RTP: Real-time Transport Protocol

Krzysztof Hebel
Multimedia Communications Laboratory
Electrical & Computer Engineering Department
University of Waterloo
February 21st 2006
Presentation Outline

- Introduction
- Overview of RTP
- RTP Packets
- RTCP (RTP Control Protocol)
- RTP Payload Format (RTP Packetization) + Network Abstraction Layer (aside)
- Conclusions
Presentation Outline

- Introduction
- Overview of RTP
- RTP Packets
- Real Time Control Protocol (RTCP)
- RTP Payload Format (RTP Packetization) + Network Abstraction Layer (aside)
- Conclusions
Introduction – OSI Model and protocol environment

- Layer 1&2 – not important
- Layer 3 – Internet Protocol IP
- Layer 4 – UDP or TCP
- Application Layer Transport - RTP
- provides transparent transfer of data between end users
TCP vs. UDP

TCP features:
- applications on networked hosts create a connection one to another
- guarantees reliable and in-order delivery of sender to receiver data
- sequence numbers for ordering received TCP segments and detecting duplicate data
- checksums for segment error detection
- acknowledgements and timers for detecting and adjusting to loss or delay
- retransmission and timeout mechanisms for error control
- unpredictable delay characteristics
- Hence: not suitable for real-time communication
TCP vs. UDP – cont’d

UDP features:
- simple, unreliable datagram transport service
- does not provide reliability and ordering guarantees
- datagrams may arrive out of order or go missing without notice
- checksum for detecting packages containing bit errors
- faster and more efficient for many lightweight or time-sensitive purposes
- obvious choice for real-time video transmission
- Reference: RFC 768, 28 August 1980
Presentation Outline

- Introduction
- Overview of RTP
- RTP Packets
- RTP Control Protocol (RTCP)
- RTP Payload Format (RTP Packetization) + Network Abstraction Layer (aside)
- Conclusions
Overview of RTP

- RTP is the Internet-standard protocol for the transport of real-time data, including audio and video. It can be used for media-on-demand as well as interactive services such as Internet telephony. RTP consists of a data and a control part. The latter is called RTCP (RTP Control Protocol).

- Reference: RFC 3550, July 2003
Overview of RTP – cont’d

- The data part of RTP is a thin protocol providing support for applications with real-time properties such as continuous media (e.g., audio and video), including timing reconstruction, loss detection, security and content identification.

- RTCP provides support for real-time conferencing of groups of any size within an internet. It offers quality-of-service feedback from receivers to the multicast group as well as support for the synchronization of different media streams.
General Scenario

- One-to-one
- One-to-many
- Many-to-many
- Local transmission (access within one machine)
- RTP packets
- RTCP (Sender and Receiver Reports)
Presentation Outline

- Introduction
- Overview of RTP
- RTP Packets
- RTP Control Protocol (RTCP)
- RTP Payload Format (RTP Packetization) + Network Abstraction Layer (aside)
- Conclusions
RTP packets

- Consist of and RTP header, optional payload headers and the payload itself
- RTP overhead = 12 bytes
- IP+UDP+RTP overhead = 20+8+12 = 40 bytes
- It is advisable to keep coded slice sizes as close to, but never bigger than, the MTU size (largest size of a packet that can be transmitted without being split/recombined on the transport and network layer), because:
  1. It optimizes the payload/header overhead relationship
  2. Minimizes the loss probability of a (fragmented) coded slice due to the loss of a single fragment on the network/transport layer and the resulting discarding of all other fragments belonging to the coded slice
- MTU sizes: ~1500 bytes for wireline IP links (max. size of an Ethernet packet), ~100 bytes in wireless environments
RTP packets - example

DVD quality video transmission:
30 frames/s, 720x480 resolution, 3 bytes per pixel
- 31,104,000 bytes/s raw rate
- 311,040 bytes/s compressed data rate (100x compression)
- MTU = 1500 bytes: $311,040/1460 = 213$ packets/s -> 319,500 bytes/s (required throughput including overhead)
- MTU = 100 bytes: $311,040/60 = 5184$ packets/s -> 518,400 bytes/s (required throughput)
RTP packets – cont’d

