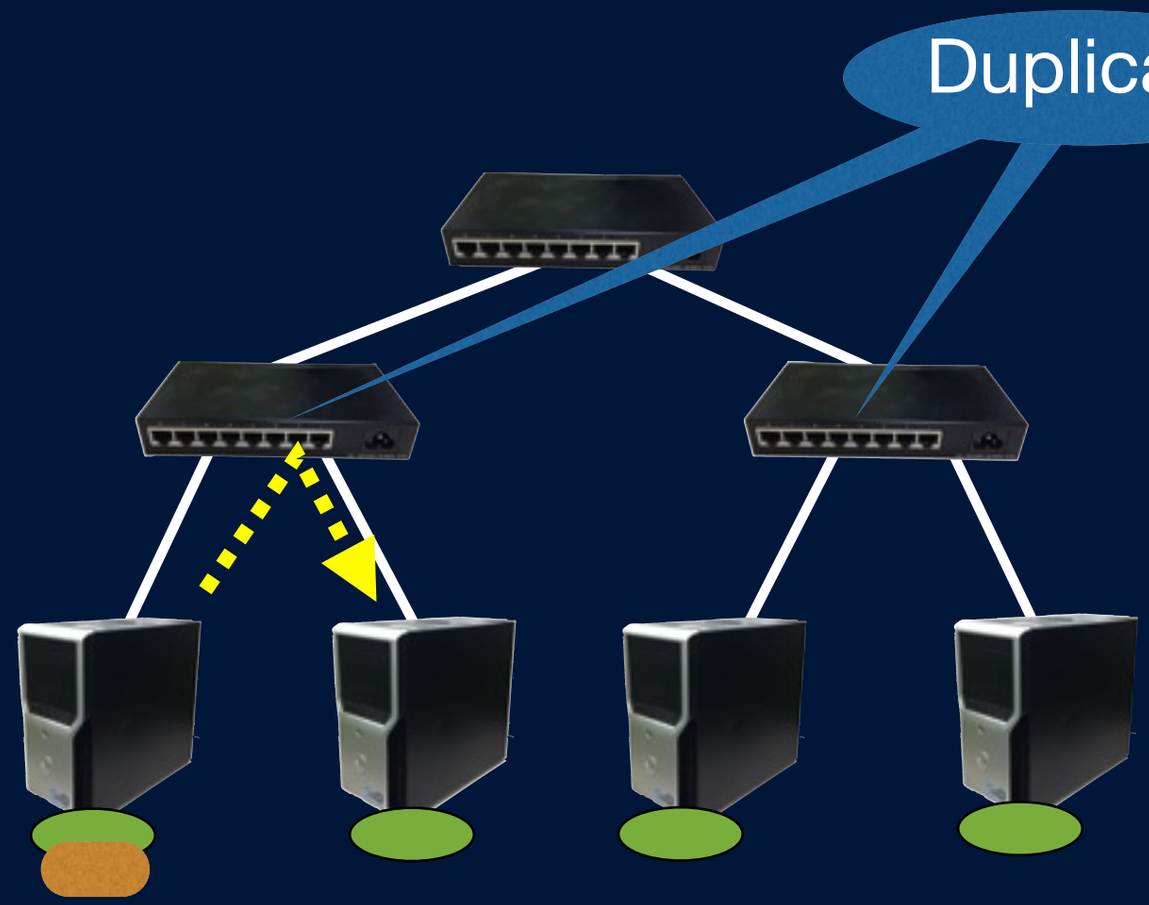


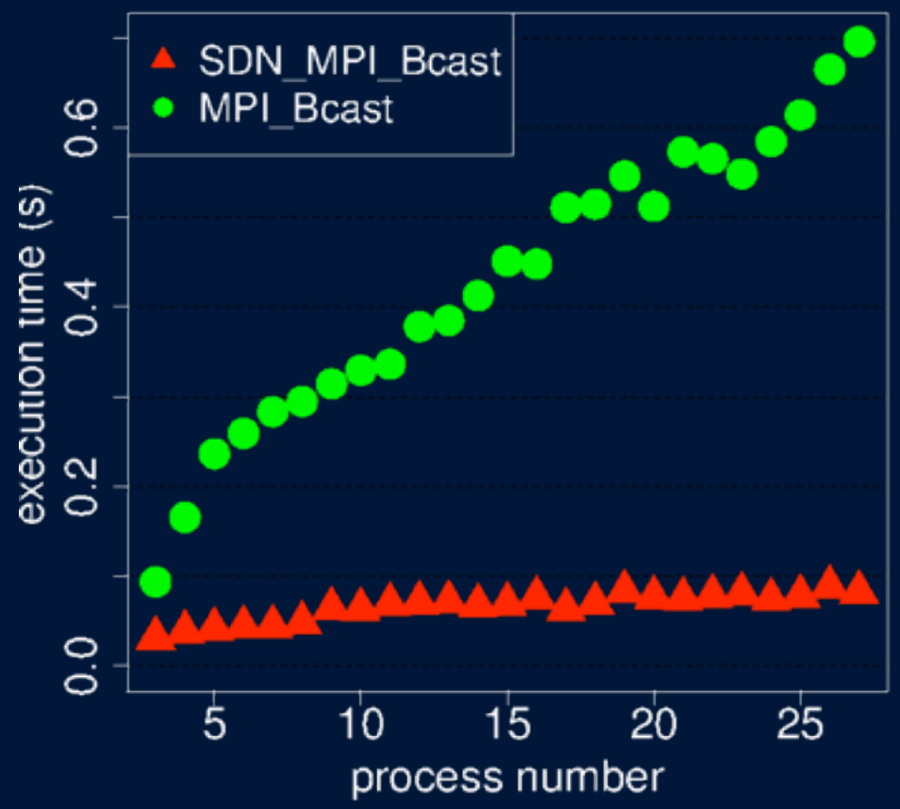


- We achieve to boost up MPI-Allreduce and MPI-Reduce leveraging SDN.

Previous works

- SDN-MPI-Bcast achieves to boost MPI-Bcast up leveraging Software Defined Networking (SDN).
  - sets multicast tree path from source to others.

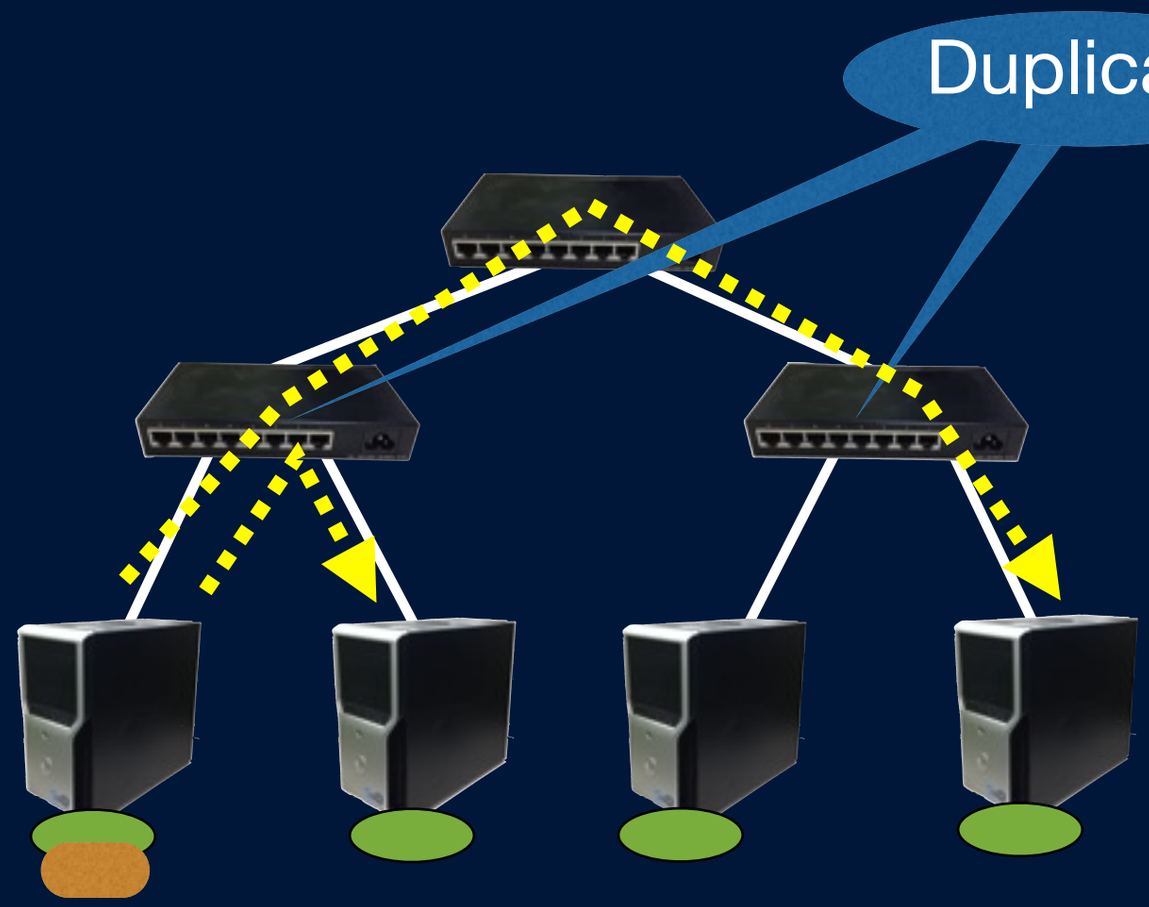


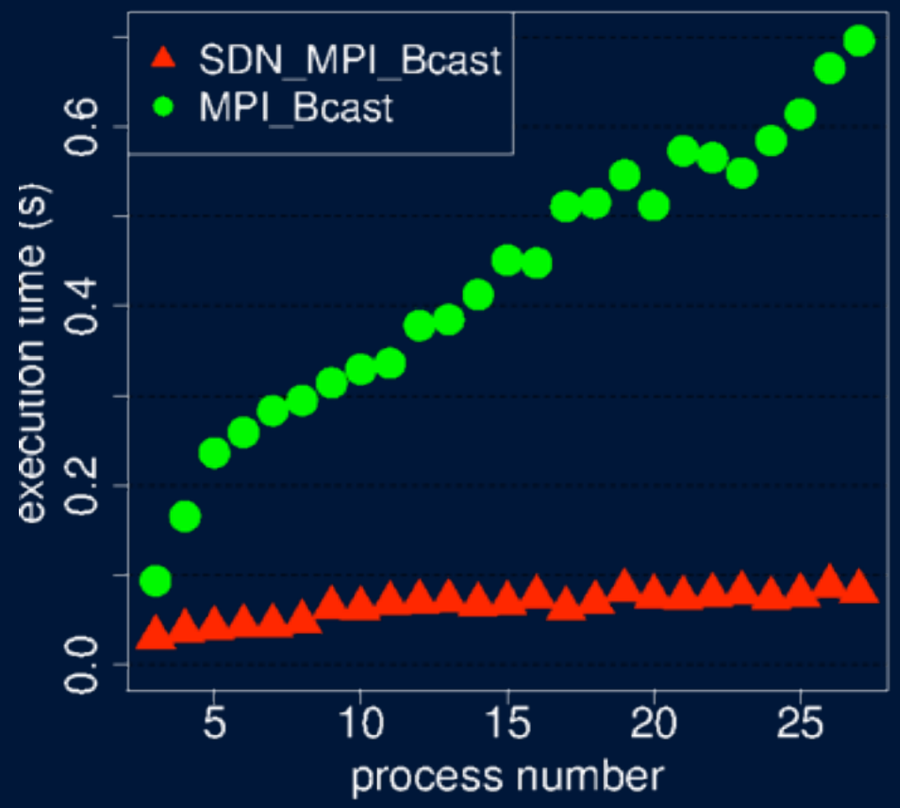


- We achieve to boost up MPI-Allreduce and MPI-Reduce leveraging SDN.

Previous works

- SDN-MPI-Bcast achieves to boost MPI-Bcast up leveraging Software Defined Networking (SDN).
  - sets multicast tree path from source to others.

