EE382C
Lecture 2

3/31/11

Homework 1

- 3.10, 5.1, 5.3, 7.3

Today

- A simple interconnection network
- Topology

Longer Term Plan

No	Date	Topic	Assignment	Read
1	29-Mar	Introduction to interconnection networks. Walk through of a simple network.		Chapters 1 & 2
2	31-Mar	Topology basics. Constraints and measures. Butterfly networks.	HW1: Topology	Chapters 3 & 4
3	5-Apr	Cube networks. Concentration and slicing.	Research Paper Assigned	Chapters 5 & 7
4	7-Apr	Non-blocking topologies.		Chapter 6
5	12-Apr	Topology overflow and wrapup. Routing basics and taxonomy.	HW2: Routing and Flow control	Chapters 8 & 9
6	14-Apr	Oblivious routing. Adaptive routing. Routing mechanics.		Chapters 10 & 11
7	19-Apr	Global adaptive routing.		Chapter 12
8	21-Apr	Flow control basics. Resources and allocation strategies. Circuit switching. Store and forward. Dropping flow control. Misrouting. Cut through. Wormhole flow control, Virtual channels.	HW3: Router architecture	
9	26-Apr	Deadlock and livelock. Principles of deadlock. Buffer deadlock and channel deadlock. Deadlock in cyclic networks. Inter-dimension deadlock. Avoiding deadlock with virtual channels. The turn model.		Chapter 14
10	28-Apr	Router microarchitecture. Basic router. Input buffers and buffer organization. Internal switch organization: crossbars, dimension-ordered, and multistage.	Project assignment	Chapter 16
11	3-May	Midterm exam, in class	Research Paper Due	
12	5-May	Router datapath components, router pipelining, router delay models.	Checkpoint 1	Chapter 17
13	10-May	Allocators. Arbiters. The allocation problem - allocating VCs to packets and bandwidth to flits. Bipartite matching. Naïve allocation. Separable allocators. Wavefront allocation.		Chapters 18 & 19
14	12-May	Network performance analysis. Analysis of networks with dropping flow control. Analysis of blocking. The effects of buffers. Simulation vs. analysis. The effect of traffic patterns. Load balance and route diversity.		Chapters 23-25
15	17-May	Reliability: Definition of Reliability and Availability. Failure mechanisms and fault models. Path diversity. Pragmatics and self-healing.	Checkpoint 2	Chapter 21
16	19-May	TBD		
17	24-May	Project Presentations		
18	26-May	Project Presentations		
19	31-May	Wrapup Lecture		

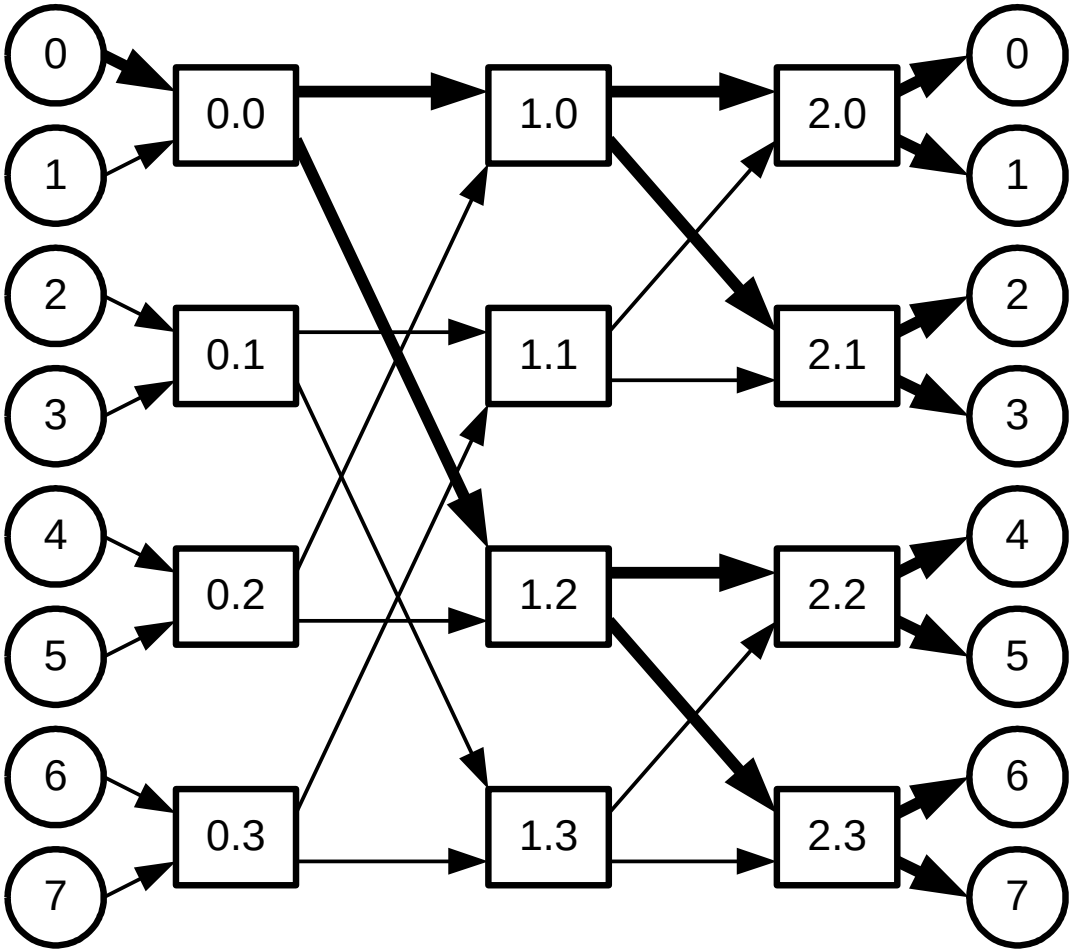
A Simple Interconnection Network

Requirements & Constraints

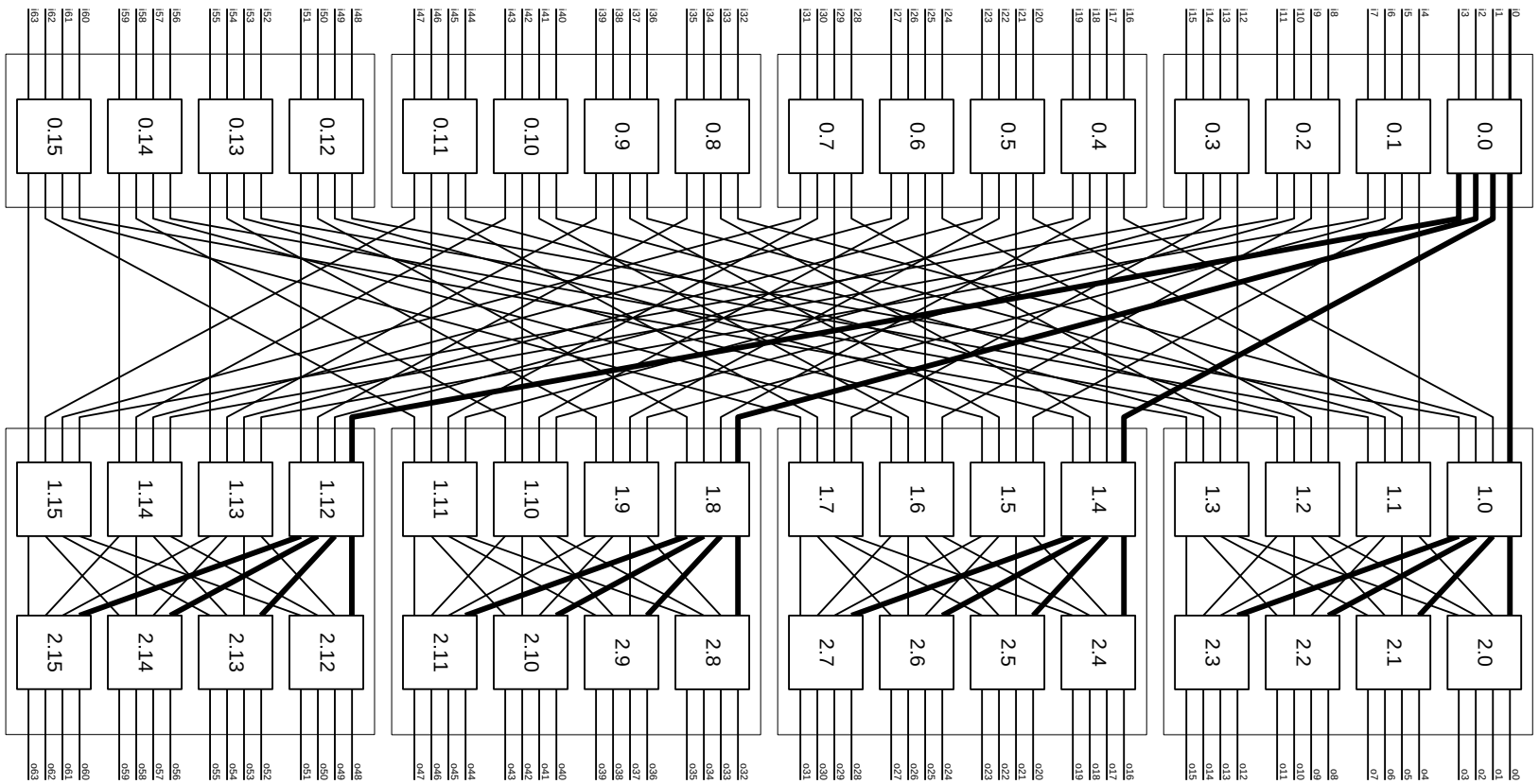
Parameter	Value
Input Ports	64
Output Ports	64
Peak Bandwidth	0.25 GByte/s
Average Bandwidth	0.25 GByte/s
Message Latency	100ns
Message Size	4–64 bytes
Traffic Pattern	random
Quality of Service	dropping acceptable
Reliability	dropping acceptable

