Performance Evaluation of iSCSI Protocol with Automatic Parallelism Tuning on SINET3 with Layer-1 On-Demand Service

○ Fumito Inoue (Osaka University)
Hiroyuki Ohsaki (Osaka University)
Yoshihiro Nomoto (NTT Corporation)
Makoto Imase (Osaka University)

Contents

- Background & motivation
  - iSCSI protocol for building IP-SANs
  - Performance degradation in long-fat networks
- iSCSI multiple connections feature
- iSCSI-APT: parallelism tuning mechanism for iSCSI
- Experiment on SINE3 with L1 on-demand service
- Experiment results
- Conclusion & future directions

Background

- Demand for high-speed remote storage
  - Remote backup for disaster recovery
  - Data replication for fast/reliable data delivery
- Emergence of iSCSI protocol
  - Protocol for building SAN on IP network
    - Low cost & easy deployment
- iSCSI performance issues in long-fat networks
  - Very low throughput due to TCP problems

Related Works

- Existing solution for improving iSCSI throughput
  - Approaches utilizing multiple links
    - Use a VPN multihoming
    - Use multiple LAN ports

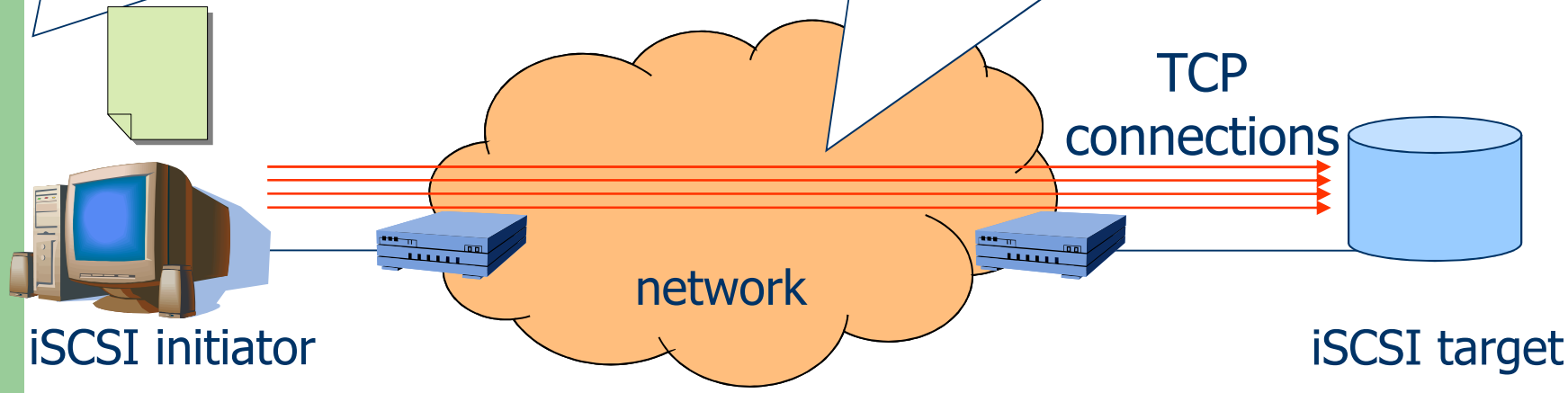
  → Force significant restrictions on network environment
  - Approaches modifying the transport protocol
    - Improve fairness among TCP connections

  → Is not realistic since a large number of iSCSI devices have already been in operation

iSCSI Multiple Connections

parallel data transfer
with multiple TCP connections

multiple TCP connections results
in greedy data transfer
→ leading high throughput?



