ECE 4750 Computer Architecture
Topic 6: Network Topology

Christopher Batten
School of Electrical and Computer Engineering
Cornell University

http://www.csl.cornell.edu/courses/ece4750
slide revision: 2014-10-15-12-21

Processors, Memories, Networks
Topology: Arrangement of Nodes and Channels

Many Potential Topologies
Topology is Constrained By Packaging

Routing: Determining Path Between Terminals

Mininal Routing vs. Non-Minimal Routing
Oblivious vs. Adaptive Routing
Deterministics vs. Randomized Routing
Flow Control: Managing Allocation of Resources

Router Microarchitecture
Evaluating A Network Implementation

Latency (seconds)

Ideal Throughput

Flow Control
Routing
Topology

Zero-Load Latency

Offered Bandwidth (bits/second)

---

Sun Niagara 2 Processor

- 8 multithreaded processors
- Single-stage crossbar connecting 8 cores to 4 L2 cache banks
- "200 GB/s" total bisection BW

ECE 4750 T06: Network Topology 9 / 12
IBM Cell Processor

- 1 general-purpose processor
- 8 processors specialized for data-parallelism
- 4 uni-directional rings
- Each ring is 128b wide at 1.6 GHz
- Network Bisection BW = 25.6 GB/s
- Total Bisection = 102.4 GB/s

MIT Raw Processor

- 16 simple RISC cores
- Two dynamically routed mesh networks (32b/channel)
- Two statically routed mesh networks for message passing (32b/channel)
- Bisection bandwidth per network is 8*32b at 400 MHz 12.8 = 12.8 GB/s
- Total bisection bandwidth is 51.2 GB/s
- Network consumes 20-30% of total chip power
**BASIC CROSS BAR TOPOLOGY**

- Single stage global crossbar difficult to scale in terms of cycle time, energy, and area
- Multi-stage topologies improve scalability but raise many other interesting challenges

**BUTTERFLY TOPOLOGY**

- k-ary n-fly
  - Router radix # stages
  - Unidirectional Channel

**2-ARY 2-FLY**

Each router is similar to a 2x2 crossbar
Torus Topologies

k-ary N-cube

Nodes in each Dimension

n-dimensional grid

Input + Output Terminal

Two unidirectional channels in opposite directions

Router

4-ary 2-cube

"Wrap around" channel

4-ary 2-cube

4-ary 2-cube

MESU

TORUS

3-ary 3-cube

MESU
Constructing $k$-ary $n$-cube from $k$ $k$-ary $(n-1)$ cubes

2-ary 1-cube

2-ary 2-cube 2-ary 3-cube

Binary $n$-cube

Hypercube

Terminology

2-ary 3-cube

4-ary 2-cube

MESU
**NODES + CHANNELS**

- BIDIRECTIONAL CHANNELS IN MESH
- UNIDIRECTIONAL CHANNELS IN CUCKY

**CHANNELS**

\[ \omega_c = \text{WIDTH} \] (# Wires)
\[ f_c = \text{frequency} \]
\[ t_c = \text{latency} \]
\[ l_c = \text{length} \]
\[ b_c = \omega_c \times f_c \] \((\text{CHANNEL BANDWIDTH})\)

- DIRECT (MESH) vs. INDIRECT (CUCKY)
  (NOT REALLY TOO IMPORTANT A DISTINCTION)

**Bisection Cuts**

- Both Cuts 4 and 5 on MESH are Bisection Cuts
- **4** is minimum bisection cut

\[ B_c = \text{Bisection Channel Count} \]
\[ \text{MESH} = 8 \]
\[ B_L = 4 \]

\[ B_c = B_c \cdot b_c \] \((\text{Bisection Bandwidth})\)
\[ \text{MESH} = B_b c \]
\[ b_L = 4 b_c \]

- Bisection BW is good way to estimate global wiring resources \((\text{ie. Technology constraint})\)
PATHS

- Hop = one element on path or a path
- Hop count = number of hops on a path
- Hop count on path from terminal A → B or mesh shown earlier is
  \[ H_2 = 5, \quad H_c = 4 \]  
  number of hops from terminal to first route 
  may or may not be included
- Minimal path = smallest hop count between 2 terminals
- Diameter = \( H_{\text{max}} \) largest minimal path between all terminal pairs

Show on mesh + Bfly

\[ H_{\text{min, min}} = 8, \quad H_{\text{max, min}} = 4 \]

- Average minimal hop count \( H_{\text{avg}} \) over all terminal pairs also is expected hop count for uniform demand

2D Mesh, \[ \frac{2}{3} k + 2 = \frac{2}{3} 4 + 2 = \frac{8}{3} + 2 = 4.3 \]

Bfly 4

PATH DIVERSITY

- Mesh illustrate multiple ways to get from upper left to lower right terminal
- Bfly was no path diversity
- Adding an extra stage can fix Bfly path diversity without impacting \( D_{\text{DS}} \) (version 2a)
| 4/350 | 600 |

Adding Extra
First Stages
↑ Pam Diversity
Traffic Patterns

Traffic matrix

Admission traffic patterns

Uniform

Random

Partition

VS. Tape

Neighbor

General

Transpose

Probability of Src 3 sending a packet to Dest 0

1/3 to every entry including sending to yourself!
**Logical to Physical Mapping**

Assume partition traffic pattern. Now are logical src/dst ids in traffic pattern mapped to physical terminal ids?