Parameter	Value
Port Width	2 bits
Signaling rate	1 GHz
Signals per chip	150
Chip cost	$200
Chip pin bandwidth	1 Gb/s
Signals per circuit board	750
Circuit board cost	$200
Signals per cable	80
Cable cost	$50
Cable length	4m at 1Gb/s

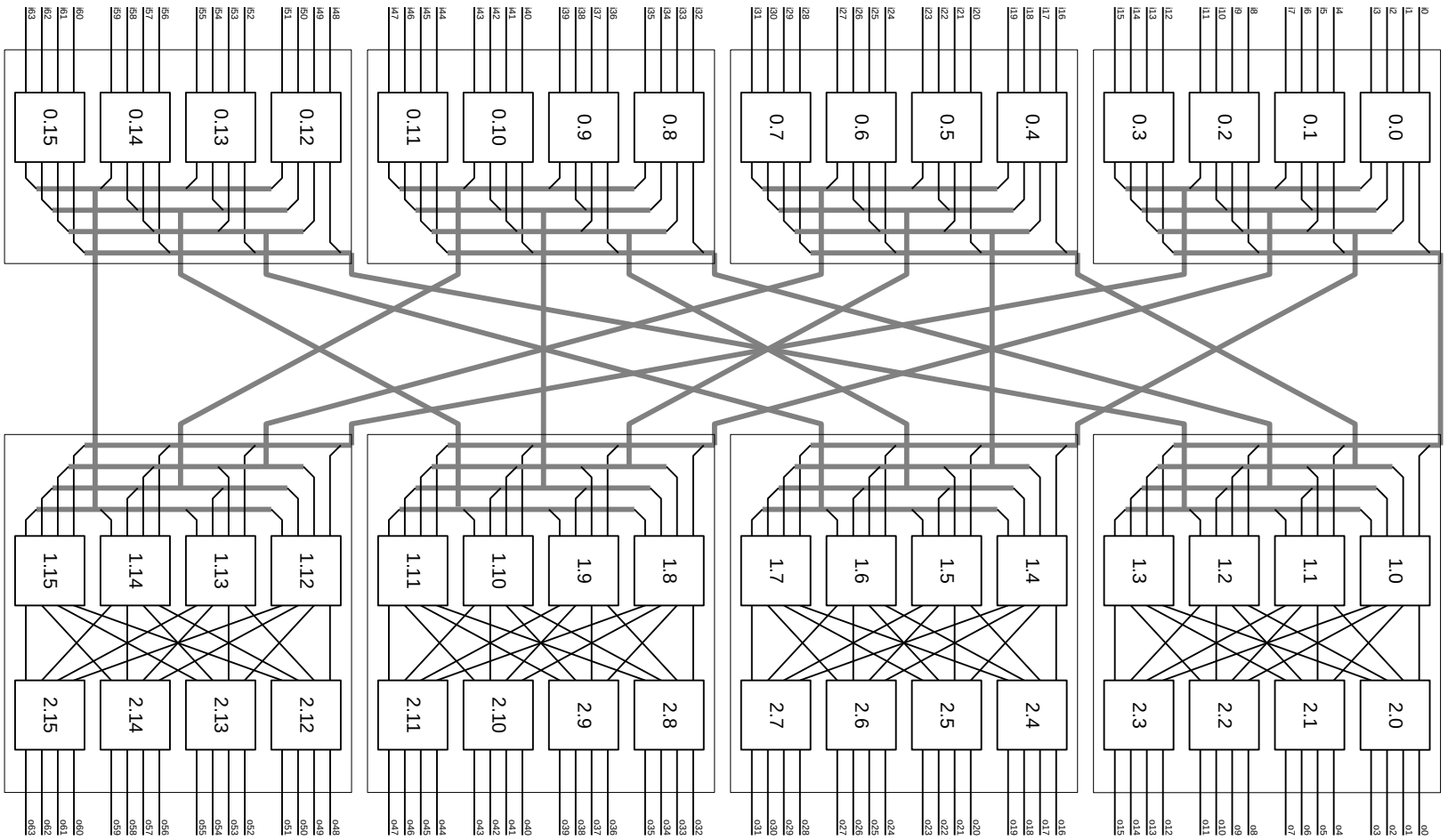
Topology – Butterfly



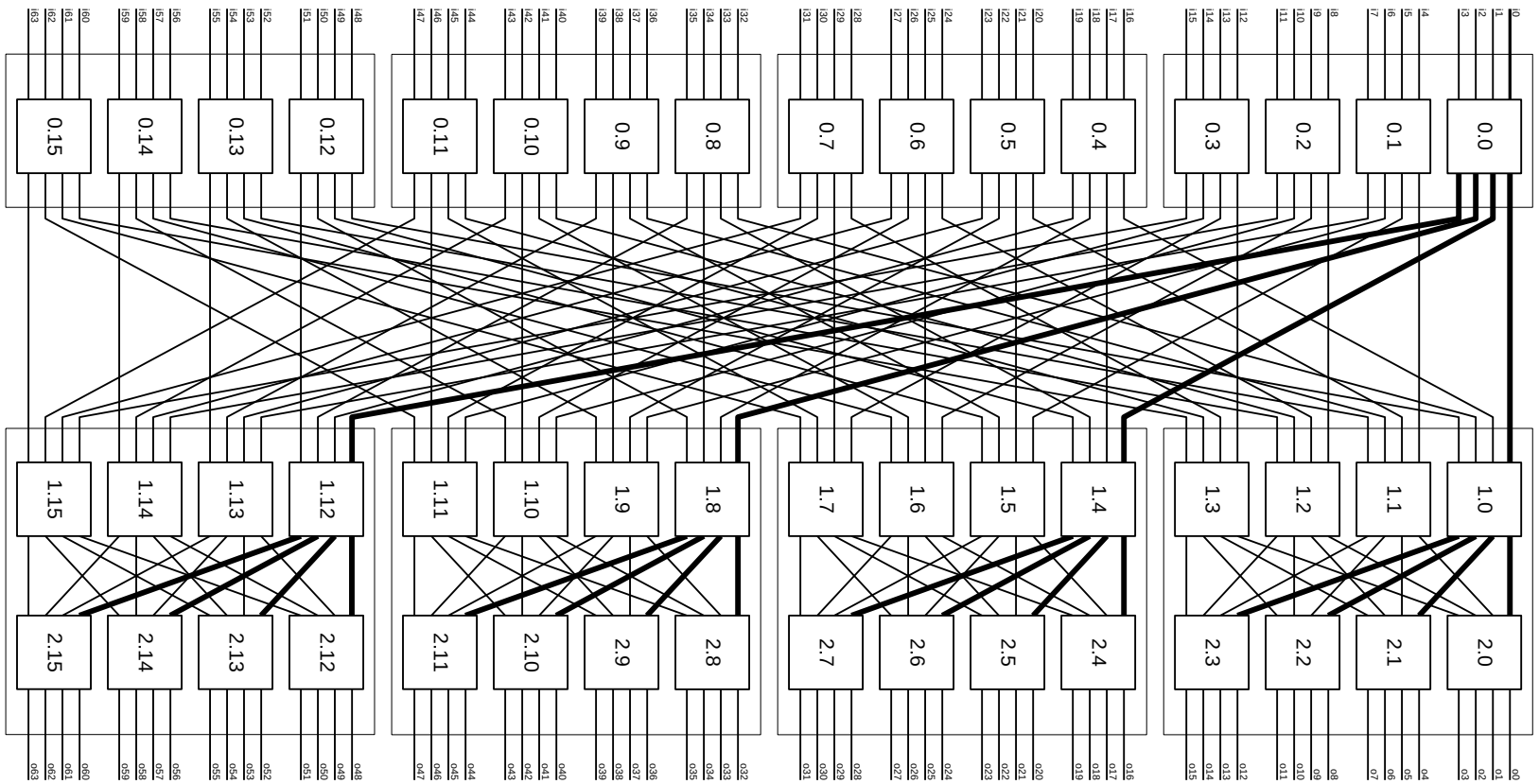
