On Modeling GridFTP using Fluid-Flow Approximation for High Speed Grid Networking

Hiroyuki Ohsaki
Department of Information Networking
Graduate School of Information Science and Technology, Osaka University, Japan

Contents

Research background

- TCP limitations
  - Uses AIMD window flow control for congestion control
    - AIMD: Additive-Increase and Multiplicative-Decrease
  - Assumes packet loss implies network congestion
  - Causes low throughput in wide-area networks
- GridFTP as a data transfer protocol for Grid
  - Solves several problems of the existing TCP

Known problems of TCP

- Low throughput
  - Significant delay at TCP's three-way handshake
  - Slow-start phase takes too long when RTT is large
  - Packet loss doesn't always indicate network congestion
  - Window flow control causes bursty packet transfer
- Non-fair bandwidth allocation
  - Realizes fairness only for homogeneous TCP connections
  - Bias against # of hops and round-trip time
- Non-negligible overhead
  - CPU overhead due to TCP protocol stack itself
  - CPU overhead and/or processing delay due to TCP checksum
  - Bandwidth overhead due to large TCP header

Research objectives

- Performance evaluation of GridFTP using an analytic approach
  - Focus on two notable features of GridFTP
    - Parallel data transfer
    - Automatic negotiation of a TCP socket buffer size
- Clarify the optimal parameter configuration of GridFTP
  - TCP socket buffer size
  - The number of parallel TCP connections

GridFTP

- Standardized in the GGF (Global Grid Forum)
- Extends the existing FTP (File Transfer Protocol)
- Supports the following features
  - Automatic negotiation of a TCP socket buffer size
  - Parallel data transfer
  - Third-party control of a data transfer
  - Partial file transfer
  - Security
  - Reliable data transfer

Automatic negotiation of a TCP socket buffer size

- Configure the TCP socket buffer size at the server
  - SBUF: client specifies the required quantity
  - ABUF: the required quantity is estimated between server and client
- Effective in networks with a high bandwidth-delay product
  - Most TCP implementations allocate a fixed size (e.g., 64 Kbytes) for each connection
  - Current implementation of GridFTP does not implement ABUF

Parallel data transfer (1)

- A single file can be transferred via multiple TCP connections
  - Alleviates the problem of the existing TCP
- May realize higher throughput since...
  - Congestion avoidance phase of TCP adopts an AIMD window control
    - Using more connections results in higher bandwidth
  - TCP socket buffer size is fixed for each connection
    - Using more connections can utilize more socket buffer
  - In slow-start phase of TCP, window size is doubled every RTT
    - Using more connections results in shorter ramp-up time

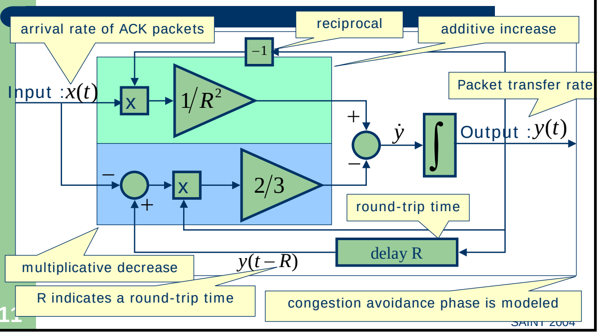
Parallel data transfer (2)

- However, if there are too many TCP connections
- May suffer lower throughput since...
  - A timeout occurs unless more than 3 duplicate ACKs arrive
    - Using more connections causes more frequent timeouts
  - TCP manages socket status for every connection
    - Using more connections requires more processing overhead
- It is necessary to determine the optimal number of connections according to network environment

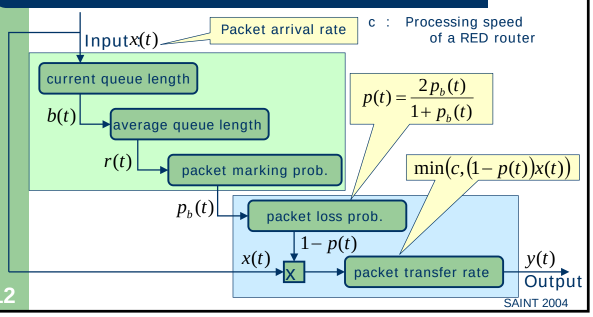
Modeling GridFTP using fluid-flow approximation