- RTP header contains the following: sequence number (used for packet-loss detection), timestamp (timing information, synchronization of media streams), payload type (identifies the media codec of the payload), marker bit (detecting the end of a group of related packets), source identifiers (contributing and synchronizing)
The RTP header has the following format:

```
0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|V=2|P|X| CC |M| PT | sequence number |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
                   timestamp
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
                   synchronization source (SSRC) identifier
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
                   contributing source (CSRC) identifiers
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
                   ....
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```
Presentation Outline

- Introduction
- Overview of RTP
- RTP Packets
- RTP Control Protocol (RTCP)
- RTP Payload Format (RTP Packetization) + Network Abstraction Layer (aside)
- Conclusions
Real-Time Control Protocol

The RTP control protocol (RTCP) is based on the periodic transmission of control packets to all participants in the session, using the same distribution mechanism as the data packets. The underlying protocol must provide multiplexing of the data and control packets, for example using separate port numbers with UDP. It is recommended that the fraction of the session bandwidth allocated to the RTCP is 5%. The primary function of this protocol is to provide feedback on the quality of the data distribution.
RTCP packets

- SR – sender report, for transmission and reception statistics from participants that are active senders
- RR - Receiver report, for reception statistics from participants that are not active senders and in combination with SR for active senders reporting on more than 31 sources
- SDES - Source description items, including CNAME (Canonical Name – RTP source identifier)
- BYE - Indicates end of participation
- APP - Application specific functions
Feedback in RTCP

- Sender and Receiver Reports (SR & RR)
- Timestamps allowing to calculate the Round-Trip Time
  \[ RTT = T4-T3+T2-T1 \]
- Packet counts
- Inter-arrival jitter (variation in delay)
- Fraction of packets lost, cumulative number of packet lost
- Number of packets expected to have been received
- Available bandwidth estimation (back-to-back packet sending)
<table>
<thead>
<tr>
<th>V=2</th>
<th>P</th>
<th>RC</th>
<th>PT=SR=200</th>
<th>length</th>
</tr>
</thead>
<tbody>
<tr>
<td>SSRC of sender</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NTP timestamp, most significant word</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NTP timestamp, least significant word</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RTP timestamp</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>sender’s packet count</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>sender’s octet count</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SSRC_1 (SSRC of first source)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>fraction lost</td>
<td></td>
<td>cumulative number of packets lost</td>
<td></td>
<td></td>
</tr>
<tr>
<td>extended highest sequence number received</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>interarrival jitter</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>last SR (LSR)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>delay since last SR (DLSR)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SSRC_2 (SSRC of second source)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>...</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>profile-specific extensions</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>V=2</td>
<td>P</td>
<td>RC</td>
<td>PT=RR=201</td>
<td>length</td>
</tr>
<tr>
<td>-----</td>
<td>---</td>
<td>----</td>
<td>----------</td>
<td>-------</td>
</tr>
<tr>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td></td>
</tr>
</tbody>
</table>

SSRC of packet sender

SSRC_1 (SSRC of first source)

Fraction lost

cumulative number of packets lost

Extended highest sequence number received

Interarrival jitter

Last SR (LSR)

Delay since last SR (DLSR)

SSRC_2 (SSRC of second source)

Profile-specific extensions
Feedback in RTCP – cont’d

- It is expected that reception quality feedback will be useful not only for the sender but also for other receivers and third-party monitors. The sender may modify its transmissions based on the feedback; receivers can determine whether problems are local, regional or global; network managers may use profile-independent monitors that receive only the RTCP packets and not the corresponding RTP data packets to evaluate the performance of their networks for multicast distribution.

- Cumulative counts in both the sender information and receiver report blocks allow to calculate differences between any two reports to make measurements over both short and long time periods, and to provide resilience against the loss of a report.
Feedback in RTCP – cont’d (2)

Using the SR and RR information we can obtain the following measurements:

- **Packet loss rate over the interval between two reception reports.** It is the difference in the cumulative number of packets lost (calculated over a given interval).
- **Number of packets expected during the interval** – it is the difference in the extended last sequence numbers received.
- **Packet loss fraction over the interval** - the ratio of the two above. This ratio should equal the fraction lost field if the two reports are consecutive, but otherwise it may not.
- **Loss rate per second** - can be obtained by dividing the loss fraction by the difference in NTP timestamps, expressed in seconds.
Feedback in RTCP – cont’d (3)

- **Number of packets received** is the number of packets expected minus the number lost.

- **Statistical validity of any loss estimates** – can be judged using the number of packets expected. For example, 1 out of 5 packets lost has a lower significance than 200 out of 1000.

- **Apparent throughput available to one receiver** – it is the number of packets received by a particular receiver times the average payload size (or the corresponding packet size), assuming that packet loss is independent of packet size.

- **Interarrival jitter** - provides a short-term measure of network congestion, it tracks transient congestion. The jitter measure may indicate congestion before it leads to packet loss.
RTCP – compound packet

- All RTCP packets must be sent in a compound packet of at least two individual packets (each periodically transmitted compound RTCP packet must include a report packet as well as the SDES CNAME)

```plaintext
if encrypted: random 32-bit integer

------------- packet -----------[------------- packet -----------][-packet-]

V          receiver          chunk  chunk
reports     item  item       item

-----------------------------------------------
R[SR #sendinfo #site1#site2] [SDES #CNAME PHONE #CNAME LOC] [BYE##why]

-----------------------------------------------
<----------------------------- compound packet ----------------------------->
<----------------------------------- UDP packet -------------------------------->

