CMPT885 High-Performance Network

Final Project Presentation

Transport Protocols on IP Multicasting

Chao Li Thomas Su Cheng Lu

{clij, tmsu, clu}@cs.sfu.ca

School of Computing Science,

Simon Fraser University

Roadmap

- Introduction
- Implementation Details
  - Centralized Approach
  - Scalable Reliable Multicast
  - Tree Based Reliable Multicast
- Discussion
- References

Introduction of Our Project

- Why we need “multicast”?
- Why we need “multicast transport protocol”?
- What we will do in our project?

Introduction of Multicast

- Unicast Vs Multicast
  - Multiple copy Vs single copy
- Datalink Layer Support for Multicast
- Network Layer Support for Multicast
  - Multicast routing
  - Group membership management
  - Group network addressing

Multicast Transport Protocols

- Why not just run TCP over Multicast?
- Multicast Transport Protocols
  - Centralized Approach
  - Unstructured Approach
    - Scalable Reliable Multicast (SRM)
  - Distributed Approach
    - Tree-based Multicast Transport Protocols (TMTP)

Centralized Multicast

- Feature
  - Rely on a central site
  - Negative acknowledgment (NACK or NAK)
  - Slotting
- Advantages
  - Simplicity and ease of implemention
- Disadvantages

SRM- Scalable Reliable Multicast

- Requires only basic IP delivery model
- Every member is responsible for loss recovery
- Allow to a wide range of group size

SRM- Scalable Reliable Multicast

- Three types of messages
  - Session message
  - Request message
  - Repair message

  Messages are multicast to the entire group

SRM- Scalable Reliable Multicast

- Session messages
  - Report the highest sequence # received by every member
  - Provide information for sender/receivers (status, # of participants, etc.)
  - Estimate host-to-host distance (needed for repair)

SRM- Scalable Reliable Multicast

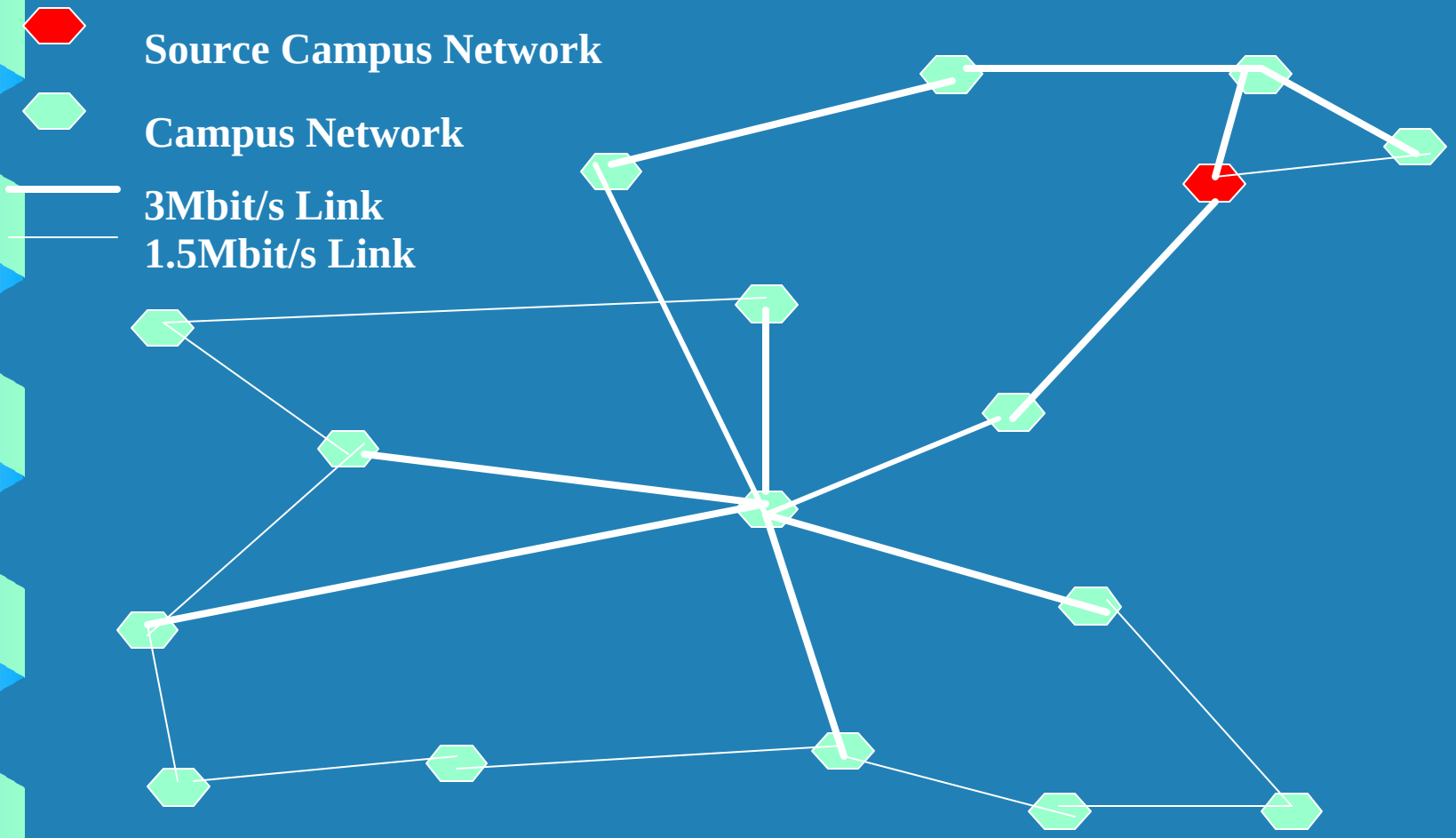
- Request messages
  - loss detected when there is a gap in the sequence #
  - A host wait a random time before multicast the request to the group
  - Any host that has a copy of the requested data can answer