Fit to packaging constraints
4-ary 3-fly



Cables

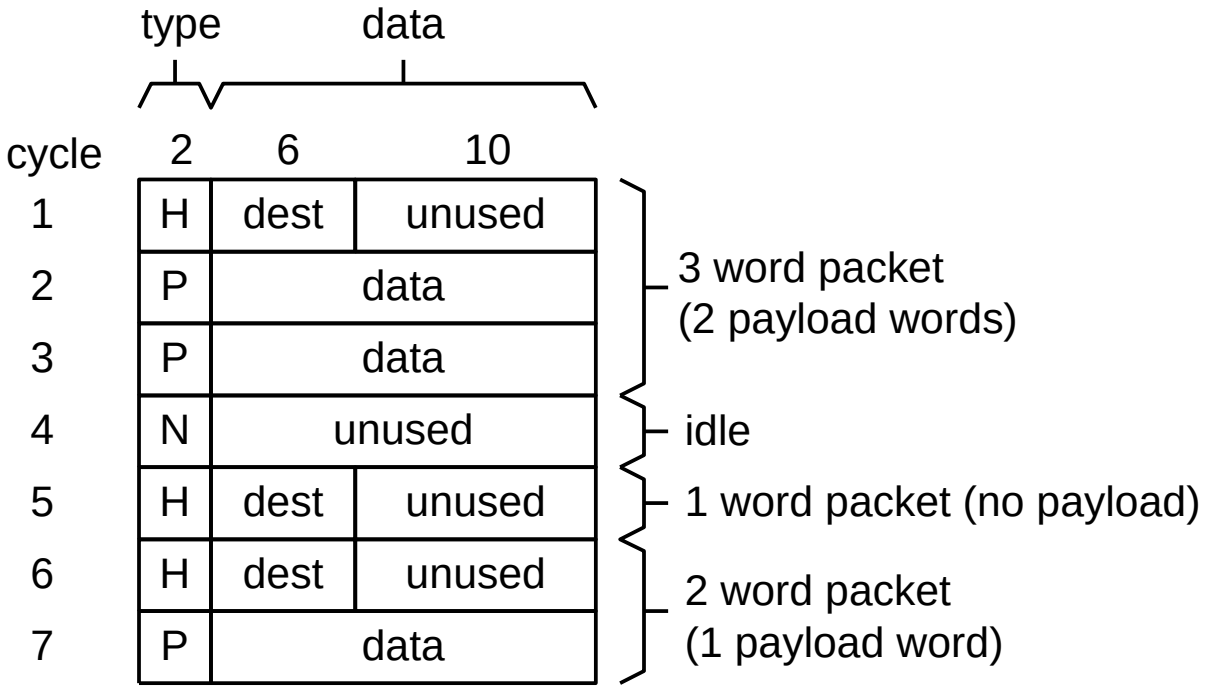

0.0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
0.10
0.11
0.12
0.13
0.14
0.15
1.0
1.1
1.2
1.3
1.4
1.5
1.6
1.7
1.8
1.9
1.10
1.11
1.12
1.13
1.14
1.15
2.0
2.1
2.2
2.3
2.4
2.5
2.6
2.7
2.8
2.9
2.10
2.11
2.12
2.13
2.14
2.15