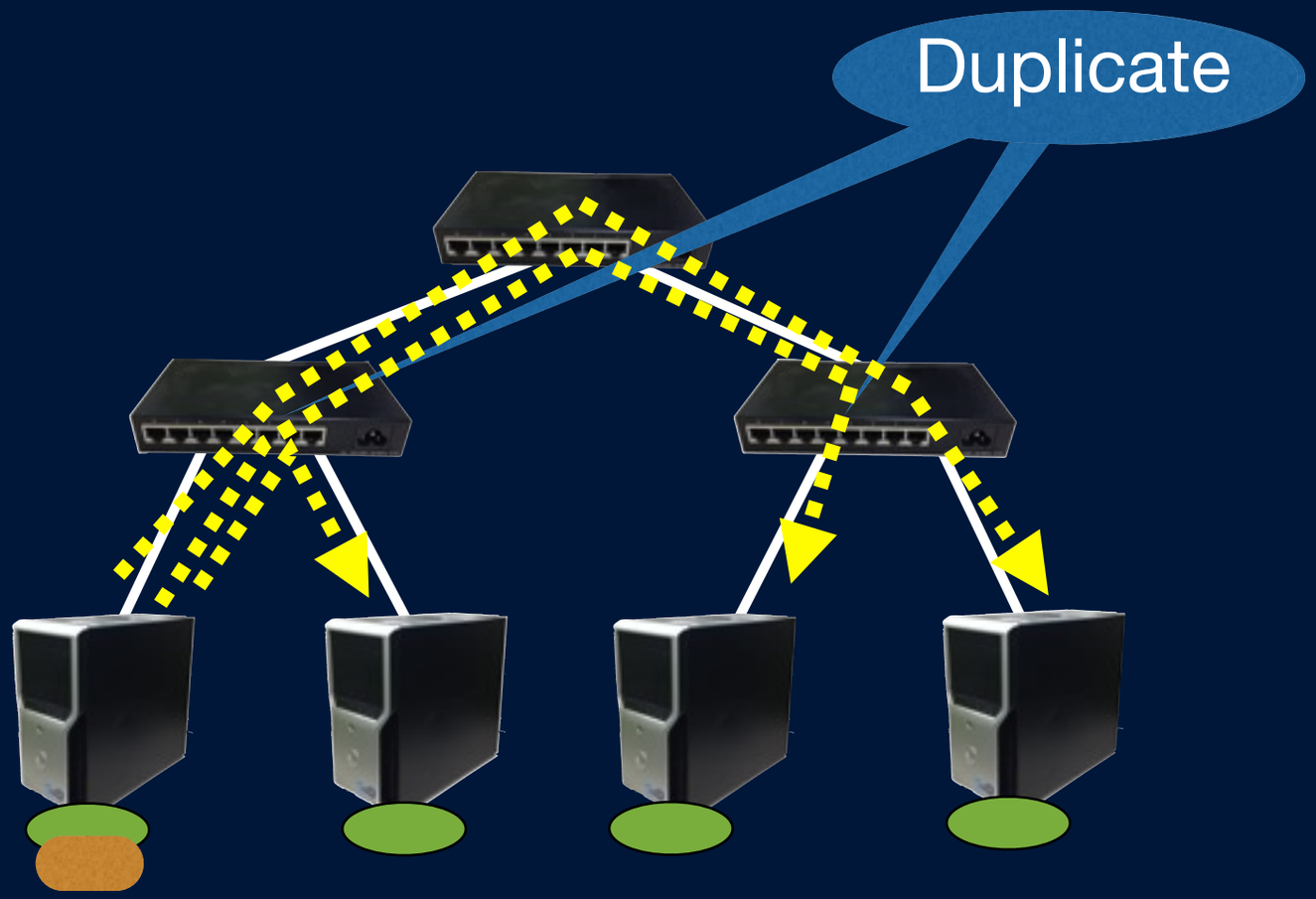


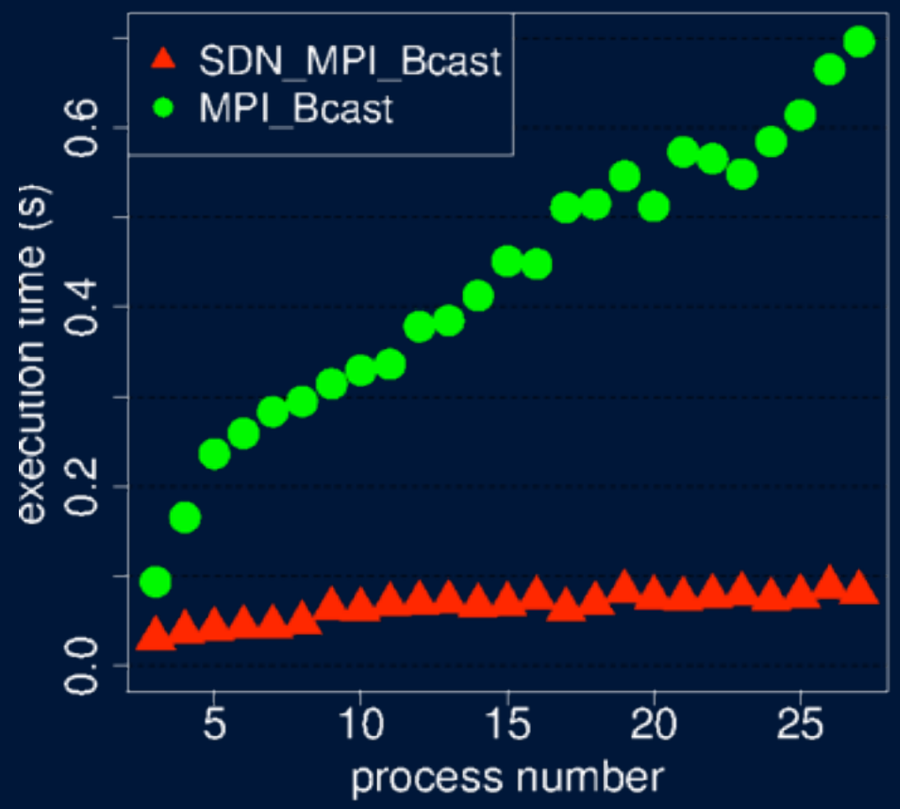


- We achieve to boost up MPI-Allreduce and MPI-Reduce leveraging SDN.

Previous works

- SDN-MPI-Bcast achieves to boost MPI-Bcast up leveraging Software Defined Networking (SDN).
  - sets multicast tree path from source to others.

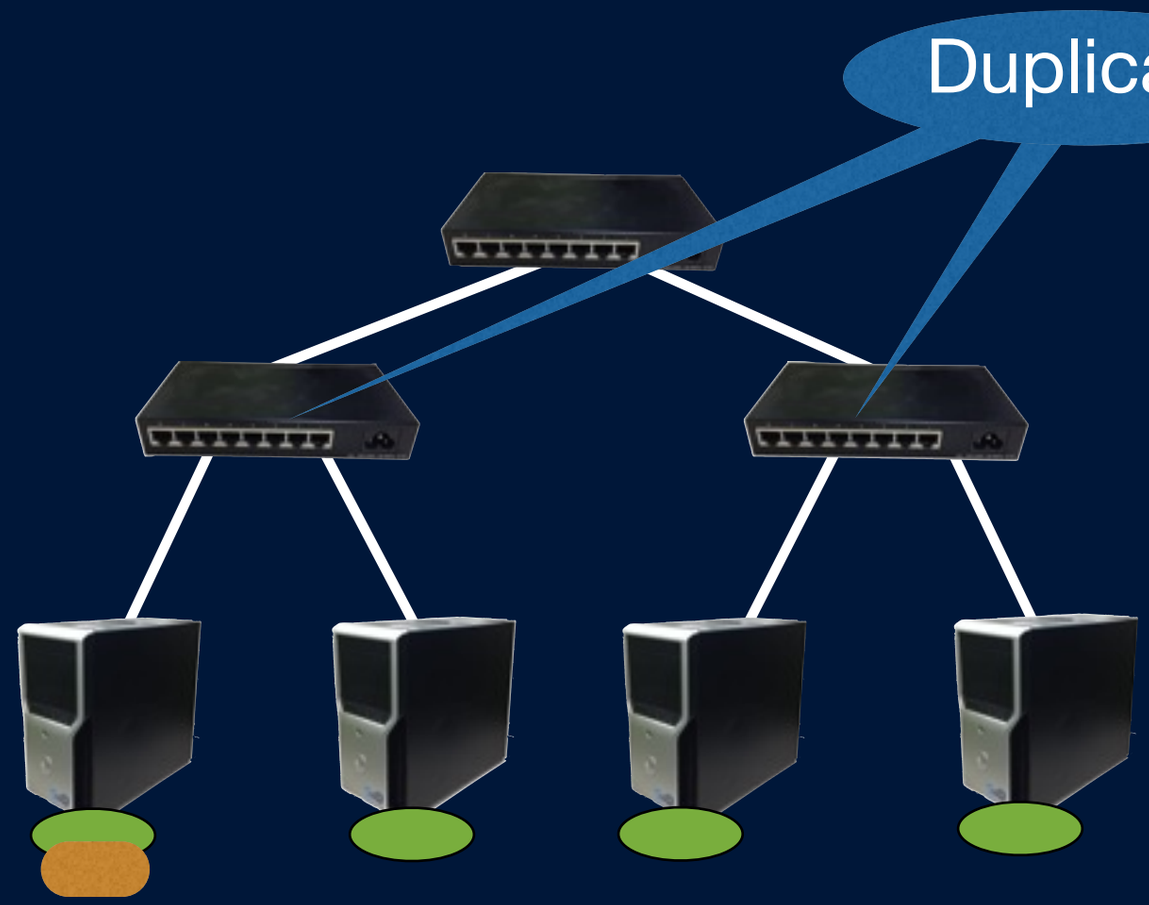


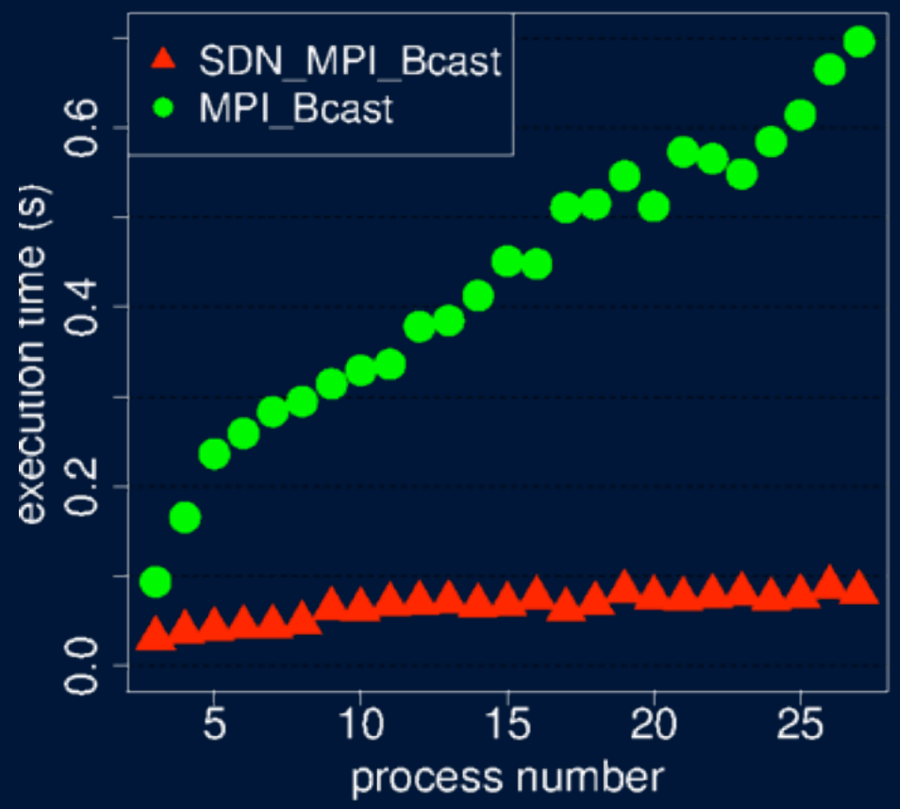


- We achieve to boost up MPI-Allreduce and MPI-Reduce leveraging SDN.

Previous works

- SDN-MPI-Bcast achieves to boost MPI-Bcast up leveraging Software Defined Networking (SDN).
  - sets multicast tree path from source to others.

