Network Traffic

#5

Traffic Characterization

- Goals to:
  - Understand the nature of what is transported over communications networks.
  - Use that understanding to improve network design
- Traffic Characterization describes the user demands for network resources
  - How often a customer:
    - Requests a web page
    - Downloads an MP3
    - Makes a phone call
  - Size/length (how long you hold network resources)
    - Web page
    - Song
    - Phone call
Traffic Characterization

- Customers request information
- Rate of requests = \( \lambda \) requests/sec
  - Calls/sec
  - Packets/sec
  - mp3’s/hour
- The volume of information requested
  - Length of the phone call (sec/call)
  - Length of movie (Bytes)
  - Size of picture (Bytes)

Traffic Characterization

- Voice Traffic: Aggregate Traffic
- Voice Traffic: Individual voice sources
- Packet Voice
- Digital Video
- Data
Voice Traffic: Aggregate Traffic

- Arrival Rate \( \lambda \)
  - Number of requests/time unit
    - Calls/sec
- Holding Time, length of time the request will use the network resources
  - Min/call

\[
\text{Average Holding Time} = \bar{T_h}
\]

Voice Traffic: Aggregate Traffic

- Traffic Intensity (load)
  - Product of the average holding time and the arrival rate
    \[ \text{T} \text{raffic Intensity} = \rho = \lambda \bar{T_h} \]
- Units of Traffic Intensity: \( \rho \) is in Erlangs
- Traffic intensity is specified for the 'Busy Hour'

Call Patterns

Volume of calls

DNHR=Dynamic Non- hierarchical routing

Network protocol that leverages the nature of traffic

Traffic

Time of Day

8 am
12 noon
5 pm
Voice Traffic: Aggregate Traffic

- A telephone line busy 100% of the time = 1 Erlang
- A telephone busy 6 min/hour is how much traffic
  - 0.1 Erlang
- 100 telephones busy 10% of the time is how much traffic
  - 10 Erlangs

Traffic is Random
- Holding time (length of a phone call)
- Interarrival time (time between calls)

Common assumptions for probability density function (pdf) for
- Holding time ~ exponential
- Interarrival time ~ exponential

Section 4.7.1 and A.1.1
Voice Traffic: Aggregate Traffic

Probability Holding Time is < t sec =

\[ P[T_h < t] = 1 - e^{-\mu t} \text { for } t > 0 \text { and } 0 \text { for } t < 0 \]

\[ \overline{T}_h = \frac{1}{\mu} \text{ Service rate } = \mu \]

Probability Interarrival Time is < t sec =

\[ P[T_I < t] = 1 - e^{-\lambda t} \text { for } t > 0 \text { and } 0 \text { for } t < 0 \]

\[ T_I = \text{Interarrival time} \]

Average Interarrival Time = 1/\lambda

Voice Traffic: Aggregate Traffic

Interarrival Histogram

Percent Items

<table>
<thead>
<tr>
<th>Items' Distribution</th>
</tr>
</thead>
<tbody>
<tr>
<td>13.15</td>
</tr>
<tr>
<td>12.05417</td>
</tr>
<tr>
<td>10.95833</td>
</tr>
<tr>
<td>9.8625</td>
</tr>
<tr>
<td>8.76667</td>
</tr>
<tr>
<td>7.670833</td>
</tr>
<tr>
<td>6.575</td>
</tr>
<tr>
<td>5.479167</td>
</tr>
<tr>
<td>4.385333</td>
</tr>
<tr>
<td>3.2975</td>
</tr>
<tr>
<td>2.191673</td>
</tr>
<tr>
<td>1.095833</td>
</tr>
<tr>
<td>0.095833</td>
</tr>
</tbody>
</table>

Traffic

Interarrival Time

0.0007591758 1.318353 2.635946 3.953539 5.271133 6.588726 7.90632

Traffic
Voice Traffic: 
Aggregate Traffic

Source 1

- $T_1$ 
- $T_h$

Time

Source 2

- $T_1$ 
- $T_h$

Time

Voice Traffic: 
Individual voice source

- Speech inactivity factor
  - Talkspurt
  - Talkspurt
  - Talkspurt
  - Silence
  - Silence
  - time

Talkspurt duration
  - Random
  - Average duration $\rightarrow$ 0.350 s to 1.3 s
  - Exponentially distributed

