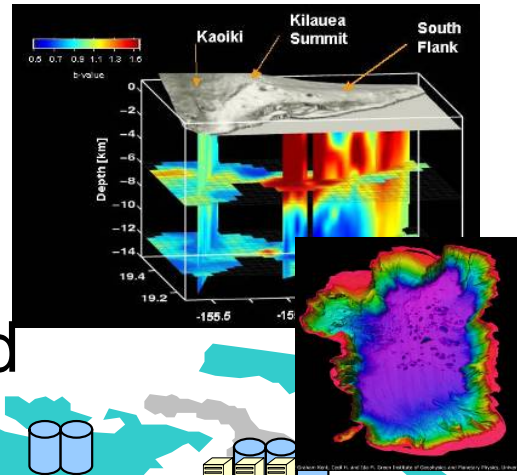
A Novel High-Performance Transport Protocol Considering Fairness with TCP in Long-distance High-speed Network

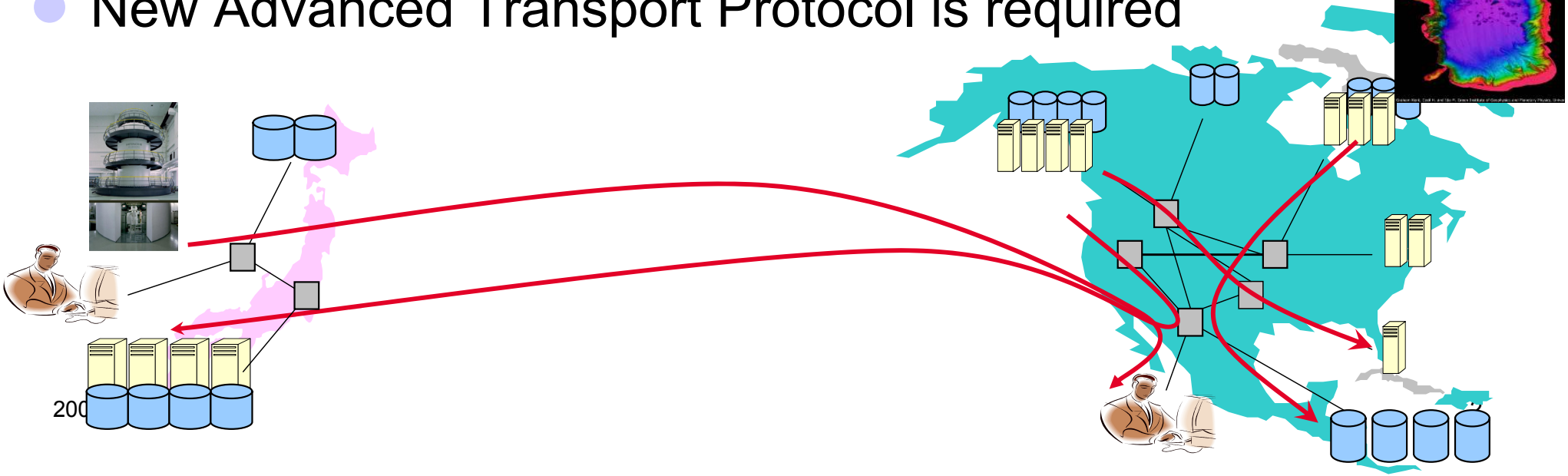
Graduate School of Information Science and Technology, Osaka University
Cybermedia Center, Osaka University
JGNII Osaka Research Center, NiCT

Fumiaki TAMESHIGE Ken-ichi BABA
Masaaki NORO Shinji SHIMOJO

Research Backgrounds

- Network bandwidth and delay is increasing rapidly
  - 10Gbps is common in high speed network by DWDM
- New applications transfer Large Volume Data
  - Handling more than terabyte in DataGrid
- TCP becomes inefficiency in LFN
  - LFN; Long Fat pipe Network
  - Because of Large Bandwidth Delay Product
- New Advanced Transport Protocol is required





High-Speed Transport Protocol (1/2)

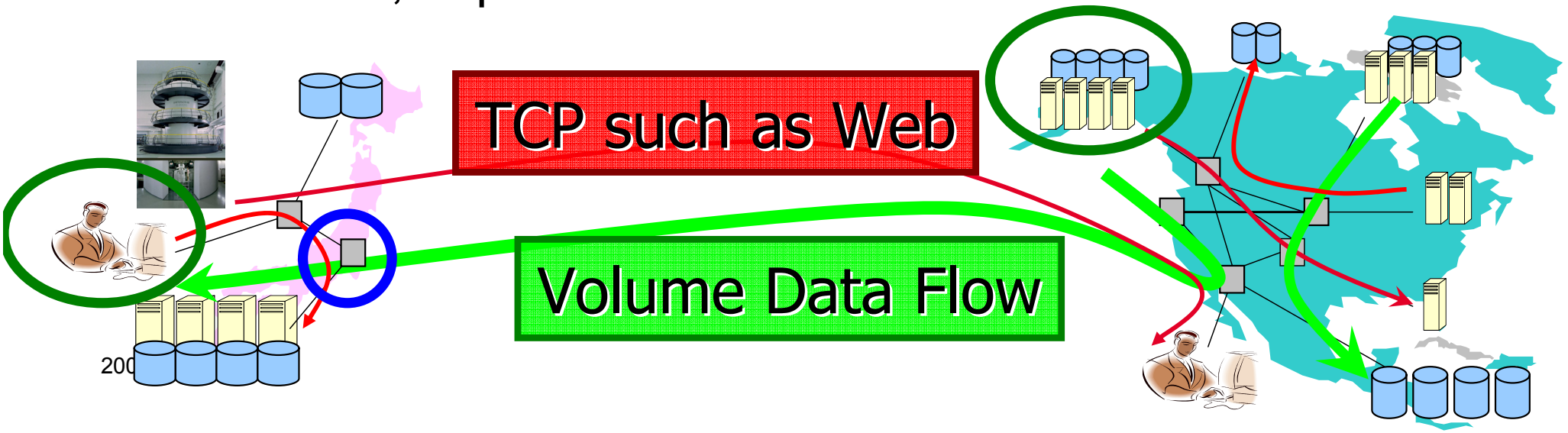
- GridFTP
  - High-Throughput with Parallel TCP connection
  - Difficult to set the adequate parameter
- HighSpeed TCP, Scalable TCP, Fast TCP
  - Improve AIMD window control in TCP
  - Difficult to realize fairness to TCP
- XCP
  - New congestion control protocol involve router
  - It costs much costs in the modification of network
- RBUDP, TSUNAMI
  - Rate-based reliable UDP
  - Protocols for dedicated network