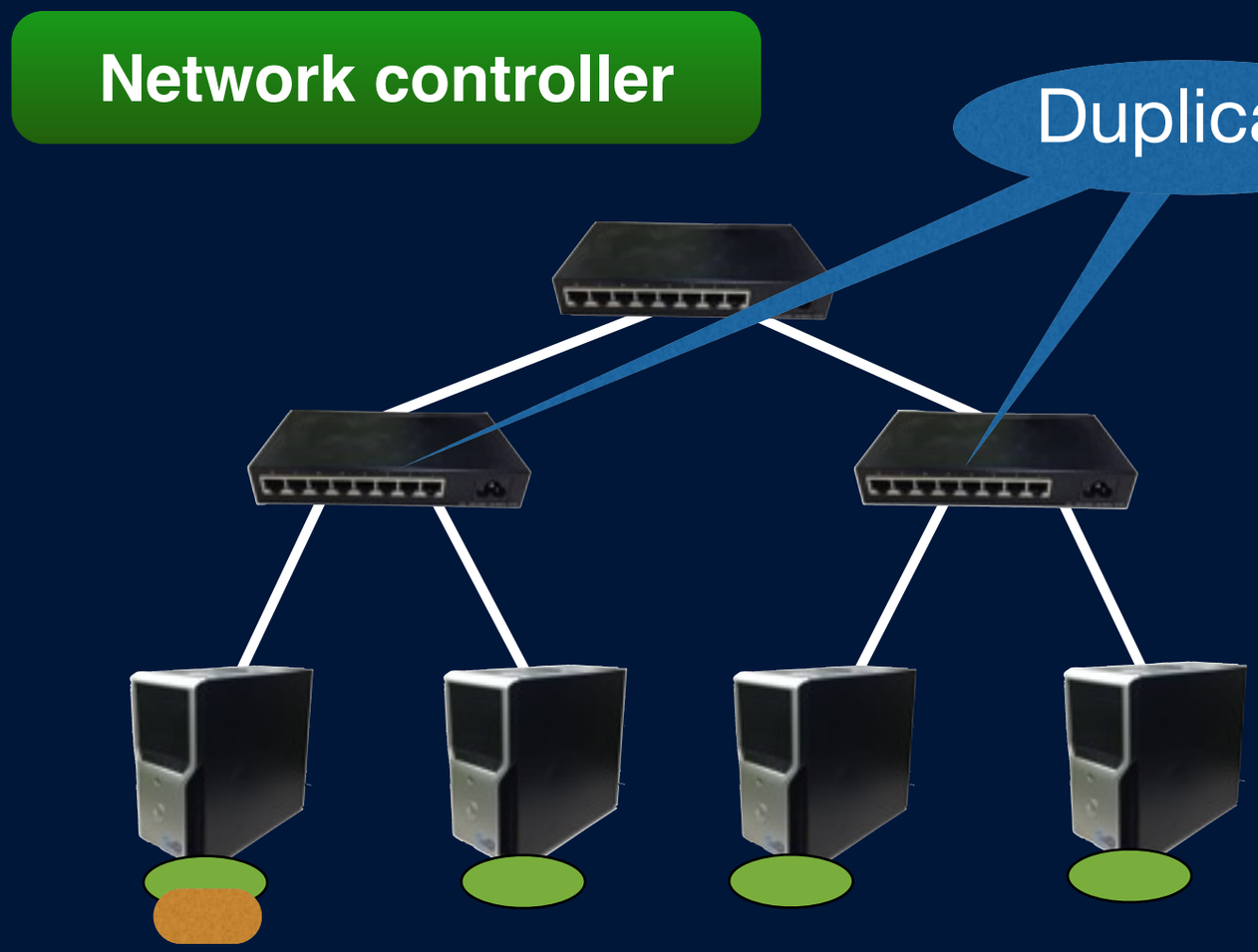


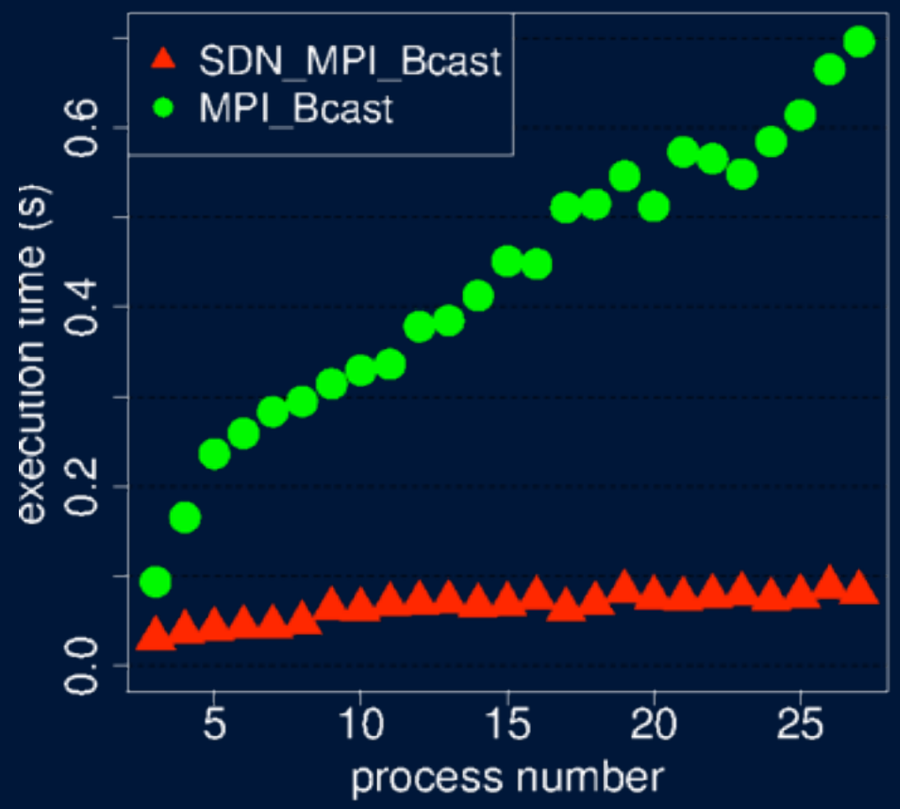


- We achieve to boost up MPI-Allreduce and MPI-Reduce leveraging SDN.

Previous works

- SDN-MPI-Bcast achieves to boost MPI-Bcast up leveraging Software Defined Networking (SDN).
  - sets multicast tree path from source to others.

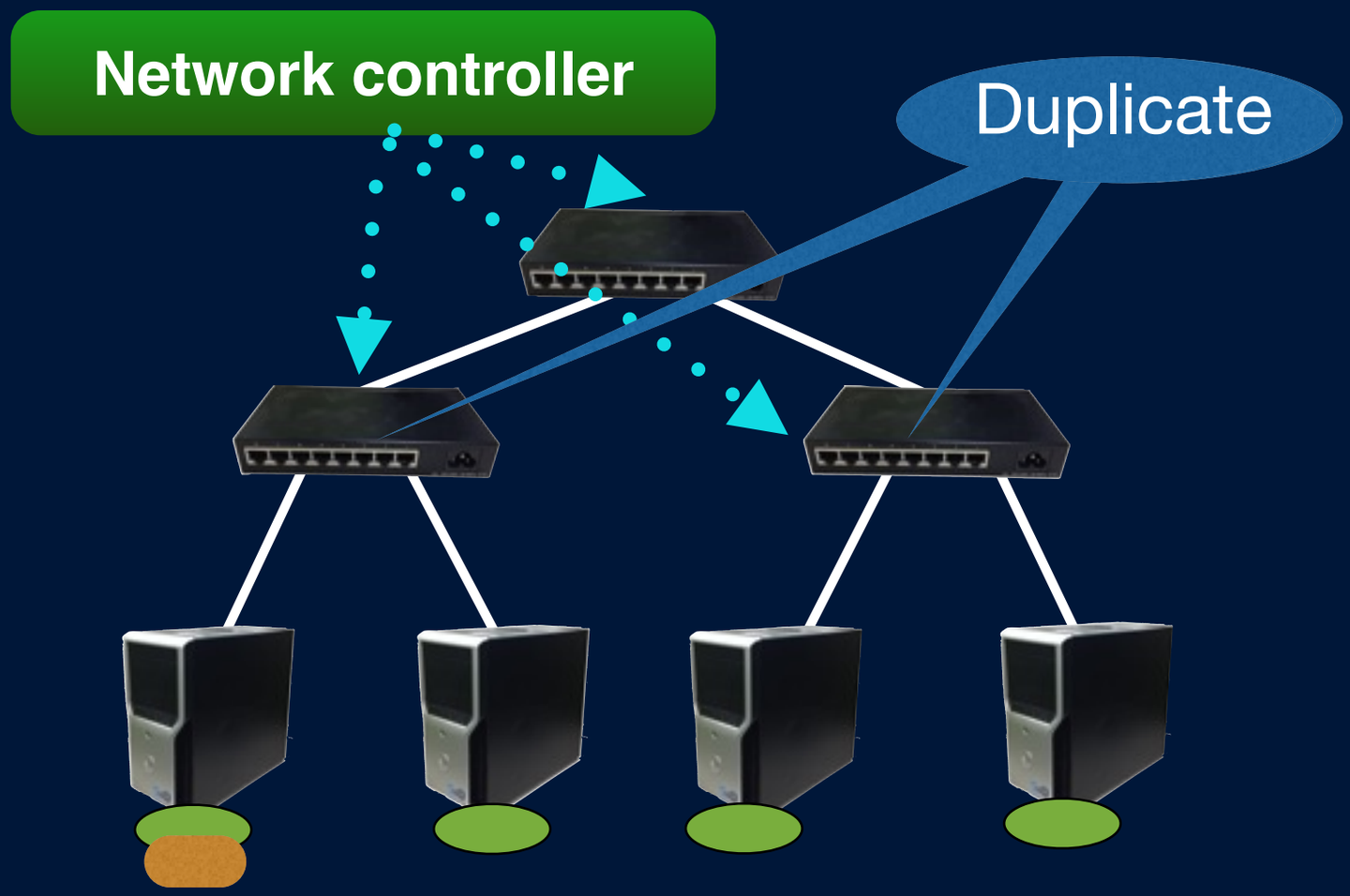


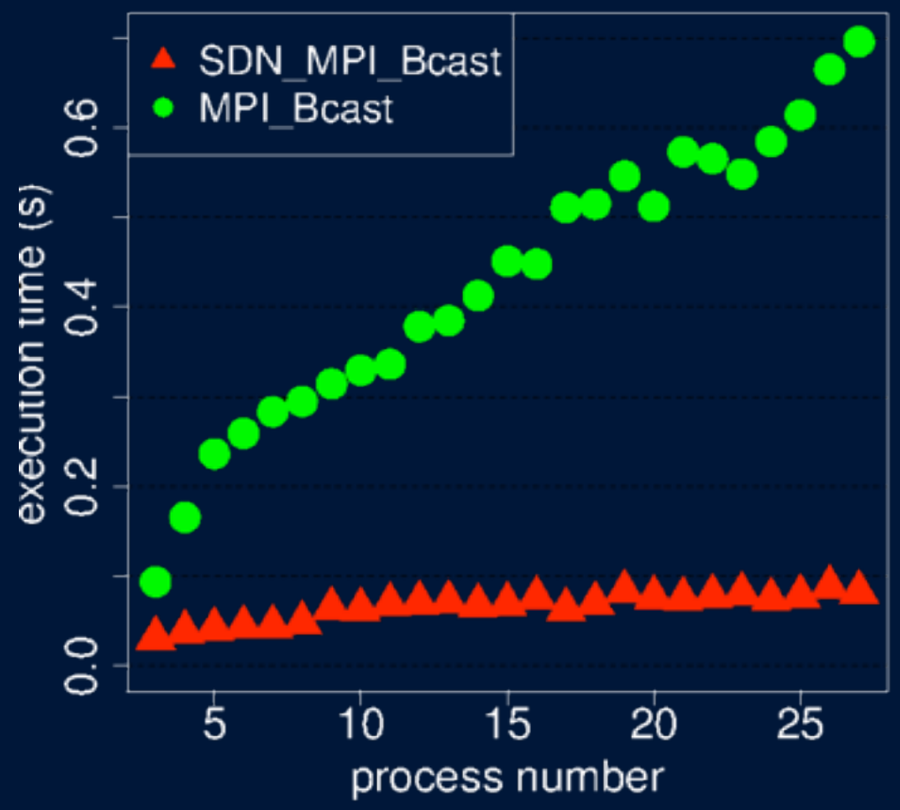


- We achieve to boost up MPI-Allreduce and MPI-Reduce leveraging SDN.

Previous works

- SDN-MPI-Bcast achieves to boost MPI-Bcast up leveraging Software Defined Networking (SDN).
  - sets multicast tree path from source to others.

