High-Performance Networking
The University of Kansas EECS 881
Architecture and Topology

James P.G. Sterbenz

Department of Electrical Engineering & Computer Science
Information Technology & Telecommunications Research Center
The University of Kansas

jpgs@eecs.ku.edu

http://www.ittc.ku.edu/~jpgs/courses/hsnets

Network Architecture and Topology
Outline

AT.1 Topology and geography
AT.2 Scale
AT.3 Resource tradeoffs
Network Architecture and Topology

Outline

AT.1  Topology and geography
   AT.1.1  Scalability
   AT.1.2  Latency
   AT.1.3  Bandwidth
   AT.1.4  Virtual overlays and lightpaths
   AT.1.5  Practical constraints
AT.2  Scale
AT.3  Resource tradeoffs
Ideal Network Model

Network Path Principle

The network must provide high-bandwidth low-latency paths between end systems.

Link Topologies

Shared Medium

- Shared medium (e.g. bus topology)

problems?
Link Topologies

Shared Medium

- Shared medium (e.g. bus topology)
  - topology constrained by medium
    - longest path
    - limited to LAN lengths due to delay

Network scalability?
Link Topologies

Shared Medium

- Shared medium (e.g. bus topology)
  - topology constrained by medium
    - longest path
    - limited to LAN lengths due to delay
- Limited scalability
  - all hosts share capacity of bus

- contention for shared medium
Shared medium (e.g. bus topology)
- topology constrained by medium
  - longest path
  - limited to LAN lengths due to delay

Limited scalability
- all hosts share capacity of bus
- contention for shared medium
  - collisions in shared medium
  - need MAC (medium access control)

Example
- original DIX Ethernet
Link Topologies
Point-to-Point Mesh

- Point-to-point links between switches
  - space division mesh

Network scalability?
Link Topologies

Point-to-Point Mesh

- Point-to-point links between switches
  - space division mesh
- Scalable:
  - add links and expand switches
  - add switches
Scalable Topologies

Mesh vs. Shared Medium

Scalability of Mesh Topologies

Mesh network technologies scale better then shared medium. Use power control and directional antennae to increase spatial reuse in shared medium wireless networks.

Latency

Network Path

Network Latency Principle

The latency along a path is the sum of all its components. The benefit of optimising an individual link is directly proportional to its relative contribution to the end-to-end latency.
Latency

Topology and Geography

- Constituents of network latency \( D = (h-1)d_s + \sum di \)
  - Geography: speed of light delay \( d_i \)
  - Topology: number of hops \( h \); switching delay \( d_s \)

### Network Diameter Principle

\[ N-1A/h \]

The number of per hop latencies along a path is bounded by the diameter of the network. The network topology should keep the diameter low.

<table>
<thead>
<tr>
<th></th>
<th>SAN</th>
<th>LAN</th>
<th>MAN</th>
<th>WAN</th>
<th>GEO</th>
<th>Mars</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dia</td>
<td>100 m</td>
<td>1 km</td>
<td>100 km</td>
<td>20 000 km</td>
<td>72 000 km</td>
<td>400×10⁶ km</td>
</tr>
<tr>
<td>RTT</td>
<td>1 μs</td>
<td>10 μs</td>
<td>1 ms</td>
<td>200 ms</td>
<td>480 ms</td>
<td>6 – 45 min</td>
</tr>
</tbody>
</table>

### Network Bandwidth Principle

\[ N-1A/b \]

The maximum bandwidth along a path is limited by the minimum bandwidth link or node, which is the bottleneck. There is no point in optimising a link that is not a bottleneck.
Overlay Networks

Abstract and Hide Underlying Network

- Overlay networks hide underlying topology
  - VPNs (virtual private networks)
  - secure overlay sessions
  - datagram overlay meshes
  - lightpaths

Network Overlay Principle

Overlay networks must provide the same high-performance paths as the physical networks. The number of overlay layers should be kept as small as possible, and overlays must be adaptable based on end-to-end path requirements and topology information from the lower layers.

Overlay Networks

Lightpath Assignment

- Lightpath assignment problem
  - preserve high-performance along overlay such that
  - no link carries more than one light path of given wavelength
Network Architecture and Topology

Network Scale

AT.1 Topology and geography

AT.2 Scale
   AT.2.1 Network engineering
   AT.2.2 Hierarchy
   AT.2.3 Bandwidth aggregation and isolation
   AT.2.4 Latency optimisation
   AT.2.5 Wireless network density
   AT.2.5 Practical constraints

AT.3 Resource tradeoffs

Network Engineering Parameters

<table>
<thead>
<tr>
<th>topology</th>
<th>degree</th>
<th># nodes</th>
<th>diameter</th>
<th>aggregation</th>
</tr>
</thead>
<tbody>
<tr>
<td>sparsely connected</td>
<td>low</td>
<td>high</td>
<td>high</td>
<td>low</td>
</tr>
<tr>
<td>densely connected</td>
<td>high</td>
<td>low</td>
<td>low</td>
<td>high</td>
</tr>
</tbody>
</table>

Network Scaling Principle

Networks should be able to scale in size while balancing switch cost against hop count.
Network Scale

**Hierarchy**

- Hierarchy: important technique to
  - control latency and aggregation
  - bound state maintained
- divide network into clusters
  - state aggregated per cluster
  - examples: Nimrod, P-NNI

**Network Hierarchy Principle**

Use hierarchy and clustering to manage network scale and complexity, and reduce the overhead of routing algorithms.

