



SplitStream: High-Bandwidth Multicast in Cooperative Environments

Muhammad Abdulla

10/14/2004

1



Outline

- Basic Concepts
 - Multicasting
 - The SplitStream approach
- Structured Overlay Network
 - Pastry
 - Scribe
- SplitStream Design
- Experimental Results

10/14/2004

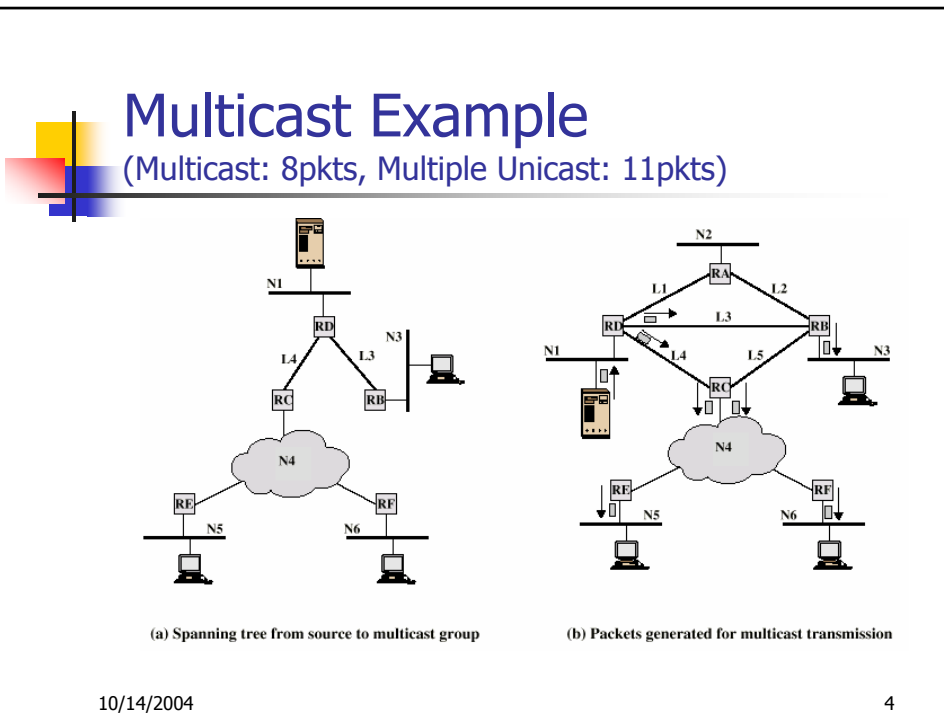
2

Multicasting

- Determine least cost path to each network that has host in group
 - Gives spanning tree configuration containing networks with group members
- Transmit single packet along spanning tree
- Routers replicate packets at branch points of spanning tree
- In Multiple Unicast: Send packet only to networks that have hosts in group

10/14/2004

3





The SplitStream Approach

- Why SplitStream? The Problem:
 - Conventional tree-based multicast is not well-suited for cooperative P2P systems.
 - ❖ The burden is carried by interior nodes: not expected in P2P systems.
 - ❖ Most peers do not have enough capacity.
- The SplitStream Solution
 - Split the content!

10/14/2004

5



SplitStream: Basic idea

- Split the content into k stripes and multicast each stripe using a separate tree.
- Load will be spread evenly across peers.
- Striping also increases resilience and robustness in P2P systems.

10/14/2004

6



Basic Challenge

- For the Protocol (SplitStream):
 - To construct a forest of multicast trees.
 - The construction overhead should be low.
 - The load should be evenly distributed among nodes.
 - Bandwidth constraints for the whole systems as well as for individual nodes should be addressed.

10/14/2004

7



Basic Challenges (Cont'd)

- For Applications:
 - Constitute stripes such that:
 - ❖ Each stripe requires approximately the same bandwidth
 - ❖ The content should be reconstructed from any subset of the stripes of sufficient size.
- How to Solve?
 - Use tree-based application-level multicasting:
 - ❖ Pastry – structured P2P overlay network protocol
 - ❖ Scribe – application-level group communication system

10/14/2004

8



Outline

- Basic Concepts
 - Multicasting
 - The SplitStream approach
- ✓ Structured Overlay Network
 - ✓ Pastry
 - ✓ Scribe
- SplitStream Design
- Experimental Results

10/14/2004

9



Pastry

- A scalable, structured P2P overlay network.
- Basic characteristics:
 - Routes each message to the node with closest nodeId to the key.
 - Normally traverses nodes with increasing shared prefix length.
 - Low delay.
 - Local convergence.

10/14/2004

10



Scribe

- An application-level group communication system built upon Pastry.
- A pseudo-random Pastry key, known as groupID is given to each multicast group.
- Group membership management is decentralized. To add a member to a group:
 - Route towards the groupID until message reaches a node in the tree.
 - Add the traversed route to the multicast tree.

