Advanced Network Architecture Research Group

Design and Implementation Experiments of Scalable Socket Buffer Tuning

Tatsuhiko Terai

Department of Informatics and Mathematical Science,

Graduate School of Engineering Science,

Osaka University

Advanced Network Architecture Research Group

Background

- Explosive increase of network traffic due to rapid increase of Internet users
  - Many improvements on network to accommodate increasing traffic
    - Link Bandwidth
    - TCP congestion control algorithm
  - Few discussions on endhost improvement
    - Ex. Busy WWW Server receives hundreds of request for document transfer per second.

The bottleneck of the data transfer processing is shifted from network to endhosts

Advanced Network Architecture Research Group

Objective

- Effective allocation of endhosts resources becomes more important
- Propose a novel architecture, SSBT (Scalable Socket Buffer Tuning)
  - High-performance and fair service for many TCP connections at the sender host
    - E-ATBT (Equation-based Automatic TCP Buffer Tuning)
    - SMR (Simple Memory-copy Reduction)

Send Socket Buffer Assignment

- Busy Internet servers (WWW, Proxy,…) handle many TCP connections which have different characteristics
  - RTT, packet loss rate, bandwidth, …
- In original method, sender host assigns fixed size of buffer to each TCP connection
  - Fixed size assignment may cause unfair and ineffective usage of send socket buffer
    - Different connections require different size of send buffer according to network conditions

Fair and effective buffer assignment considering network condition is needed for improving endhost performance

E-ATBT
(Equation-based Automatic TCP Buffer Tuning)

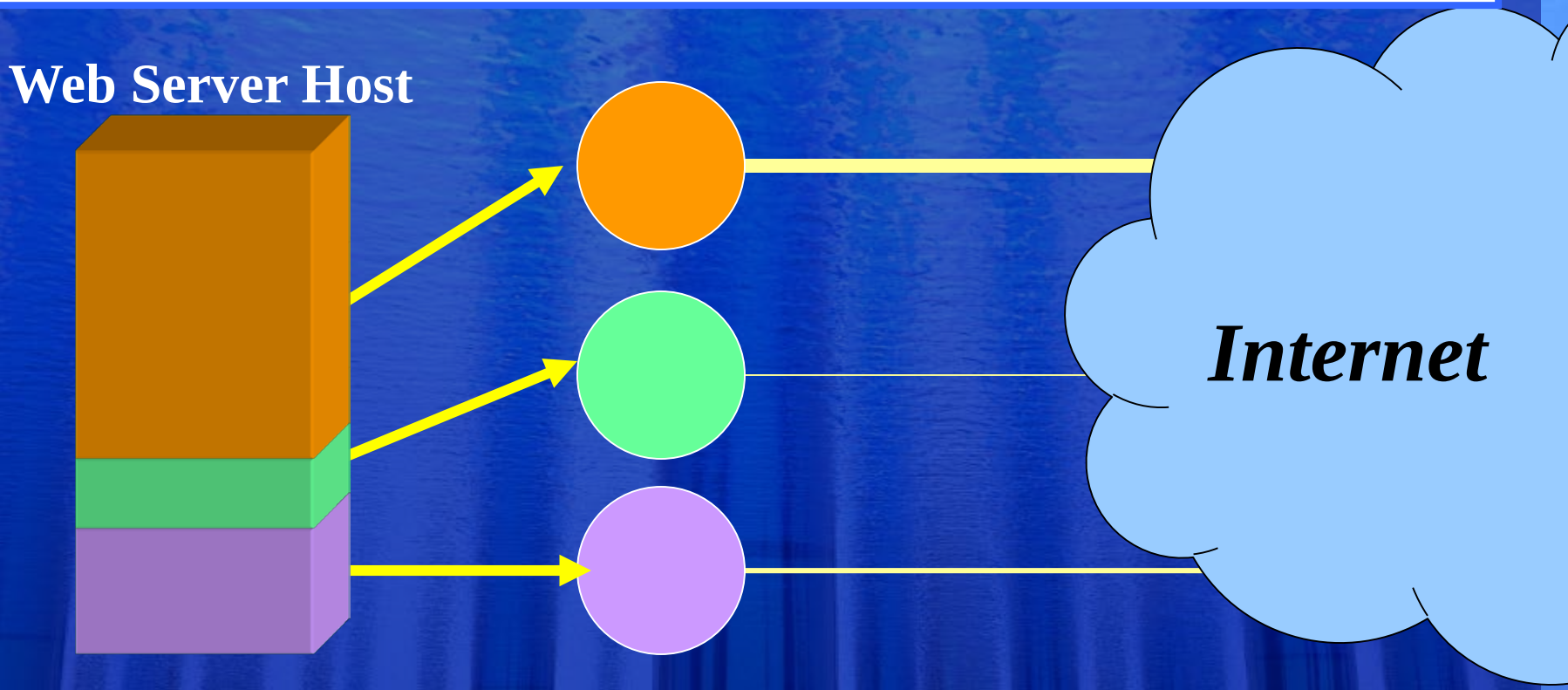
Advanced Network Architecture Research Group