SRM- Scalable Reliable Multicast

- Repair messages
  - Wait a random time before multicast the repair packet to the group

Simulation Environment

Topology(from[8][9])



Configuration Details

Background traffic	HTTP Each campus network has one web server, other nodes are clients # of Sessions 800 # of page components 3 TCP/FTP page component size 12KB Senders/receivers spared all over Session interval 10s burst size 80,000 byte

The credit goes to Velibor Markovski

More Configuration Details

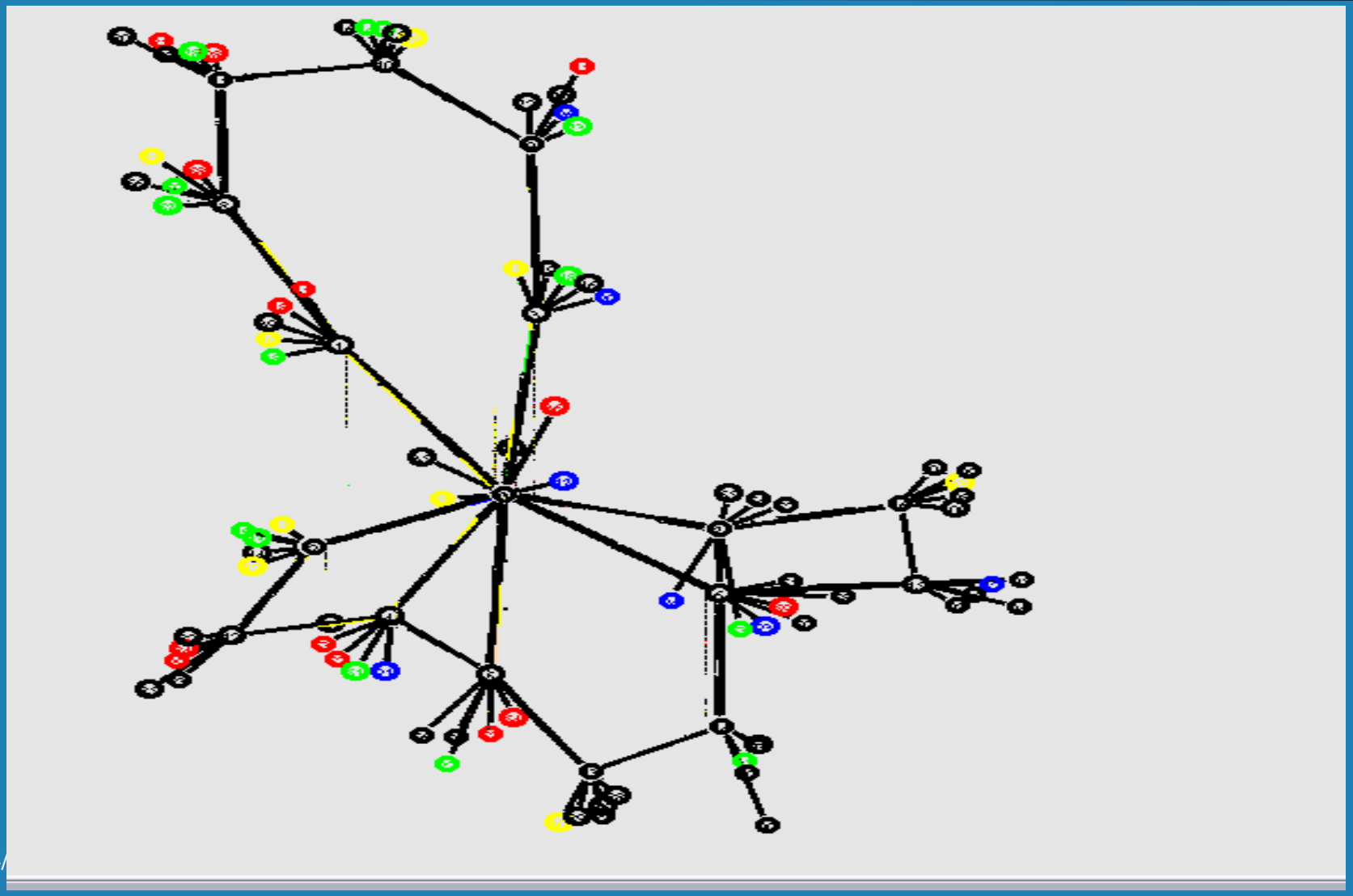
Multicast traffic	CBR traffic (0.04s) 1 sender and 9 receivers 512 packet size Start at 10s for 200sec. Session message frequency 2s(SRM default)