---

**Aggregation, Isolation, Latency**

- Physical hierarchy to
  - limit number of hops and control aggregation
  - example: Internet (NSFNET) prior to 1994
**Network Scale**

**Aggregation, Isolation, Latency**

- **Hierarchy:** bandwidth isolation
  - isolate bandwidth in different subnet layers
  - exploit locality at network edge
    - e.g. within campus or enterprise network
  - aggregate bandwidth in network core

**Hierarchy to Aggregate and Isolate Bandwidth**

*Use hierarchy to manage bandwidth aggregation in the core of the network, and to isolate clusters of traffic locality from one another.*

- **Hierarchy:** latency engineering
  - control network diameter and latency
  - engineer low diameter at each level
  - shallow hierarchy ≈ 3 – 5 levels
    - home network or LAN
    - access or campus network
    - MAN
    - WAN backbone

**Hierarchy to Minimise Latency**

*Use hierarchy and cluster size to minimise network diameter and resultant latency.*
Wireless Network Density

- Transmission power gives tradeoff between
  - transmission range
  - degree of connectivity

**Wireless Density Control to Optimise Degree and Diameter**

Use density control and long link overlays to optimise the tradeoff between dense low-diameter and sparse high-diameter wireless networks.

**Practical Constraints**

- Policy-based routing
- Network provider deployment
  - complex topologies
  - peering not high-performance
  - many hops/POP

**Administrative Constraints Increase the Importance of Good Design**

Policy and administrative constraints distort the criteria that govern the application of many high-performance network design principles. The importance of good (principled) design is elevated when these constraints are present.
Network Architecture and Topology

Resource Tradeoffs

AT.1 Topology and geography
AT.2 Scale
AT.3 Resource tradeoffs
   AT.3.1 Bandwidth, processing, and memory
   AT.3.2 Latency as a constraint
   AT.3.3 Relative scaling with high speed
   AT.2.4 Active networking

Resource Tradeoffs

Resources and Constraints

• Network is a collection of resources and constraints:
  P processing  M memory
  B bandwidth  E energy or power
  L latency

• Relative composition
  – must be balanced to optimise cost and performance
  – determine topology, engineering, and functional placement

Network Resource Tradeoff & Engineering Principle

Networks are collections of resources. The relative composition of these resources must be balanced to optimise cost and performance, and to determine network topology, engineering, and functional placement.
Resource Tradeoffs

Resources and Constraints

- Network is a collection of resources and constraints:
  - P: processing
  - M: memory
  - B: bandwidth
  - E: energy or power
  - L: latency

- Relative composition
  - Must be balanced to optimise cost and performance
  - Determine topology, engineering, and functional placement

- Objective function to determine optimal mix
  \[ f(\pi(P), \beta(B), \mu(M), \varepsilon(E), \lambda(L)) \]

Example 3.5: Content Cache Location

- Example:
  - Content cache location
  - $(B) = \infty$, $(M) = 0$
  - Where goes the content?

⇒ where goes the content?
• Example: content cache location
  \( B = \infty \), \( M = 0 \)
  \( \Rightarrow \) everyone has local copy of everything
  \( B = 0 \), \( M = \infty \)
  \( \Rightarrow \) where goes the content?

  single copy on server: stream every time demanded
Resource Tradeoffs
Example 3.5 Content Cache Location

• Example: content cache location
  $(B) = \infty$, $(M) = 0$
  ⇒ everyone has local copy of everything
  $(B) = 0$, $(M) = \infty$
  ⇒ single copy on server: stream every time demanded
  *is this realistic?*
Resource Tradeoffs

Example 3.5 Content Cache Location

- Example: content cache location
  - plot cost of bandwidth vs. depth in network distribution tree

- plot cost of memory
Example 3.5 Content Cache Location

- Example: content cache location
  - plot cost of bandwidth
  - plot cost of memory

what next?

\[ l_{\text{opt}} = \min[\beta(B) + \mu(M)] \]
  - optimal level \( l \)
Resource Tradeoffs

**Example 3.5 Content Cache Location**

- Example: content cache location
  - plot cost of bandwidth
  - plot cost of memory

$$l_{opt} = \min[\beta(B) + \mu(M)]$$
  - optimal level $l$
  - latency constrains level
  - maximum distance from client

$$l_{opt} = \max(l_{Lmax}, l_{opt})$$

Resource Tradeoffs

**Relative Scaling**

- Relative scaling of resources is important
  - if all resources/constraints scale *uniformly*, no change
  - speed of light remains constant
    - latency becomes relatively more important
    - bandwidth-$\times$-delay product requires increasing $M$
  - technology scales non-uniformly
    - example: Moore’s law increase in $P$
      - enables IP lookup/classification in hardware, active networking

Resource Tradeoffs Change and Enable New Paradigms

*The relative cost of resources and constraints changes over time due to non-uniform advances in different aspects of technology. This should motivate constant rethinking about the way in which networks are structured and used.*
**Resource Tradeoffs**

**Active Networking**

- Decreasing cost of processing and memory
  - enables more computation in the network nodes
- Active networking
  - strong AN
    - users inject *capsules* of code
    - executed on nodes
  - moderate AN
    - network service providers provisions protocols and services
    - may be dynamically provisioned with active packets

---

**High-Performance Networking**

**Acknowledgements**

The material in these foils comes from the textbook supplementary materials:

- Sterbenz & Touch,
  *High-Speed Networking: A Systematic Approach to High-Bandwidth Low-Latency Communication*
  http://hsn-book.sterbenz.org