High-Speed Transport Protocols (2/2)

- UDT (UDP based Data Transfer)
  - Application level protocol for volume data transfer
  - Over UDP with reliability and congestion control
  - Receiver sends ACK and NAK periodically to inform the sender of lost information
  - UDT can share bandwidth almost equally to TCP in small network; eg 100Mbps/10ms
- Problem of UDT
  - Basic performance and Impact of TCP do not investigate fully
  - A lot of packet losses may occur when multiple UDT flows coexist in the same path

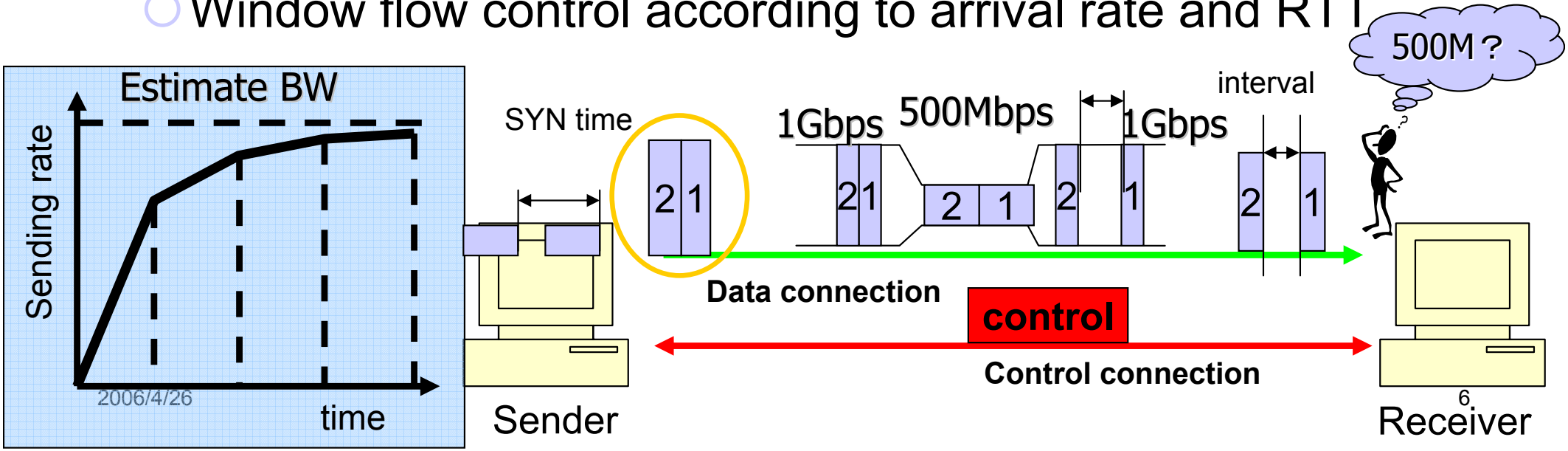
Research Target

- Goal
  - To realize effective high-speed data transfer in LFN
    - Prevent the influence to other competitive flow such as TCP
    - Stable high Throughput with Volume data flows
- Approach
  1. UDT-g; A new protocol in End node side
  2. RED-i; Improvement of RED in network side



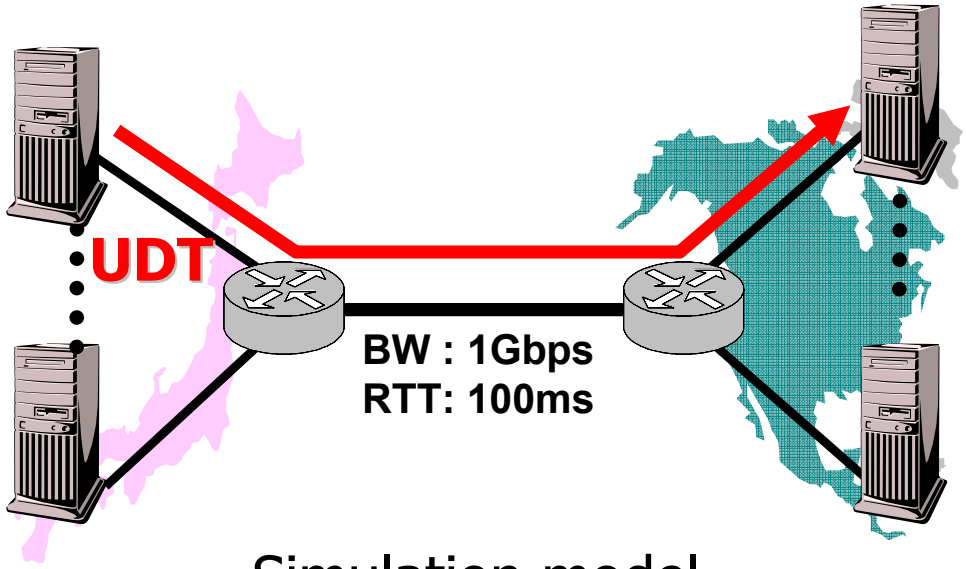
Data Transfer by UDT

- Bandwidth Estimation by Packet Pair
  - Data packets are sent periodically back to back
  - Receiver measures the interval of arrival packets
  - Packet pair estimates the link capacity in non-congestion
- Rate control and Flow control
  - Rate control with adjusting the interval of sending packet
  - Approaching the estimate bandwidth by Decreasing AIMD
  - Window flow control according to arrival rate and RTT