too many TCP connections
result in low throughput

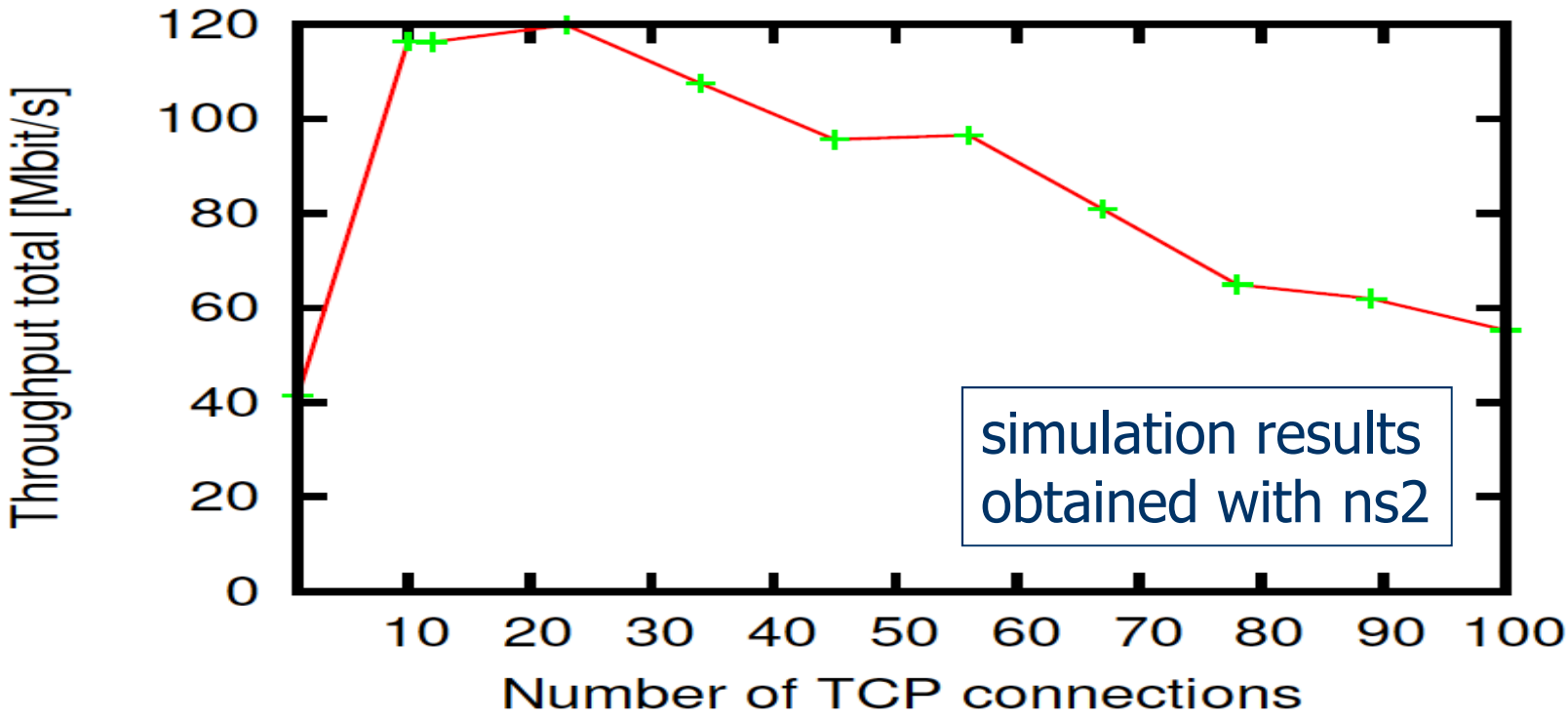
→

# of TCP connections must be
adjusted according to
network status

[10] T. Ito, H. Ohsaki, and M. Imase, "On parameter tuning of data transfer protocol GridFTP in wide-area Grid computing," in *Proceedings of Second InternationalWorkshop on Networks for Grid Applications (GridNets 2005)*, pp. 415–421, Oct. 2005.