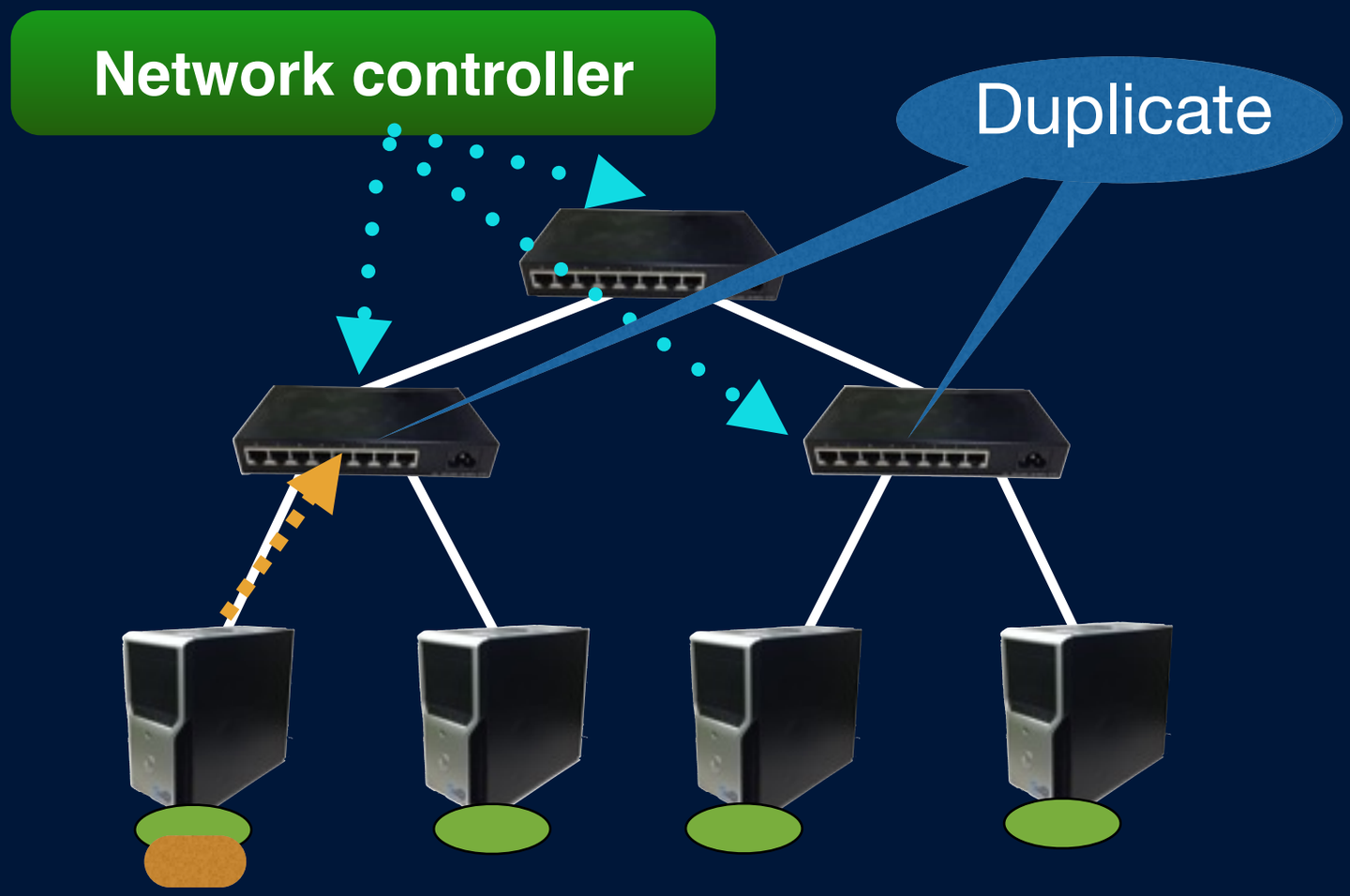


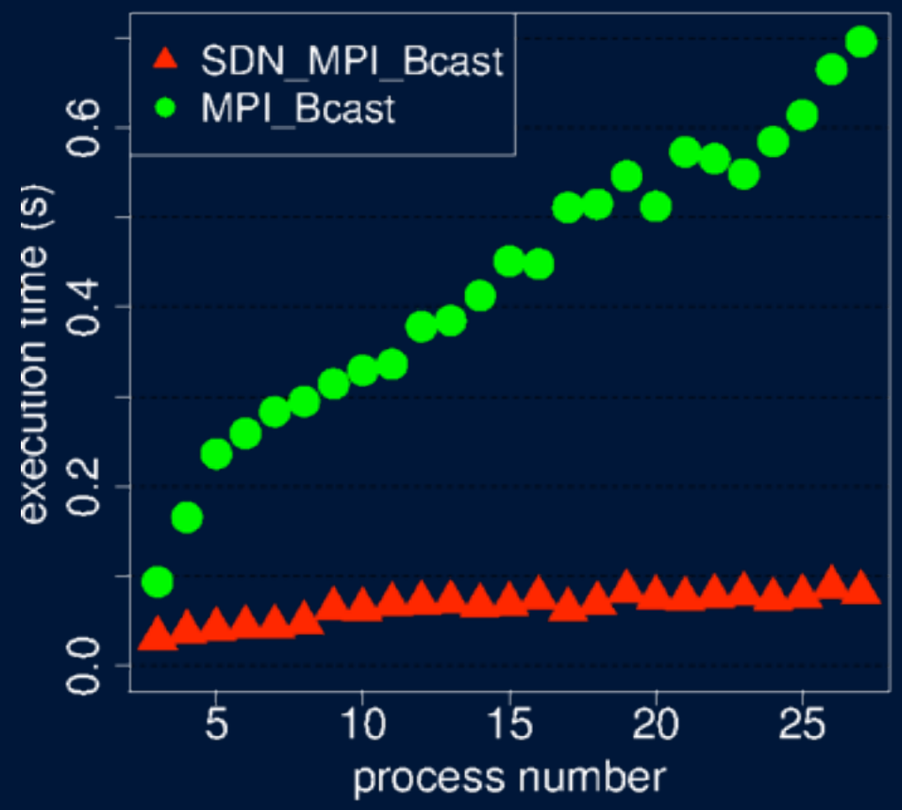


- We achieve to boost up MPI-Allreduce and MPI-Reduce leveraging SDN.

Previous works

- SDN-MPI-Bcast achieves to boost MPI-Bcast up leveraging Software Defined Networking (SDN).
  - sets multicast tree path from source to others.

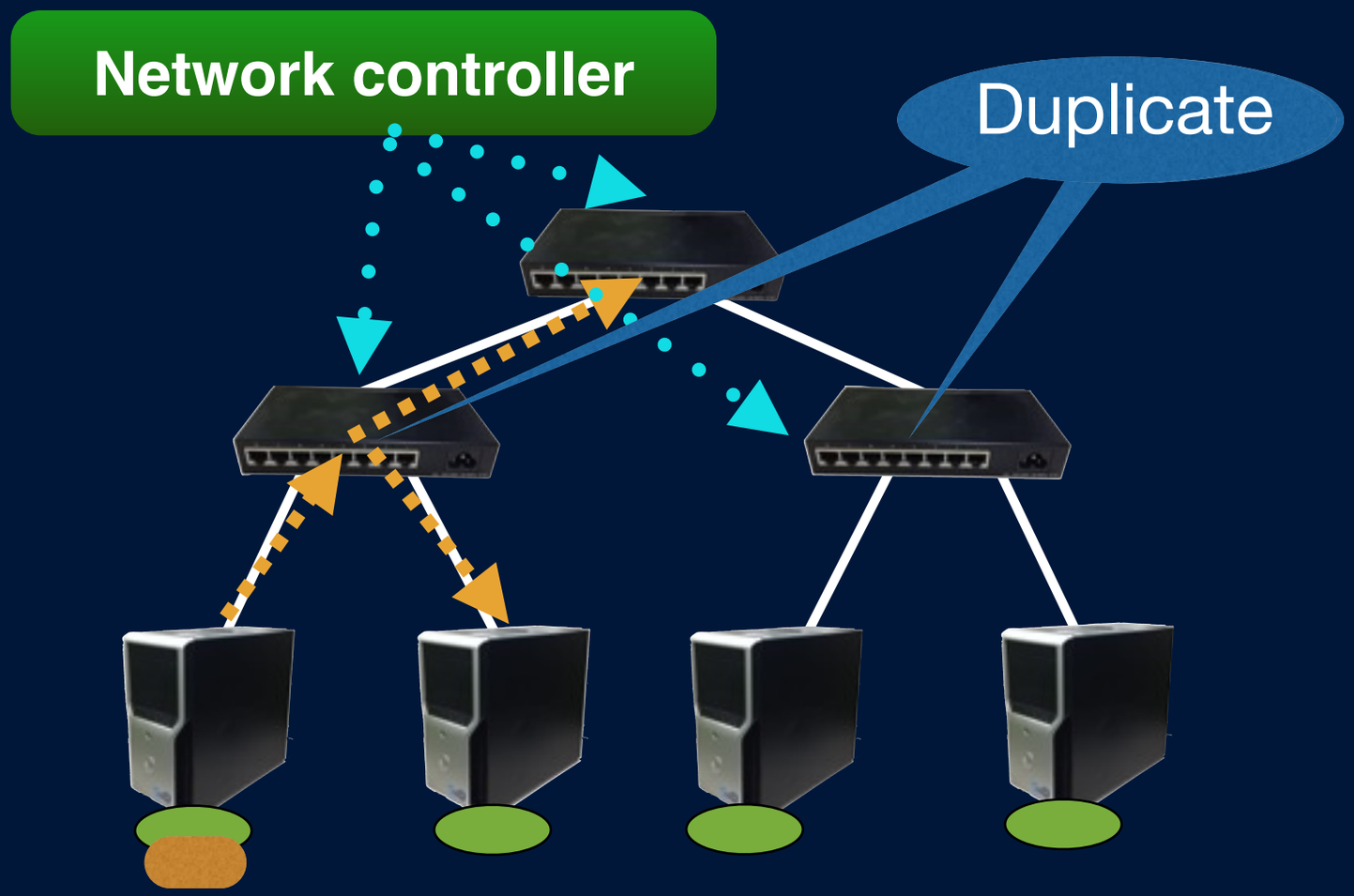


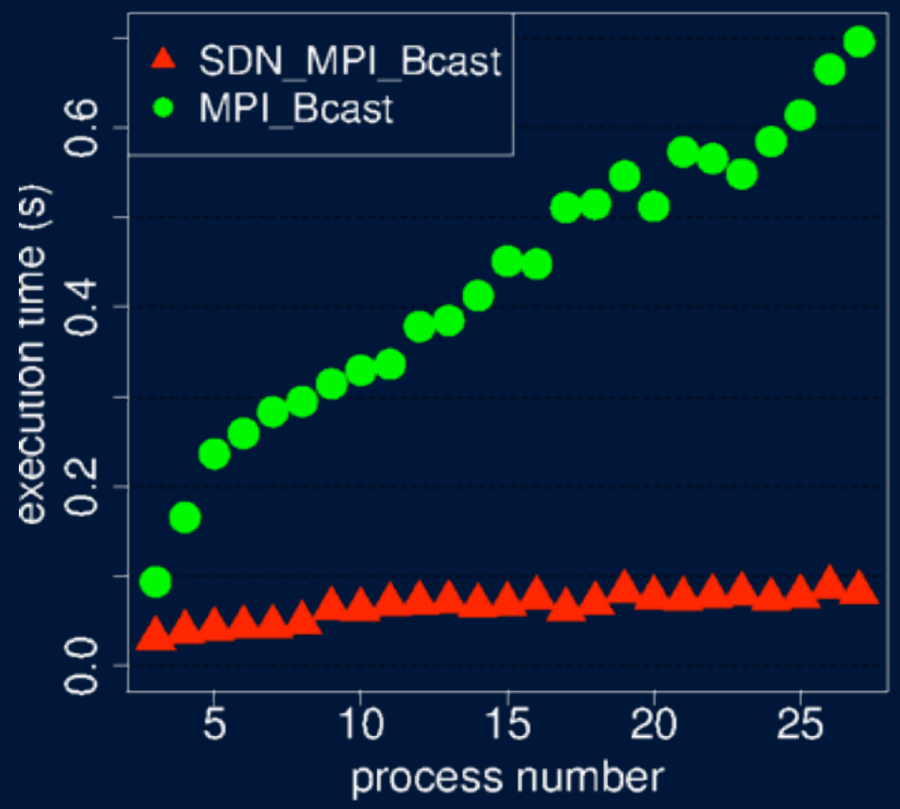


- We achieve to boost up MPI-Allreduce and MPI-Reduce leveraging SDN.

Previous works

- SDN-MPI-Bcast achieves to boost MPI-Bcast up leveraging Software Defined Networking (SDN).
  - sets multicast tree path from source to others.