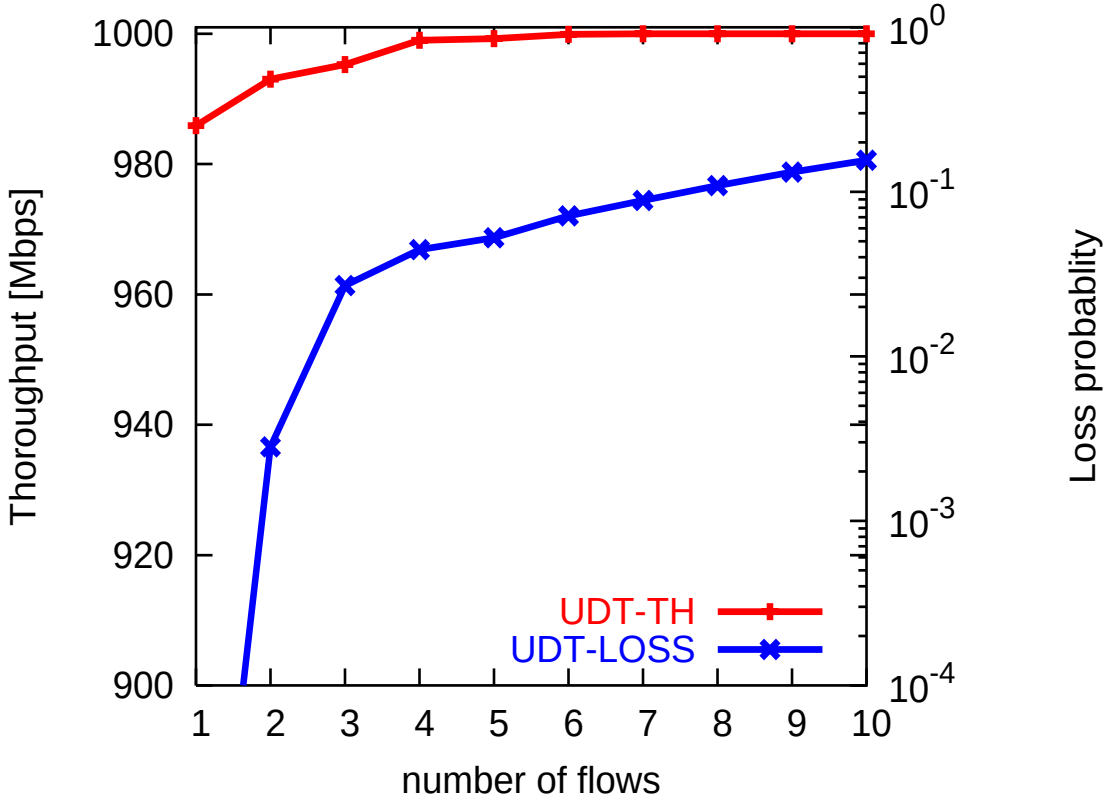


Basic property and Problem of UDT

- UDT can use up link capacity in spite of number of flow
- Packet loss probability increases rapidly as the increase in the number of UDT flow
- Each UDT flows estimate more than actual available bandwidth, so UDT is aggressive to other flows



Simulation model



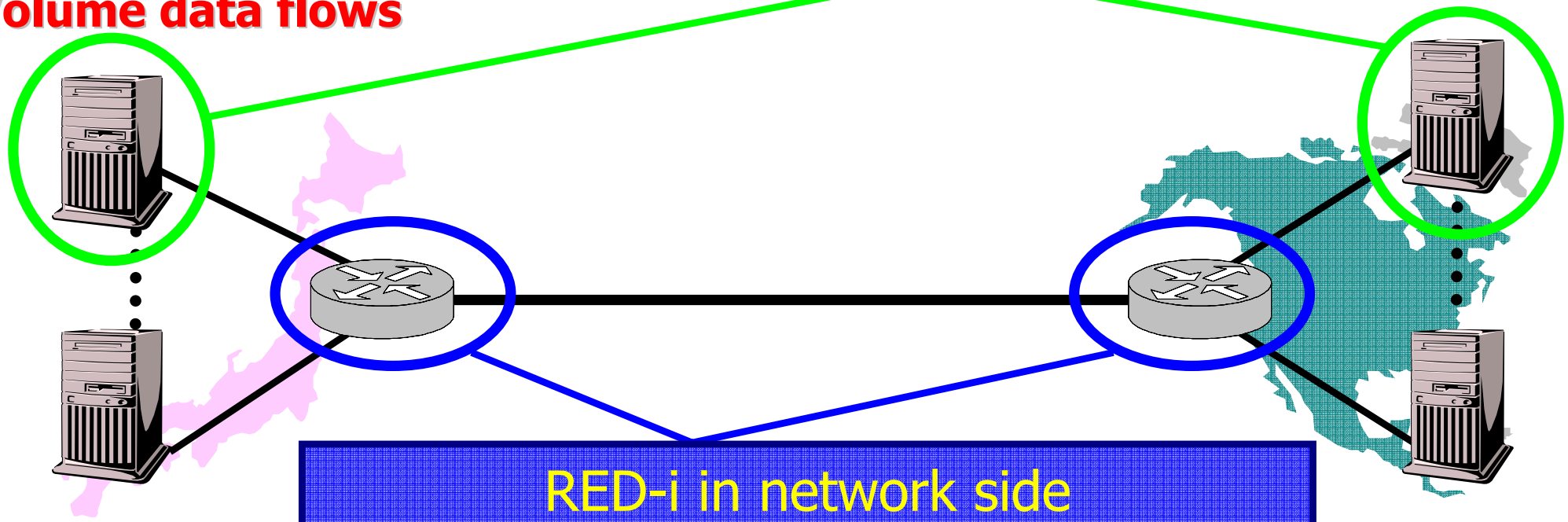
Proposed Control Method
combined approach; UDT-g and RED-i