Routing – By Destination Tag

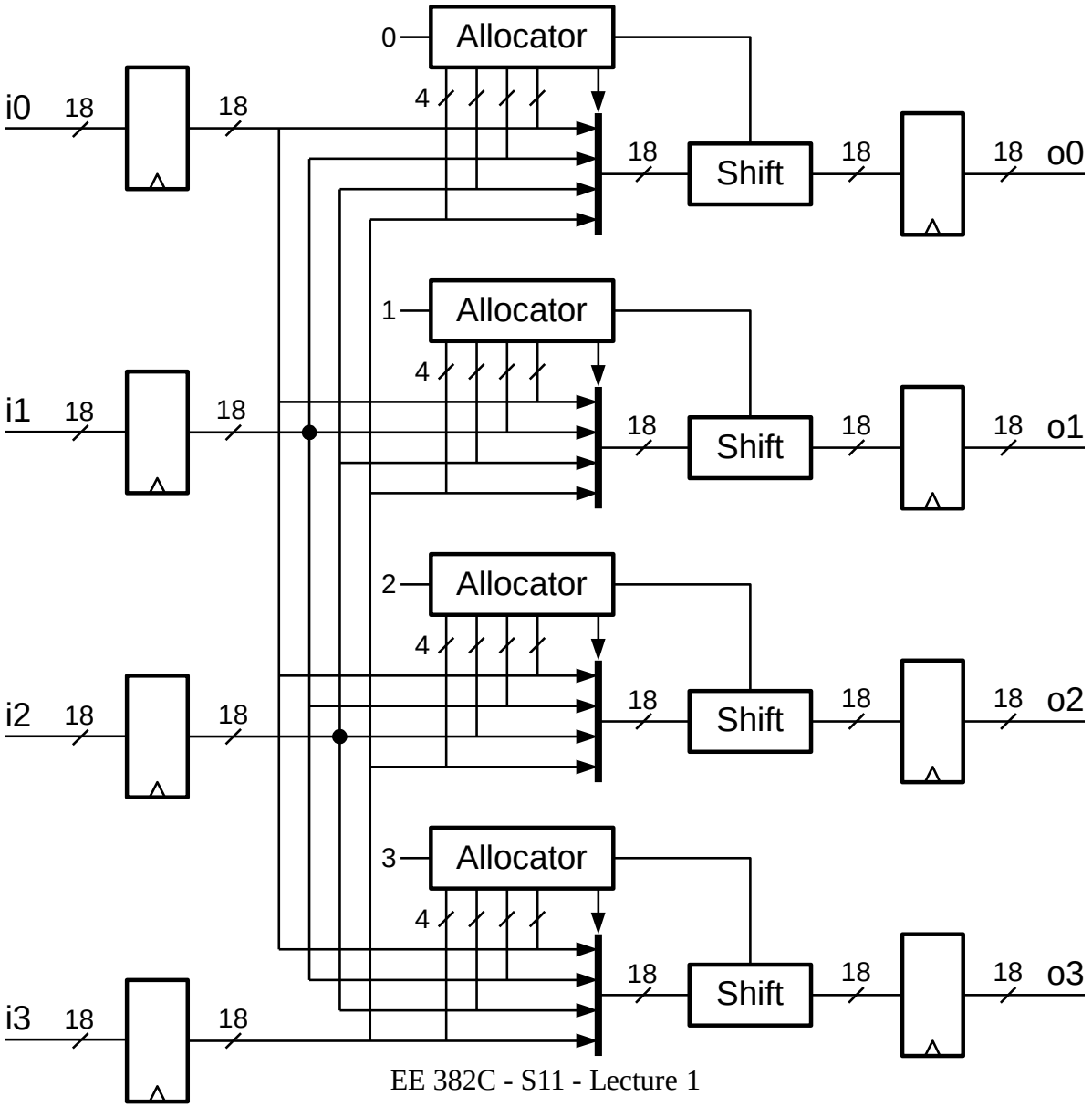


Flow Control

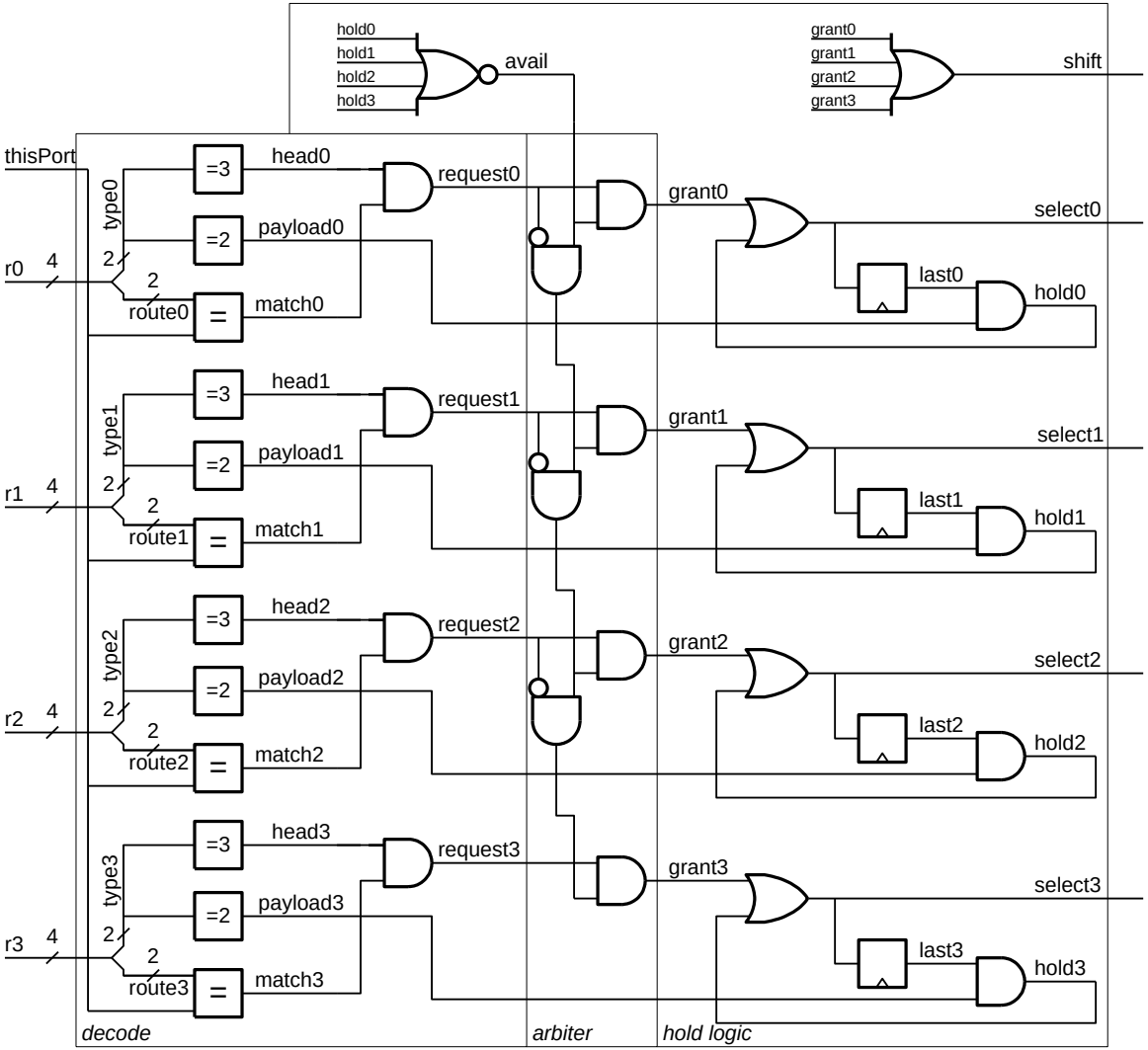
- Drop on conflict
  Assume higher level retransmit protocol



Router Architecture



Allocator


hold0
hold1
hold2
hold3
avail
grant0
grant1
grant2
grant3
shift
thisPort
r0
r1
r2
r3
4
type0
type1
type2
type3
2
=3
=2
=
head0
payload0
match0
route0
request0
grant0
select0
last0
hold0
head1
payload1
match1
route1
request1
grant1
select1
last1
hold1
head2
payload2
match2
route2
request2
grant2
select2
last2
hold2
head3
payload3
match3
route3
request3
grant3
select3
last3
hold3
decode
arbiter
hold logic