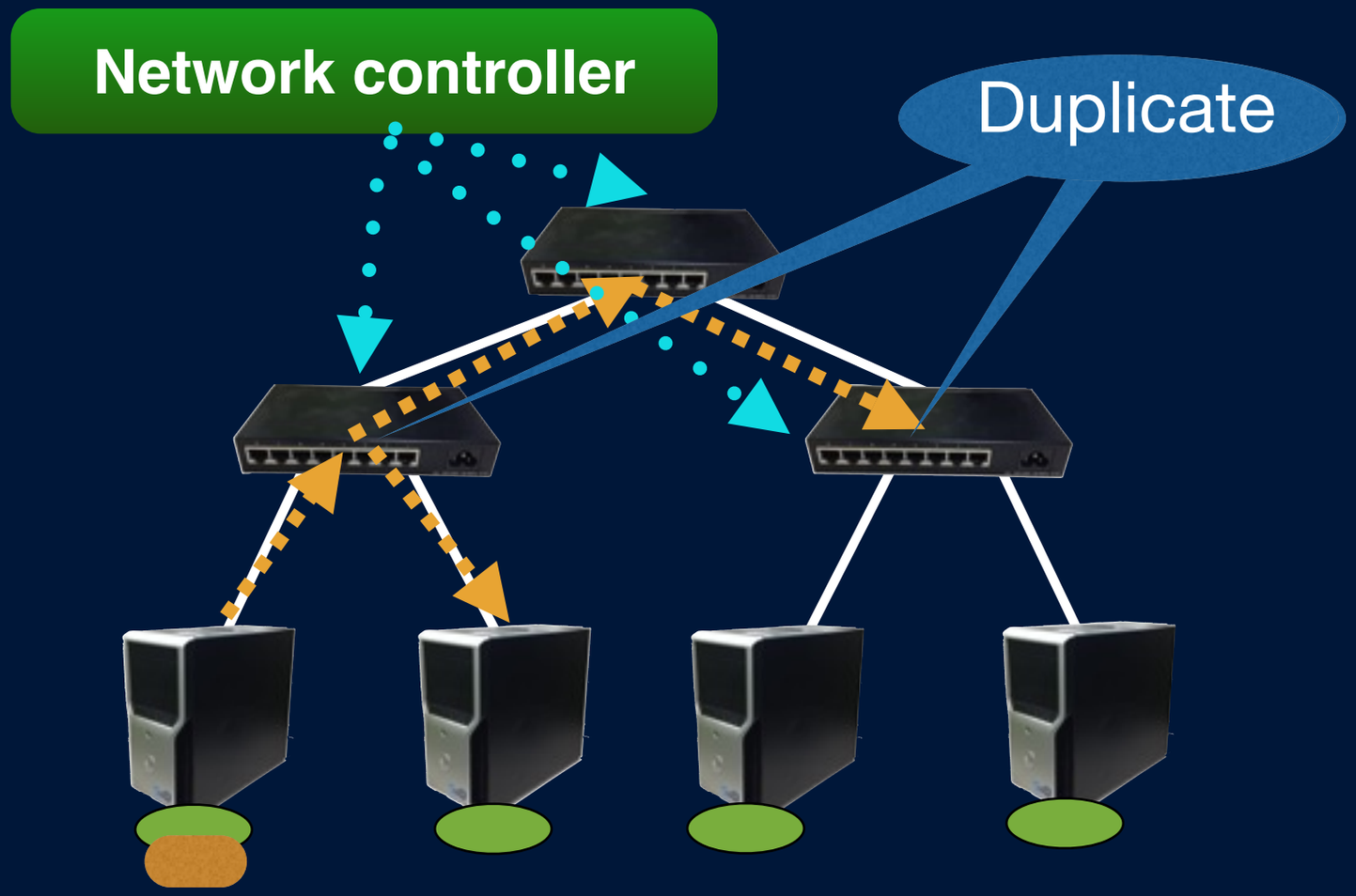


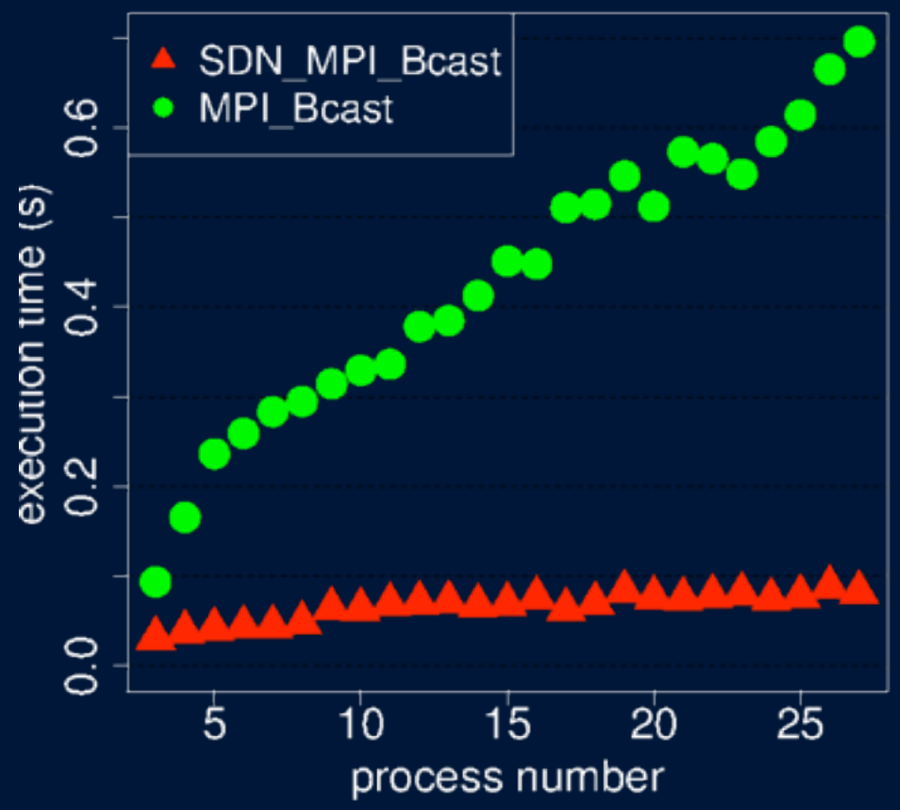


- We achieve to boost up MPI-Allreduce and MPI-Reduce leveraging SDN.

Previous works

- SDN-MPI-Bcast achieves to boost MPI-Bcast up leveraging Software Defined Networking (SDN).
  - sets multicast tree path from source to others.

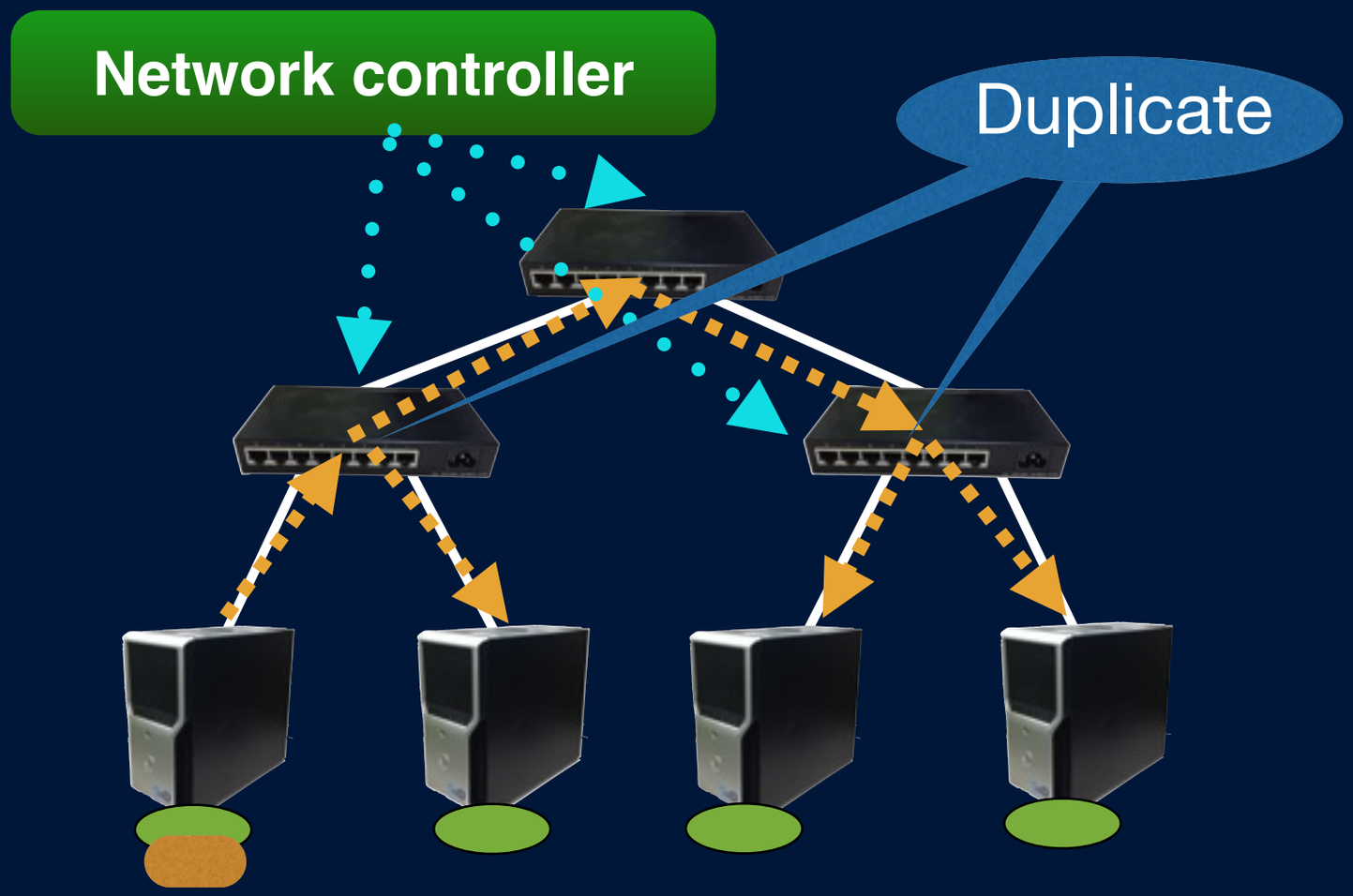


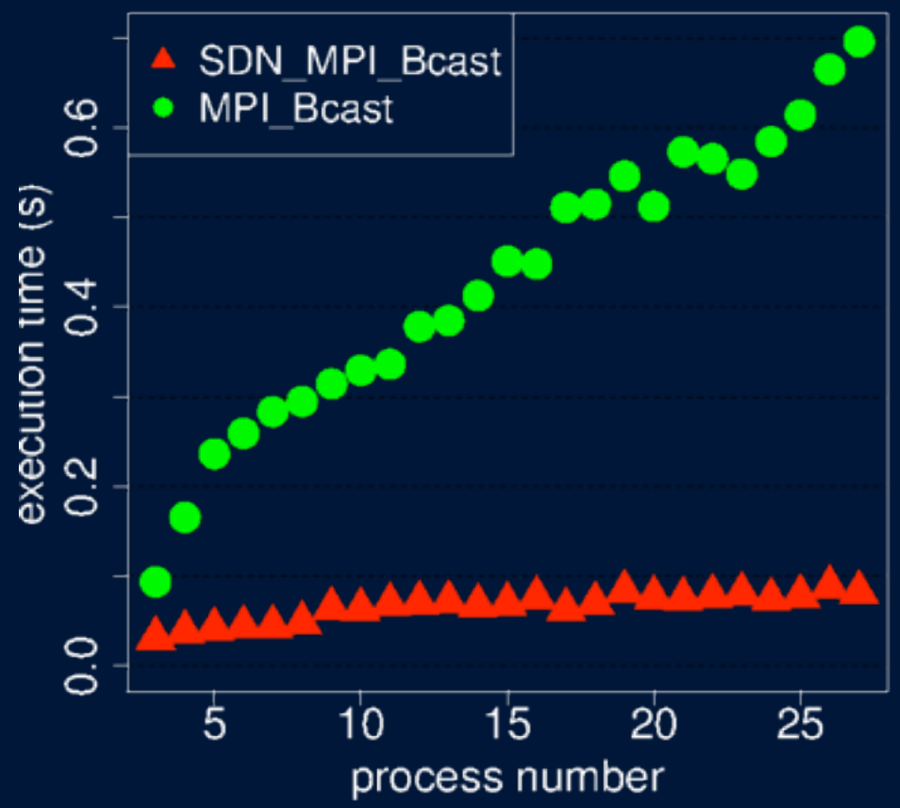


- We achieve to boost up MPI-Allreduce and MPI-Reduce leveraging SDN.

Previous works

- SDN-MPI-Bcast achieves to boost MPI-Bcast up leveraging Software Defined Networking (SDN).
  - sets multicast tree path from source to others.

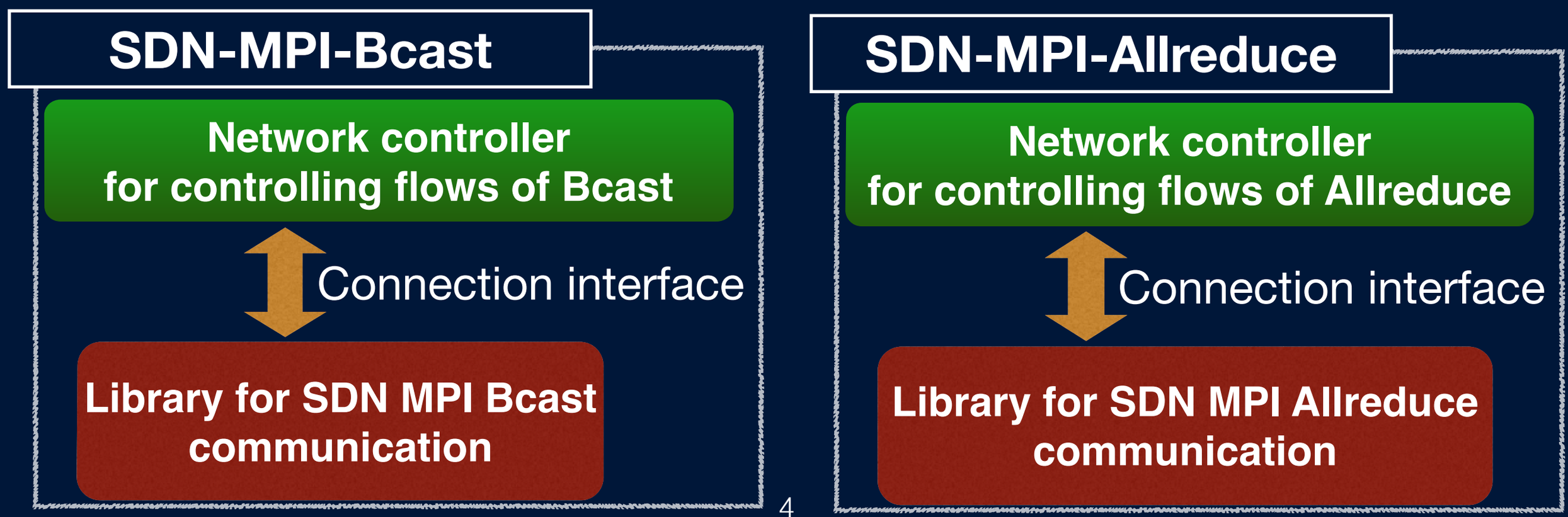




- We achieve to boost up MPI-Allreduce and MPI-Reduce leveraging SDN.

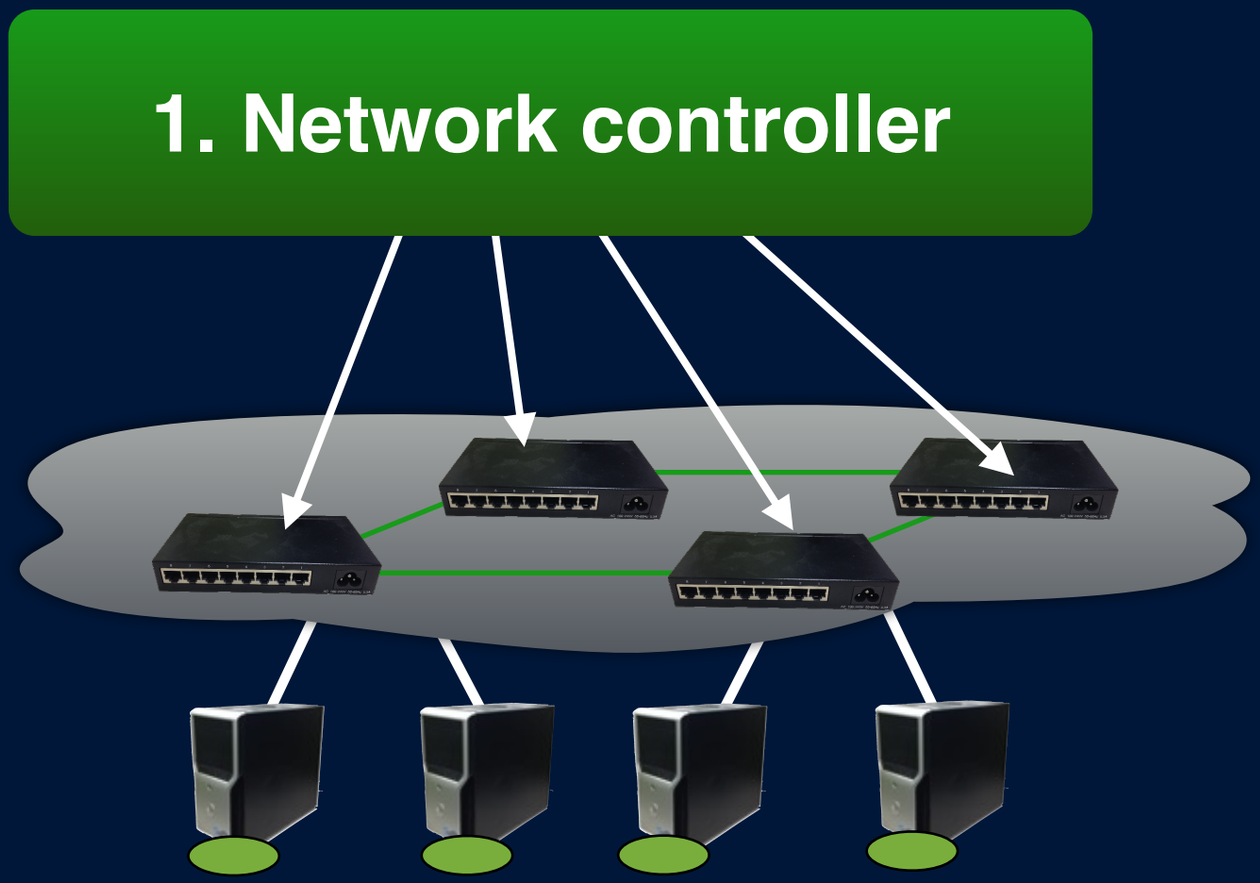
Problem of previous works

- **Efficient and extendable general SDN MPI framework does not exist.**
    - **Each boosting methods are developed individually.**
        - **Hard to include new boosting methods.**