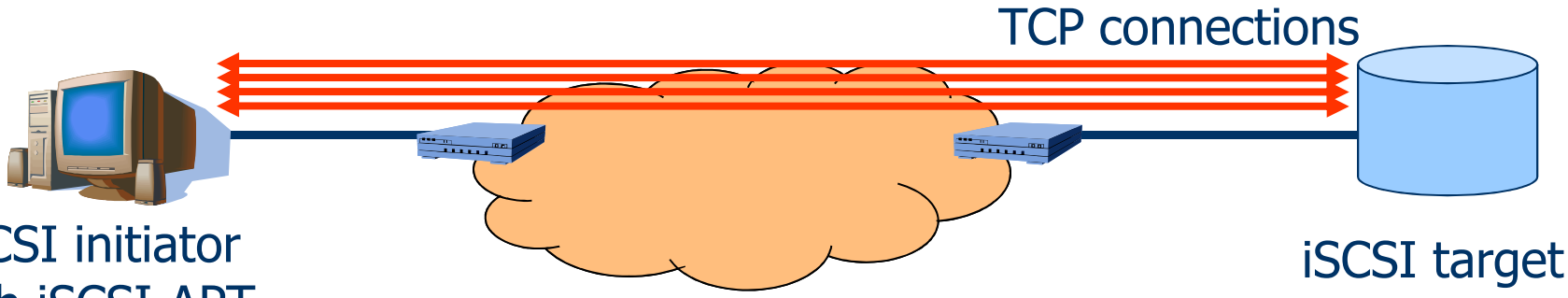
Throughput vs. # of TCP Connections

- Aggregate TCP throughput is a convex function for # of TCP connections

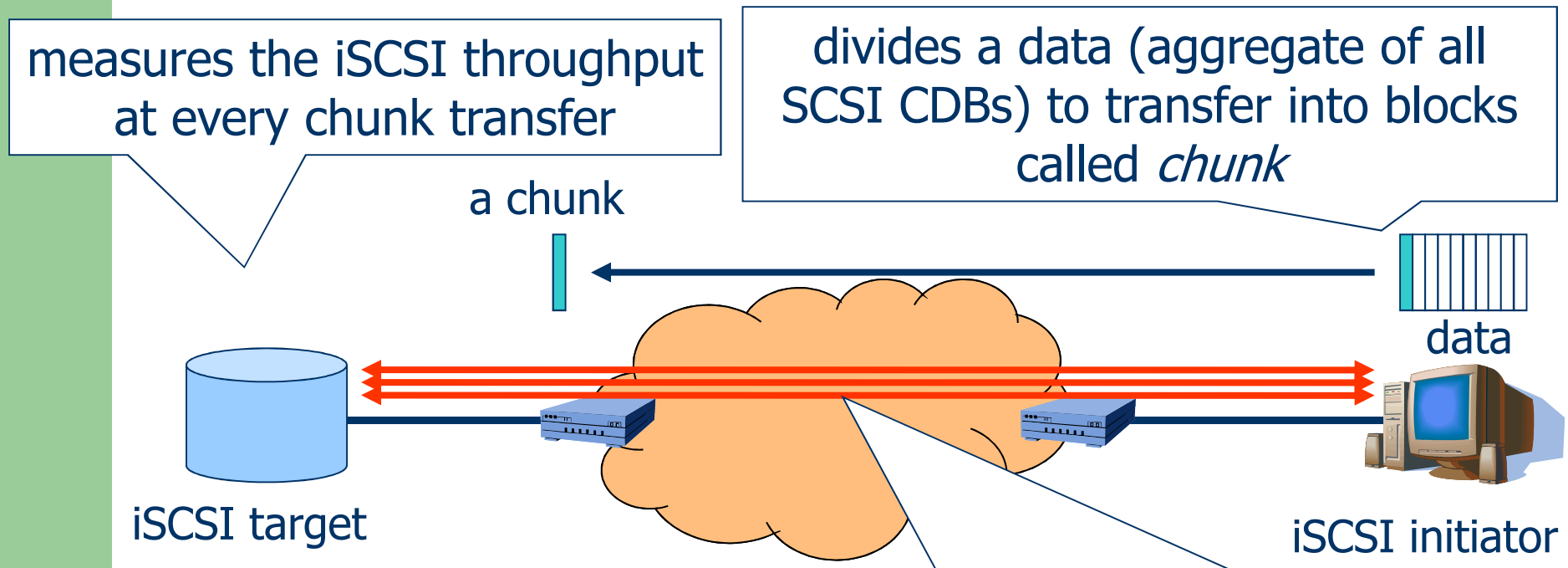


iSCSI-APT: Automatic Parallelism Tuning Mechanism for iSCSI Protocol

- Automatically optimize # of iSCSI connections
  - Utilize the fact TCP throughput is a convex function
    - Numerically solve a maximization problem
  - Measure throughput & RTT in a passive way
- Maximize iSCSI throughput in long-fat networks

