Adaptive Video Multicasting

Presenters:
Roman Glistvain, Bahman Eksiri, and Lan Nguyen

Outline
- Approaches to Adaptive Video Multicasting
  - Single rate multicast
  - Simulcast
  - Active Agents
  - Layered multicasting
- Detailed description of layered multicasting
  - RLM
    • uses current uniform packet drop routers
  - RLMP
    • extension of RLM
    • requires priority based packet dropping.
- Comparison of various approaches
- Concluding remarks

Aproaches for video multicasting

Classification of various video multicasting protocols:
- Single rate transmission vs Multi-rate transmission
- Place where adaptation is performed:
  • source/destination (end nodes)
  • intermediate nodes

Single rate multicast
- Single rate adaptation (end nodes)
- Active agents (intermediate nodes)

Multirate
- Simulcast (end nodes)
- Layered adaptation (end nodes)
- Active agents (intermediate nodes)

Single rate adaptation

Feedback based system
- The source adjusts the sending rate based on the feedback it receives from the end nodes.

Problems with this approach
- Unfair distribution of bandwidth
  • Small capacity receivers suffer from congestion
  • High capacity receivers are underutilized
- Feedback implosion problem
  • Receivers may end up sending the feedback at the same time.

Single Rate Adaptation (cont)

Practical systems
- INRIA IVS video conferencing system
- Uses polling rounds to poll receivers with congestion level bigger than the one it knows about.
- Avoids feedback implosion problem

Simulcast

- Single rate multicast and multiple unicast are two extremes
- The intention of simulcast is to bridge the gap between the 2 groups

Algorithm:
- Different bit-rate streams are sent over different multicast groups
- Allows for feedback based source rate adjustment in each group
Simulcast (cont)

Evaluation of Simulcast
- Advantages
  - Better granularity of receiver fairness in heterogeneous environment
- Disadvantages
  - Same stream is broadcast multiple times (with different bit rates)
  - Worse bandwidth utilization than single rate multicast

Practical algorithm
- DSG (destination set grouping)
  - 3 streams (low, medium, high quality)
  - Feedback based source rate control for each stream
  - Receiver can subscribe to different streams under different network conditions

Real Time protocol

- RTP (Real Time Protocol)
  - Universal framework for transferring real time data
    - multicast/unicast
    - Allows for active network elements
    - Timing and synchronization
- RTCP (Real Time Transport Control Protocol)
  - Provides feedback about
    - Jitter
    - Packet loss
    - Other parameters which can be used for rate control.

Services of RTP protocol

- Application level datagram protocol.
- Allows for active/passive gateways with various functionality
- Active Gateway – acts as a source
  - RTP Mixer (picture)
    - Combines multiple RTP streams into single stream
    - Changes the synchronization information
- Passive gateway – transparent to end nodes
  - RTP Translator
    - Translates formats/semantics

Active Agents
- Embedded into intermediate network elements
  - Adjust the dataflow to match the conditions of local network segment using feedback.
- Services offered by typical active agents:
  - Video transcoding (rate adjustment)
  - RTP Mixer
  - RTP translator

Usage of active agents
- Large networks are partitioned into small segments.
- Active agents are responsible for a segment of the network.

Evaluation of active agents

Advantages of Active Agents
- Very fine tuned adjustments for various network conditions
- A lot better adaptation than end-to-end based methods

Disadvantages of Active Agents
- Implementation complexity
- Inability to deal with high video traffic rates.
- Deployment problems
  - Usually the number of media gateways is very limited (due to high cost)
  - The operators have to decide where to place them in the network (which can be quite dynamic).

Layered adaptation

- Video is encoded into different streams.
  - Basestream is required to play back the video
  - Each additional enhancement layer refines the quality of the video
- Straightforward approach:
  - Packet prioritization – higher layer have lower packet priority
- Receiver driven approach (RLM)
  - Transmit various layers of video over different multicast groups
  - Let receivers subscribe to the layers they see fit.
TCP Friendliness
- Single rate multicast, Simulcast
  - Unfair
- Layered adaptation
  - RLM
    - unfair to TCP traffic in short run (because of join experiments) but fair in long run
  - RLMP
    - routers can decide what to do with low level priority traffic and assign it lower priority then TCP
- Active Agents
  - Function as routers/bridges – can adjust the multicast utilization based on user priorities.

Internet Heterogeneity and scale
- Problem
  - React to congestion
  - Network capacity for multicasting
  - Heterogeneous receivers
  - Locally degrading quality
  - Adding and dropping Layers (joining and leaving multicast groups)

RLM Protocol
- Basic idea.
  - Based on Layered Adaptation.
  - Different layers are multicast in different multicast groups.
  - Completely receiver driven.

Adapting method
- On congestion, drop a layer.
- On spare capacity, add a layer.