Silence period
  - Random
  - Average duration $\rightarrow$ .58s to 1.6s
  - Exponentially distributed
Voice Traffic:
Individual voice source

- Digital Speech Interpolation (DSI) \([\text{Section 5.7.2}]\)
  - Uses “silence detection”
  - Multiplex at the talkspurt level
  - View as call set up at talkspurt level
  - ~Doubles the capacity
  - Analog version called “Time Assignment and Speech Interpolation (TASI)”
  - Packet Voice with **silence detection** effectively does DSI
  - **Effectively VoIP does DSI**

Voice Traffic:
Individual voice source

- Signal redundancies \(\Longleftrightarrow\) Voice coding
- Pulse code modulation (G.711) PCM
  - 8bits/sample @ 8000 samples/sec
  - 64kb/s
- Adaptive Differential PCM 32kb/s
- Linear Predictive 2.4 to 16 kb/s
- For Voice over IP: rate < 8kb/s
  - G.723.1 is emerging as a popular coding choice. G.723 is an algorithm for compressed digital audio over telephone lines.
## Comparison of popular CODECs

<table>
<thead>
<tr>
<th>Compression scheme</th>
<th>Compressed rate (Kbps)</th>
<th>Required CPU resources</th>
<th>Resultant voice quality</th>
<th>Added delay</th>
</tr>
</thead>
<tbody>
<tr>
<td>G.711 PCM</td>
<td>64 (no compression)</td>
<td>Not required</td>
<td>Excellent</td>
<td>N/A</td>
</tr>
<tr>
<td>G.723 MP-MLQ</td>
<td>6.4/5.3</td>
<td>Moderate</td>
<td>Good (6.4)</td>
<td>High</td>
</tr>
<tr>
<td>G.726 ADPCM</td>
<td>40/32/24</td>
<td>Low</td>
<td>Good (40)</td>
<td>Very low</td>
</tr>
<tr>
<td>G.728 LD-CELP</td>
<td>16</td>
<td>Very high</td>
<td>Good</td>
<td>Low</td>
</tr>
<tr>
<td>G.729 CS-ACELP</td>
<td>8</td>
<td>High</td>
<td>Good</td>
<td>Low</td>
</tr>
</tbody>
</table>

There is no "right CODEC". The choice of what compression scheme to use depends on what parameters are more important for a specific installation. In practice, G.723 and G.729 are more popular that G.726 and G.728.

For details of other VoIP Codecs see: [http://www.zytrax.com/tech/protocols/voip_rates.htm](http://www.zytrax.com/tech/protocols/voip_rates.htm)

---

## Voice Traffic:
### Individual voice source

- Example: How many calls can be supported on a system with the following parameters?
  - TDM
  - Coding rate/voice channel = ADPCM @ 32 Kb/s
  - DSI
  - Line rate = 1.536 Mb/s
    (note a T1/DS1 line is 1.544 Mb/s)

- Number of ADPCM channels = \[
\frac{1.563 \text{ Mb/s}}{32 \text{ Kb/s}} = 48
\]
- With DSI you get 2 calls/channel = 96
Voice Traffic: Packet Voice

- Example: Parameters for a packet voice system
  - 1 source
  - Sample rate = 8000 samples/sec (ITU G.711)
  - 8 bits/sample (1 byte/sample)
  - 8 ms/packet \(\rightarrow\) *Critical parameter*
  - Packet size (bytes/packet) = \((8\text{ms/packet})*(8000\text{ bytes/sec})=64\text{ Bytes}\)
    [assuming no overhead bytes]
  - Link rate = 10 Mb/s
  - Clocking time/packet (or Holding time/packet) = 
    \((64\text{bytes/packet})*8\text{bits/byte})/(10\text{ Mb/s})= 51.2\text{us}\)

Voice Traffic: Packet Voice

<table>
<thead>
<tr>
<th>51.2 us</th>
<th>Transmit at a Constant Bit Rate (CBR)</th>
</tr>
</thead>
<tbody>
<tr>
<td>8 ms</td>
<td>8 ms</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>51.2 us</th>
<th>Receive with variable interpacket arrival times</th>
</tr>
</thead>
<tbody>
<tr>
<td>X ms</td>
<td>time</td>
</tr>
</tbody>
</table>