- Provide fair and effective send socket buffer assignment
  - Estimate an ‘expected’ throughput of each TCP connection by monitoring three network
  - *p (packet loss rate), rtt (Round Trip Time), rto (Retransmission Time Out)*
- Determine assigned buffer size from the estimated throughput
- Max-Min fairness policy for re-assigning the excess buffer
  - Re-assigned to the connections need more buffer the required buffer size of those connection

E-ATBT Method for assigning the send socket buffer

Advanced Network Architecture Research Group

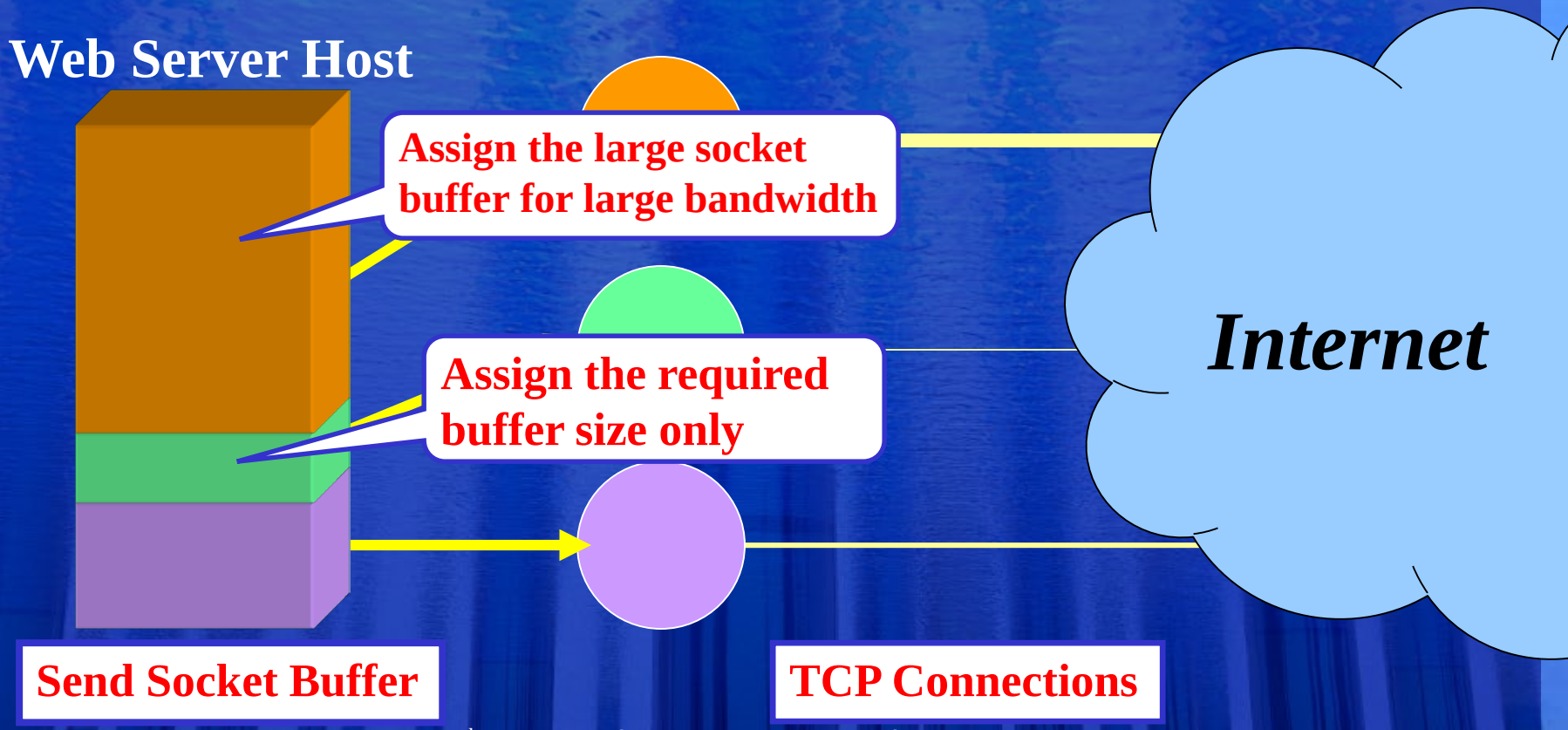
In E-ATBT, assigned buffer size is determined from the estimated throughput based on the mathematical analysis method



