Part 1: Lecture 3
Beyond TCP

Summary of last time

TCP congestion control phases
TCP state machine
Differences between TCP stack implementations

You discuss with your colleague next to you what are the limitation of TCP stacks and make a list of suggestions for improvement.

Your list

- Multicast support
- Encryption
- Multiple streams
- Support for real-time applications
- Optimization of fairness
- …
Limitations of TCP

According to RFC 4960 – Stream Control Transport Protocol:

1. Not everybody needs reliable data transfer and strict order-of-transmission delivery of data.
2. Streams may force application to mark boundaries of messages or PUSH to get whole messages delivered.
3. TCP sockets have limited scope and don’t exploit nodes with multiple IP addresses for high available transfers.
4. Vulnerability to SYN attacks

Possible solutions

Create completely new protocols.
Provide reliability on top of UDP.
Further tweak/improve TCP.

SCTP

RFCs 4960 defines SCTP
Stream Control Transmission Protocol
How do you setup a landline phone call?

SS7 is an out-of-band protocol that separates the voice path and the signaling path.

It is possible to run SS7 over IP networks, if one can:
1. Minimize end-to-end delay.
2. Guarantee short failover time in case of network failures.

SCTP Multi-homing

Every IP address of the peer is considered as a path.

All paths are continuously supervised and initially confirmed.

One path, the so-called primary path, is used for initial data transmission.

In the case of (timer-based) retransmissions an alternate path is used.

Load-sharing is not part of RFC 4960 but subject of ongoing research.

SCTP data transfer

- Multiplexing and demultiplexing for ordered delivery of messages.

- Only data sent within the same stream is delivered in sequence relative to that stream.
  - This minimizes the impact of head of line blocking in case of message loss.
SCTP message format

SCTP packet

- SCTP Header
- SCTP Chunk
- SCTP Chunk
- SCTP Chunk

Chunks may contain:
1. Control messages
   a. Applicable to the association
   b. Applicable to one of the streams
2. Data exchanged on a stream

Header:
1. Identifies the association
2. Contains security and verification details

SCTP header format

- Source port 16 bits
- Destination port 16 bits
- Verification tag
- Checksum

- Source and destination ports are the same as for the other IP transport protocols;
- The verification tag helps to protect from attacks; it allows to validate the sender.

SCTP chunk formats

<table>
<thead>
<tr>
<th>Chunk type</th>
<th>Chunk flags</th>
<th>Chunk length</th>
<th>Data</th>
<th>Padding</th>
</tr>
</thead>
</table>

- Data is interpreted according to the chunk type;
- Chunks for the whole association or for individual streams are processed in order

SCTP chunk types

- INIT, INIT-ACK, COOKIE-ECHO, COOKIE-ACK
- HEARTBEAT, HEARTBEAT-ACK
- DATA, SACK
- SHUTDOWN, SHUTDOWN-ACK, SHUTDOWN-COMPLETE
- ERROR, ABORT
- FORWARD-TSN
- ASCONF, ASCONF-ACK
- AUTH

SCTP Association

- Initiator
- Responder
SCTP data

- The DATA chunk contains:
  - TSN – Transmission Sequence Number
  - stream number
  - stream sequence number

<table>
<thead>
<tr>
<th>Chunk type</th>
<th>Chunk flags</th>
<th>TSN</th>
<th>Stream Sequence Number</th>
<th>Stream Identifier</th>
<th>Payload Protocol Identifier</th>
<th>User data</th>
</tr>
</thead>
</table>

16 bits 16 bits

SCTP acks

- Uses the selective acknowledgement idea from TCP:
  - Cumulative TSN - transmission sequence number - plus gap acknowledgments blocks, for received data blocks.

<table>
<thead>
<tr>
<th>Received</th>
<th>Missing</th>
<th>Received</th>
<th>Missing</th>
</tr>
</thead>
<tbody>
<tr>
<td>30 29 28 27 26 25 24 23 22 21 20 19 18 17 16 15 14 13 12 11 10 9 8 7 6 5 4 3 2 1</td>
<td>36 35 34 33 32 31 30 29 28 27 26 25 24 23 22 21 20 19 18 17 16 15 14 13 12 11 10 9 8 7 6 5 4 3 2 1</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

TSN = 36

Start gap ack offset = 4
End gap ack offset = 9

TSN = 36

Start gap ack offset = 12
End gap ack offset = 14

Availability of implementations

- Integrated in FreeBSD 7.
- For Linux: Part of 2.6 kernels and even back-ported to 2.4 kernels.
- Integrated in Solaris 10.
- For BSD-Unix, Linux, Solaris, Mac OS X, HP-UX and Windows: sctpLib (userland library).
- Several commercial implementations.
- Integrated in almost all SS7 nodes.

Evolutionary approach

- MPTCP - MultiPath TCP

- Tomorrow at 10:30am we will have a guest lecture from Benno Overeinder from NLNet labs on the topic.

- How do SCTP and MPTCP differ?
- What are the pros and cons of the two approaches?
VoIP and videoconferencing

Voice over IP

- Voice and multimedia sessions over the Internet/IP
- What’s the main advantage? Lower costs than on PSTN.