X not equal 8ms because of network delays
If X is too big packet may arrive too late for play out
Voice Traffic: Packet Voice

- Packet voice looks like a steady flow or Constant Bit Rate (CBR) traffic
- However, voice can be Variable Bit Rate or VBR
  - "silence detection"
  - Variable rate coding
- Problem: After going through the network the packets will not arrive equally spaced in time. Thus playback of packet voice must deal with variable network delays

Assume network delay is uniformly distributed between [25 ms, 75 ms]
- Same as having a fixed propagation delay of 25 ms with a random network delay uniformly distributed between [0 ms, 50 ms]
- Note receiver will run out of bytes to play back after 8 ms.
- Solution:
  - Buffer 50 ms (or 8 packets or 2.8 Kbits)
  - Worst case, receiver will run out of data just as a new packet arrives
Voice Traffic: Packet Voice

- New problem: networks delays are unknown and maybe unbounded
- A voice packet may arrive at 85 ms and be too late to be played back
  - Late packets are dropped
  - Last packet may be played out in dead time
- Packet voice (video) schemes must be able to deal with variable delay and packet loss
  (Should voice packets be retransmitted?)

VoIP Quality

ITU-T Recommendation G.114-One-way transmission time, May 2003
Voice Traffic: Packet Voice

G.723.1 is a voice coding standard, linear prediction compression algorithm.

VoIP- Delay budget
Factors in End to End Delay

- Assumption: maximum delay from mouth-to-ear needs to be on the order of 200 - 300 ms

ITU G.114
- < 150 ms acceptable for most applications
- [150ms, 400 ms] acceptable for international
- > 400 ms unacceptable

From: http://www.protocols.com/papers/voip2.htm
VoIP- Delay budget
Factors in End to End Delay

- Example: Delay Budget (depends on assumptions)
  - Formation of VoIP packet at TX ~ 30 ms
    - 20 ms of voice/packet is default for Cisco 7960 router
  - Other VoIP packet processing ~70 ms
    (see: http://www.rmav.arauc.br/pdf/voip.pdf)
  - Propagation ~10 ms
  - Network Delays ~10 ms
  - Extraction of VoIP packet at Receiver ~30 ms
  - Jitter Buffer ~ 100 ms
    - Compensates for variable network delay
  - Total 250 ms

- Possible trade-offs:
  - Jitter Buffer vs voice packet loss
  - VoIP packet size vs length of jitter buffer


---

Video: *Analog video*

- Bandwidth ~ 4 Mhz
- Uncompressed rate 64 Mb/s
- Components of the signal
  - Luminance
  - Chrominance
  - Audio
  - Synchronization
Digital Video: MPEG

- Moving Pictures Experts Group
- Compresses moving pictures taking advantage of frame-to-frame redundancies
- MPEG Initial Target: VHS quality on a CD-ROM (320 x 240 + CD audio @ 1.5 Mbits/sec)

Digital Video: MPEG

- Converts a sequence of frames into a compressed format of three frame types
  - I Frames (intrapicture)
  - P frames (predicted picture)
  - B frames (bidirectional predicted picture)
Digital Video: MPEG

Exploits frame to frame redundancies

Frame sizes for talking head video.

Frame sizes for action video.

Each frame would be transported using multiple packets

From: Transmission of MPEG-2 Video Streams over ATM, Steven Gringeri et al., IEEE Multimedia, 1998
MP3- MPEG Layer 3 Audio

- MPEG specifies a family of three audio coding schemes, Layer-1, -2, -3.
- Each Layer has increasing encoder complexity and performance (sound quality per bitrate).
- The three codecs are compatible in a hierarchical way, i.e., a Layer-N decoder is able to decode bit stream data encoded in Layer-N and all Layers below N.
- The MP3 compression algorithm is based on a complicated psycho-acoustic model.
- The majority of the files available on the Internet are encoded in 128 kbits/s stereo.
- A high quality file is 12 times smaller than the original.
- CDs can be created that contain over 160 songs and can play for over 14 hours on a PC.
- Music can be efficiently stored on a hard disk and then directly played from there.