UDT-g in End node side

Estimating Available Bandwidth

Stable Rate Control

Volume data flows



RED-i in network side

Improving packet loss probability

Prevent the increasing of queuing delay

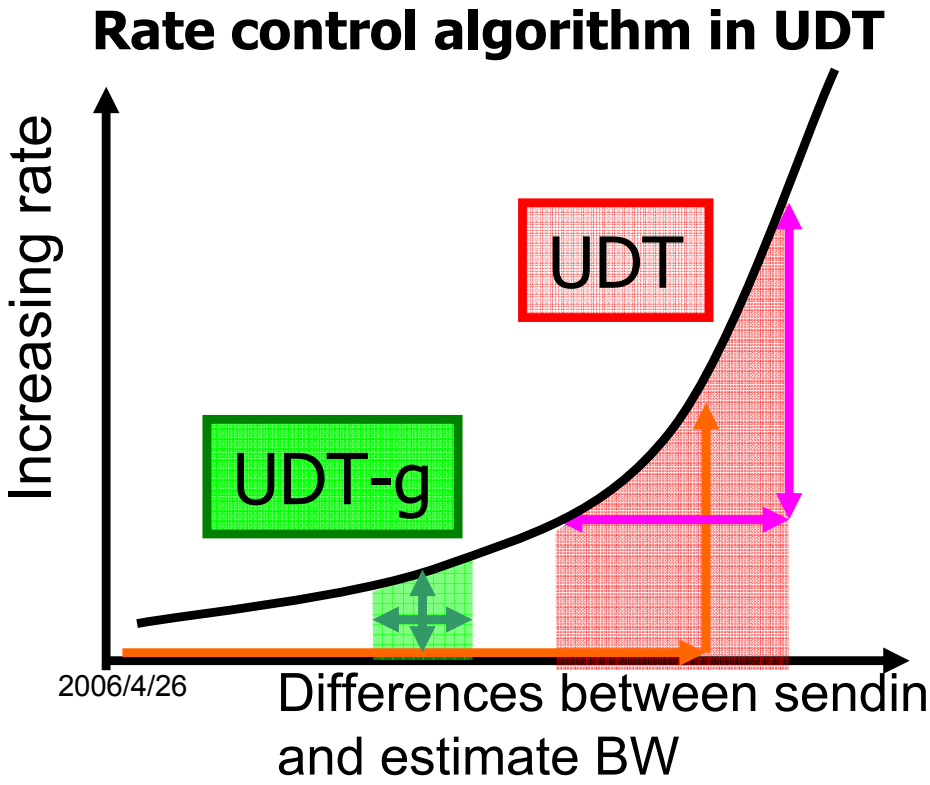
New Bandwidth Estimation in UDT-g

- Conventional：UDT
  - EWMA calculations with low constant smoothing value
    - Overestimate bandwidth by packet pair
    - Delay congestion detection in long-distance network
- Proposal：UDT with gentle Bandwidth Estimation
  - To control estimate bandwidth and proper rate control
    - UDT-g compare the average value with current measurements
    - Calculation lower value in congested network on purpose
  - Reduce aggressive packet sending in congestion
    - α decide trade-off between efficiency and fairness

```
if (ave_bw ≦curr_bw)
    ave_bw = ave_bw*0.875 + curr_bw*0.125;
else
    ave_bw = ave_bw*(1−α) + curr_bw*α;   (0.125<α<1)
```

Relationship between UDT-g and Rate control with Decreasing AIMD

- Estimate Available Bandwidth rather than Link Capacity
  - UDT-g enable to control differences between sending rate and estimate Bandwidth

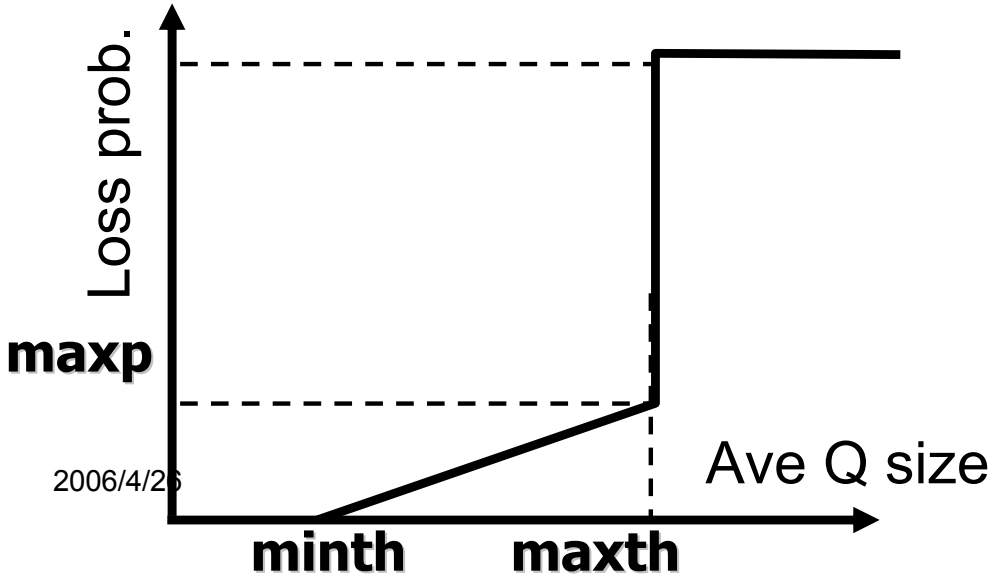


UDT-g can inhibit range of increasing rate and band of fluctuation

Stable rate control in congestion

Queue mechanism of the router

- Drop Tail: widely implemented with FIFO queue
  - Consecutive packet loss in congestion
  - Difficult to set queue size in LFN
- RED: Random Early Detection
  - Dropping packet probabilistically before buffer over flow
  - Keep average queue size small, improve fairness among flows