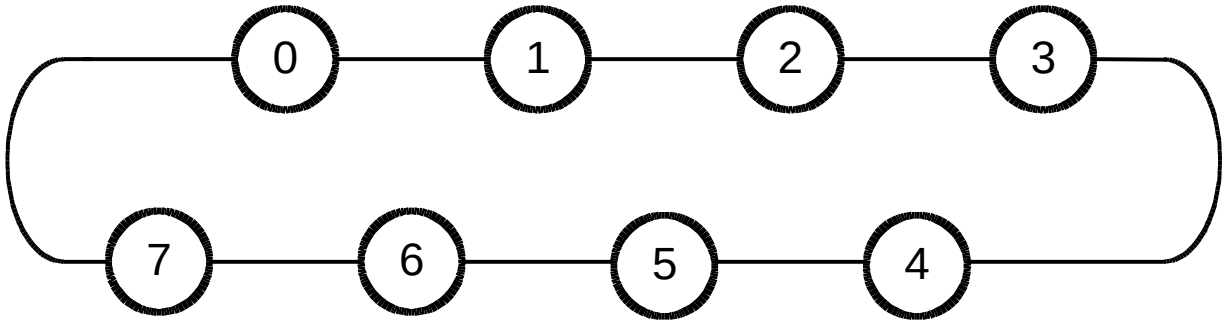
Topology Basics

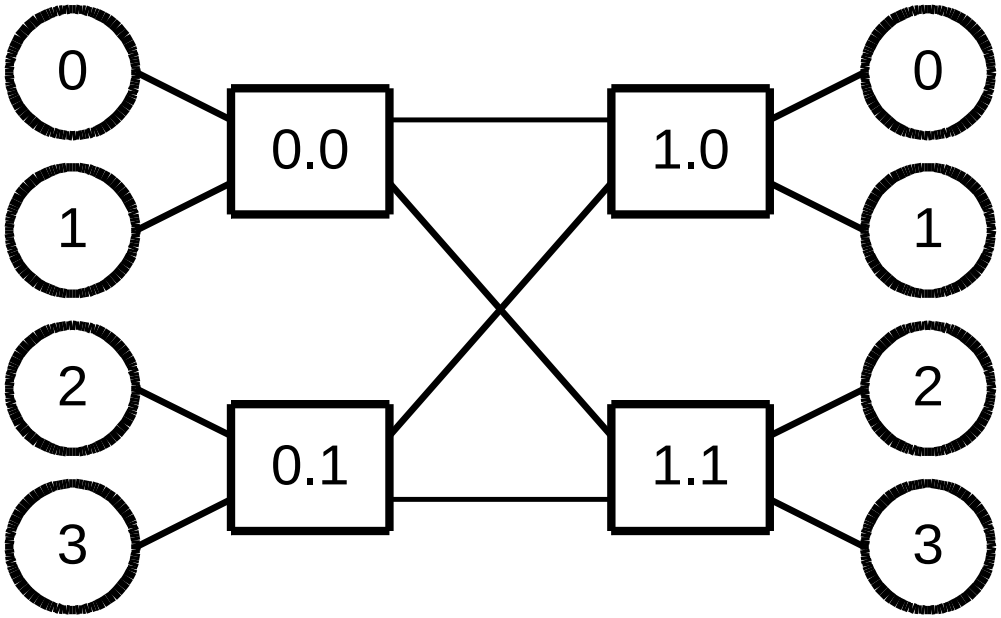
Why different topologies for these two machines?

Cray T3D, 1995

Cray Black Widow, 2007

Sample topologies



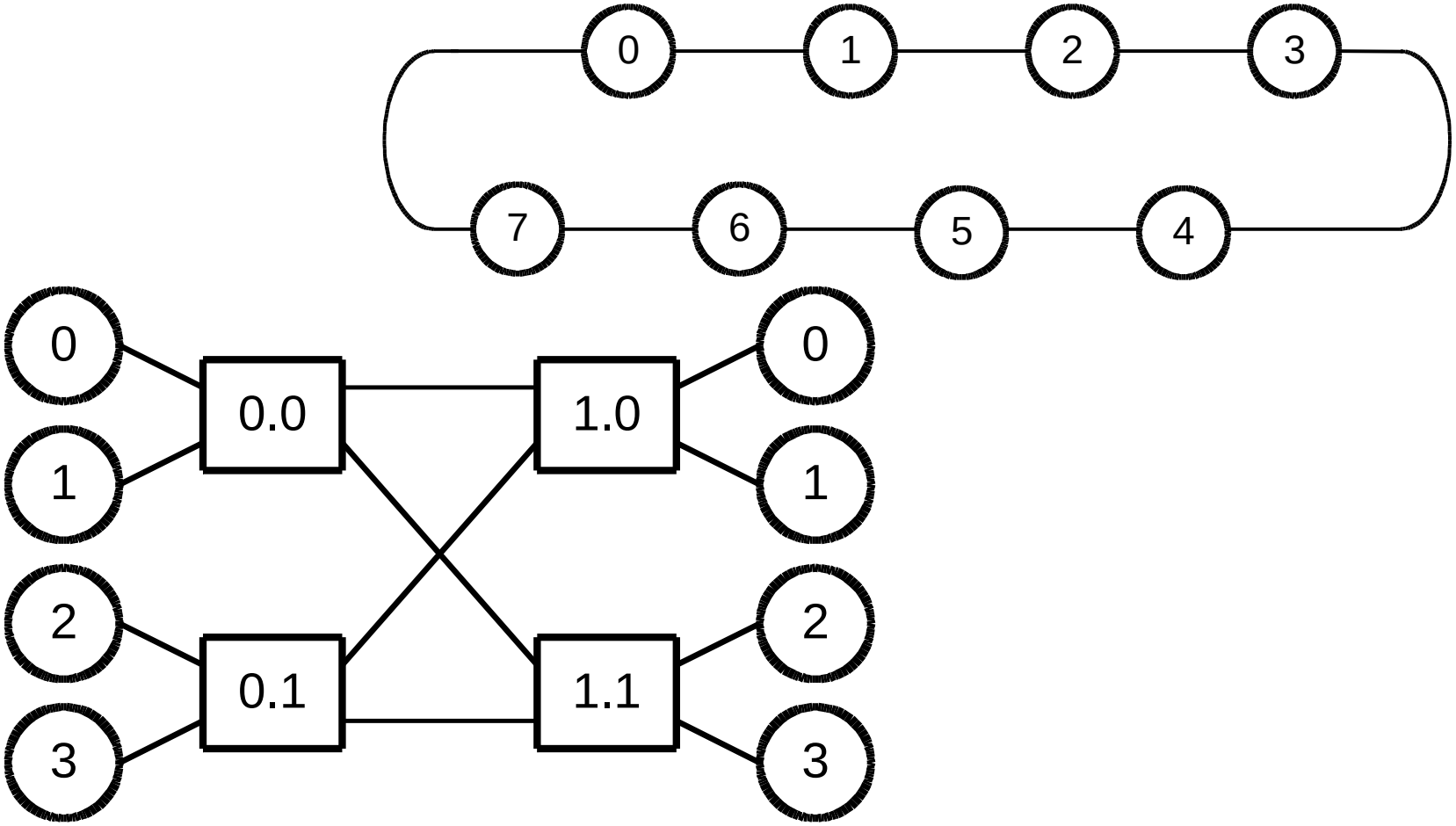


Nomenclature

- $I = (C, N^*)$
- Set of nodes N^*
  - $N \subseteq N^*$, set of terminal nodes
- Degree, δ
  - $\delta = \delta_I + \delta_O$
- For each channel, $c \in C$
  - Bandwidth, $b_c = f \times w_c$
  - Latency, $t_c = l_c/v$

Example

–What are N*, N, C, δ for these networks?