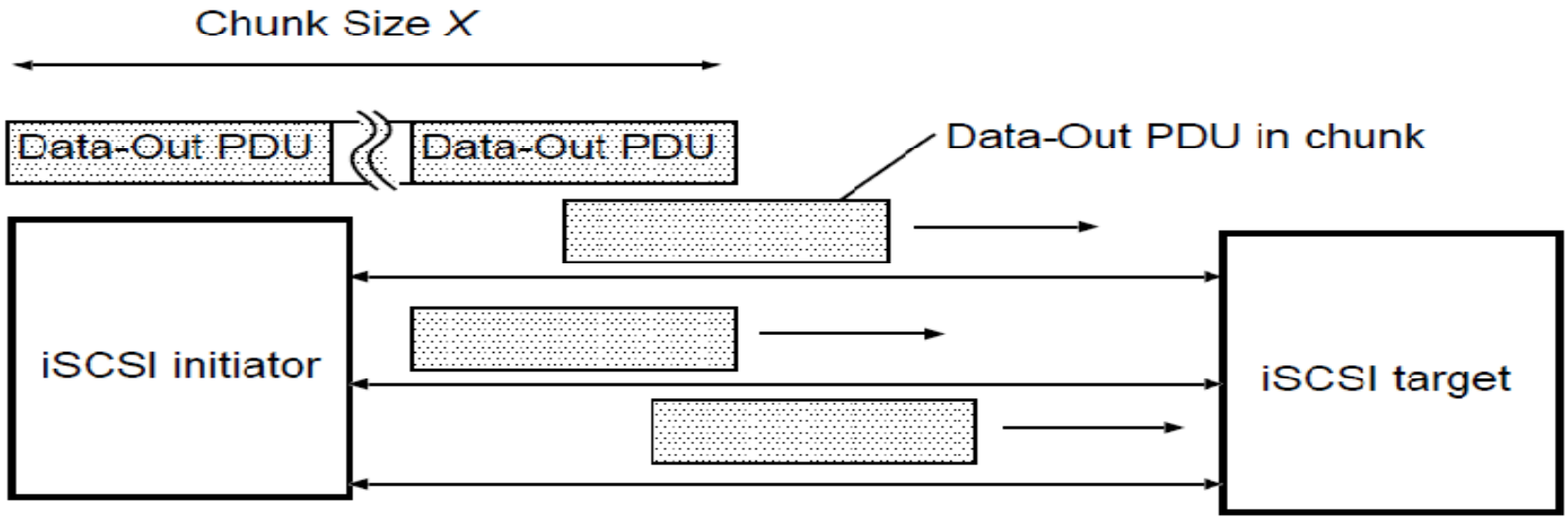


Basic Ideas of iSCSI-APT



Overview of Chunk Transfer

- iSCSI-APT transfers data as a series of blocks called chunk
  - Chunk transfer is realized by transferring multiple iSCSI PDU (Protocol Data Unit)



iSCSI-APT Advantages

- Applicable to diverse network environments
  - No assumption on network configurations/underlay protocols
    - Just assume concavity of TCP throughput
- iSCSI initiator side realization
  - Compatible with any existing iSCSI targets
    - Many iSCSI targets have been in operation
- Effective for continuous transfer of large volume data
  - E.g., Remote backup, data replication

Research Objective

- Performance evaluation of iSCSI-APT on SINET3
  - SINET3 is a next-generation information communication network
  - Experiment on a **real** long-fat network
    - Use our iSCSI-APT implementation
    - Quite different from network emulator

→ Demonstrate that iSCSI-APT operates quite effectively on a real long-fat network

