Advanced Computer Networks

Ibrahim Matta

IPN
InterPlaNet

facebook

What to expect?

- Increase understanding of fundamentals and design tradeoffs

- Discuss latest developments and research issues

- Naming & addressing, routing, connection management, flow / congestion control, queue management
- Architectures: extensions, overlays & clean-slate

- Modeling and correctness/performance analysis

Background?

- Basic networking
  - TCP/IP protocols and Internet principles
- Some mathematical sophistication
  - Basic probability and statistics

High Performance Networking

- TCP/IP extensions and performance-sensitive protocols and applications
  - *E.g.* features implemented in **Cisco IOS**, overlay architectures, convergent architectures, private (enterprise) networks, data-center networks

- Integrated Services (IntServ) and RSVP
- Differentiated Services (DiffServ)
- Multi Protocol Label Switching (MPLS)
- Traffic Engineering (or QoS/CoS Routing)
- Content Distribution Networks (CDN), *e.g.* Akamai
- Peer-to-Peer Networks (P2P), *e.g.* BitTorrent

- And clean-slate architectures

Ibrahim Matta – CS @ BU

Our Focus

- Learn about the fundamentals (and history) so
  - you don't re-invent the wheel!!
    e.g., algorithms for scheduling and routing used at different levels of the architecture
  - think about what's wrong so you don't repeat it, or if you ignore it, at least, know that you are ☺

- Learn about the design & dynamics of networks (and computing systems in general)!
  - *e.g.,* adaptations by the users and system/network

Ibrahim Matta – CS @ BU

A Feedback Control System
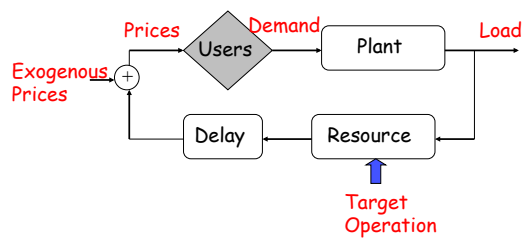
Ibrahim Matta – CS @ BU

How to achieve High Performance?

- Enhancements to datagram delivery
  - or should we use circuits?

- Original goal:
  - a robust communication system that can survive nuclear attacks [Paul Baran, 1960-64]
  - *"Both the US and USSR were building hair-trigger nuclear ballistic missile systems ...long-distance communication networks at that time were extremely vulnerable ...That was the issue. Here a most dangerous situation was created by the lack of a survivable communication system."* (Baran in Abbate, 10)

Ibrahim Matta – CS @ BU

Baran's Design: ARPANET then the Internet

- Packet switching technology
- Totally distributed
  - all nodes are equal
- Robust
  - adequate physical redundancy
  - adaptive routing
  - priority forwarding to transit over new packets
- Ends tolerate and recover from errors

Ibrahim Matta – CS @ BU

The Internet: Primary (original) Requirements

- Multiplexing
  - → packet switching
- Survivability (robustness)
  - → end-to-end, stateless net, datagram
- Service generality
  - → TCP, UDP, ... over IP
- Diverse network technologies
  - → "best-effort" IP

Ibrahim Matta – CS @ BU

The Internet:
End-to-End Principles

- *A function that can be entirely accomplished in an end node is left to that node, and the communication state is kept only in that node*
  → "fate-sharing", e.g. TCP

- The network is built with no knowledge of, or support for, any specific app or class of apps
  – Occam's razor: "the simplest of competing theories/models is preferred to the more complex" [Merriam-Webster]

Ibrahim Matta – CS @ BU

The Internet:
Secondary / Later Requirements

- Distributed management
  → two-tiered routing
- Security
  → encryption
- Mobility
  → mobile IP
- Resource allocation
  → fairness, QoS

Ibrahim Matta – CS @ BU

The Internet:
Other (new) requirements

- Accountability → value-based pricing
- Trust → firewalls, traffic filters
- Less sophisticated users → proxies
-

- **E2E principles are often broken! Are they?**
  – *e.g.*, web caches, proxies, *etc.* do application-specific processing within the net
  – **OK only as low-cost performance enhancements!**

Ibrahim Matta – CS @ BU

Different Approach to Reliability

- Phone system
- System reliability
  - every component reliable → minimal downtime
- Tightly controlled
  - signaling and access control
    - separate control plane
    - service predictability in data plane
  - → end-specific state inside the net (circuit-switched, hard state)
    - circuit switches simpler than IP routers!
- Later, adaptive routing of calls

Ibrahim Matta – CS @ BU

Applications, applications, applications

- Real-time: voice, video, emergency control, stock quotes, ...
- Non-real-time (or best-effort): telnet, ftp, …

- **Real-time apps have timing requirements:**
  - hard with deterministic or guaranteed requirements:
  no loss, packet delay less than deadline, difference in delays of any 2 packets less than jitter bound, …
  Note: reducing jitter within the Net reduces buffers needed to absorb delay variation at receiving host
  - soft with statistical or probabilistic requirements:
  no more than x% of packets lost or experience delay greater than deadline, …

Ibrahim Matta – CS @ BU

Is end-to-end control (ala TCP) enough?

- **Problem:** with common FCFS schedulers at routers, delay and delay variance increase very rapidly with load

- For an M/M/1 model:
  average delay = 1 / [ServiceRate - ArrivalRate]
  = 1 / [ServiceRate (1 - Load)]
  delay variance = 1 / [ServiceRate^2 (1 - Load)^2]

- As load increases, buffer overflows and router starts dropping packets

Ibrahim Matta – CS @ BU

Is end-to-end control (ala TCP) enough?