10/14/2004

11



Outline

- Basic Concepts
 - Multicasting
 - The SplitStream approach
- Structured Overlay Network
 - Pastry
 - Scribe
- ✓ SplitStream Design
- Experimental Results

10/14/2004

12

SplitStream Properties

- General condition for forest construction:
 - *It should be possible to connect nodes such that each node receives no more stripes than its desired in degree and its out degree is no more than its capacity.*

- Necessary Condition

$$\sum_{\forall i \in N} I_i \leq \sum_{\forall i \in N} C_i$$

- Sufficient Condition

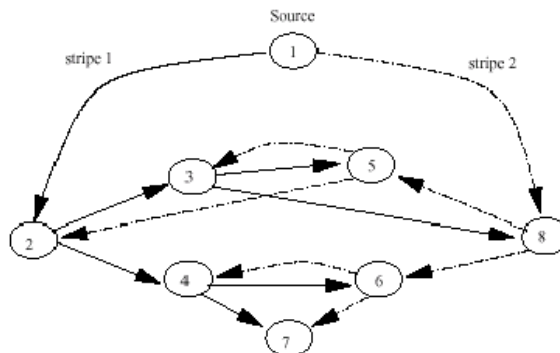
$$\forall i: C_i > I_i \Rightarrow I_i + T_i = k$$

10/14/2004

13

SplitStream Properties (Cont'd)

- A simple example:



10/14/2004

14



SplitStream Properties (Cont'd)

- Conditions may be strengthened to address free-riders:
 - Require that a node's forwarding capacity be larger than or equal to its indegree.
 - To address free-riding, SplitStream may be changed to force stronger constraints at node.

10/14/2004

15



Building interior-node-disjoint trees

- Each node in a set of trees is interior node in at most one tree and leaf node in the other trees.
- Exploiting Pastry's properties, we can create groupIDs that differ in the most significant bits to get k to have a disjoint set of interior nodes.
- The parameter k is normally chosen to be 2^b (b is the bits for each digit of an ID in Pastry) so that nodes will have equal chance of becoming an interior node in some tree.

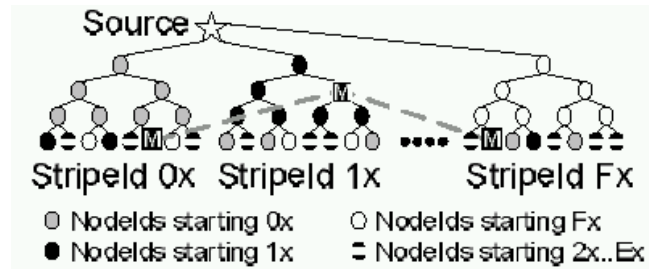
10/14/2004

16



Building interior-node-disjoint trees (Cont'd)

- SplitStream's forest construction:



10/14/2004

17



Limiting Node Degree

- Forest of interior-node-disjoint trees satisfies the nodes' inbound bandwidth constraints.
- But it does not necessarily satisfy outbound bandwidth constraints.
- Scribe's built-in mechanism to limit nodes' outdegree is not guaranteed to work in SplitStream.

10/14/2004

18



Locating Parents

- Main objective: handle the case where a node that has reached its maximum outdegree receives a join request from a prospective child:
 - The node looks for children in stripes that do not share a prefix.
 - Selects a child to drop.
 - The orphaned child will look for a new parent:
 - ❖ Attempts to find closest to stripeID.
 - ❖ The sibling either adopts or rejects.
 - ❖ If not successful, use spare capacity group.

10/14/2004

19



Spare Capacity Group

- A special group of nodes that have spare capacity (fewer children than its forwarding capacity).
- Scribe delivers the message to the nearest (physically) node in spare capacity group through anycast (In anycast, there are multiple possible recipients, but message is delivered to one of them.).

10/14/2004

20



Spare Capacity Group (Cont'd)

- The node in the spare capacity group that received the message checks two conditions:
 - It receives one of the stripes the orphan seeks to receive.
 - There will not be a cycle if it accepts the orphan.
 - If these two tests succeed, the node takes the orphan as a child.

10/14/2004

21



Outline

- Basic Concepts
 - Multicasting
 - The SplitStream approach
- Structured Overlay Network
 - Pastry
 - Scribe
- SplitStream Design
- ✓ Experimental Results

10/14/2004

22



Experimental Results

- Network simulation environments:
 - GATech
 - Mercator
 - CorpNet
- Performance Metrics
 - Forest construction overhead
 - Forest multicast performance
 - Resilience to node failures

10/14/2004

23



Experimental setup

- SplitStream Configuration
 - Number of bits for each digit $b = 4$.
 - Leaf set size $l = 16$.
 - Number of streams per SplitStream multicast channel is $k = 2^4 = 16$.
 - Six different configurations for node degrees, in the form of $X \times Y$ (where x is the desired indegree, y is the capacity):
 - ❖ $16 \times 16, 16 \times 18, 16 \times 32, 16 \times NB$
 - ❖ $d \times d, \text{Gnutella}$

10/14/2004

24



Forest construction overhead

- Include Figure 8 from paper.
- Include Table 1 from paper.

10/14/2004

25



Forest multicast performance

- Include Figure 12 from paper.
- Include Table 3 from paper.
- Include Table 4 from paper.

10/14/2004

26



Resilience to node failures

- Include Table 6 from paper.
- Include Figures 15 and 16 from paper.

10/14/2004

27



Conclusions

10/14/2004

28