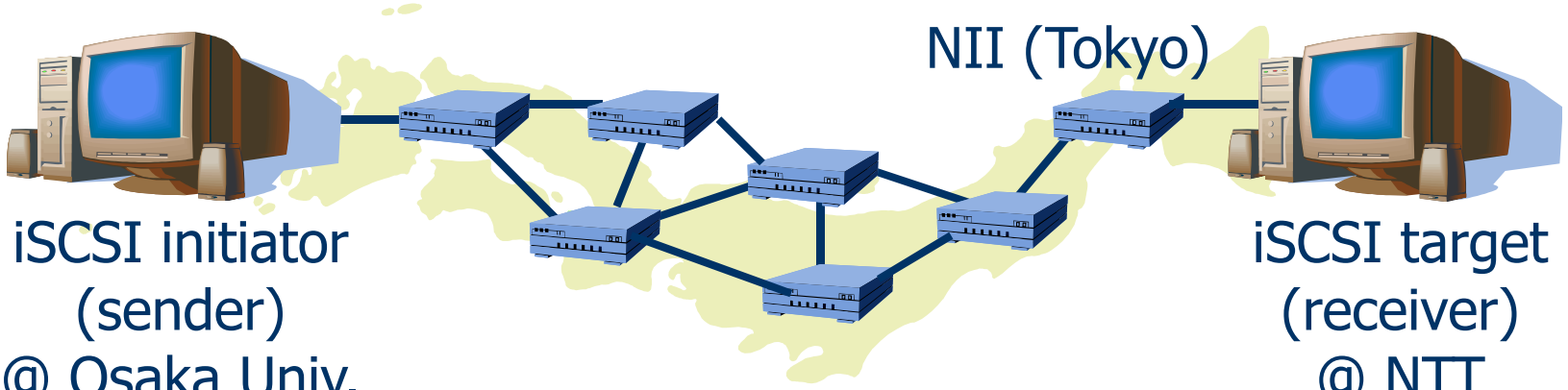
SINET3 Overview

- A next-generation information communication network
  - Connects universities and research institutions
  - Has 1--20 Gbit/s link bandwidth between the edge and core nodes
  - Experimentally provides Layer-1 On-Demand Bandwidth Allocation service
    - **User** (not administrator) can make reservation **instantaneously**