Devices

- Softphones
  - A software program for making telephone calls over the Internet
- VoIP phones

In the future:
- Netbooks and smartphones, that support 5G handoff (IEEE 802.21)
  - Allow roaming between 802.11 networks and (3G) cellular networks.

Real-time interactive applications

- Skype
- Polycom
- Net2Phone

PC to PC
PC to phone
Videoconference
With webcam
Video conferencing

- Reduced travel costs
- Easier access to remote experts
- More personal than a voice conference

Sampling at constant rate

Which problems will you have?
How will transport the data over the network?

Delay and jitter

Streaming stored video:

1. video recorded (e.g., 30 frames/sec)
2. video sent
3. video received, played out at client (30 frames/sec)

Cumulative data
Network delay

Streaming: at this time, client playing out early part of video, while server still sending later part of video
How does delay occur?

A packet needs to be transmitted
- packet needs to be processed
- packet needs to travel on link
- packets need to wait its turn

Four sources of packet delays

$d = d_{\text{trans}} + d_{\text{prop}} + d_{\text{proc}} + d_{\text{queue}}$

$d_{\text{trans}}$: transmission delay
- $L$: packet length (bits)
- $R$: link bandwidth (bps)
- $d_{\text{trans}} = L/R$

$d_{\text{prop}}$: propagation delay
- $d$: length of physical link
- $s$: propagation speed in medium (~2x10^8 m/sec in optical fiber)
- $d_{\text{prop}} = d/s$

$d_{\text{proc}}$: nodal processing
- check bit errors
- determine output link
- typically < msec

$d_{\text{queue}}$: queueing delay
- time waiting at output link for transmission
- depends on congestion level of router

End-to-end delay

$d = \sum \left( \frac{L}{R_i} + \frac{d_i}{s} + Q_i(t) \right)$
Streaming stored video: revisited

- constant bit rate video transmission
- variable delay
- client video reception
- constant bit rate video playout at client
- client playout delay

- client-side buffering and playout delay: compensate for network-added delay, delay jitter

Loss tolerance

- delay loss: IP datagram arrives too late for playout at receiver
  - delays: processing, queuing in network, end-system (sender, receiver) delays
  - typical maximum tolerable delay: 150 ms
- network loss: IP datagram lost due to network congestion (router buffer overflow)
- loss tolerance: depending on voice encoding, losses concealed, packet loss rates between 1% and 10% can be tolerated.
Multimedia protocols at a glance

In the PSTN
- SS7
- TDM

In the Internet
- SIP
- H.323
- RTCP
- UDP
- IPv4/IPv6
- RSVP
- TCP/SCTP
- RTP

Call management
- Transport of voice/video

Call management
- TDM
- Transport of voice/video
SIP versus RTP

- Initiate a session between two endpoints.
- Does not carry any voice or video data itself.

Transfer the traffic (voice or video):

Real-time Transport Protocol

- RTP is a top-level transport protocol used for real-time applications. Think of delivery of voice and video data:
  - Lightweight: one single message
  - Runs over another transport protocol
  - It supports multicast.

- Accompanied by RTCP - Real-Time Transport Protocol:
  - A management protocol
  - Allows endpoints to exchange information about data flows
  - Used by RTP to determine how to tune its behavior

RTP features

- Runs on top of UDP:
  - No guarantee of reliability
  - No guarantee of packet ordering

- It uses timestamps, sequence numbering, and delivery confirmation for each packet sent.

- It supports error-correction schemes for increased robustness and basic security options for encrypting packets.
**RTP header**

- **Sequence number**: A simple counter that starts at a random value when a session begins; it increments by one with each RTP message sent in that session; it provides a mechanism for the receiver to resequence packets that arrive out of order and to detect missing elements.
- **Timestamp**: It carries the time index for the first sample of the RTP message. The receiver can use this field to reassemble the information stream with the appropriate timing.

**RTCP functions**

1. **Integrated media synchronization**
   - i.e. when video and audio are transmitted on different streams
2. **QoS reports**
   - i.e. number of lost packets, jitters
3. **Participation reports**
   - i.e. when a participant is leaving the call
4. **Participation details**
   - i.e. information about the source, email address, sender names.
- each RTP session: typically a single multicast address; all RTP/RTCP packets belonging to session use multicast address
- RTP, RTCP packets distinguished from each other via distinct port numbers. RTCP port number = RTP port number + 1
- To limit traffic, each participant reduces RTCP traffic as number of conference participants increases.
- Typically 5% of the session bandwidth.

RTCP message types

- SDES (Sender Descriptor) message:
  - used by an application to join an RTP session;
- BYE message:
  - used when an application leaves the session;
- SR (Sender Report) message
- RR (Receiver Report) message
  - allow traffic monitoring;
- APP (application) message
Session Initiation Protocol

- SIP is described in RFC 3261 (2002)
- An application layer protocol
  - used instead of SS7 to initiate and terminate voice and video calls over IP networks (VoIP)
- It is independent of the underlying transport protocol
  - It can run on UDP, TCP, SCTP, RTP

<table>
<thead>
<tr>
<th>Call Control Application</th>
<th>Media Application</th>
</tr>
</thead>
<tbody>
<tr>
<td>SIP</td>
<td>RTCP, RTP, TCP, SCTP, UDP</td>
</tr>
</tbody>
</table>

SIP creates an infrastructure of network hosts (called SIP proxy servers) to which user agents can send registrations, invitations to sessions, and other requests.