More experiment data will be collected on

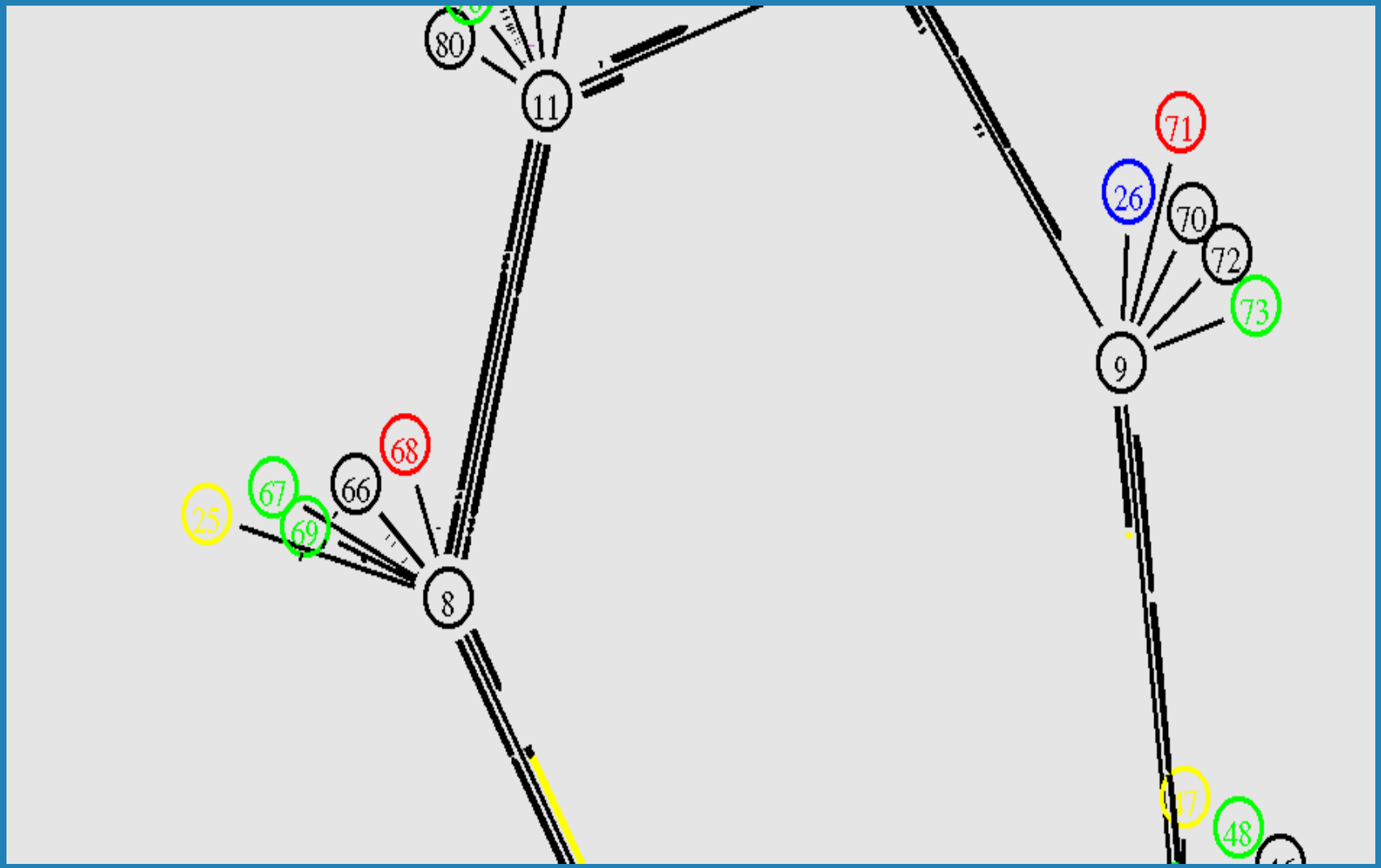
