Mesh-Based Content Routing Using XML

Alex C. Snoeren, Kenneth Conley, and David K. Gifford
MIT Laboratory for Computer Science

Presented by: Jie Mao
CS295-1 Fall 2005

Outline
- Introduction
- Resilient mesh networks
- Algorithms and protocols
- Evaluation
- Conclusion

Motivation
- Mesh-based overlay networks
- Reliable multicasting
- Content-based routing
- Time-critical data transmission
- Heterogeneous clients
Motivation

- Question:
  - How to achieve low latency while reliably distribute data streams? – packet loss-recovery mechanism

- Assumption
  - The increased transport costs can be justified by the value of reliable and timely data delivery

- Approach
  - Loss-tolerant encoding – packet diversity
  - Channel diversity
  - Sender diversity

Basic Idea

- Construct an acyclic (n-1) resilient content distribution mesh - n is configurable
- Distributed publish-subscribe network comprised of XML routers
- An overlay network that transports XML streams
- Clients join network by specifying an XML query that describes the XML packets they want to receive

Why XML?

- Advantages
  - Interpretation of client data needs in terms of well-defined XML queries
  - Aid network scheduling
  - Tools and standards make robust applications
  - In-network processing

- Is increased data volume a disadvantage?
  - No. Data compression eliminates it

Contributions

- XML Routing
  - Support arbitrary content routing

- Mesh-based Overlay Network
  - Multiple redundant paths
  - Better latency performance than tree-based approach

- Diversity Control Protocol
  - Source-independent sequence number
  - Reduces latency and improves reliability
Outline

- Introduction
- Resilient mesh networks
  - Mesh network components
  - Router content configuration
  - Clients join
- Algorithms and protocols
- Evaluation
- Conclusion

Mesh network components

- Root routers
- Internal routers
- Clients
- XML Combining routers
  - Merge XML feeds from different sources into a single feed

Router Content Configuration

- Static configuration
  - Internal routers carry all of the XML packets
  - Clients have wider choice of routers to be their parents
  - Requires fixed bandwidth capacity throughout the mesh

Router Content Configuration

- Dynamic configuration
  - Routers only carry packet stream to service their children
  - Routers forward combined queries to its parents
  - Bandwidth saving for partial interests
  - Latency for mesh construction and repair
Client Joining
- Composing an XML query
- Contacting $n$ existing routers
  - Router discovery? – explained later
- Sending XML query
- Receiving XML stream

Outline
- Introduction
- Resilient mesh networks
  - Algorithms and protocols
    - XML router core
    - Diversity control protocol (DCP)
    - Mesh formation and maintenance
- Evaluation
- Conclusion

XML Router Core
- Input component
  - Maintaining DCP connections
  - Mesh initialization and reconfiguration algorithms
  - Data decompression
  - TCP compatibility

XML Router Core
- XML switch
  - Matching packets to queries
  - Forwarding packets to questing links
  - Efficient state machine
**XML Router Core**

- Output component
  - Forwarding packets using DCP
  - Handling join requests
  - Link-based data compression
  - TCP compatibility

**Diversity Control Protocol (DCP)**

- Same stream of packets is sent to receiver by multiple sources
- Receiver can reassemble a packet stream from diverse senders using the first error-free packet received from any source.

**DCP Sequencing**

- Requirements for in-order packet stream reassembly
  - A total ordering of packet identifiers
  - Packet identifiers only associate with packet content, not be sender specific
  - Packet identifiers must be selected at root routers and remain identifiable throughout the mesh
  - Receiver identifier processing must admit gaps

**DCP Sequencing**

- Sequence Number
  - Monotonically increasing 32-bit application serial number (AN)
  - DCP packet header

```
<p>| | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>AN</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
```

*AN: Address Number*
DCP is reliable

- Retransmission
  - Packets are buffered and transmitted in-order
  - Receiver discards or buffers packets by current AN
  - Send negative acknowledgment (NACK) with current AN
  - Positive acknowledgment with current AN
  - Requested at each hop
  - NACK implosion problem? – Not likely

Mesh formation and maintenance

- Adding routers and clients
  - Mesh primitives
    - Join(Q), Children(Q), Parents
  - Automatic parent selection for (n-1) resilience
    - S = set of root routers
    - For each node in S, send a join request and remove the node from S
    - If a node accept the join, add it to the parent set P. If n nodes are in P, quit
    - If a node declines the join, ask it for a list of its children, and add them into S
    - If S is not empty, go to Step 2

Mesh formation and maintenance

- Mesh repair
  - A node actively joins a new parent when one of its parents fails
  - Acyclic mesh – router’s level number
  - Cold re-initialization

Outline

- Introduction
- Resilient mesh networks
- Algorithms and protocols
- Evaluation
  - DCP performance
  - XML routing performance
  - Experience with air traffic control data
- Conclusion
DCP performance

- Two XML router implementation
  - Full-featured, multi-threaded Java implementation
  - High-performance prototype based on Click
- Experimental design
  - Four machines, 2 Linux, 2 FreeBSD
  - Each node requests the entire set of XML stream
  - Emulation of link loss

---

DCP performance

- Redundancy reduces loss exponentially

---

DCP performance

- Latency

---

XML routing performance

- Dual processors
- 100Mb Ethernet
- 1000Mb Ethernet
- 262-byte packets
Experience with air traffic control data

- Aircraft Situational Display to Industry (ASDI)
- ASDI format and XML format
  - ASDI Format:
    - 153014022245CCZVTZ UAL1021 512 290 4928N/12003W

Experience with air traffic control data

- ASDI format and XML format
  - XML Format:
    - \(<xml version="1.0"/>\>
    - \(<messageid>153014022245CCZVTZ</messageid>\>
    - \(<flight>\>
      - \(<id>UAL1021</id>\>
      - \(<flightleg status="active">\>
        - \(<speed type="ground">512</speed>\>
        - \(<altitude type="reported" mode="plain">290</altitude>\>
      - \(<coordinate>\>
        - \(<lat>4928N</lat>\>
        - \(<lon>12003W</lon>\>
      - \(<flightleg/>\>
    - \(<flight/>\>

Outline

- Introduction
- Resilient mesh networks
- Algorithms and protocols
- Evaluation
- Conclusion
Conclusion

- Three key ideas
  - XML router
  - Resilient overlay network
  - Diversity communication protocol

Future work

- Protocol refinement
- Additional functionality
- DCP self-tuning
- XML router extensions

Questions?