Application Lecture 2: Network Performance

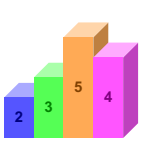
7COM1030 – Multicast and Multimedia Networking

Dr. Xianhui Che (Cherry)
x.che@herts.ac.uk
School of Computer Science
University of Hertfordshire, UK

University of Hertfordshire UH

Performance Management

- *How smoothly is the application running?*
- The goal is to measure and make available various aspects of network performance so that internetwork performance can be maintained at an acceptable level.
- Examples of performance variables:



Throughput

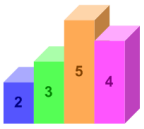
Response time

Network utilization

Performance Variables

The most common performance variables that are used to present network status include:

Response time



Throughput

Network utilization

• ETE Delay • Transmission Delay	• Current Throughput • Maximum Throughput • Goodput	• Network Efficiency • Network Utilization • Link Utilization • Saturation Point

Response Time

- Typical measurement:
  - **End-to-End Delay (ETE)** is the response time in the application level.
  - **TCP or UDP Transmission Delay** is the measurement of network delay in the transport level.
- ETE delay is normally used to evaluate network performance. This measurement is widely used in the industrial field.
- Network delay in the transport layer is used to observe the protocol performance. This measurements are mostly used in the academic and research field.

Acceptable ETE Delay

- ETE delay should be maintained at ≤ **ms** level. Each protocol may have different requirement for the ETE delay.
- **Case Study:**
  - The size of a typical webpage is about 50kb.
  - Surveys have shown that internet users will start to lose interests if the web page is not displayed after 10 seconds; and after 20 seconds if the page is still not displayed properly, most users will give up and close the page connection.
  - Hence, to achieve the best performance, the ETE delay needs to be kept under 0.2ms;
  - To achieve acceptable performance, the ETE delays should be 0.4ms at most.

Throughput

- Throughput is the data transfer rate. Measures the data amount being transferred per time unit.

Unit	Symbol	10^x	2^x
Kilo-bits per second	Kb/s	10^3	2^{10}
Mega-bits per second	Mb/s	10^6	2^{20}
Giga-bits per second	Gb/s	10^9	2^{30}
Tera-bits per second	Tb/s	10^{12}	2^{40}
Peta-bits per second	Pb/s	10^{15}	2^{50}
Exa-bits per second	Eb/s	10^{18}	2^{60}
Zetta-bits per second	Zb/s	10^{21}	2^{70}
Yotta-bits per second	Yb/s	10^{24}	2^{80}

Throughput Measurement

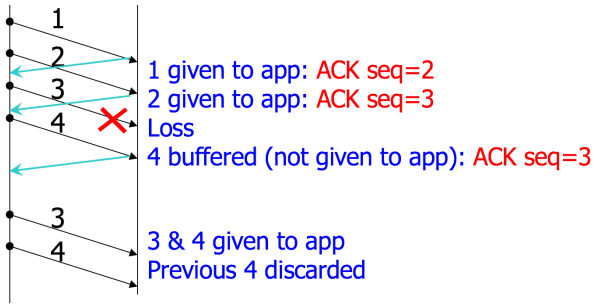
- Usual measurements are:
  - What is the current throughput of a network that runs a certain traffic profile?
  - What is the maximum throughput given a certain network topology and link capacity?
- **Current Throughput** is determined by the following elements in the network:
  - Traffic profile of applications (services)
  - Congestion situation
- **Maximum Throughput**, a.k.a. Network Capacity, gives the limit for the data transporting in the network. Two ways to obtain this value:
  - By pushing the simulation to the extreme (recommended)
  - By measurement based on equation (not exactly accurate)

Goodput

- **Goodput** is an effective measurement for the network efficiency.
- It only measures the useful bits throughput during an end-to-end transmission, excluding protocol overheads, re-transmission, acknowledgement, etc.
- Goodput is always smaller than throughput.
- Goodput stats can be collected at the transport layer.
- A good technique to collect TCP goodput is to measure the ACK sequence number.

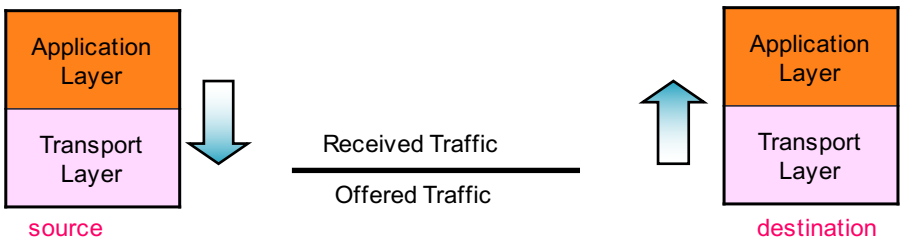
Measuring TCP Goodput

- TCP transmission uses a numbering method for sending and acknowledging. ACK number is the sequence number of last successfully received data plus one, in other words, the sequence number of the incoming data that is expected next. Reception of duplicate ACK triggers the TCP re-transmission.



Network Efficiency

- Network Efficiency is the ratio of successfully delivered traffic (from destination transport layer to destination application layer) to the offered traffic (from source application layer to source transport layer) in an End-to-End transmission.

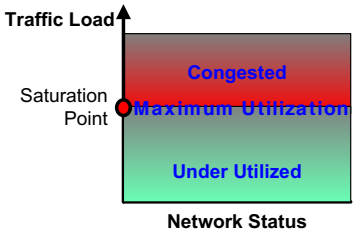


Network Utilization

- Can be calculated by

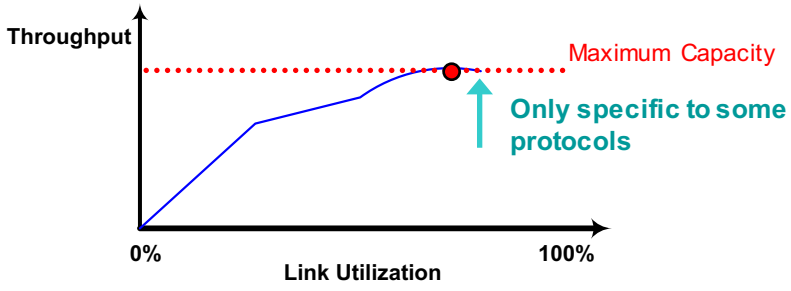
  current throughput/maximum throughput

- An alternative is to measure the link utilization, which is much easier and more straightforward.
- A good network solution is to try to achieve the maximum network utilization so that network resources are not wasted.
- The network reaches saturation state once the maximum network utilization level is met.



Link Capacity vs. Network Capacity

- Throughput increases with the link utilization.
- The network normally reaches the maximum capacity before the links reach the maximum utilization. The reason for that is the network will be very likely congested before the link reaches its full utilization. When the maximum network utilization has been met, it is not for sure the case of link utilization.



Questions?

- **Email:** x.che@herts.ac.uk
- **Office:** LB218
- **Tel:** 01707 286206