Digital Video: MPEG

- Compression ranges:
  - 30-to1
  - 50-to-1
- MPEG is evolving:
  - MPEG 1
  - MPEG 2
  - MPEG 4
  - MPEG 7
Digital Video: MPEG-4

- Initially for audio-video coding for “low bit-rate” channels,
  - Internet
  - Mobile applications
- Now used for kb/s to 10’s Mb/s video
- MPEG-4 is a significant change from MPEG-2
- Scalability is a key feature of MPEG-4
- MPEG-4 contains a Intellectual Property rights (IPR) management infrastructure

Digital Video: MPEG-4

- Object based: Audio-visual objects (AVO)
- AVO are described mathematically and given a position in 2D or 3D space
- Viewer can change vantage point and update calculations done locally
- No distinction between “natural” and “synthetic” AVOs: treats two in an integrated fashion
- Each AVO is represented separately and becomes the basis for an independent stream
- Each AVO is reusable, with the capability to incorporate on-the-fly elements under application control
- Content transport with QoS for each component
Data Traffic:  
General Characteristics

- Highly variable  
- Not well known  
- Likely to change as new services and applications evolve.

Data Traffic:  
General Characteristics

- Highly bursty, where one definition of burstyness is:

$$Burstyness = \frac{Peak \ rate}{Average \ rate}$$
Example: During a typical remote login connection over a 19.2kb/s modem a user types at a rate of 1 symbol/sec or 8 bits/sec and then transfers a 100 kbyte file. Assume the total holding time of the connection is 10 min.

What is the burstyness of this data session?

The time to transfer the file is

\[
\frac{800,000 \text{ bits}}{19,200 \text{ b/s}} = 41 \text{ sec.}
\]

So for 600 - 41 sec = 559 sec.

the data rate is 8 bits/sec or

\[4,472 \text{ bits were transferred in 559 sec.}\]

Thus in 600 sec. 4,472 + 800,000 bits were transferred, yielding a average rate of:

\[804,472 \text{ bits/600 sec} = 1,340 \text{ bits/sec.}\]

The peak rate was 19.2 Kb/s so the burstyness for this data session was:

\[19,200/1,340 = 14.3\]
Data Traffic: General Characteristics

### Asymmetric Nature of Interactive Traffic

- **User Burst**
- **Think Time**
- **Computer Burst**
- **Idle Time**

This Asymmetric property has lead to asymmetric services
Data Traffic: General Characteristics

- In Time Division Multiplexing (TDM) user must wait for turn to use link.
- Statistical Multiplexing (Stat Mux)
  - Note high burstness leads to “long” idle times
  - By transmitting the ‘bursts’ on demand the link can be efficiently shared.
  - To help insure fairness break the ‘burst’ into packets and transmit on a packet basis

Data Traffic: General Characteristics

- Element length
  - Message
  - Packet
  - Cell
- Arrival rate
  - Message/sec
  - Packets/sec
  - Cells/sec
Data Traffic:
General Characteristics

- Traffic intensity (< 1 with one server)
  \[ \rho = \frac{\lambda}{T_h} \]

  where

  \[ T_h = \frac{\text{Average Packet Length in Bits}}{\text{Link Capacity in Bits / sec}} = \frac{L}{C} \]

  Average Packet Length in Bits = \( L \)
  
  Link Capacity in Bits / Sec = \( C \)

Data Traffic:
General Characteristics

- Standard Assumptions
  - Message length has an exponential pdf
  - Interarrival time has an exponential pdf

Data was taken from special traces in http://www.nlanr.net/
Data was captured at the Internet Uplink of the University of Auckland by the
Wand Research group in the year 2000. The tap was installed on an OC-3 link.
KU/ITTC has collected aggregate traffic data from Sunflower Datavision.

From the Internet into Datavision
Mean = 8.876 Mb/s.
Maximum = 18.952 Mb/s

From Datavision out to the Internet
Mean = 5.133 Mb/s.
Maximum = 12.093 Mb/s
Data Traffic: Conclusions

- Very bursty
- Problems with traffic modeling
  - Rapidly evolving applications
  - Complex network interactions
- Issues:
  - Do models match “real” traffic flows?
  - Are the performance models based on specific traffic assumption robust

Conclusions

- Network traffic defines the demands for network resources
- Network traffic is dynamic
  - Changes with the deployment of new application
  - Time of day
- Models for network traffic are continuing to evolve