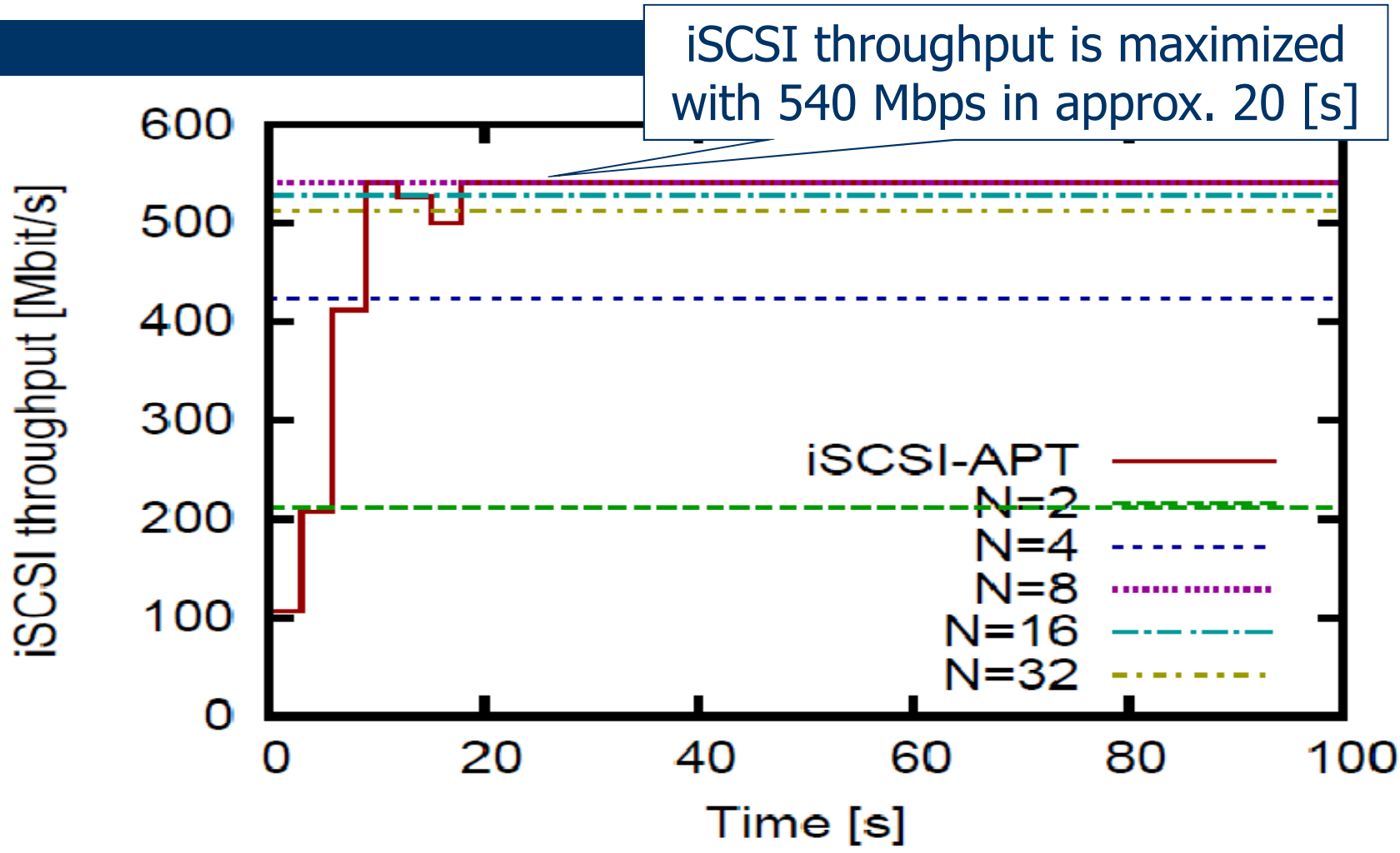
  → Flexible experiment setup

Performance Evaluation of iSCSI-APT with L1 On-Demand Service

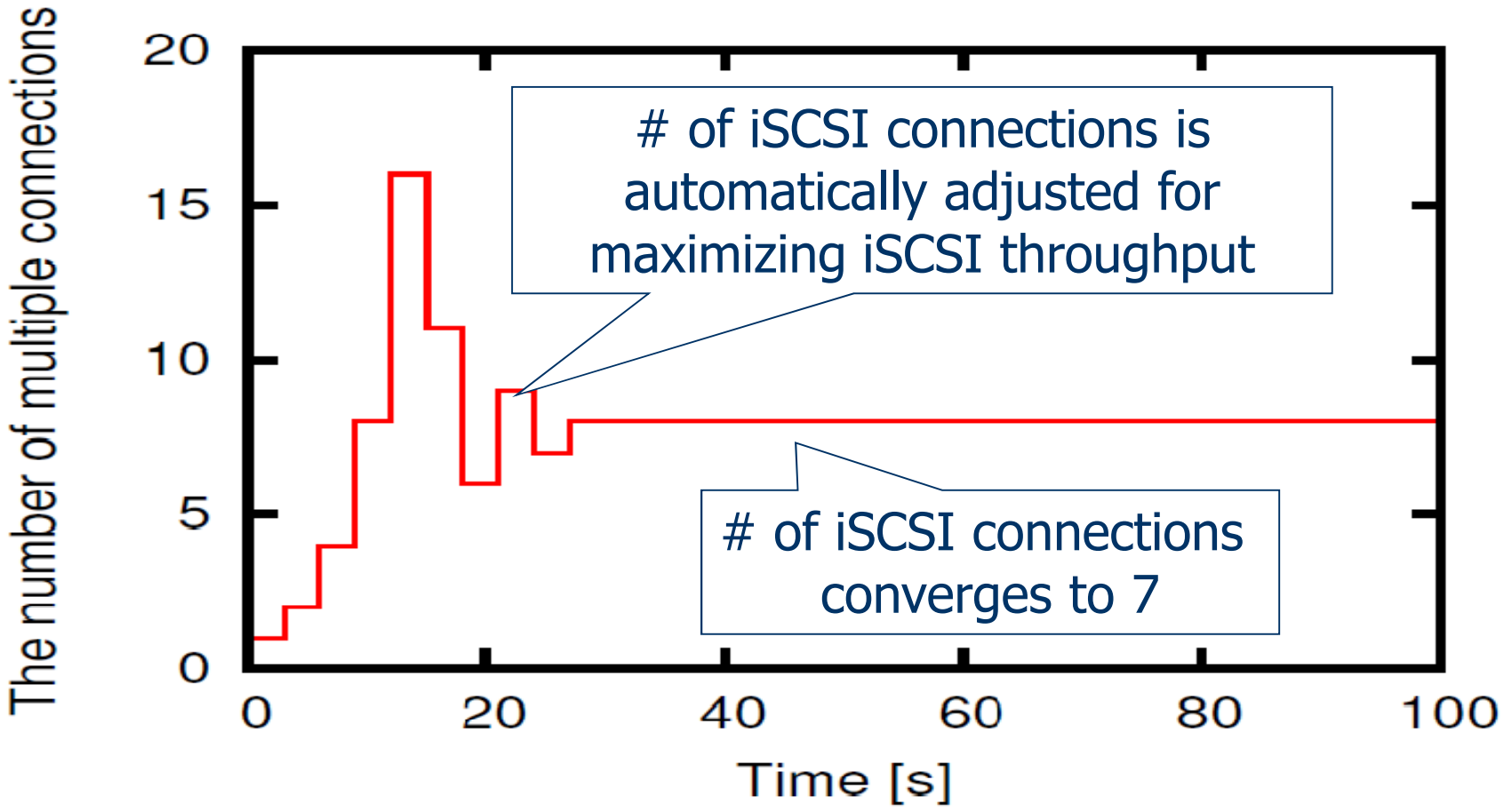
- Network configuration
  - NTT (Tokyo) --- NII (Tokyo) --- Osaka Univ. (Osaka)
  - Allocate end-to-end L1 path
- Data transfer with iSCSI-APT
  - Osaka Univ. (Osaka) -> NTT (Tokyo)