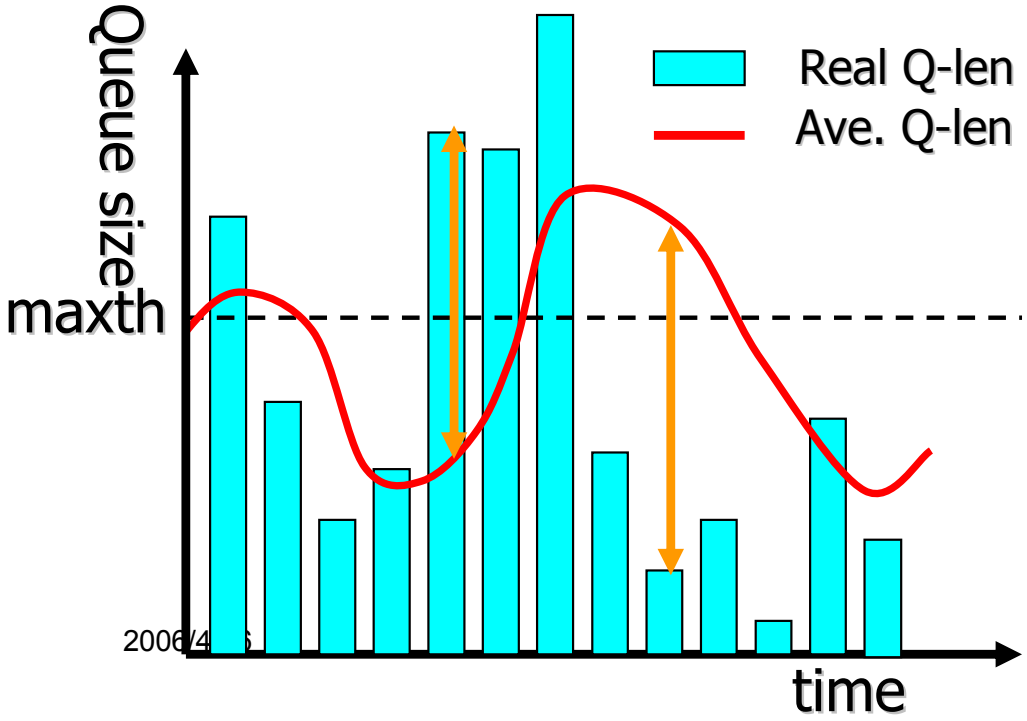


Prevent consecutive packet drop by using RED in LFN

Improvement of packet loss probability & queuing delay

Improvements of RED

- Conventional : RED (Random Early Detecition)
  - RED cannot drop packet in appropriate manner with average queue size
  - Large differences between average and instantaneous queue size
    - Actual queue size increase or decrease drastically in LFN
- Proposal : RED based instantaneous queue length
  - RED-i expects random packet drop even if in long-distance network
  - RED-i can react drastic changes of queue size and is sensitive to congestion