Capacity inference
- How to find out if there is a spare capacity?
  - Join experiments
- How to avoid transient congestion caused by Join experiments?
  - Adding layers at well chosen times
  - Learning mechanism

RLM Adaptation
- Join experiments cause congestion
- Using join-timer
- Adaptive detection-time
Scaling RLM

- System scales poorly if each receiver independently carry out adaptation algorithm.
- Avoid scaling problem by scaling down join-experiment rates proportional to the group size.
- Shared learning mechanism to avoid too long converge time

Scaling RLM - continued

- Broadcasting join-n layer experiment to the group
- Group to watch the affect of new join-experiment
- In case of congestion, all the group members to scale back the nth join-timer.
- The learning process is conservative
- Overlapping problem and how/when to avoid it
- Shared learning process determines what does not work rather than what does work.

Complete Algorithm

1) On stable congestion:
   - Drop a layer, increase a join timer for the problematic layer
2) When the join timer expires:
   - Multicast a message that join experiment has started on a certain layer
   - Add that layer
   - Wait for the detection timer to expire
   - If there is a congestion during the experiment – goto 1, otherwise stay with the current layer of subscription.

Evaluation of the RLM algorithm

- Packet loss over a certain period of time.
- Convergence rate – to reach the optimal throughput
- The two metrics will have to be used together to evaluate the performance of the algorithm
  - Acceptably low loss rates and fast convergence times imply a well functioning system.
- Simulation terms
  - Latency Scalability
    - How different network latencies affects the loss rate of the algorithm.
  - Session size
    - How the number of different receivers affect the loss rate, convergence time of the algorithm
  - Bandwidth Heterogeneity
    - How the algorithm performs in presence of large quantities of receivers with different bandwidth constraints.
  - Superposition
    - Performance of the algorithm when multiple sessions share the same link.

Simulation results

- Latency scalability
  - The higher the network latency – the higher the packet loss due to the fact that it takes a very long time to detect the result of a join experiment
- Session Size
  - Doesn’t have a huge impact on the loss rates
  - It does increase convergence times due to the fact that more receivers might initiate join experiments in parallel
- Bandwidth heterogeneity
  - The protocol works well with a big amount of receivers with different bandwidth constraints.
- Superposition
  - Aggregate link utilization close to 1, bandwidth allocation can be unfair.

Packet dropping policies

- RLM: uniform dropping (e.g. FIFO)
  - Packets are dropped at all the layers
  - Performance adjustment is done by receivers
  - Performs poorly under congestion (why?)
  - Unfair allocation of bandwidth among receivers.
- Priority-based dropping
  - Each layer associates with a priority, set by the sender
  - The dropping is done by routers based on priorities
  - Maintains reasonable quality under congestion and bursty traffics (how?)
  - Stable and “fair” allocation of bandwidth
Priority vs uniform dropping

Drawbacks of priority dropping

- **Additional complexity**
  - Routers need to examine the priority header
  - Requires additional packet scheduling policy
- **Fairness**
  - Achieved only if all sessions use the same layering (priority) scheme
- **Unnecessary traffic**
  - Caused by receivers who subscribe to all the layers and "forget" to unsubscribe when not needed.

Combine RLM w/ Priority (RLMP)

- **Goals**
  - Maintains the simplicity of RLM
  - Achieves stability and fairness
- **Consists of 2 parts:**
  - Network portion: priority dropping at routers
  - Receiver portion: based on RLM

Network portion of RLMP

- The sender encodes video into multiple layers BUT does not assign the priorities
- It’s the receiver who decides the priority
- When a receiver joins a layer, it tells the routers its desired priority level.
- At routers, each outgoing link is attached with a priority, which is the maximum of all the priorities requested by all the downstream receivers

Receiver portion of RLMP

- Receivers subscribe to multiple $N$ layers (RLM)
- $N$ changes according to network conditions
- $N-1$ first layers are high priority. The $N$th layer is low priority
- Purpose of the $N$th low priority layer
  - Reserves unused bandwidth
  - Guesses the network conditions by the packet loss rate at the $N$th layer
Finding optimal $N$

- Adaptively add and drop layers based on packet loss rates
- Adding a layer
  - Join layer $(N+1)$ at low priority
  - Re-join layer $N$ at high priority
- Dropping a layer
  - Leaving layer $N$
  - Re-join layer $(N-1)$ at low priority

Features of RLMP

- **Fairness**
  - Can be achieved without the requirement that all sessions use the same priority scheme
  - Receivers who share the same upstream link, also share the same priority
- **Stability**
  - Characteristic of priority-based
  - By the use of a long term loss estimator

Example 1

- Pure priority-based with no congestion

Example 2

- Pure priority-based with congestion

Example 3

- RLMP with no congestion

Example 4

- RLMP with congestion
Summary RLM vs RLMP

- **RLM**
  - Simplicity
  - Performs well under CBR traffic sources
  - Unfairness
  - Performs poorly under bursts and congestion
- **RLMP**
  - More complexity
  - Fairness
  - Stability: maintain reasonable under network bursts and congestion.

Drawback of RLMP

- Do not support more than two levels of priorities (why?)

Conclusion

- No perfect algorithms for video multicast distribution
- Various approaches with various degree of complexity
- Requires end-to-end or network based QoS support to achieve a better degree of quality, stability and fairness.
- Complexity cost