E-ATBT Method for assigning the send socket buffer

Advanced Network Architecture Research Group

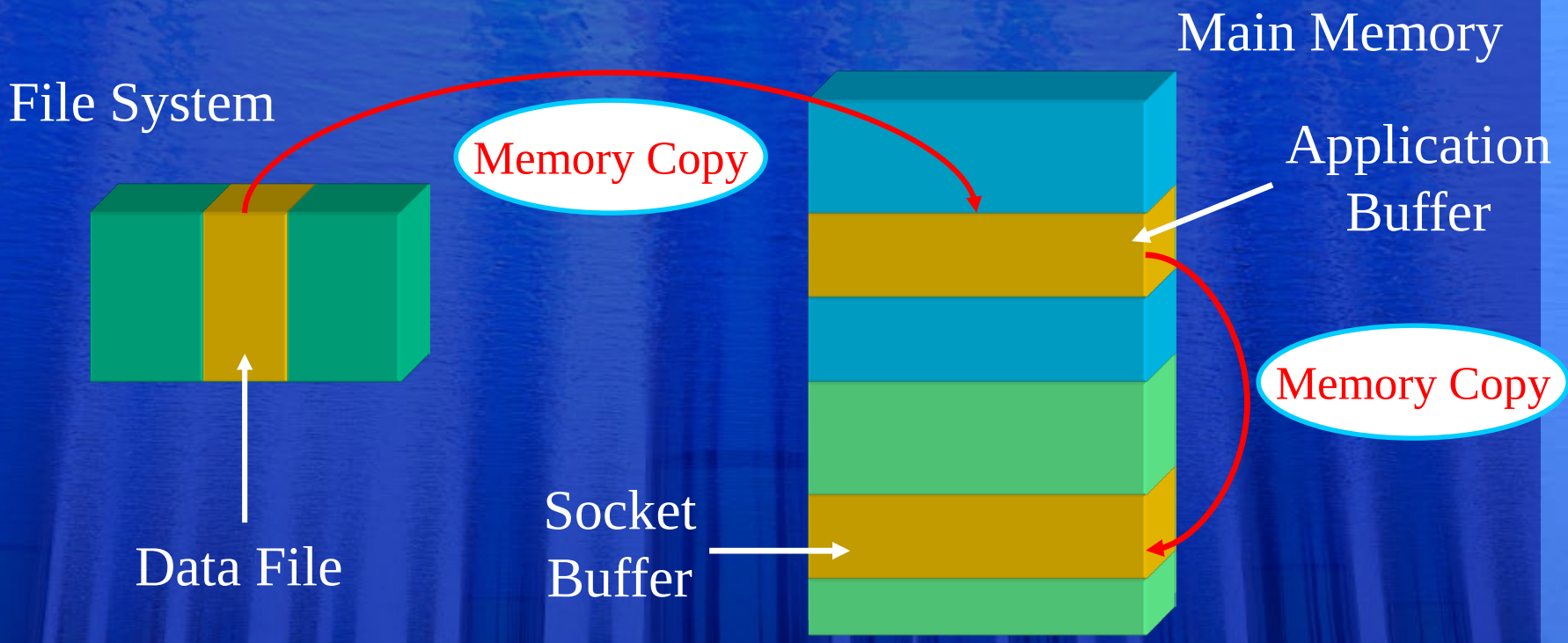
In E-ATBT, assigned buffer size is determined from the estimated throughput based on the mathematical analysis method