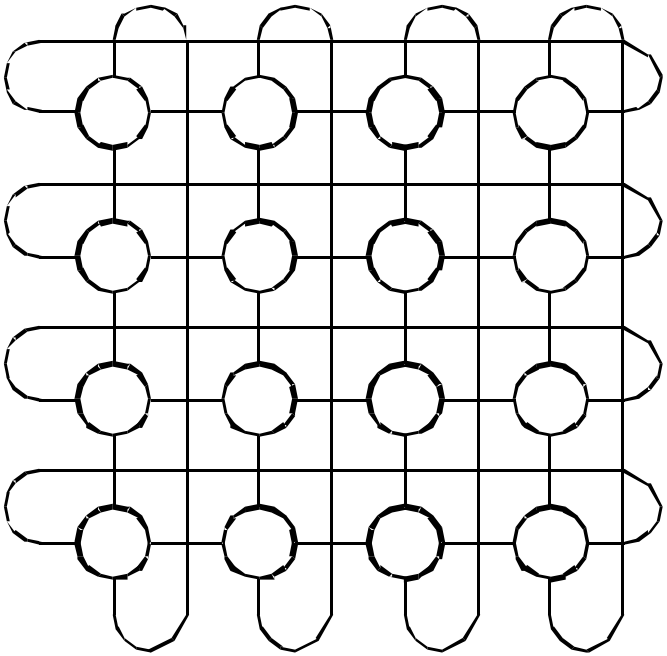


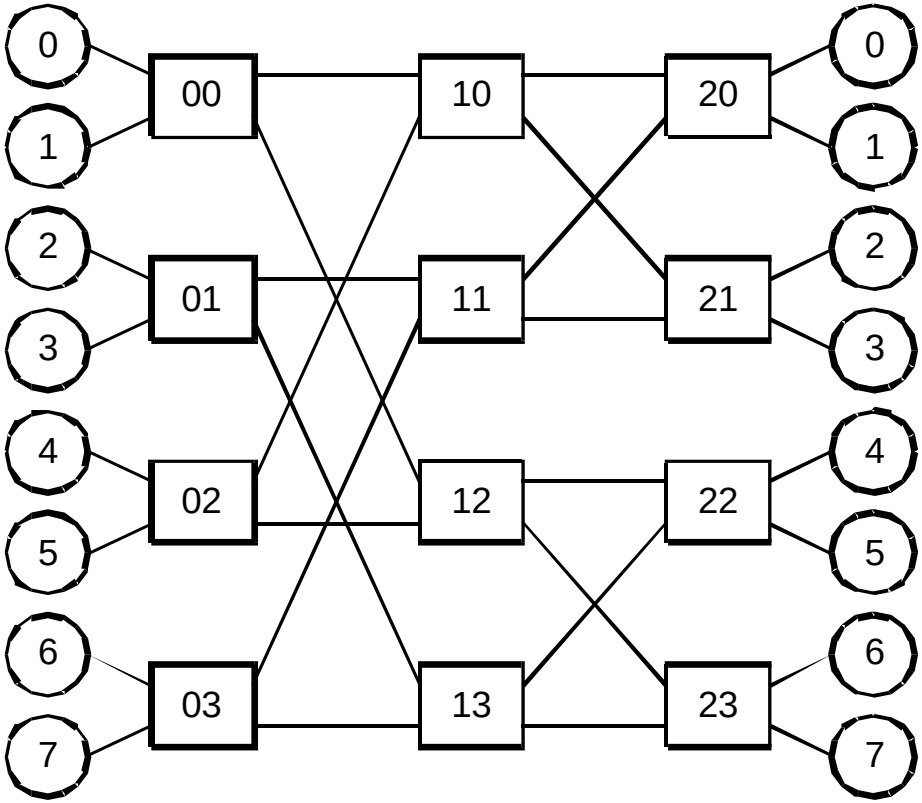
Bisection

- Cut of the network: c(N1, N2)
- Bisection is an equal cut: |N1| ~ |N2|

- Channel bisection, B_C = min |c(N1, N2)|

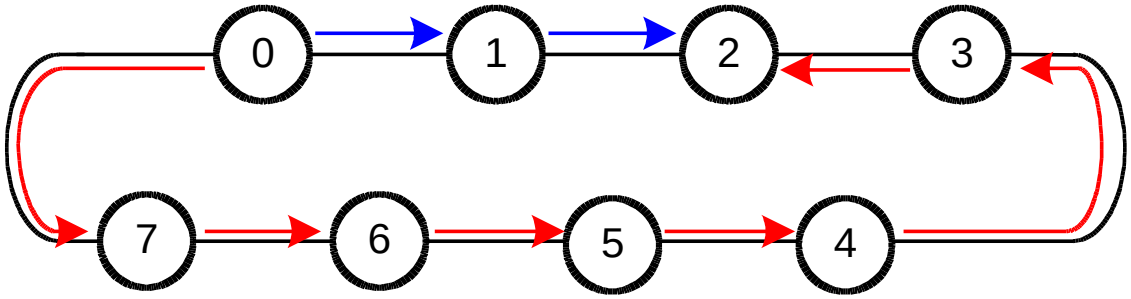
What is the Channel Bisection of these networks?





Path length

- For packets from terminal 0 to 2,
  - Path, P = { (0,1), (1,2) }
- Hop count, H = |P|



- Longest minimal path in network is called H_{max}
- For N terminals with switches of degree δ,
  - $H_{max} \geq \log_{(\delta/2)} N$

Traffic matrix

- Each entry $\lambda_{s,d}$ corresponds to fraction of traffic
  - from source *s* to destination *d*

$$\Lambda = \begin{bmatrix} \lambda_{s,d} \end{bmatrix}$$

- Admissable traffic matrices
- Permutation matrix, Π
- Random traffic

Some Example Matrices

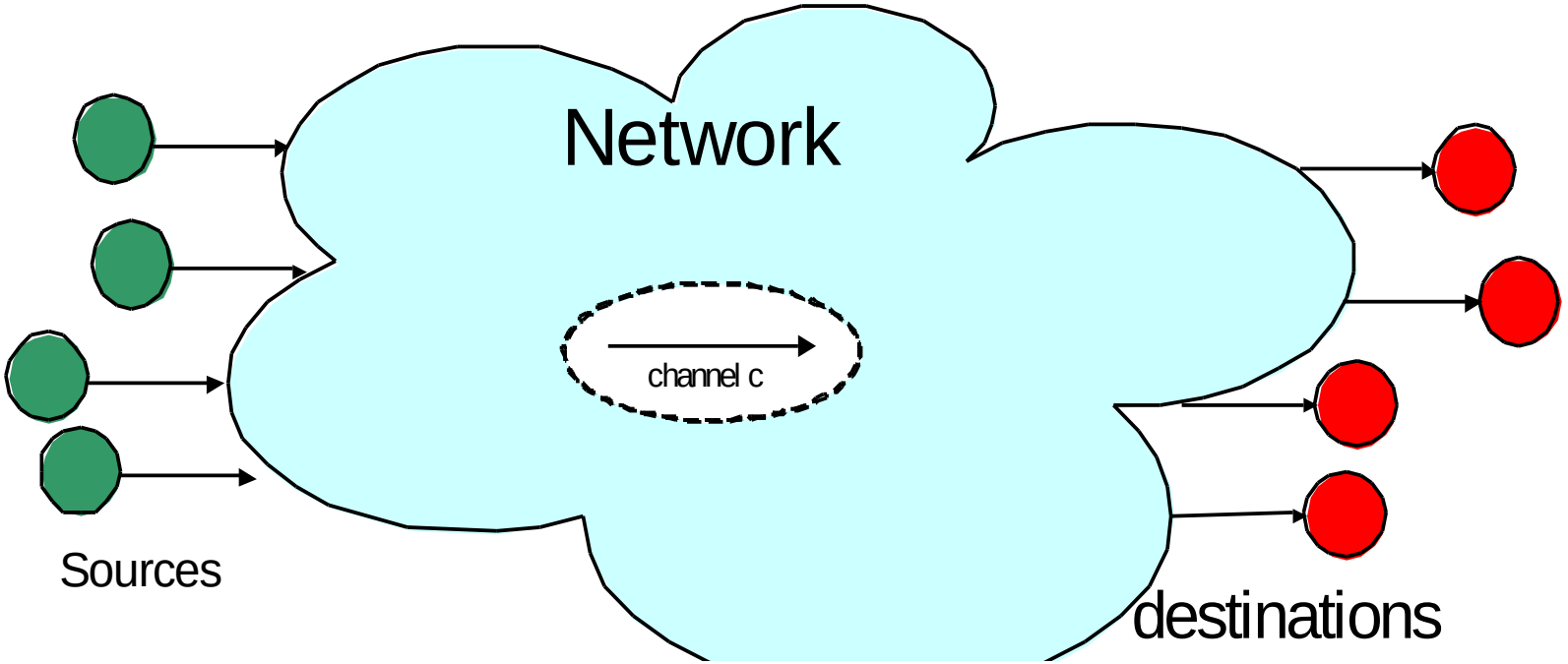
.25	.25	.25	.25
.25	.25	.25	.25
.25	.25	.25	.25
.25	.25	.25	.25