- Individually model building blocks of a network
  - SISO (Single-Input Single Output) continuous-time systems
    - Congestion control mechanism of TCP
    - RED router
    - Propagation delay of a transmission link
  - GridFTP is modeled by interconnecting these continuous-time systems

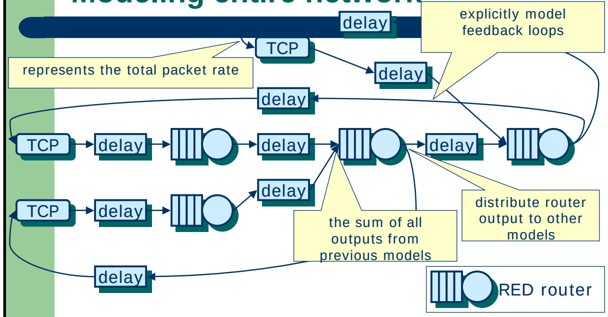
Modeling congestion control mechanism of TCP



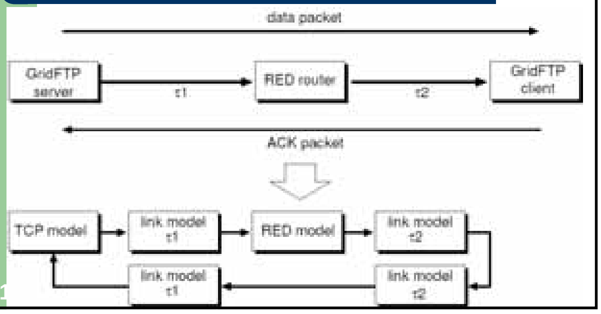
Modeling RED router



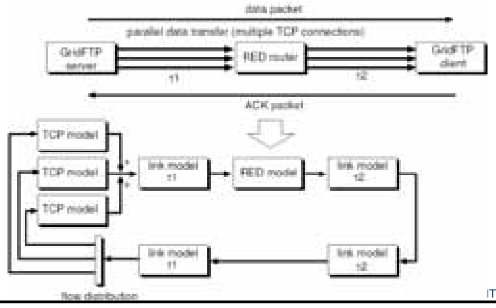
Modeling entire network



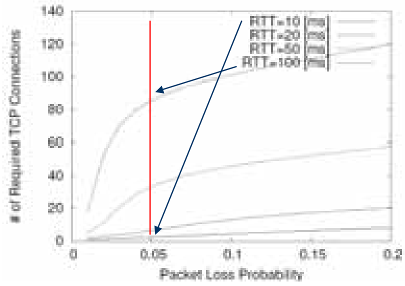
Modeling GridFTP (case of a single connection)



Modeling GridFTP (case of multiple connections)



Numerical example (# of TCP connections for full link utilization)



Network bandwidth: 1.0 [packet/ms]

Conclusion

- Modeling GridFTP using fluid-flow approximation
  - Model building blocks of a network as SISO continuous-time systems
    - Congestion control mechanism of TCP
    - RED router
    - Propagation delay of a transmission link
  - GridFTP is modeled by interconnecting these models
- Clarify the optimal parameter configuration of GridFTP
  - TCP socket buffer size
  - The number of parallel TCP connections

Future work

- Performance evaluation of GridFTP
  - Steady state analysis
    - e.g., throughput, packet loss probability, round-trip time
  - Transient state analysis
    - e.g., stability, ramp-up time, convergence time, and overshoot
  - Fluid-based network simulation
- GridFTP improvements
  - Dynamic control parameter tuning
    - e.g., TCP socket buffer size, # of parallel TCP connections
  - Dynamic GridFTP server selection
    - Choose a set of GridFTP servers using our analytic results