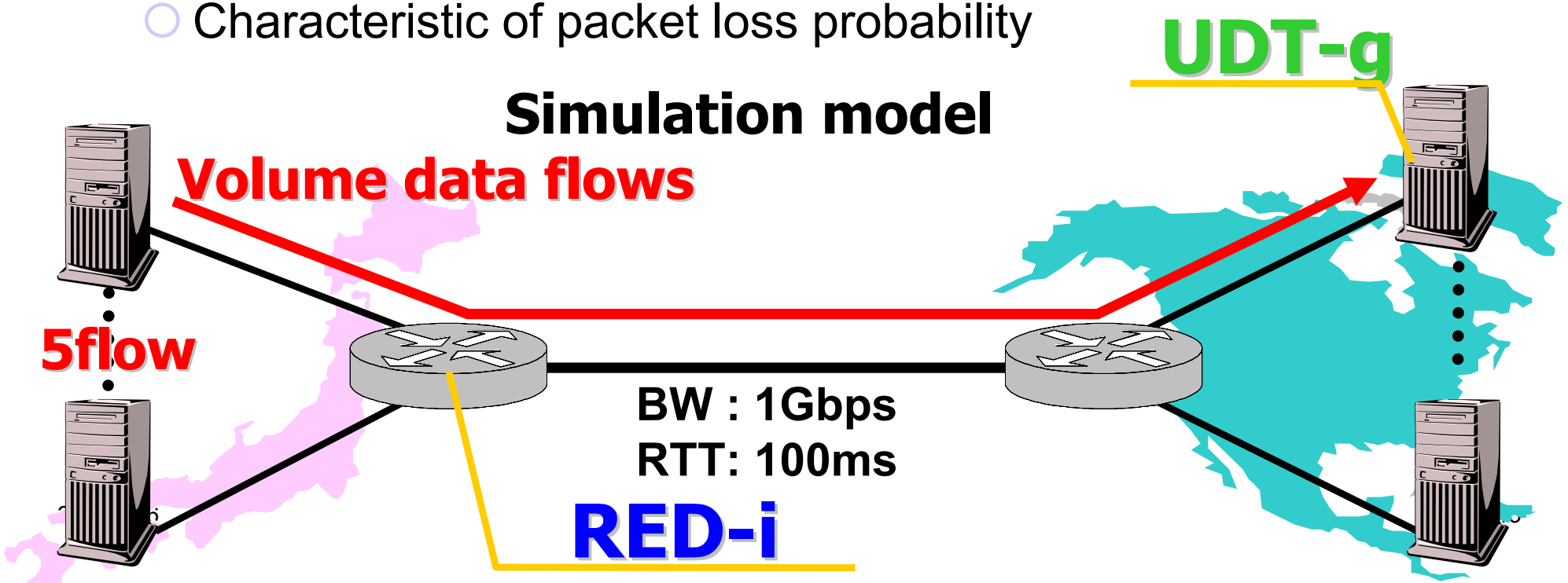


Average queue size does not keep up with instantaneous queue size

Probabilistic packet drop works effectively even if in LFN

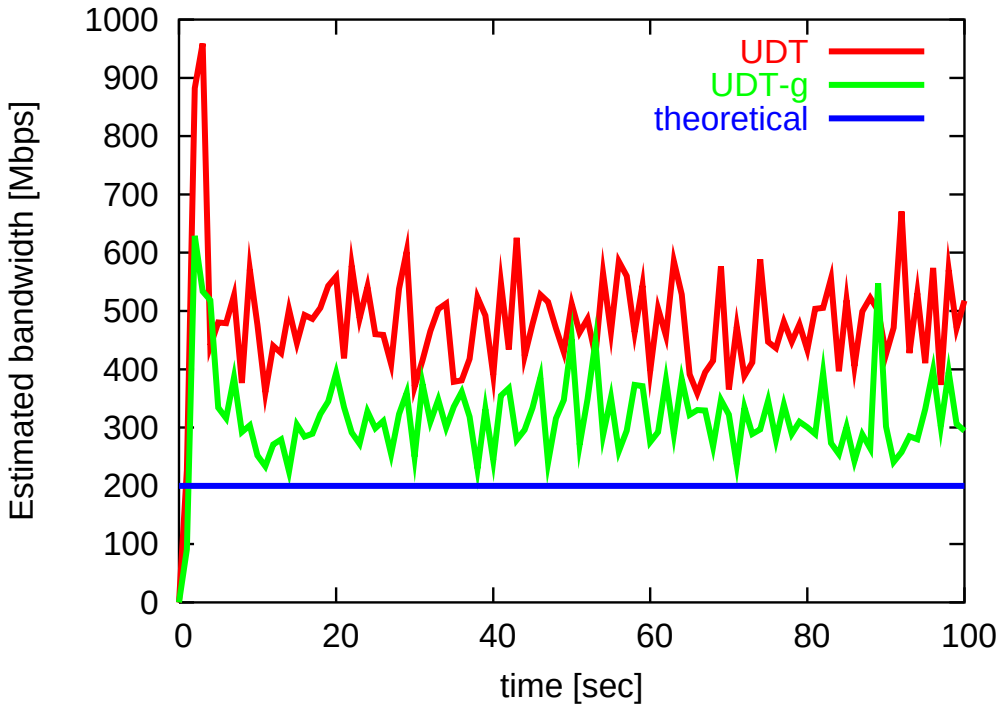
Performance Evaluation multiple volume data flows

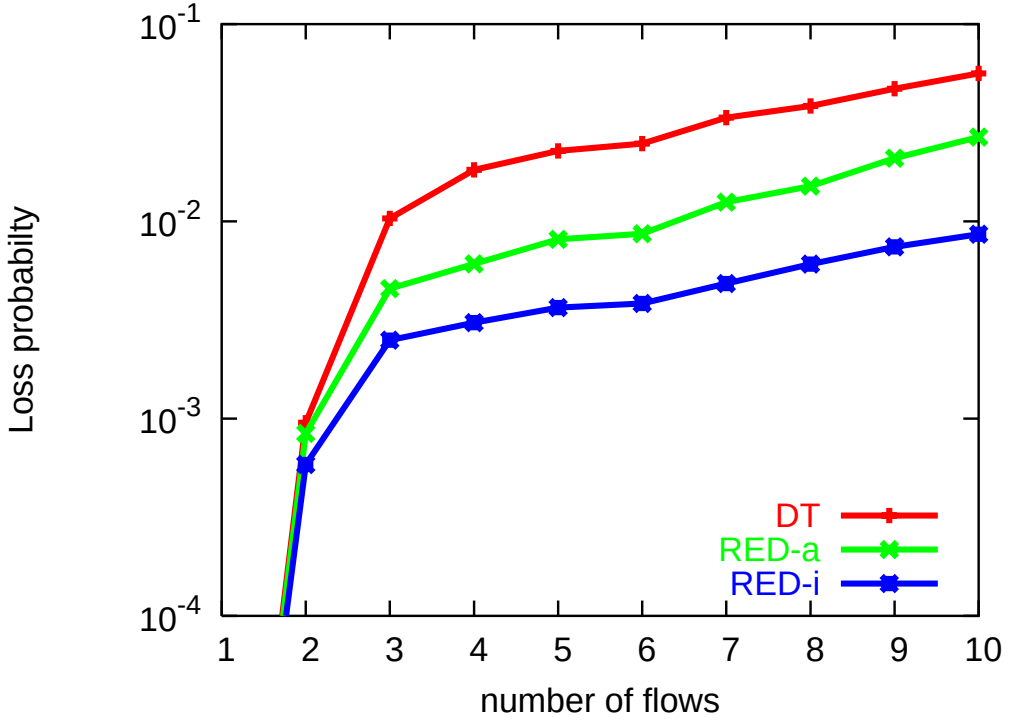
- High performance volume data transfer in LFN
  - Intercontinental volume data transfer over 1Gbps
    - more than terabyte class;
    - Relatively small number of end node; e.g. 1～10
- Performance measure
  - Characteristic of estimate bandwidth in volume data flow
  - Characteristic of packet loss probability