Mapping can turn any logical permutation pattern into any other or none. Given permutation pattern usually assumes a specific mapping.
**Performance: Throughput**

- Bottleneck channel will limit aggregate throughput.

Channel load ($N_c$) is amount of traffic that crosses channel $c$ if each input injects one unit of traffic according to given traffic pattern.

**Example Traffic Pattern**

$src \, i \rightarrow dest \, i+1$

Channel load ranges from 0-2.

Max channel load ($N_{max}$) is 2. These are the bottleneck channels that will limit throughput.

Alternative way to think about $N$. Channel load is ratio of data demand from channel to data injected by one terminal.
IDEAL THROUGHPUT

\[ \Theta_{\text{term}} = \frac{6e}{\eta_{\text{max}}} \]

\[ \Theta_{\text{tot}} = N \cdot \frac{6}{\eta_{\text{max}}} = 8 \cdot \frac{6}{2} = 48 \]

Often interested in ideal throughput under UNLOAD TRAFFIC.

Ring Example:

Two unidirectional channels

Because symmetric, focus on this channel.

Non-minimal path 1 \rightarrow 0

INCLUDER? Depends on floating algorithm.

Sometimes ideal throughput is for ideal routing algo

Sometimes for ideal routing algo

USE MINIMAL ROUTES

For now:

\[ N = 3 \cdot \frac{4}{7} = 0.75 \]

But this does not assume ideal routing. Ideal routing would load balance paths of length 2, send 0.5 traffic counterclockwise and 0.5 traffic clockwise.
- Ideal routing will result in the lowest $\gamma_{\text{max}}$ for a given traffic pattern.
- For any example and uniform random traffic,

$$\gamma_{\text{max}} = 0.5, \quad \Theta_{\text{term}} = \frac{6}{0.5} = 12, \quad \Theta_{\text{tot}} = 4 \cdot \frac{6}{0.5} = 86$$

**Small Diffy Examples**

1. $\gamma_{\text{max}} = 1, \quad \Theta_{\text{term}} = 6$

2. $\gamma_{\text{max}} = 1, \quad \Theta_{\text{term}} = 6$

3. $\gamma_{\text{max}} = 2, \quad \Theta_{\text{term}} = 6 \frac{1}{2}$

4. $\gamma_{\text{max}} = 4, \quad \Theta_{\text{term}} = 6 \frac{1}{4}$

Each tick = $\frac{1}{4}$ unit.
More generally for uniform random traffic

- On average, with uniform random traffic, half the traffic crosses the direction.
- Out of total units of traffic, N/2 cross direction.
- Ideal routing will evenly balance load across directions (this was an issue in Ring example).

\[ \text{N}_{\text{max}} = \frac{N/2}{Bc} = \frac{N}{2Bc} \]

\[ \text{Q}_{\text{term}} = \frac{6}{N/2Bc} = \frac{2Bc}{N} = \frac{2Bc}{2Bc} \]

\[ \text{Q}_{\text{tot}} = N \frac{2Bc}{N} = 2Bc \]

**Performance: Latency**

\[ 36 \text{ b/cycle} = \text{packet length} \]

\[ 8 \text{ b/cycle} = b_a \]
$$T = T_{\text{head}} + \frac{L}{b}$$

Serialization Latency

5 Head Pmt Latency
Includes tc, t_r, hop count, + contention

$$T = H_r t_r + H_c t_c + \frac{L}{b}$$

Latency due to router hps
Latency due to channel hops
Serialization latency

Zero Load Latency (no contention)
Four ways to improve latency

\[ T_\phi = H_r t_r + H_c t_c + \frac{L}{6} \text{ wide channels or shorter msg}s \]

Shorter routes Faster routes Faster channels

Avg latency vs offered BW

\[ T_\phi \] offered BW

Flow control

Ideal routing

Topology preparation
**Activity**

**MESH**

**SMESH**

**Assume:**

- \( L = 1286 \)
- \( b_c \) for MESH is \( 326 \) / cycle
- \( b_c \) for SMESH is \( 166 \) / cycle
- \( t_c = 3 \)
- \( t_c = 1 \)

**Which topology can achieve higher ideal throughput under uniform random traffic?**

1. \( \Theta_{term,\text{MESH}} > \Theta_{term,\text{SMESH}} \)
2. \( \Theta_{term,\text{MESH}} < \Theta_{term,\text{SMESH}} \)
3. \( \Theta_{term,\text{MESH}} = \Theta_{term,\text{SMESH}} \)

**Calculate zero load latency under uniform random traffic**
**MESH**

\[ T_C = 6 \quad T_B = 6 \cdot 3.2 = 19.2 \text{ b/cycle} \]

\[ \Theta_{cm} = \frac{2 \cdot T_B}{N} = \frac{2 \cdot 19.2}{6} = 6.4 \text{ b/cycle} \]

but \( T_C \) is just 32 b/cycle! Limiting by worst terminal injection backwash!