SDN MPI framework

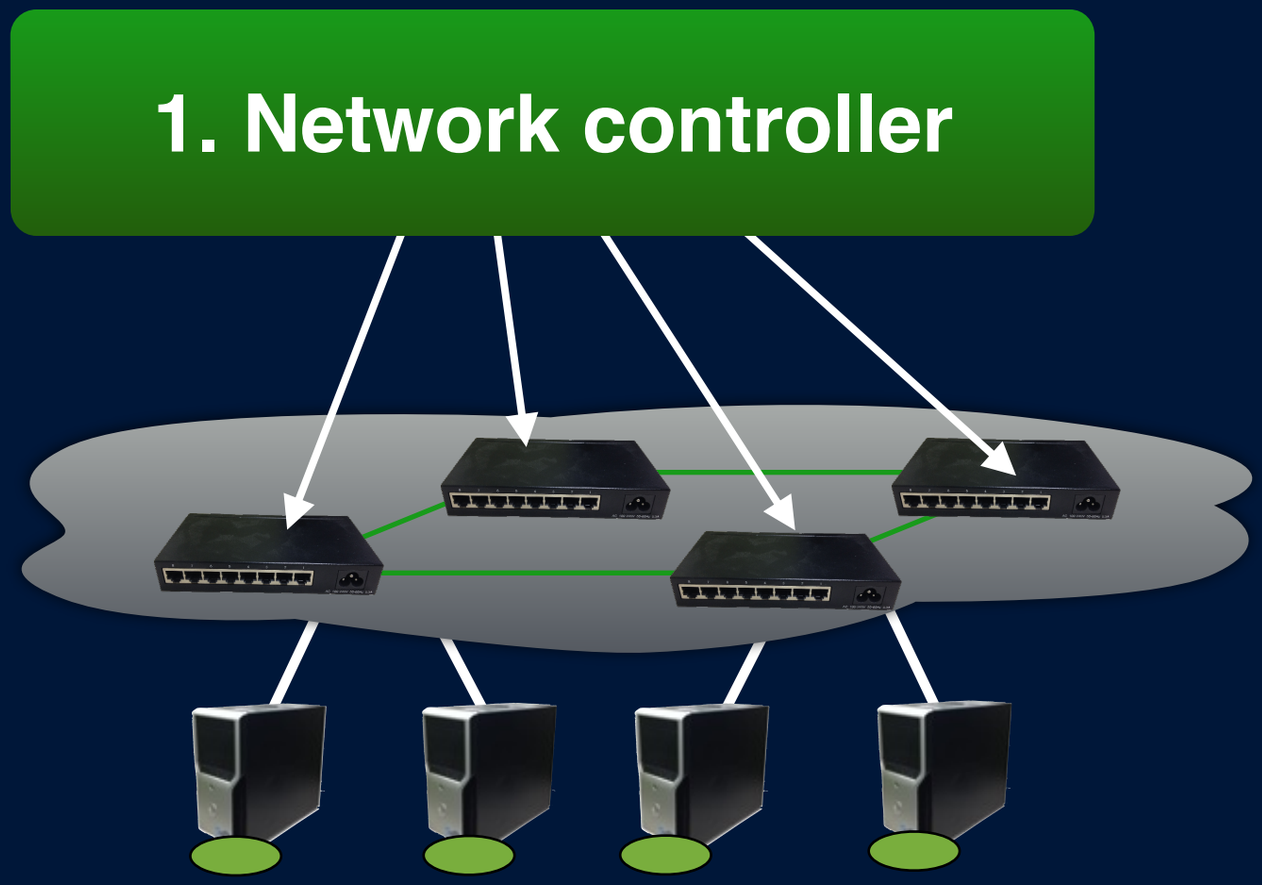
- **SDN MPI framework that includes previous work's methods and expendable for new boosting methods.**



SDN-MPI-Bcast
SDN-MPI-Allreduce
other flows

SDN MPI framework

- SDN MPI framework that includes previous work's methods and expendable for new boosting methods.



Bcast module

Allreduce module

… module

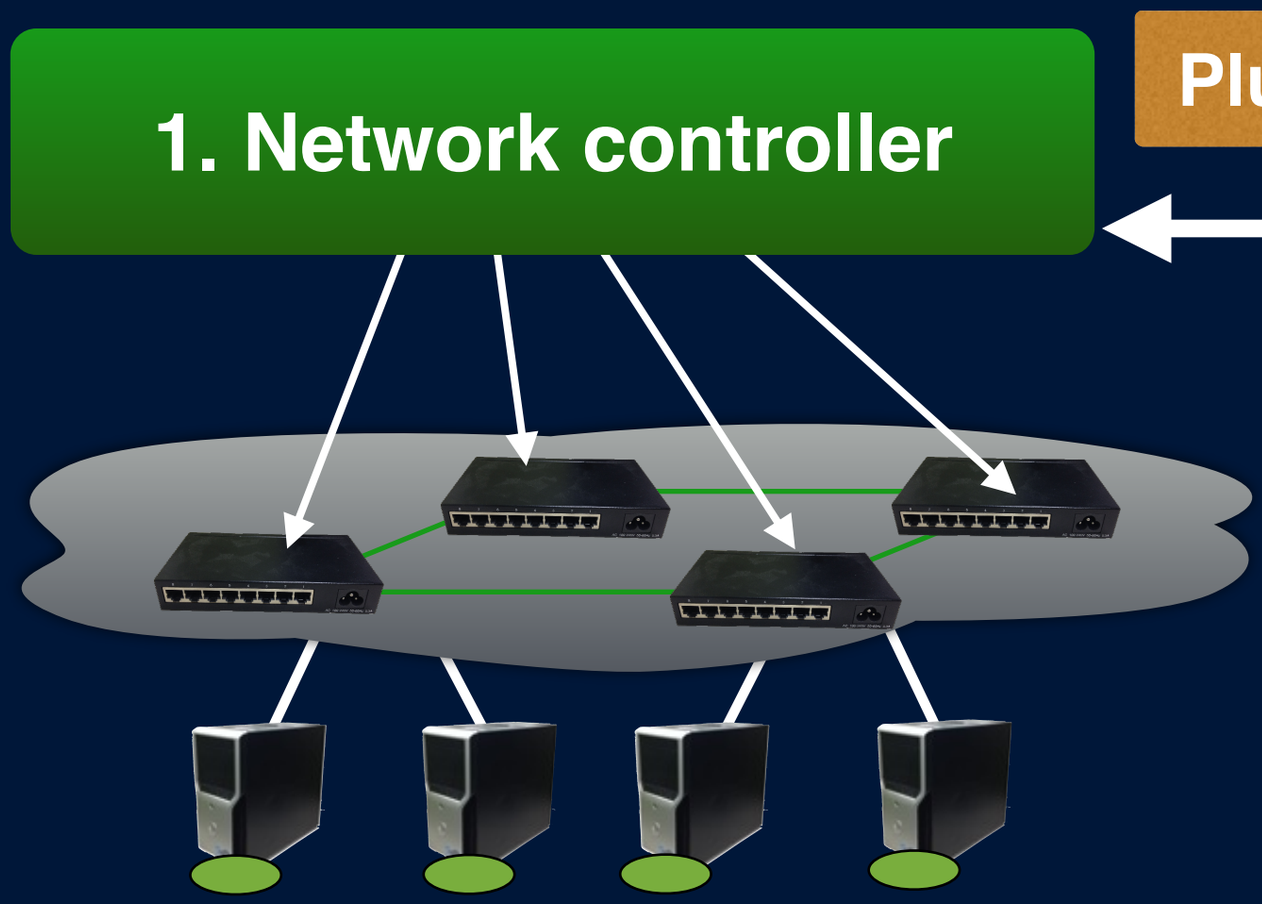
SDN-MPI-Bcast

SDN-MPI-Allreduce

other flows

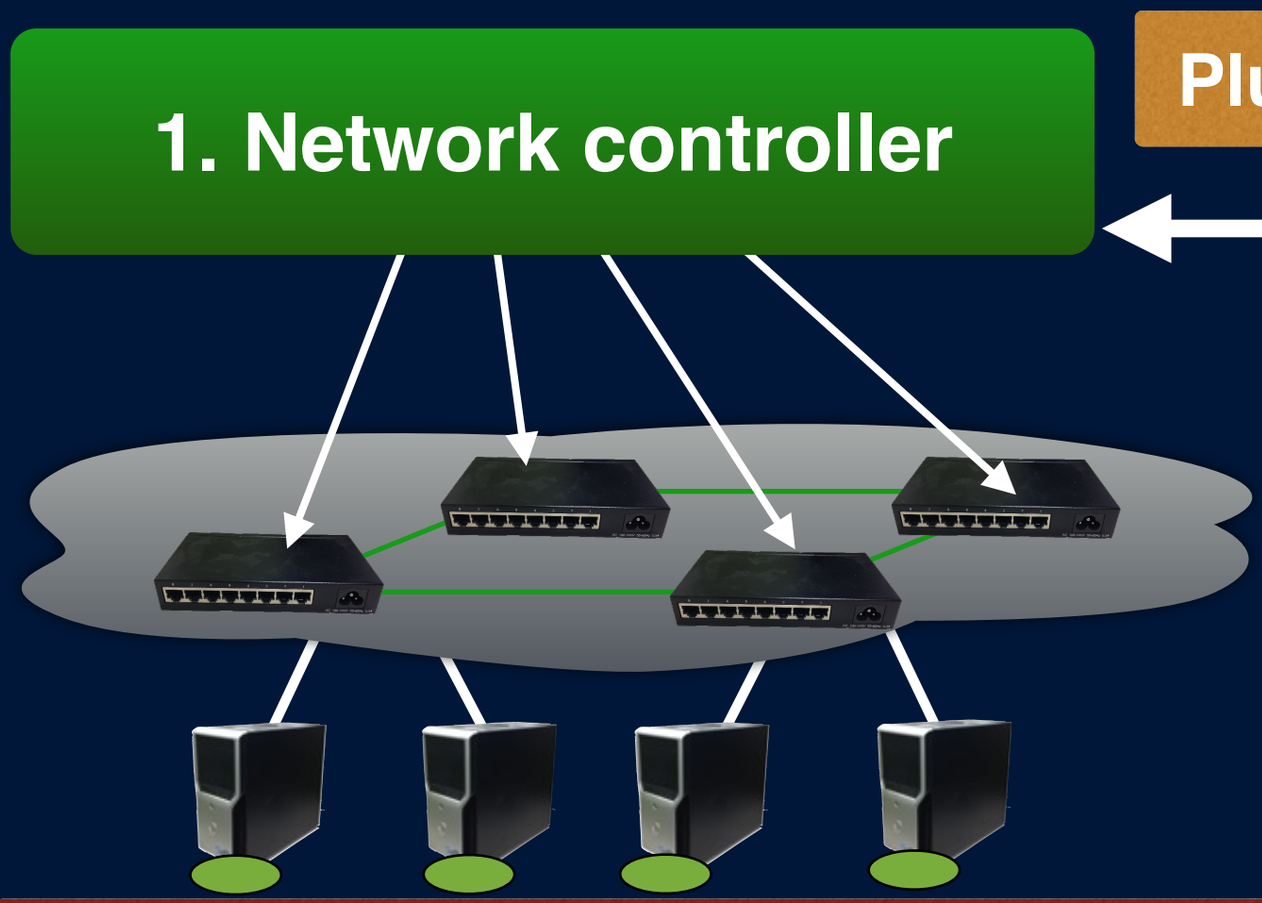
SDN MPI framework

- SDN MPI framework that includes previous work's methods and expendable for new boosting methods.