-Larger group size

-Star war Trace

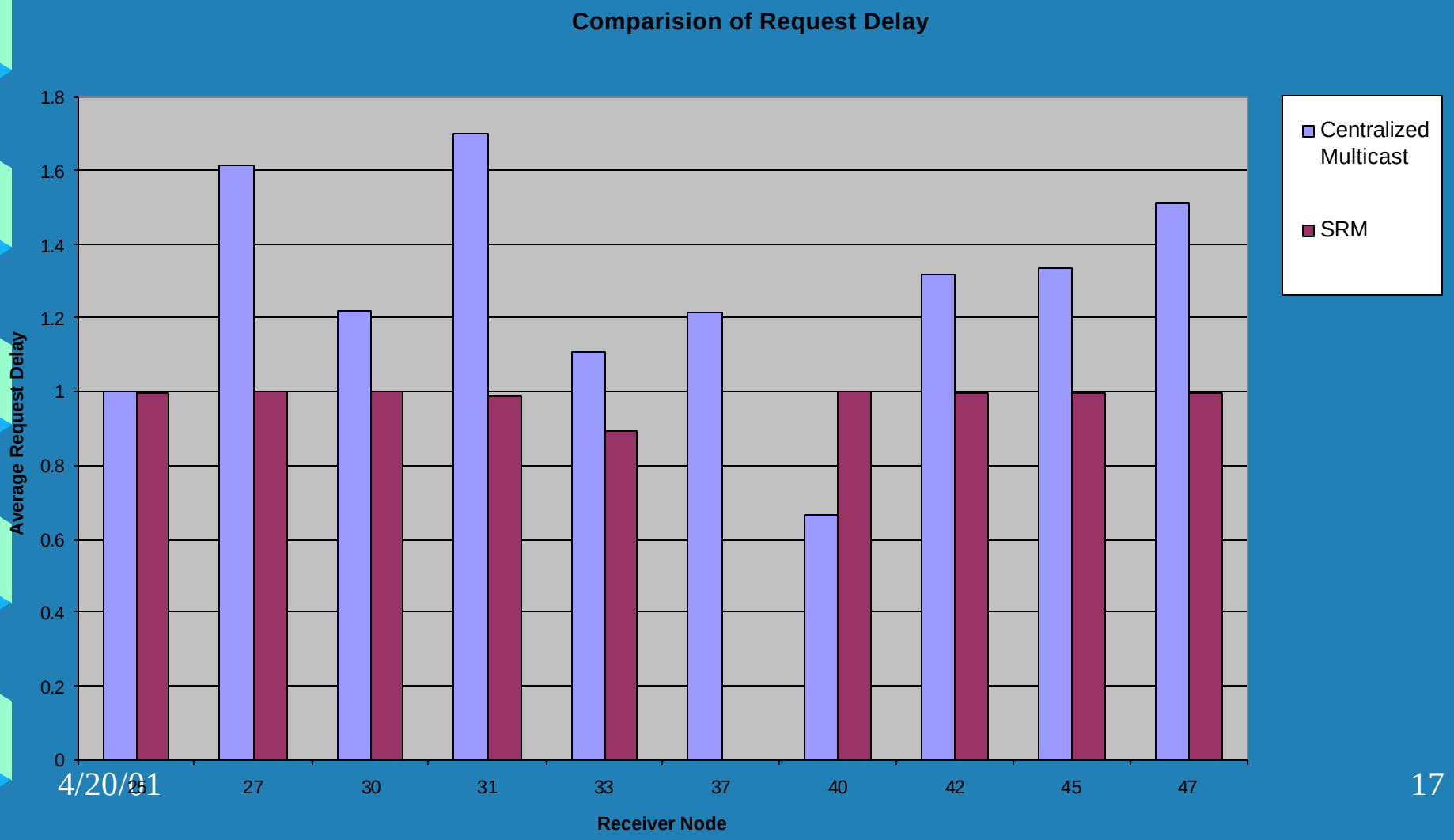
Snapshot(topology)