Achievable \( \Theta_{cm, \text{min}} = 32 \text{ b/cycle} \)

---

**SMESH**

\[ T_C = 10 \quad T_B = 10 \cdot 16 = 160 \text{ b/cycle} \]

\[ \Theta_{cm} = \frac{2 \cdot T_B}{N} = \frac{2 \cdot 160}{6} = 53.3 \text{ b/cycle} \]

but \( T_C \) is just 16 b/cycle! Again limiting by worst terminal injection backwash!

Achievable \( \Theta_{cm, \text{min}} = 16 \text{ b/cycle} \)

---

**MESH**

Min Bifurcation<br>**CUT**

---

**SMESH**

Min Bifurcation<br>**CUT**
**MESY**

First calculate $T_{\phi}$ under uniform random traffic

<table>
<thead>
<tr>
<th>SRC</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>1</td>
<td>2</td>
<td>3</td>
<td>2</td>
<td>3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>4</td>
<td>3</td>
<td>2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>1</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>3</td>
<td>2</td>
<td>3</td>
<td>2</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>3</td>
<td>2</td>
<td>3</td>
<td>2</td>
<td>1</td>
<td></td>
</tr>
</tbody>
</table>

$1 \times 6 = 6$

$2 \times 14 = 28$

$3 \times 12 = 36$

$4 \times 4 = 16$

$86/36 = 2.39$

$T_{\phi} = H_{ctn} + H_{cte} + \frac{L}{b}$

$= 2.39 \times 3 + 1.39 \times 1 + 128/36$

$= 7.17 + 1.39 + 4$

$T_{\phi} = 12.56$ cycles

**MSER**

First calculate $T_{\phi}$ under uniform random traffic

<table>
<thead>
<tr>
<th>SRC</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>1</td>
<td>2</td>
<td>3</td>
<td>2</td>
<td>3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>3</td>
<td>2</td>
<td>2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>2</td>
<td>3</td>
<td>1</td>
<td>2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>3</td>
<td>2</td>
<td>3</td>
<td>2</td>
<td>1</td>
<td></td>
</tr>
</tbody>
</table>

$1 \times 6 = 6$

$2 \times 22 = 44$

$3 \times 8 = 24$

$74/36 = 2.06$

$T_{\phi} = H_{ctn} + H_{cte} + \frac{L}{b}$

$= 2.06 \times 3 + 1.06 \times 1 + 128/16$

$= 6.18 + 1.06 + 8$

$T_{\phi} = 15.24$ cycles
SMESH

MESHE

Traffic!

Achievable term
(6th order)
Calculating $\mathcal{H}_{\text{MW}}$

2-ary 2-cube
Circles = terminals
Squares = routers
"Terminal channels"

Calculating $\mathcal{H}_{\text{MW},i}$

<table>
<thead>
<tr>
<th>src</th>
<th>0123</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>1232</td>
</tr>
<tr>
<td>1</td>
<td>2123</td>
</tr>
<tr>
<td>2</td>
<td>3212</td>
</tr>
<tr>
<td>J</td>
<td>2321</td>
</tr>
</tbody>
</table>

$1 \times 4 = 4$ including $i \rightarrow i$
$2 \times 8 = 16$
$3 \times 4 = 12$ excluding $i \rightarrow i$

$32/16 = 2$
$28/12 = 2.3$

Calculating $\mathcal{H}_{\text{MW},i}$ with terminal channels

<table>
<thead>
<tr>
<th>src</th>
<th>0123</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>2343</td>
</tr>
<tr>
<td>1</td>
<td>3243</td>
</tr>
<tr>
<td>2</td>
<td>4323</td>
</tr>
<tr>
<td>J</td>
<td>3432</td>
</tr>
</tbody>
</table>

$2 \times 8 = 8$ including $i \rightarrow i$
$3 \times 8 = 24$
$4 \times 4 = 16$ excluding $i \rightarrow i$

$40/16 = 2.5$
$40/12 = 3.3$

Calculating $\mathcal{H}_{\text{MW},i}$ without terminal channels

<table>
<thead>
<tr>
<th>src</th>
<th>0123</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0121</td>
</tr>
<tr>
<td>1</td>
<td>1012</td>
</tr>
<tr>
<td>2</td>
<td>2101</td>
</tr>
<tr>
<td>J</td>
<td>1210</td>
</tr>
</tbody>
</table>

$1 \times 8 = 8$ including $i \rightarrow i$
$2 \times 8 = 16$
$3 \times 8 = 24$
$4 \times 8 = 32$

$16/16 = 1$
$16/12 = 1.3$

Equation from book is $\frac{nk}{3}$ for even $n$

$\frac{nk}{3} = \frac{2.2}{3} = 1.33$

This is for $\mathcal{H}_{\text{MW},i}$ without terminal channels + ignoring $i$ sensors' stuff