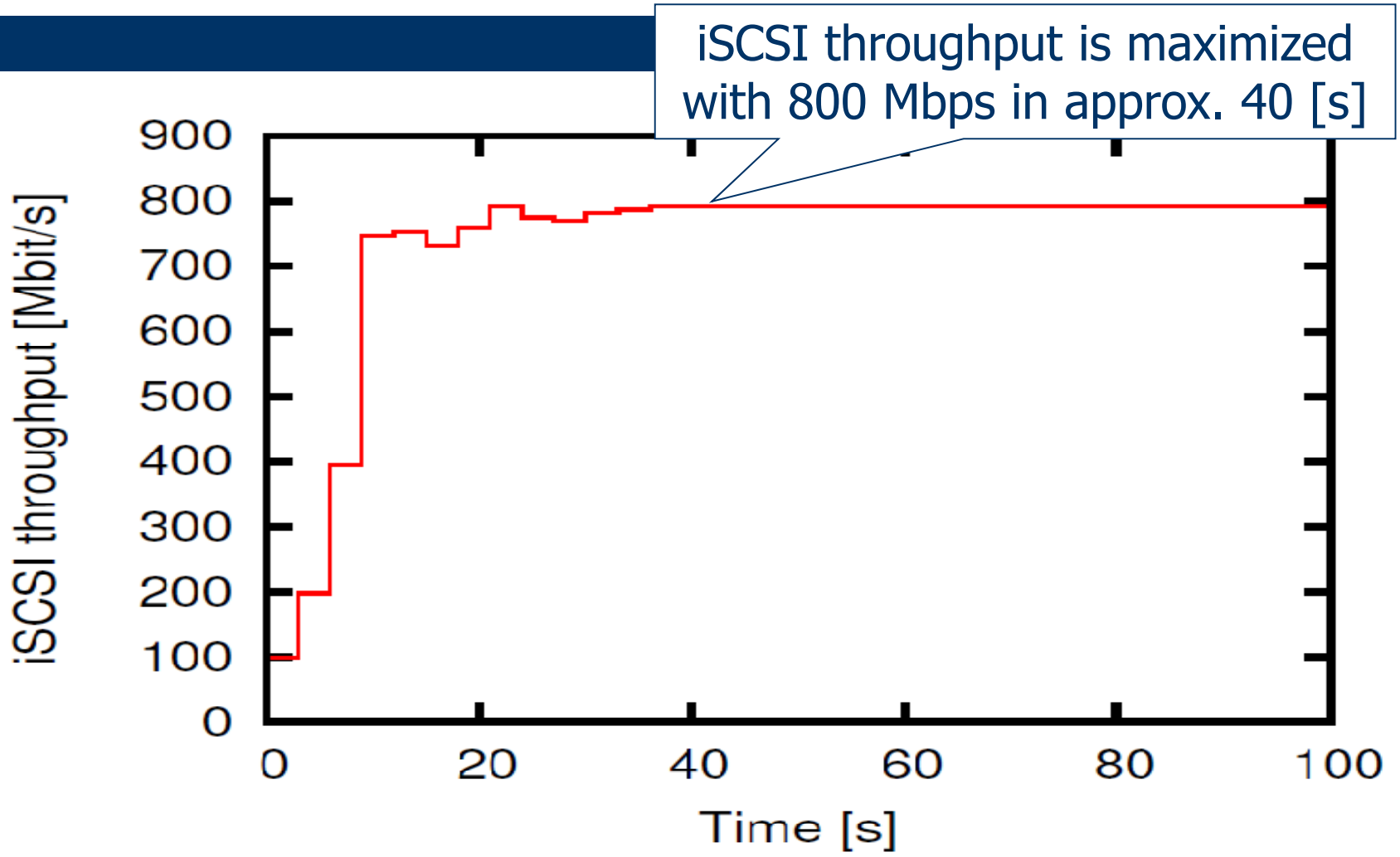
Evolution of iSCSI Throughput (Case of 600Mbps L1 Path)