SMR
(Simple Memory-copy Reduction)

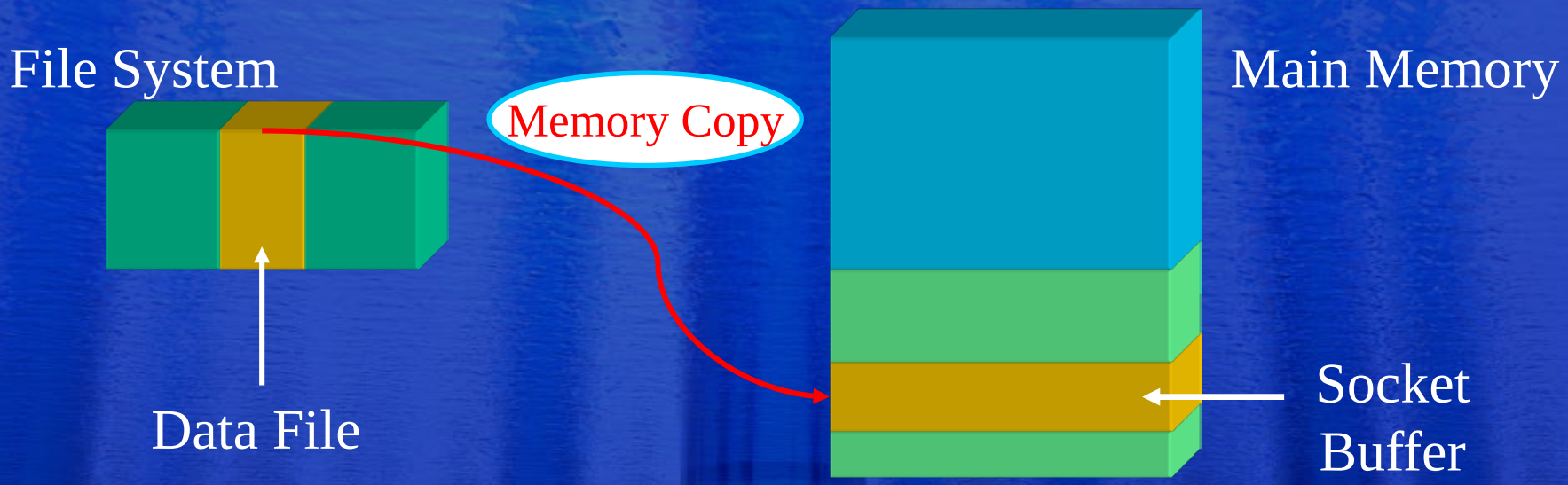
Advanced Network Architecture Research Group

- Original mechanism needs two memory copy operations. Memory copy is large overhead on endhost processing



SMR
(Simple Memory-copy Reduction)

Advanced
Network
Architecture
Research
Group



SMR scheme reduces the data transfer overhead at the sender host by reducing the redundant memory copy operation