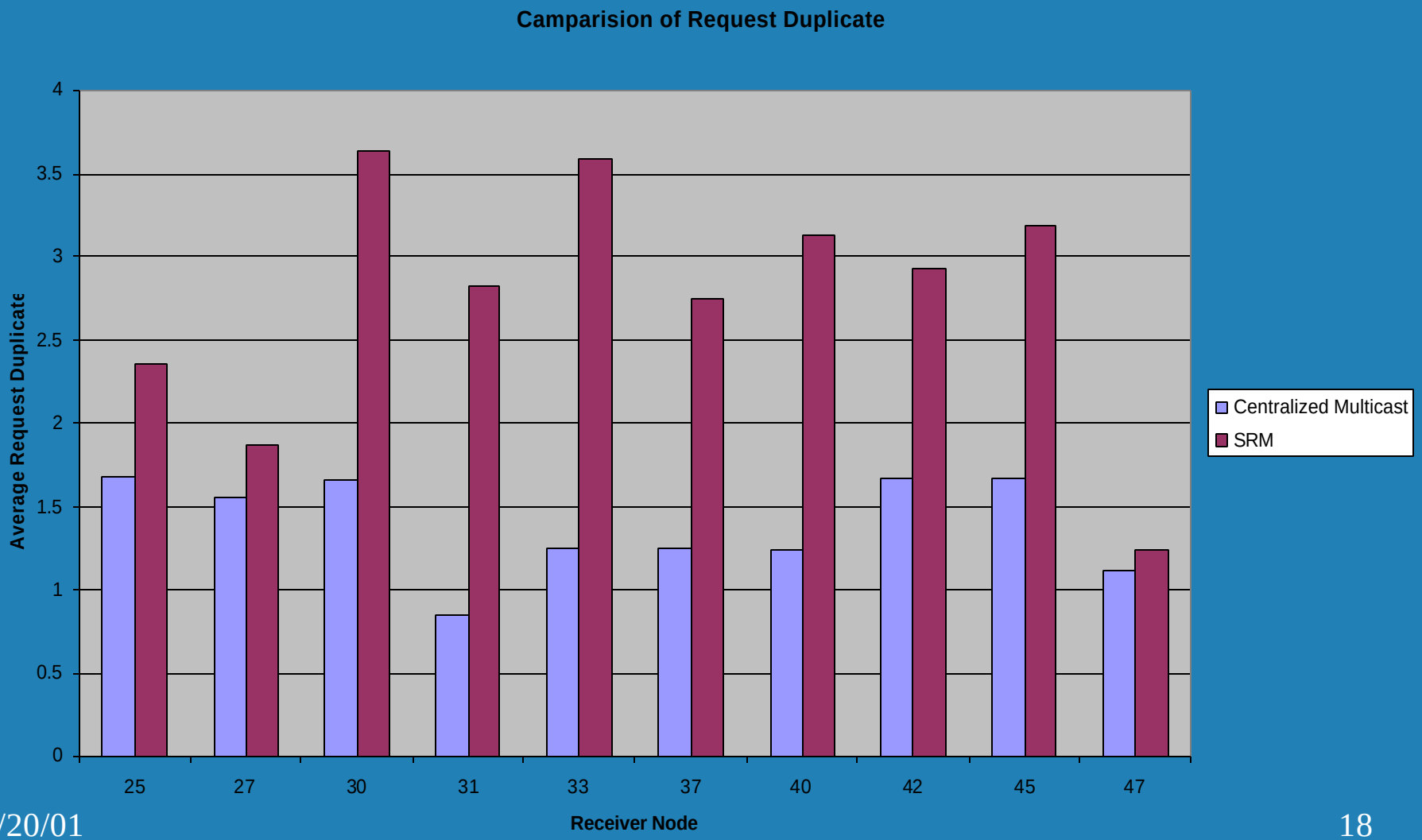
Snapshot(router)