#: SSRC/CSRC identifier
```

Figure 1: Example of an RTCP compound packet
Presentation Outline

- Introduction
- Overview of RTP
- RTP Packets
- RTP Control Protocol (RTCP)
- RTP Payload Format (RTP Packetization) + Network Abstraction Layer (aside)
- Conclusions
RTP payload format for H.264

- Employs the native NAL (Network Abstraction Layer) interface, based on NAL units (NALUs)
- NALU – byte string of variable length that contains syntax elements of a certain class (coded slice, type A, B, C data partition or a sequence or picture parameter set)
- NALU header – type (5-bit field, types 1-12 currently defined by H.264), NRI (employed to signal the importance of a NALU for the reconstruction process), forbidden bit (specified to be 0, network elements can set it to 1 when they identify bit errors in the NALU)
- Reference: RFC 3984, February 2005

```
+------------------+
| 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
+------------------+
|                | NRI | Type |
+------------------+
```
Aside: Network Abstraction Layer (NAL)

- H.264 makes a distinction between a Video Coding Layer (VCL) and a Network Abstraction Layer (NAL). The output of the encoding process is VCL data (a sequence of bits representing the coded video data) which are mapped to NAL units prior to transmission or storage.
- Each NAL unit contains a Raw Byte Sequence Payload (RBSP), a set of data corresponding to coded video data or header information.
- A coded video sequence is represented by a sequence of NAL units that can be transmitted over a packet-based network or a bitstream transmission link or stored in a file.

| NAL header | RBSP | NAL header | RBSP | NAL header | RBSP |
Aside – cont’d – RBSP types

- Parameter Set – global parameters for a sequence (picture dimensions, video format, macroblock allocation map)
- Supplemental Enhancement Information
- Picture Delimiter – boundary between video pictures
- Coded slice – header and data for a slice, this RBSP unit contains actual coded video data
- Data Partition A, B or C – Data Partitioned slice layer data (A – header data for all MBs in the slice, B – intra coded data, C – inter coded data)
- End of sequence
- End of stream
- Filler data
Packetization Design Constraints

- Low overhead, so that MTU sizes of 100 bytes (or less) to 64 kbytes (maximum size of an IP packet – rarely used because of other MTU constraints) are feasible
- It should be easy to distinguish important from less important RTP packets, without decoding the bit stream carried in the packet

Table 2. Example of NRI values for coded slices and coded slice data partitions of primary coded reference pictures

<table>
<thead>
<tr>
<th>NAL Unit Type</th>
<th>Content of NAL unit</th>
<th>NRI (binary)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>non-IDR coded slice</td>
<td>10</td>
</tr>
<tr>
<td>2</td>
<td>Coded slice data partition A</td>
<td>10</td>
</tr>
<tr>
<td>3</td>
<td>Coded slice data partition B</td>
<td>01</td>
</tr>
<tr>
<td>4</td>
<td>Coded slice data partition C</td>
<td>01</td>
</tr>
</tbody>
</table>
Payload specification should allow the detection of data that became undecodable due to other losses, without a need to decode the bit stream.

- It should support NALU fragmentation into multiple RTP packets.
- It should support NALU aggregation – more than one NALU to be transported in a single RTP packet (the NALU size is then limited to 65535 bytes – as opposed to single NALU packet).

Table 1. Summary of NAL unit types and their payload structures

<table>
<thead>
<tr>
<th>Type</th>
<th>Packet</th>
<th>Type name</th>
<th>Section</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>undefined</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1-23</td>
<td>NAL unit</td>
<td>Single NAL unit packet per H.264</td>
<td>5.6</td>
</tr>
<tr>
<td>24</td>
<td>STAP-A</td>
<td>Single-time aggregation packet</td>
<td>5.7.1</td>
</tr>
<tr>
<td>25</td>
<td>STAP-B</td>
<td>Single-time aggregation packet</td>
<td>5.7.1</td>
</tr>
<tr>
<td>26</td>
<td>MTAP16</td>
<td>Multi-time aggregation packet</td>
<td>5.7.2</td>
</tr>
<tr>
<td>27</td>
<td>MTAP24</td>
<td>Multi-time aggregation packet</td>
<td>5.7.2</td>
</tr>
<tr>
<td>28</td>
<td>FU-A</td>
<td>Fragmentation unit</td>
<td>5.8</td>
</tr>
<tr>
<td>29</td>
<td>FU-B</td>
<td>Fragmentation unit</td>
<td>5.8</td>
</tr>
<tr>
<td>30-31</td>
<td>undefined</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Presentation Outline

- Introduction
- Overview of RTP
- RTP Packets
- RTP Control Protocol (RTCP)
- RTP Payload Format (RTP Packetization) + Network Abstraction Layer (aside)
- Conclusions
Conclusions

- RTP provides powerful instruments for adaptive video transmission
- Potential applications include wireless links
- Optimization can be done within the frames of the protocol specification (loosely defined packet sizes and RTCP communication frequency)
Thank you for your attention!
Questions?