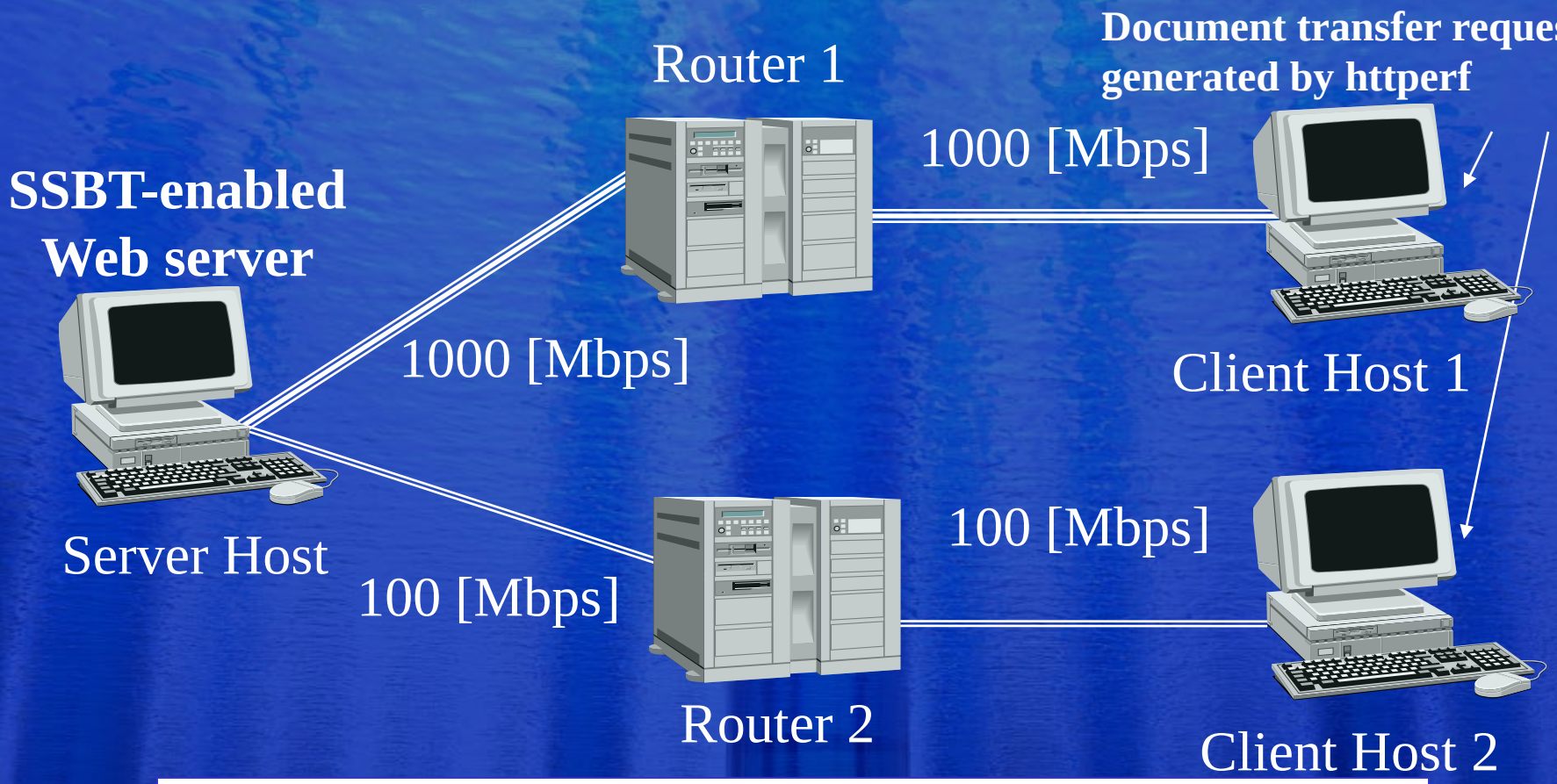
Implementation Experiments

- Fair buffer assignment among different connections
  - Time-dependent behavior of the assigned buffer size
- Web server performance evaluation
  - Average performance gain of the SSBT scheme
    - Considering realistic web access traffic [1]
      - Document size distribution
      - Idle time distribution of requests
      - Embedded documents distribution