Simulation Results(Request delay)



Simulation Results(Request duplicate)



Tree-based Multicast Transport Protocol (TMTP)

We are Implementing this protocol in ns2 using C++ and TCl

TMTP – key features

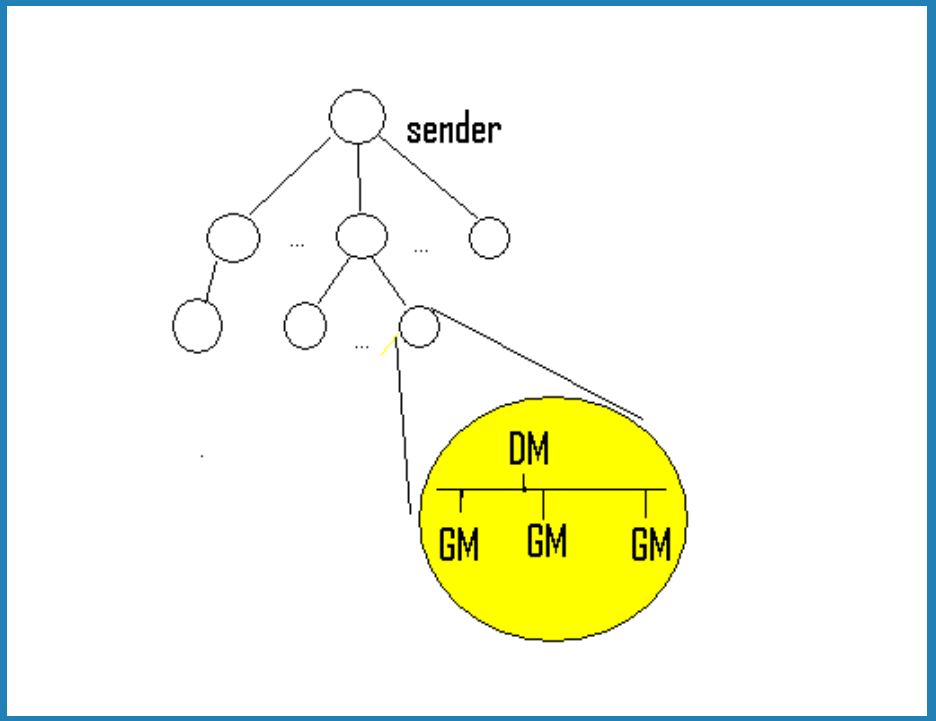
- It exploits the IP multicast for packet routing and delivery.
- It dynamically organize the participants into hierarchical control tree.
- It achieves scalable reliable multicast via hierarchical control tree used for flow and error control.

TMTP – hierarchical control tree

- It organizes the participants into a hierarchical domain.
- A domain manager(DM) acts as a representative of each domain.
- Every DM performs two roles for the responsibility for reliability –

  inter-domain and

  intra-domain .



TMTP –control tree management

- The control tree grows and shrinks dynamically in response to additions and deletions to and from the multicast group.

TMTP– Join Tree

- A new DM executes an *expanding ring* search to join the control tree.
- The JoinTree algorithm Utilizes a time-to-live value (TTL).

```
While (NotDone){
  Multicast a SEARCH-PARENT msg
  Collect respondses
  If (no responses)
     Increment TTL   // try again
  Else
  select closet respondent as parent
  send JOIN-REQ to parent
  wait for JOIN-CONFIRM reply
  if (JOIN-CONFIRM received)
     NotDone = False
     Else  //try again
}
```

TMTP – Leave Tree

- A DM only leave the tree after its last local member leaves the group.
- Internal managers is complicated.
- Leaf managers is straightforward.

```
If ( I am a leaf manager)
    send LEAVE-TREE to parent
    receive LEAVE-CONFIRM
    terminate
Else   // I am an internal manager
    Fullfill all pending obligations
    send FIND-NEW-PARENT to children
    receive FIND-NEW-PARENT from all children
    send LEAVE-TREE to parent
```

TMTP – error control(1)

An important concept:

- *limited scope multicast messages*

  – It restricts the scope of a multicast message.