1	0	0	0
0	0	1	0
0	1	0	0
0	0	0	1

0	0	1	0
0	0	0	1
1	0	0	0
0	1	0	0

1	0	0	0
0	0	0	0
1	0	0	0
0	1	0	1

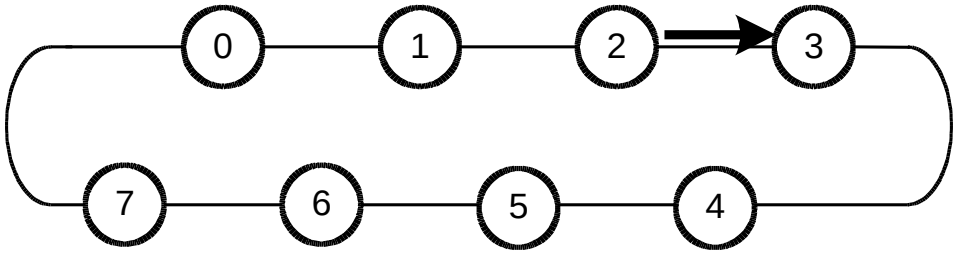
Channel load



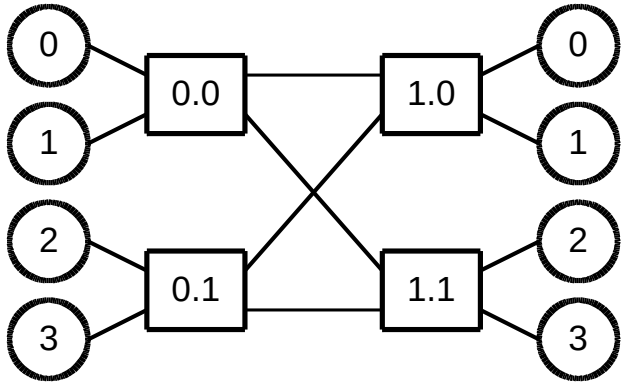
Channel load, γ_c

Channel load examples

- 8-ary 1-cube
  - Pick channel (2,3)
  - Load from
    - node 7: 1/16
    - node 0: 3/16
    - node 1: 5/16
    - node 2: 7/16
    - Total =



- Butterfly, load γ

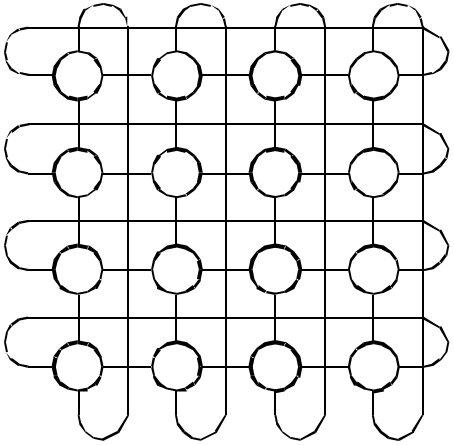


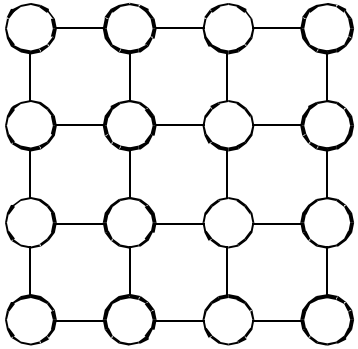
- Throughput, $\theta = b/\gamma_{max}$

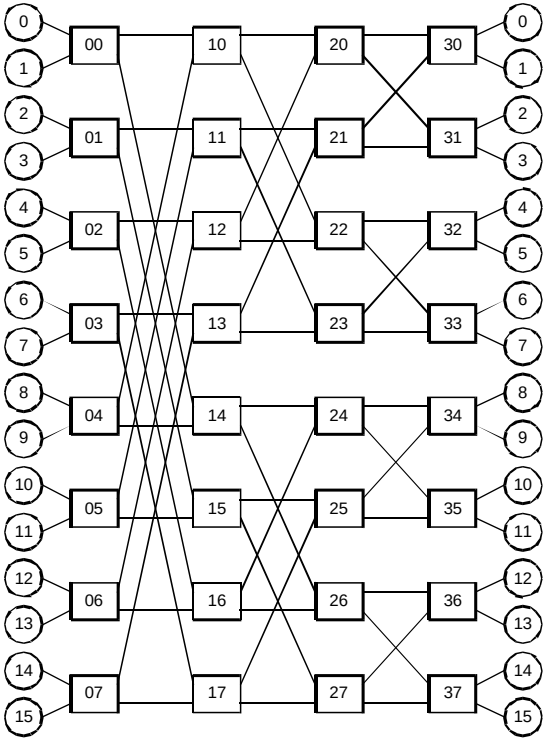
Channel Load Bounds

- Bisection:
  - $\gamma_B = N/2B_C$
- Average hopcount
  - $\gamma_{max} = NH_{avg}/C$

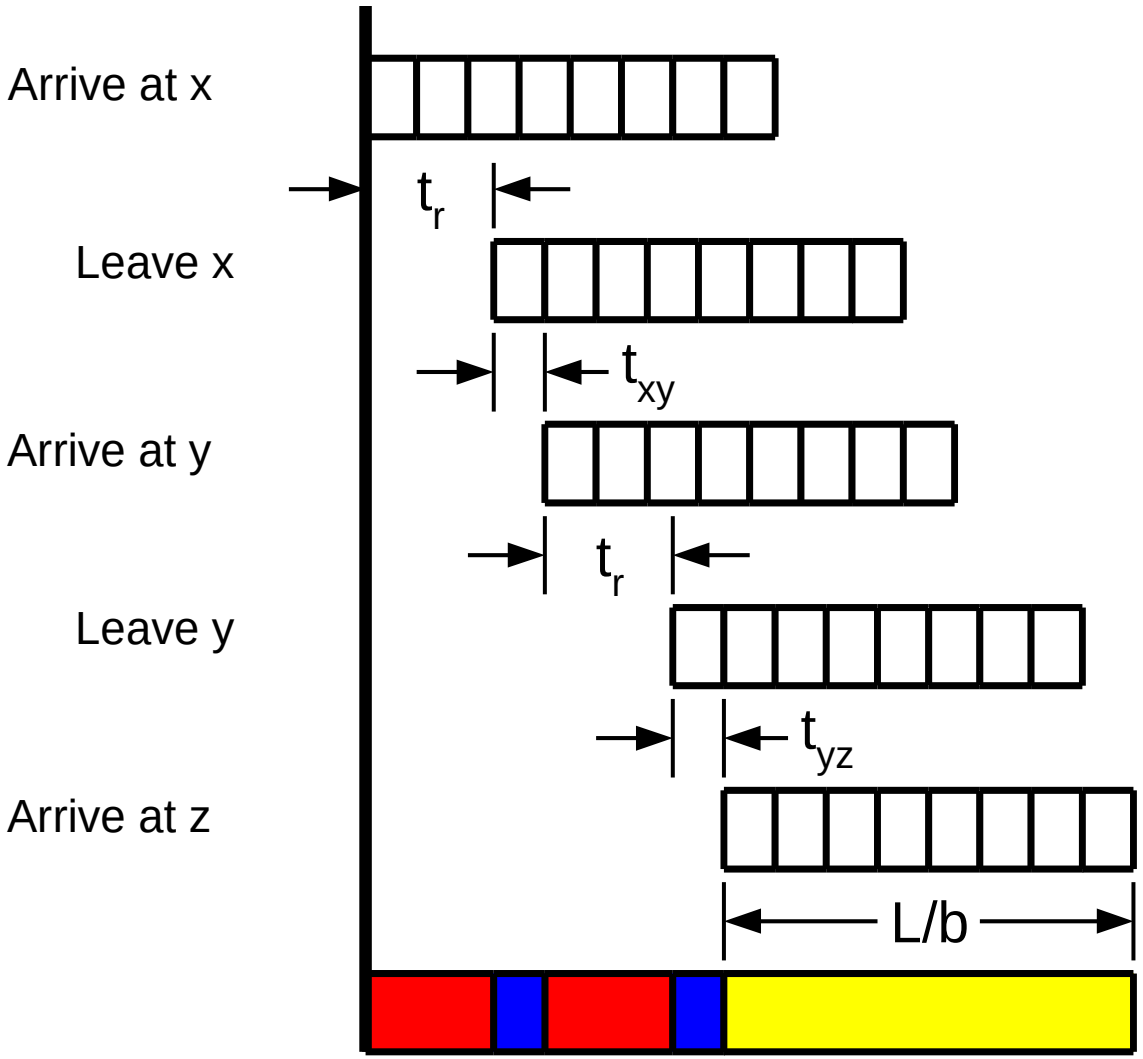
What are the channel load bounds on these networks?