---

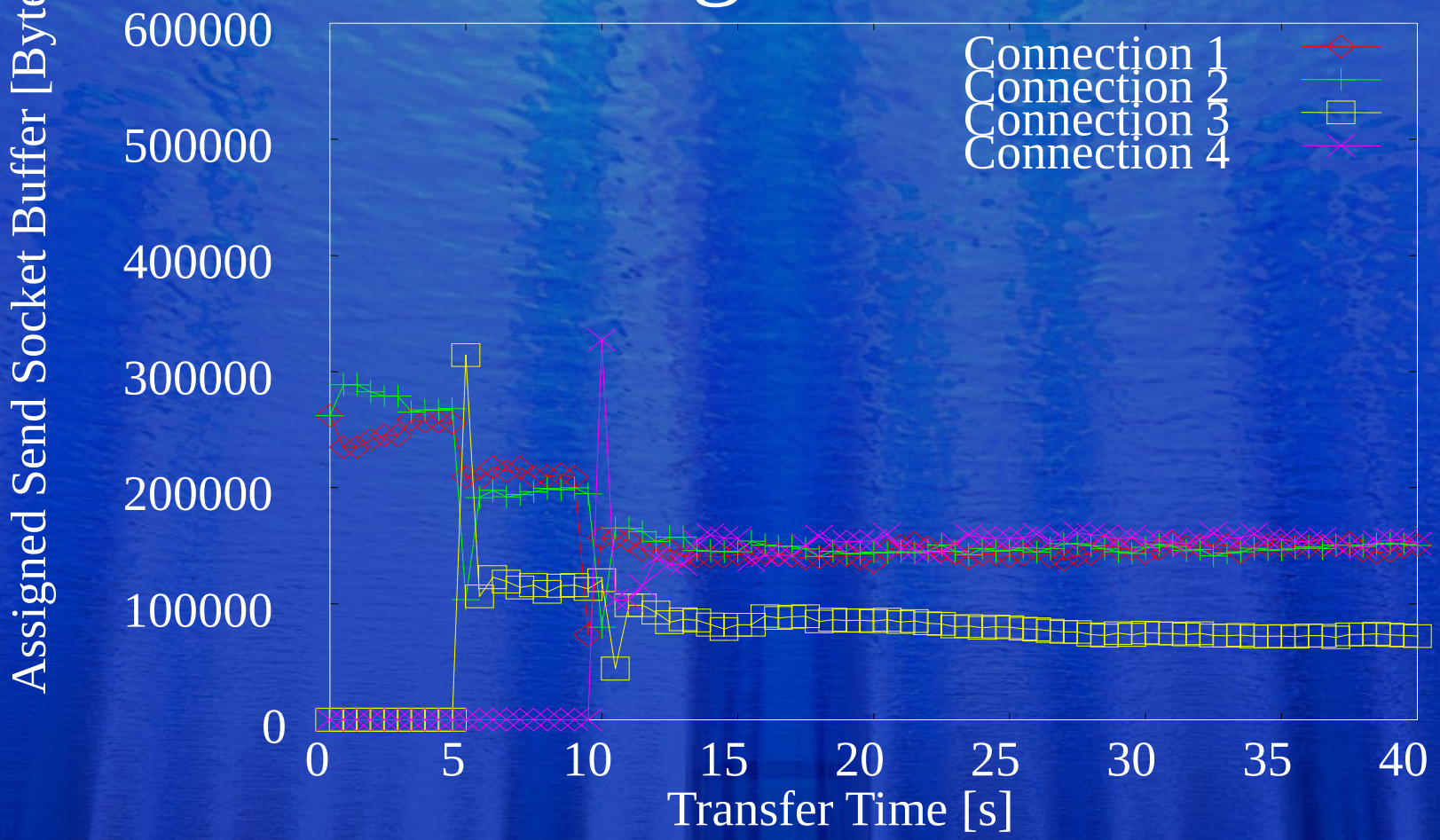
[1] P.Barford and M.Crovella, “Generating Representative Web Workloads for Network and Server Performance Evaluation”, in *Proceedings* of ACM SIGMETRICS ’98, 1998

Network topology



Each client generate the requests for document transfer to Web server and measure the data transfer time