  – It sets the TTL value in IP header: multicast radius.

TMTP – error control(2)

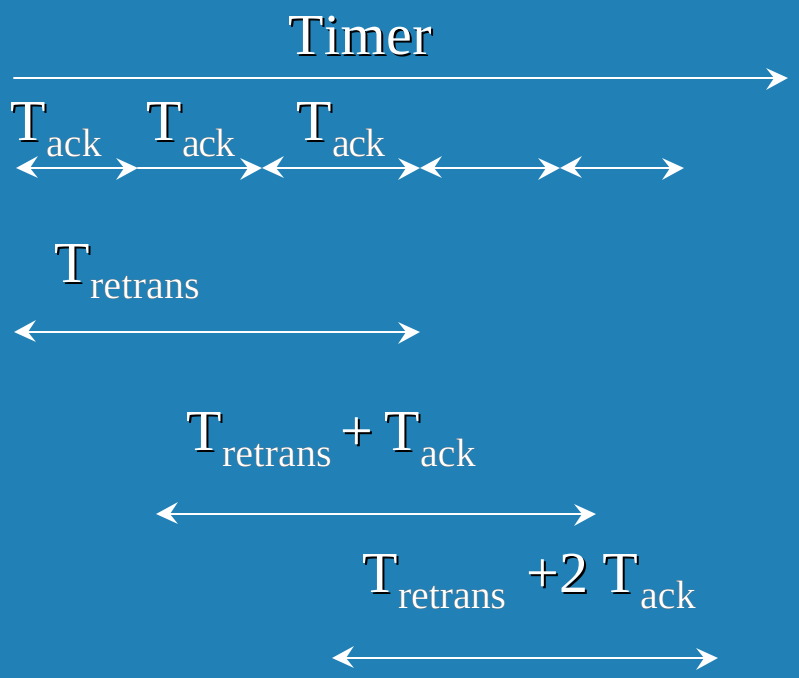
TMTP uses 2 error control techniques:

- sender initiated approach:
  - Periodic positive ACK for receiver,
  - Timeout,
  - Retransmissions (limited scope multicast).
- receiver initiated approach:
  - Negative ACK (NACK) to sender,
  - NACk is restricted and suppressed,
  - Retransmissions (limited scope multicast).

TMTP –Flow control

- Window-based flow control
- 2 timers:

  $T_{retrans}$ and T_{ack}
- $T_{retrans} = n * T_{ack}$, n=3



Test environment (2)

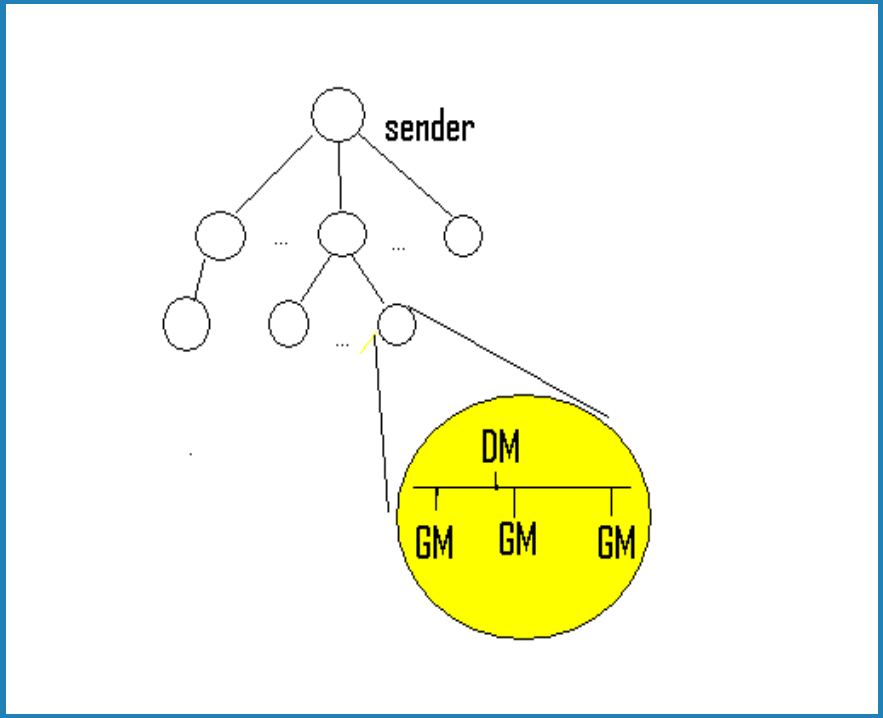
- We use standard IP multicast;
- We implement a sender-initiated reliable multicast transport agent using ns2;
- T_{ack} = RTT, $T_{retrans}$=3*T_{ack}
- multicast traffic