Multiple Volume data flows in LFN

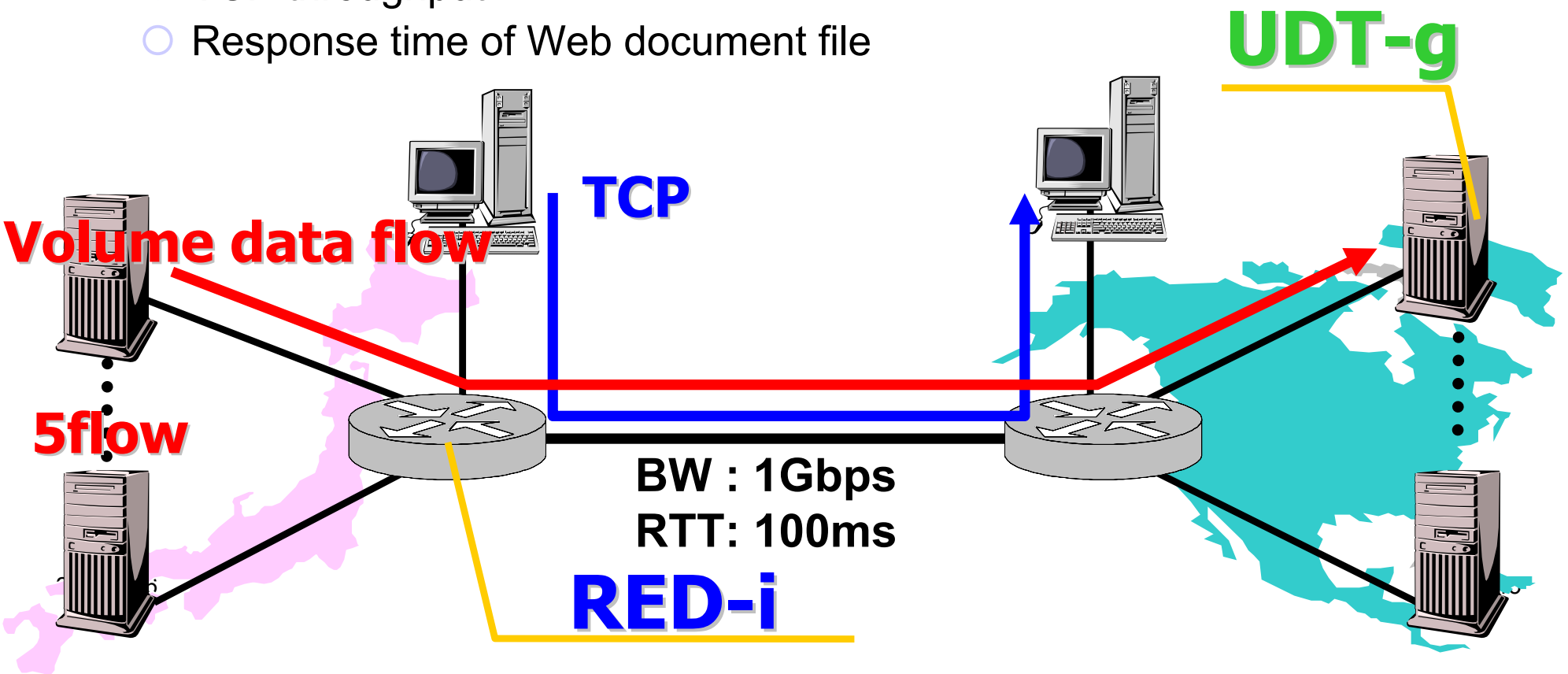
- UDT-g
  - Controls excess estimate bandwidth and calculate proximity available bandwidth
- RED-i
  - reduces average queue size compared with DT
  - Improves packet loss probability when multiple flows coexist