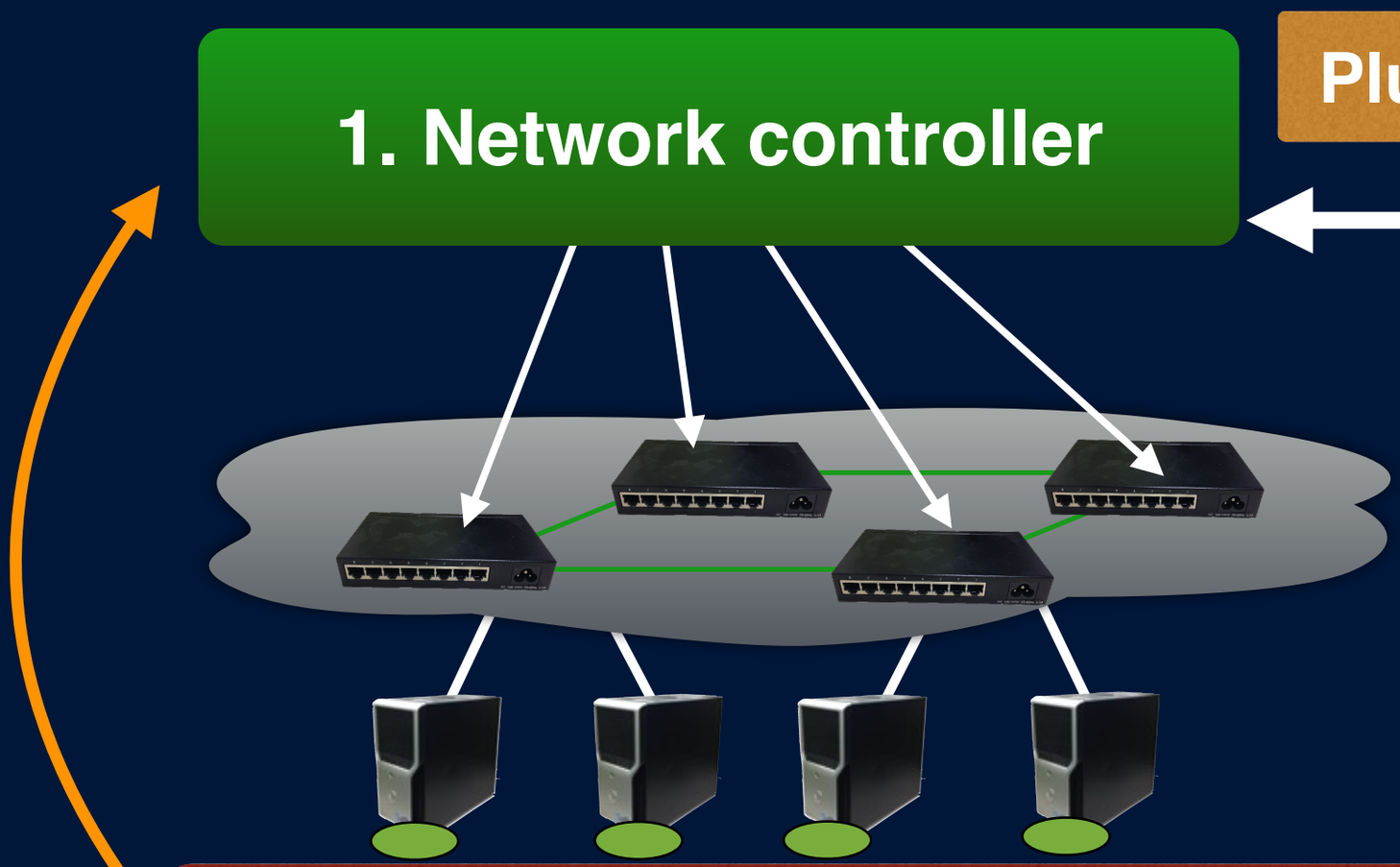
SDN MPI framework

- SDN MPI framework that includes previous work's methods and expendable for new boosting methods.



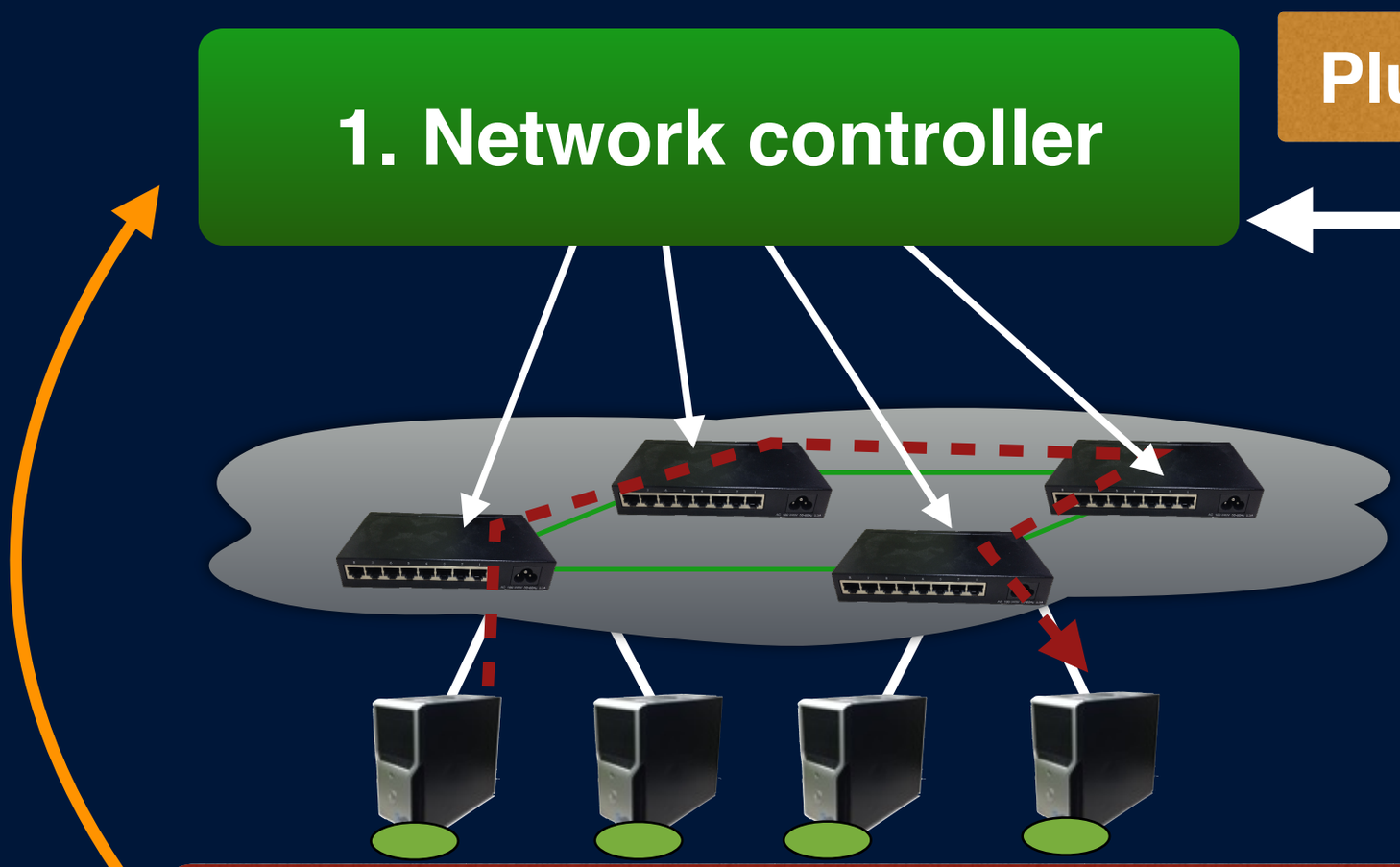
SDN MPI framework

- SDN MPI framework that includes previous work's methods and expendable for new boosting methods.



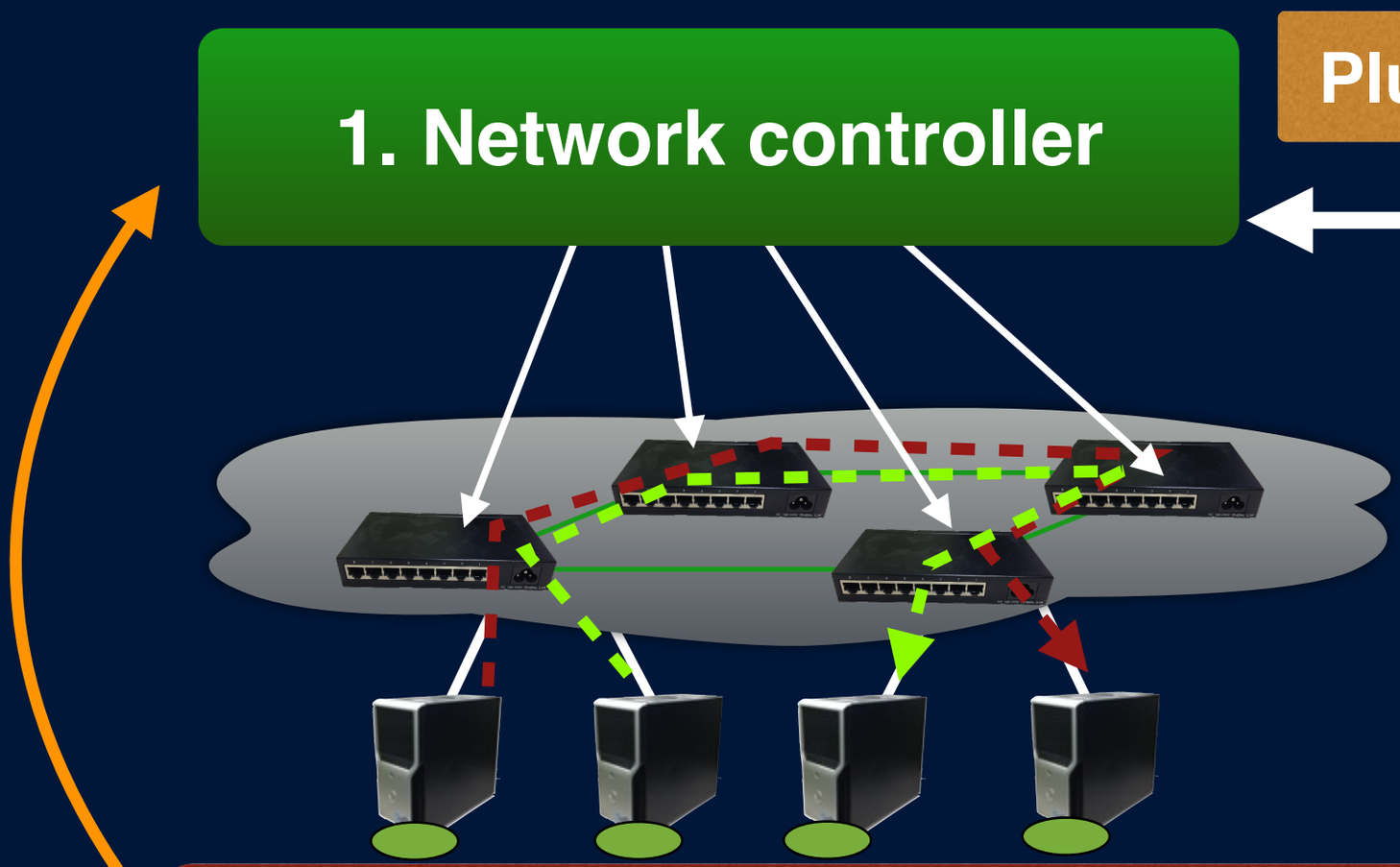
SDN MPI framework

- SDN MPI framework that includes previous work's methods and expendable for new boosting methods.



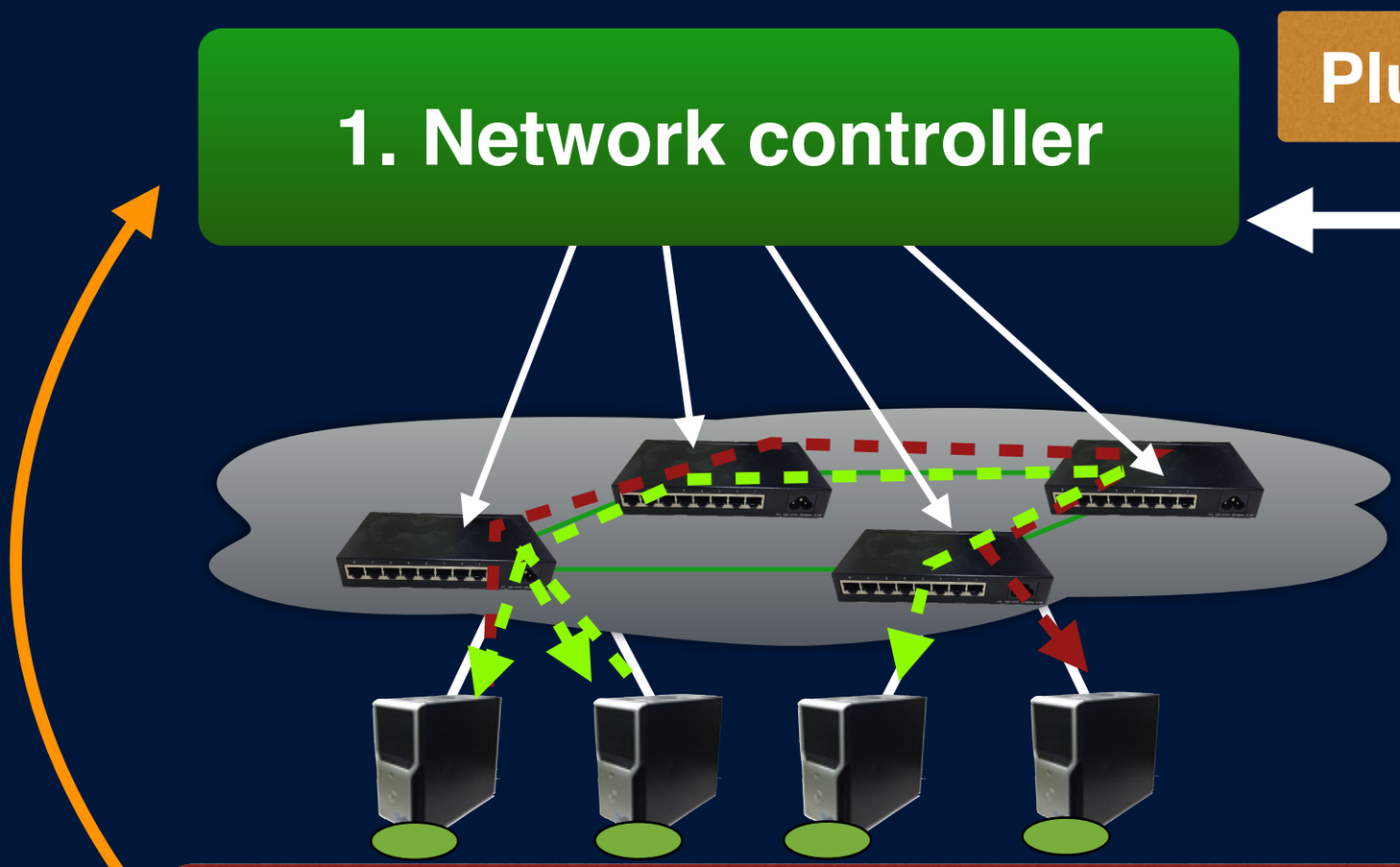
SDN MPI framework

- SDN MPI framework that includes previous work's methods and expendable for new boosting methods.