Performance measure

- End-to-end delay
- Packet loss
- Bandwidth consuming

Some comments

- It is not flexible
- periodically JoinTree to change hierarchy



Reference

[1] M. Banikazemi. “IP Multicasting: Concepts, Algorithms, and Protocols, IGMP, RPM, CBT, DVMRP, MOSPF, PIM, MBONE,”; http://charly.kjist.ac.kr/~dwlee/homepage/ipmulitcast.htm

[2] B. Williamson. “Developing IP Multicasting Networks Volume 1 - Distance Vector Multicast Routing Protocol and Multicast Open Shortest Path First,” pp. 106-127 pp.194-211, Cisco Press, 2000.

[3] E. C. Douglas, “Internetworking with TCP/IP Volume 1 - Internet Multicasting (IGMP),” pp. 289-302, Prentice-Hall, Inc., 1995.

[4] S. Deering, D. Estrin, D. Farinacci, V. Jacobson, C. G. Liu, and L. Wei, “An Architecture for Wide-Area Multicast Routing,” *ACM SIGCOMM '94*, vol. 24 no. 4, pp. 126-135.

[5] S. Floyd, V. Jacobson, C.-G. Liu, S. McCanne. and L. Zhang. “A Reliable Mulitcast Framework for Light-weight Sessions and Application Level Framing,” *ACM SIGCOMM’ 95*, Aug. 1995, pp. 342-356.

- [6] S. Armstrong, A. Freier, K. Marzullo, “RFC 1301: Multicast Transport Protocol,” Feb. 1992, http://www.cis.ohio-state.edu/cgi-bin/rfc/rfc1301.html
- [7] B. Whetten, G. Taskale. “An Overview of Reliable Multicast Transport Protocol II,” *IEEE Network*, Jan/Feb 2000, pp. 37-47; www.komunikasi.org/pdf/multicast/reliable-multicast-transport-protocol-ii.pdf
- [8] M. T. Lucas, B. J. Dempsey, and A. C. Weaver. “MESH: Distributed Error Recovery for Multimedia Streams in Wide-Area Multicast Networks,” In *Proceedings of IEEE International Conference on Communication (ICC '97)*, pp. 1127-1132, June 1997.
- [9] M. T. Lucas. “Efficient Data Distribution in Large-Scale Multicast Networks,” Ph.D. Dissertation, Department of Computer Science, University of Virginia, May 1998.
- [10] M. Goncalves and K. Niles, “IP Multicasting, Concepts and Applications,” New York: McGraw-Hill, 1998, pp. 91-116, pp. 273-287, pp. 305-324.

- [11] ns-2 network simulator: http://www.isi.edu/nsnam/ns
- [12] Star Wars trace in ns format: http://www.research.att.com/~breslau/vint/trace.html
- [13] C. Hanle and M. Hofmann, “Performance Comparison of Reliable Multicast Protocols using the Network Simulator ns-2,” *Proceedings of IEEE Conference on Local Computer Networks (LCN)*, Boston, MA, USA, October 11-14, 1998.
- [14] V. Markovski, “Simulation and Analysis of Loss in IP Networks – Simulation scenarios,” M. Sci. Thesis, Department of Engineering Science, Simon Fraser University, Oct. 2000, pp. 24-30.
- [15] R. Yavatkar, J. Griffioen, and M. Suda. “A Reliable Dissemination Protocol for Interactive Collaborative Application,” In *Proceedings of the ACM Multimedia ’95 Conference*, Nov. 1995.

Individual contribution

- Chao Li: topology construction
  centralized approach
- Thomas Su: topology construction
  SRM
  background traffic
- Cheng Lu: building a TMTP agent into ns2

Thanks