SIP supports:

- Session setup
- Session management
- User capabilities
- User location
- User availability

SIP identifies

- SIP uses URI - Universal Resource Identifiers - to identify the communicating resource.
- Two schemes are defined:
  - Sip:
  - Sips:
    - When secure transmission is required, i.e. SIP messages are transported over TLS.

```plaintext
sip:user:password@host:port
sips:user:password@host:port
sip:alice@atlanta.com;transport=tcp
sip:alice@192.0.2.4
```
SIP servers

- Distributed functions across servers or consolidated in one node:
  - Proxy server
    - Intermediary program that acts as both a UA (User Agent) server and UA client to make requests on behalf of other UA clients.
  - Location server
  - Redirect server
  - Registrar server

Direct SIP call

- SIP call with proxy

- SIP messages
  - Text based, with syntax similar to HTTP

- SIP request
- SIP response

- INVITE, ACK, BYE, CANCEL, OPTIONS, REGISTER, PRACK, SUBSCRIBE, NOTIFY, PUBLISH, INFO, REFER, MESSAGE, UPDATE

- 1xx—Informational Responses
- 2xx—Successful Responses
- 3xx—Redirection Responses
- 4xx—Client Failure Responses
- 5xx—Server Failure Responses
- 6xx—Global Failure Responses
A closer look

**Start line**

```
HTTP/1.1 200 OK
```

**SIP headers**

```
Date: Sun, 20 Mar 2011 12:00:00 GMT
Content-Type: application/udp
```

**Message body**

```
Name translation and user location

1. caller wants to call callee, but only has callee’s name or e-mail address.
2. need to get IP address of callee’s current host:
   - user moves around
   - DHCP protocol
   - user has different IP devices (PC, PDA, car device)
3. result can be based on:
   - time of day
   - caller
   - status of callee

Registrar and Location Server

- Registrar: Registering the users and passing the information collected to the location server
- Location server: The entity in the SIP network that maintains the mapping between SIP addresses and IP address.

**Registrar and Location Server**

- Registrar: 192.0.2.100
- Location Server: 192.0.2.101

Caller jim@umass.edu calls keith@upenn.edu

```
REGISTER sip:192.0.2.101:5060 SIP/2.0
Via: SIP/2.0/UDP 192.0.2.100;branch=z9hG4bK26789
Max-Forwards: 70
From: <sip:jim@umass.edu>;tag=3932932932
To: <sip:keith@upenn.edu>;tag=3932932932
Call-ID: 192.0.2.101-3932932932
CSeq: 1 REGISTER
Content-Type: application/udp
Expire: 180
Content-Length: 0
```
H.323 is another signaling protocol for real-time, interactive applications.

- 1996: ITU-T published Version 1 of Recommendation H.323 in “Visual Telephone Systems and Equipment for LANs which provide a non-guaranteed Quality of Service”
  - not designed for the Internet!
  - only local calls, small number of users!
- 2009: current version
  - Operates well on WANs
  - Widely adopted also in large installations

**H.323 protocol stack**

**H.323 components**
A **Terminal** is an endpoint which provides for real-time, two-way communications with another H.323 terminal, GW, or MCU. This communication consists of speech only, speech and data, speech and video, or speech, data and video.

The **MCU (Multipoint Control Unit)** consists of:
1. a mandatory Multipoint Controller (MC) for call signaling, conference control
2. an optional Multipoint Processor (MP) switching/mixing of media stream

A **Gateway** provides for real-time, two-way communications between terminals belonging to networks with different protocol stacks.

The **Gatekeeper** provides address translation and controls access to the network resources for H.323 terminals, GWs and MCUs.

**Gatekeeper functionalities**

1. **Address Translation**
   - Translates H.323 IDs (such as p.grosso@uva.nl) and E.164 numbers (0031205257533) to endpoint IP addresses

2. **Admission control / bandwidth control**
   - Every call within the zone gets authorized by the GK
   - Admission requests (destination address, bandwidth) sent to GK

3. **Call control**
   - e.g. call transfer, call forwarding bus

**H.323 zones**

- Endpoints do register themselves at a GK
- All H.323 endpoints registered to a single GK build an **H.323 zone**:
  - H.323 zones are independent of physical network topology
  - Each zone has only one GK (exception: Alternate GKS)

**H.323 versus SIP**
Home reading

For the test on Apr. 12 read:
• DSCP
URL: http://tools.ietf.org/html/rfc2474
(up to section 4.3 “Summary” included)

Literature

Chapter 7 – Multimedia networking (for SIP and H.323)
Chapter 7 – Transport over IP (for SCTP and RTP)