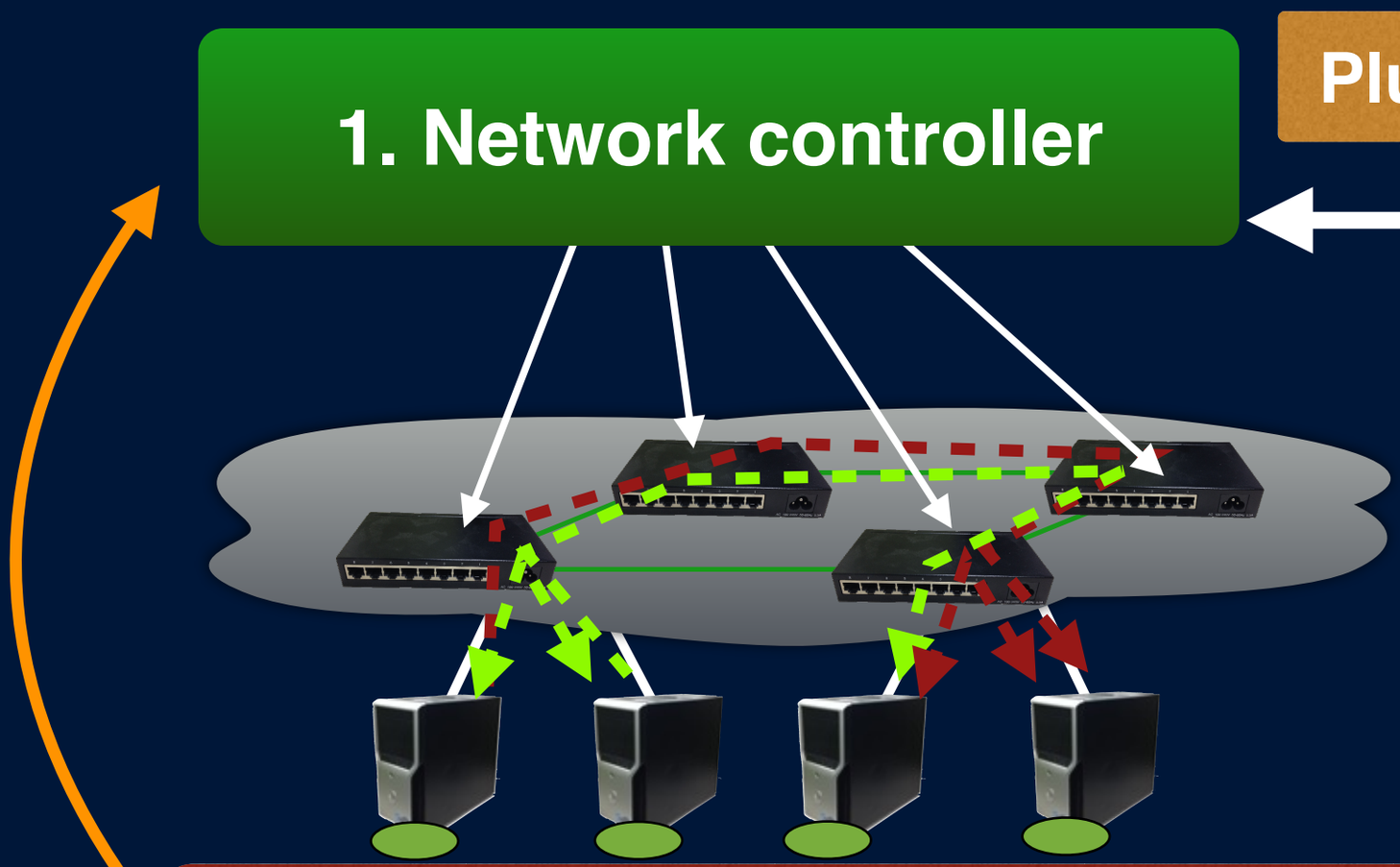
SDN MPI framework

- SDN MPI framework that includes previous work's methods and expendable for new boosting methods.



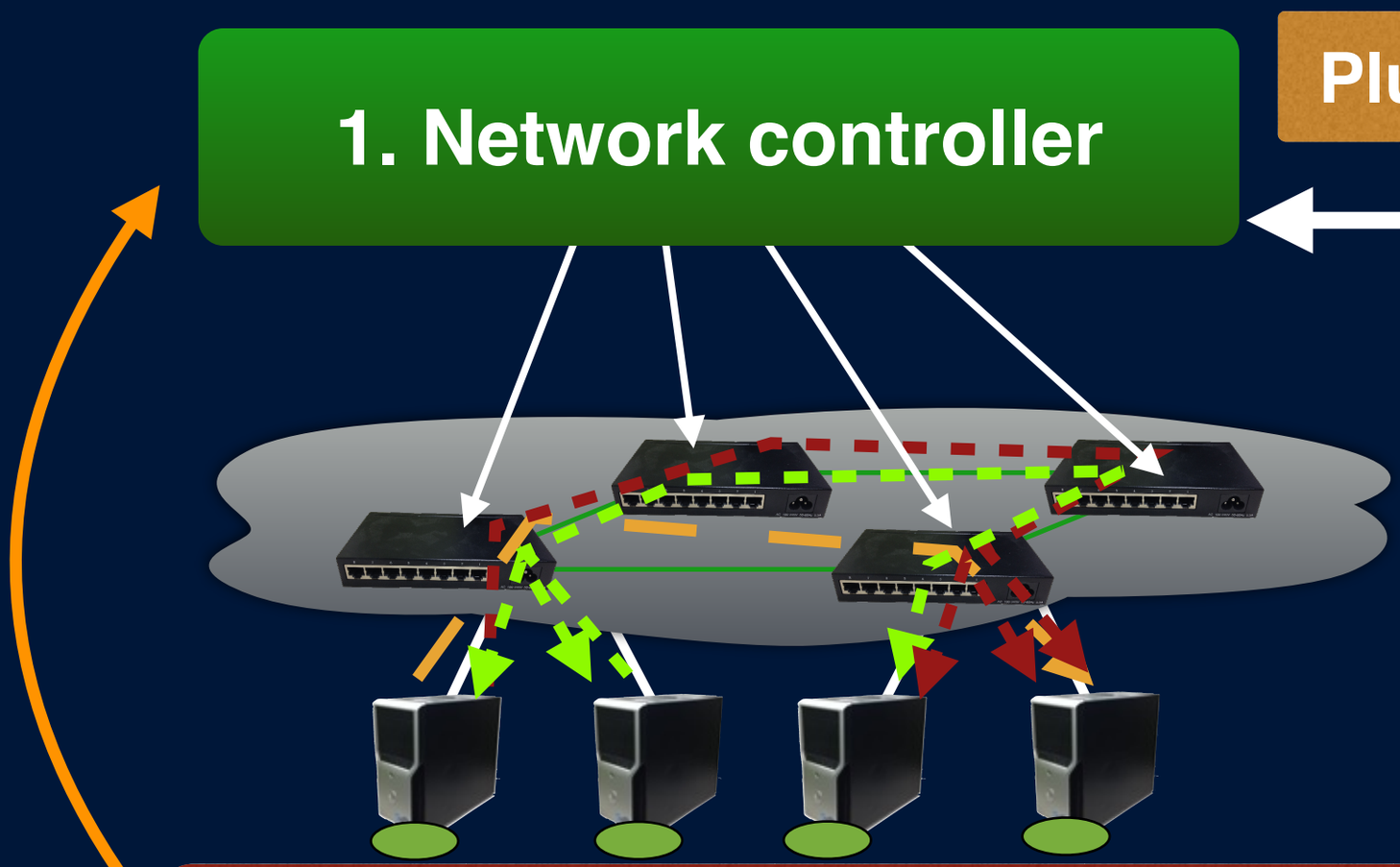
SDN MPI framework

- SDN MPI framework that includes previous work's methods and expendable for new boosting methods.



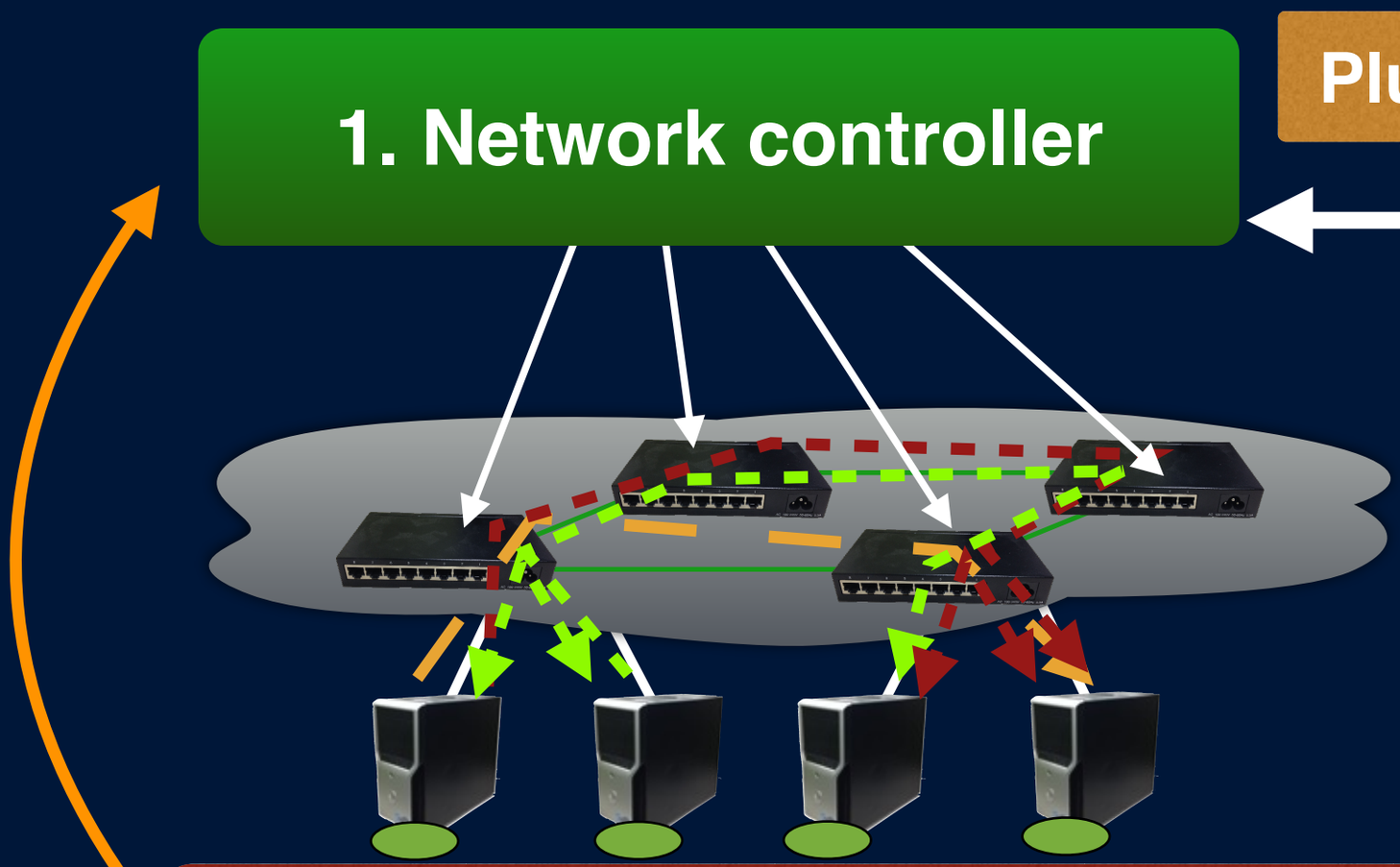
SDN MPI framework

- SDN MPI framework that includes previous work's methods and expendable for new boosting methods.



SDN MPI framework

- SDN MPI framework that includes previous work's methods and expendable for new boosting methods.



Conclusion

- **Introduced general SDN MPI framework which is able to include previous works easily.**

Please come to our poster for more detailed information