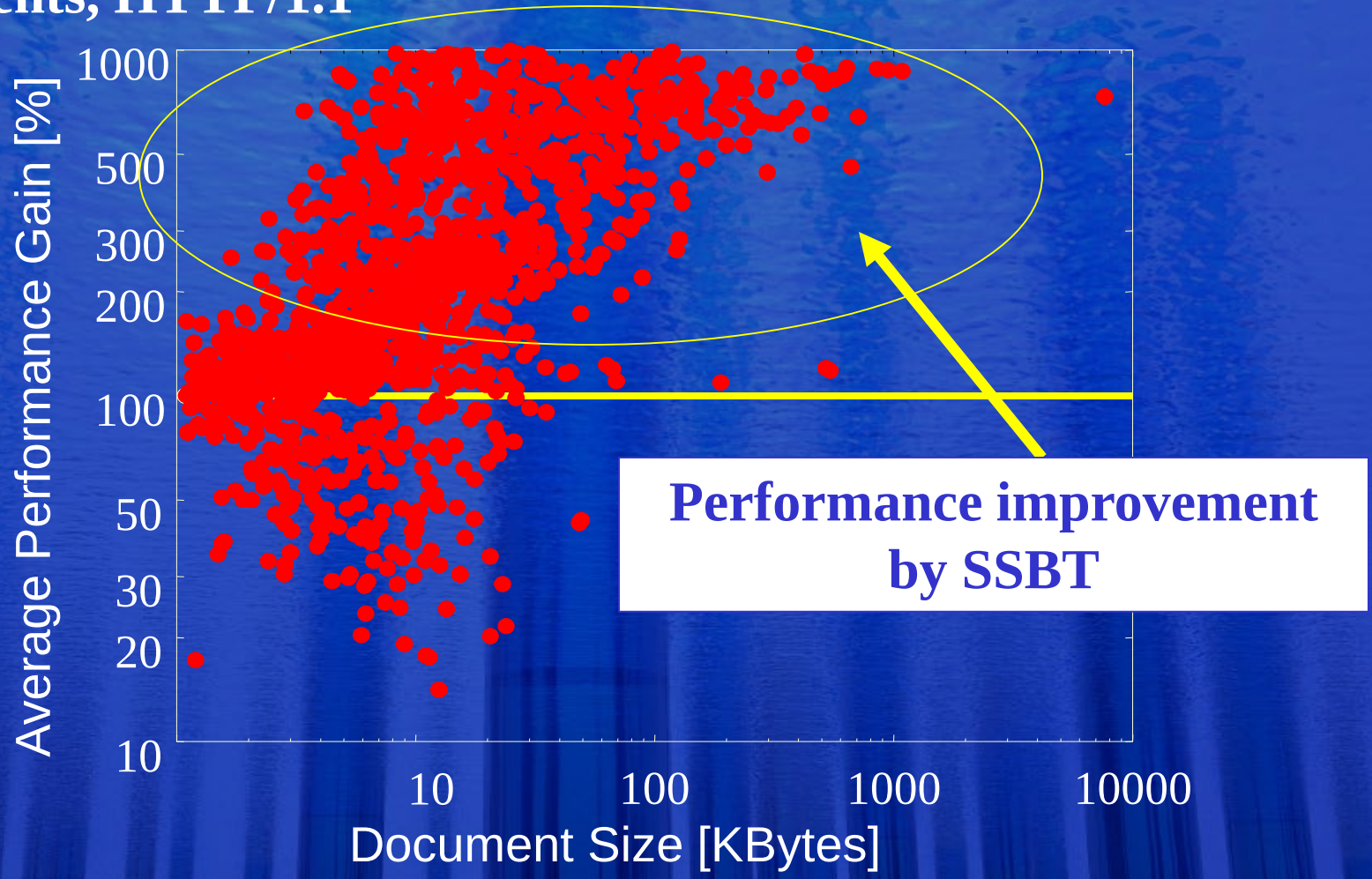
Buffer assignment result



E-ATBT can provide the stable and fair buffer assignment

Web server performance

600Clients, HTTP/1.1



Conclusion Remarks and Future Works

Advanced Network Architecture Research Group

- Proposed SSBT for utilizing the send socket buffer effectively and fairly
- Confirmed the effectiveness of the SSBT algorithm through implementation experiments and shown SSBT can improve the overall performance of a server
- New resource management scheme for Internet busy server (Ex. HTTP Proxy server)
  - Enhanced E-ATBT for proxy servers
  - Manage the persistent TCP connections

Advanced
Network
Architecture
Research
Group