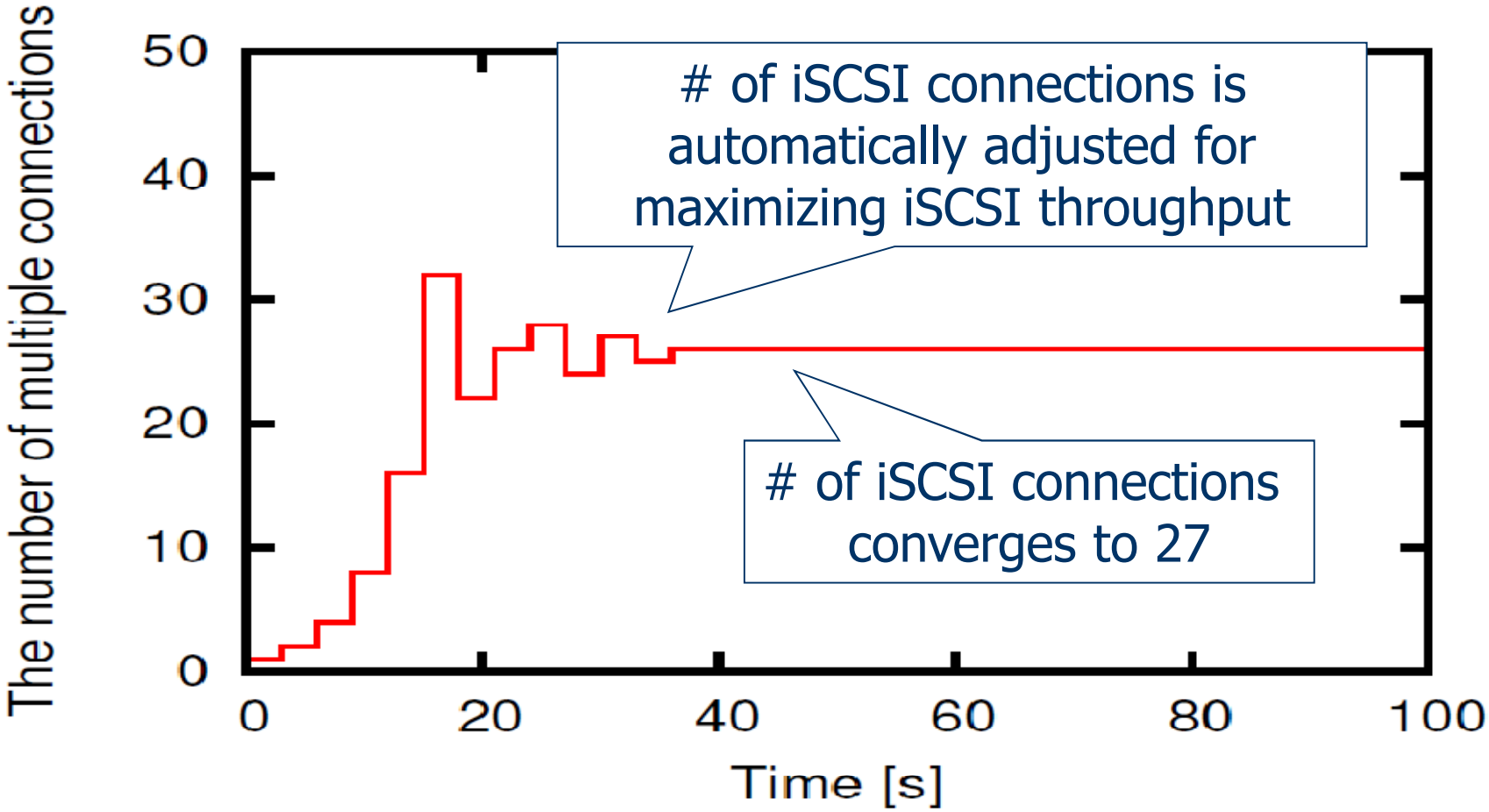
Evolution of # of iSCSI Connections (Case of 600Mbps L1 Path)



Evolution of iSCSI Throughput (Case of 900Mbps L1 Path)



Evolution of # of iSCSI Connections (Case of 900Mbps L1 Path)



Conclusion

- Explained overview of our iSCSI-APT
  - A mechanism which maximizes iSCSI throughput
    - Adjusts the number of multiple connections so that iSCSI throughput is maximized
    - Operates only at an iSCSI initiator
- Performance evaluation of iSCSI-APT on SINE3
  - Many scenarios thanks to L1 on-demand service
  - Successfully showed effectiveness of iSCSI-APT

Future Works

- Further experiments
  - Experiment with 10 Gbps interface
  - Performance evaluation among multiple sites
  - Experiment with dynamic bandwidth allocation
- Enhance iSCSI-APT algorithm
  - By taking account of several characteristic of the iSCSI protocol

Thank you for your kind attention