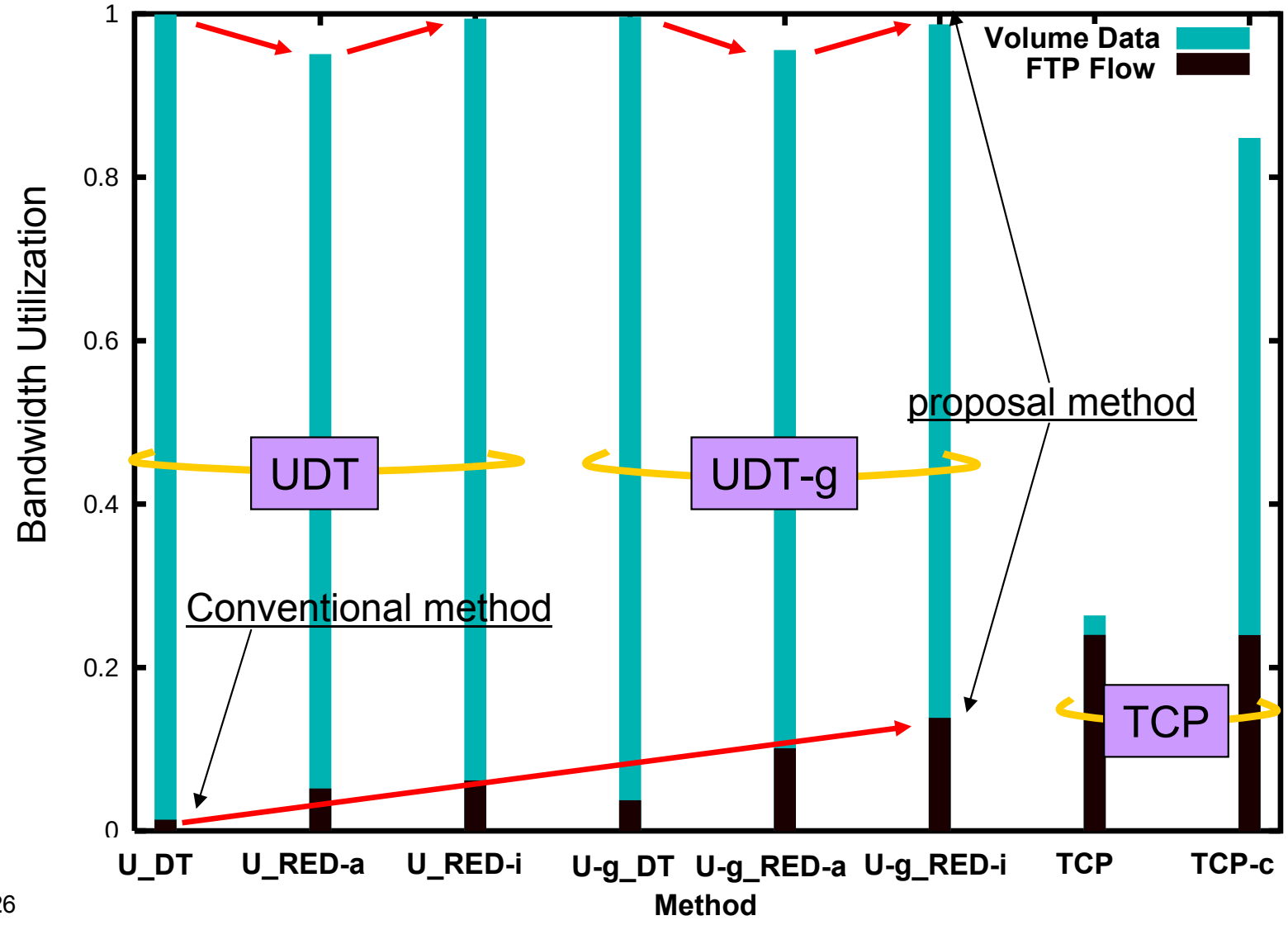


Influences of TCP flows

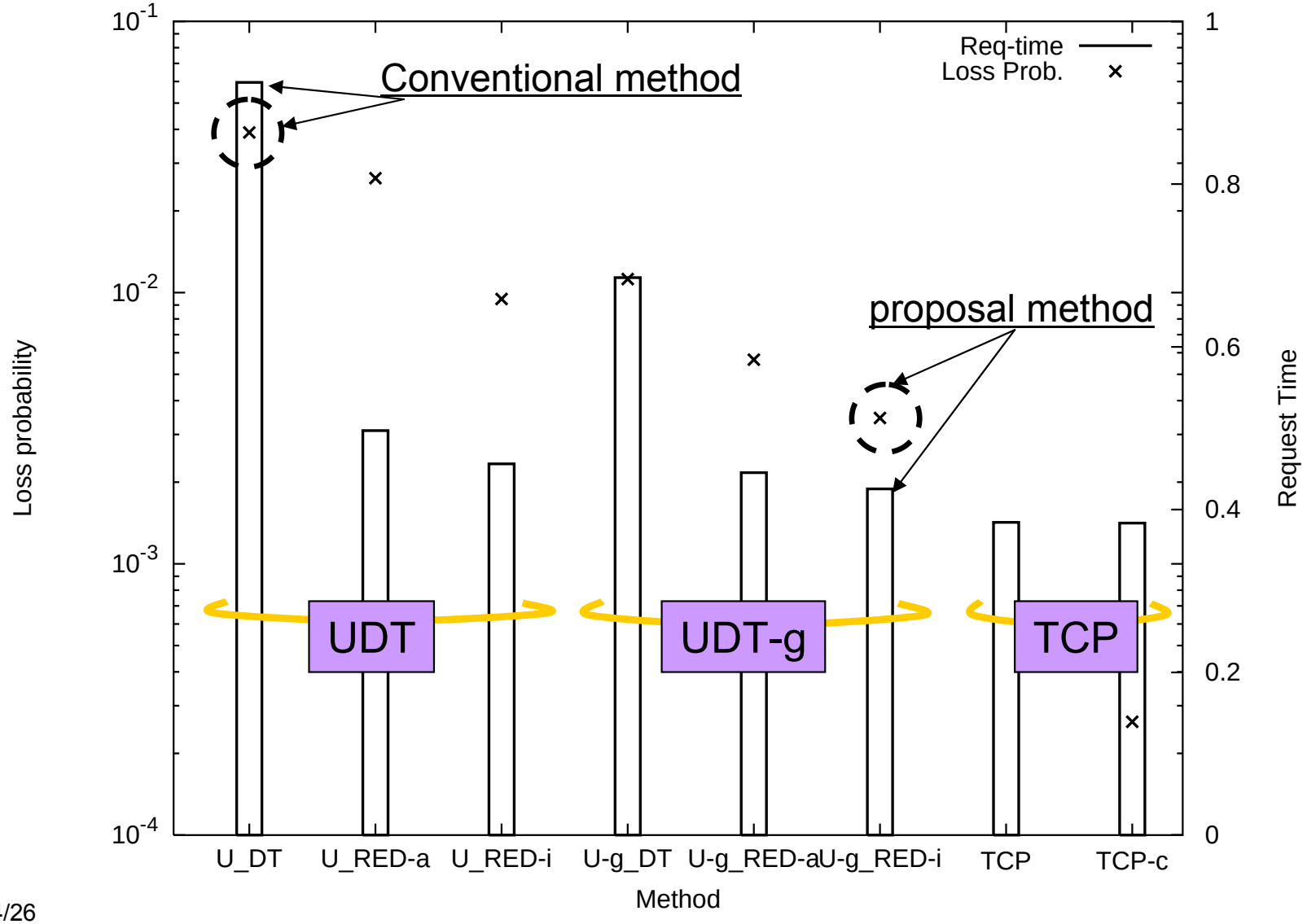
- Competitive TCP flow
  - Long lived TCP：FTP flow
  - Short lived TCP：HTTP flow
- Performance measure
  - Utilization of Bandwidth
  - TCP throughput
  - Response time of Web document file



FTP flow and Volume data flow



HTTP flow and Volume data flow



Conclusion

- Effective Volume data transfer in LFN
  - TCP cannot fill up with bandwidth
- Proposed approach; UDT-g and RED-i
  - New Bandwidth Estimation to realize stable rate control in congested network
  - Improvement of packet loss probability and control queue size in bottleneck link
- Feature work
  - Further investigation of effective high volume data transfer and active queue management in LFN