- **Solution:** TCP reduces load (slow start and congestion avoidance algorithm)
- 2 TCP users on different hosts sharing the same bottleneck may get different share of the bandwidth (uncontrolled unfairness)
  - ➔ *users should not trust the network*
- Some users may not "play by the rules" and reduce their sending rates upon congestion, *i.e.* not TCP-friendly sources like a voice or video UDP-based application
  - ➔ *network should not trust the users*

Ibrahim Matta – CS @ BU

The Erosion of Trust

"The simple model of the early Internet –
a group of mutually trusting users attached to a
transparent network – is gone forever."

"Making the network more trustworthy, while the
end-points cannot be trusted, seems to imply more
mechanism in the center of the network to enforce
"good" behavior."

[David Clark & Marjory Blumenthal, 2000]

Ibrahim Matta – CS @ BU

Economics, economics, economics

"It is in the nature of private enterprise to separate
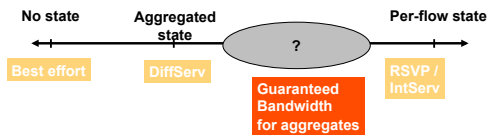users into different tiers with different benefits and
price them accordingly."

"Low prices and ease of use are becoming more
important than ever, suggesting growing appeal of
bundled and managed offerings over do it yourself
technology."

[Clark and Blumenthal, August 2000]

Ibrahim Matta – CS @ BU

Tradeoffs, tradeoffs, tradeoffs

- Can we tradeoff some state for service predictability?
- Maintain survivability and flexibility
  - quick recovery from failures
  - "run over anything"
  - support for many applications

No state | Aggregated state | ? | Per-flow state

Best effort | DiffServ | Guaranteed Bandwidth for aggregates | RSVP / IntServ
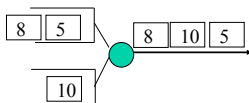
Ibrahim Matta – CS @ BU

Network Engineering

- Resource provisioning
- Traffic routing/engineering

- Architectural Enhancements
  - e.g. MPLS and Class-based Weighted Fair Queuing (CBWFQ) by Cisco, Juniper, Linux, etc.

Ibrahim Matta – CS @ BU

WFQ

- WFQ provides isolation and delay guarantees
- FQ simulates fair bit-by-bit RR by assigning packets priority based on finishing times under bit-by-bit RR
  - Approximation error bounded by

    max_pkt_size / capacity

- WFQ can assign different weights to different flows

| 8 | 5 |

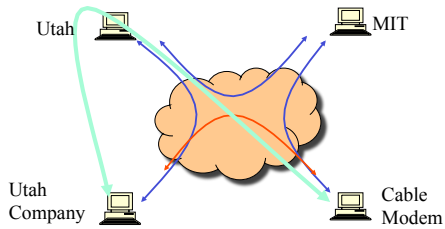| 8 | 10 | 5 |

| 10 |

Ibrahim Matta – CS @ BU

Effective Bandwidth

- Allocated bandwidth should depend on traffic characteristics and requirements

- Consider a discrete-time model with unit service rate and arrival process with mean R and variance V. To satisfy a delay bound D:
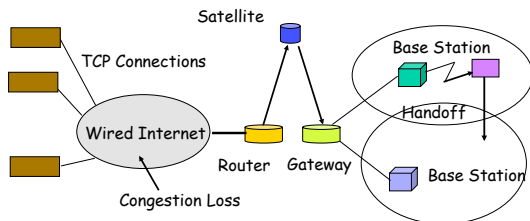
$$R + V/(2D - 1) < 1$$

Ibrahim Matta – CS @ BU

Overlays over IP or underlays

Utah

MIT

Utah Company

Cable Modem

- Overlay routes around Internet failures:
  - Outages: configuration/operational errors, fiber cuts, etc.
  - Performance failures: severe congestion, denial-of-service attacks, etc.

Ibrahim Matta – CS @ BU

Challenged Internet!

Satellite

Base Station

TCP Connections

Handoff

Wired Internet

Router   Gateway

Base Station

Congestion Loss

- High error rates, large delays, low capacities, …
- Radio, underwater, deep space, … links!

Ibrahim Matta – CS @ BU

Syllabus (sample questions)

- Introduction & Performance
  - where should we place a new functionality?
  - how do we describe performance?

- Naming & Addressing
  - what objects should we name?
  - what is the relationship between name and address?

- Connection Management
  - should we do hard state or soft state?

Syllabus (sample questions)

- Congestion control
  - do competing TCP flows converge to fairness and efficiency? if so, how close and how fast?
  - why have there been many recent TCP versions? how do they compare?
  - how do we design congestion-sensitive applications?
  - which packet should a router drop in times of congestion?
  - how does dropping affect performance?

Syllabus (sample questions)

- Routing
  - what makes a routing protocol unstable / inefficient?
  - how can we exploit the network structure to efficiently compute multi-criteria paths?

- Scheduling
  - which packet to serve next so as to provide fairness and service guarantees to flows?
  - what kind of guarantees to which flows?

Syllabus (sample questions)

- Architectures
  - how can we put together a scalable architecture that provides predictable performance?
  - how do protocols used by application entities (in overlays) differ from those used by routers?
  - how does the arch support wireless extensions?
  - what is the cost of transitioning to a new arch?

Ibrahim Matta – CS @ BU

Syllabus

Grand objective: *design and manage a (virtual or physical) network that delivers "desired" service quality*

Ibrahim Matta – CS @ BU

What is Quality? Service?

- Quality encompasses data loss, induced delay or latency, consistency in delays (jitter), efficient use of network resources, …

- Service means end-to-end communication between applications (e.g., audio, video, Web browsing), from a class of applications (e.g., TCP, UDP), …

- Tension between scalability and granularity of service

Ibrahim Matta – CS @ BU