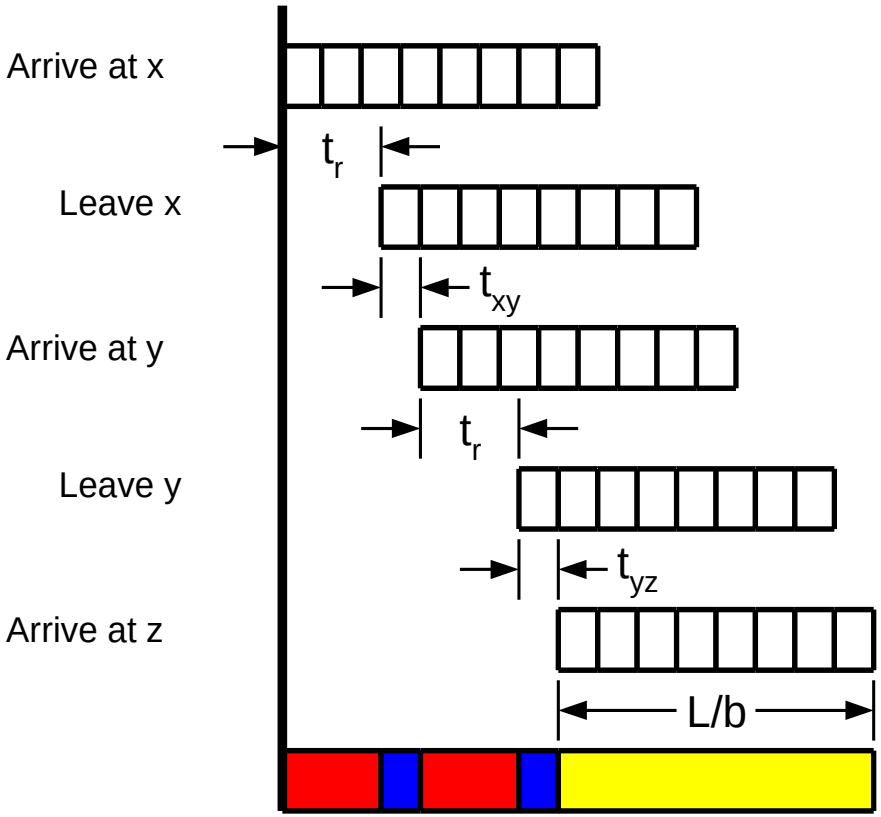


Latency


Arrive at x
t_r
Leave x
t_{xy}
Arrive at y
t_r
Leave y
t_{yz}
Arrive at z
L/b

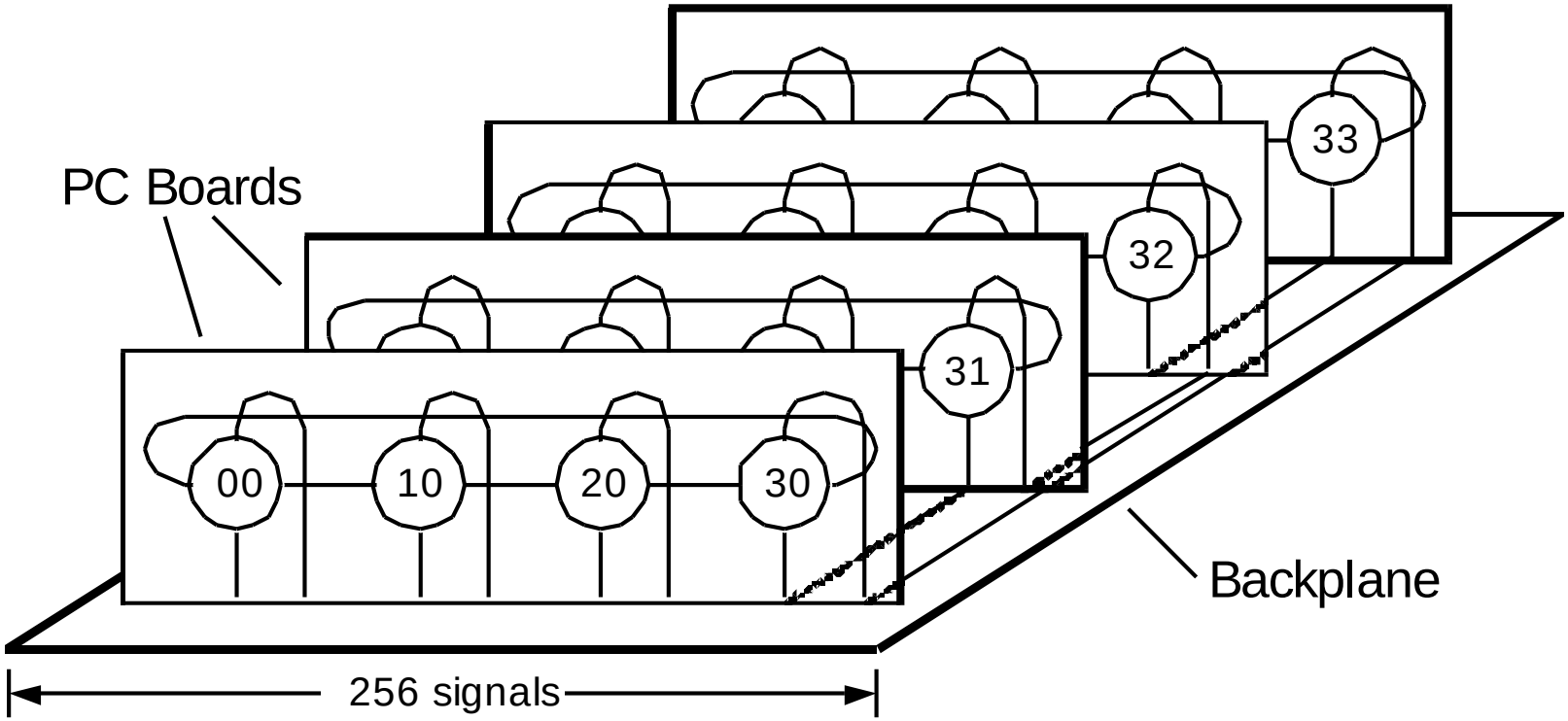
Latency

- Total latency, $T = T_h + T_s$
- Serialization latency, $T_s = L/b$
- Header Latency, $T_h = Ht_r + D/v + T_c$
- Zero-load latency, $T_0 = Ht_r + D/v + L/b$



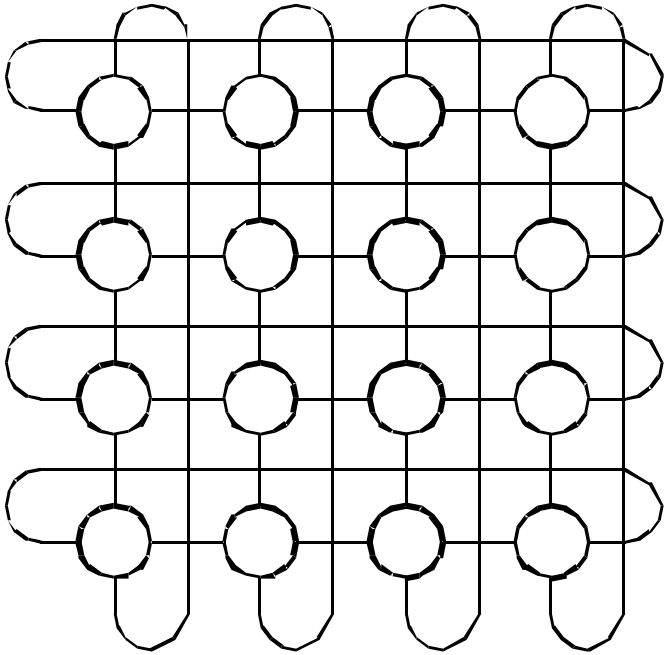
Packaging

- Bandwidth constraints applied at different levels



Packaging

- Concerned with 2 constraints
  - B_n is constraint at bottom level
  - B_s is constraint across cut of the system
  - (Middle levels may add constraints)

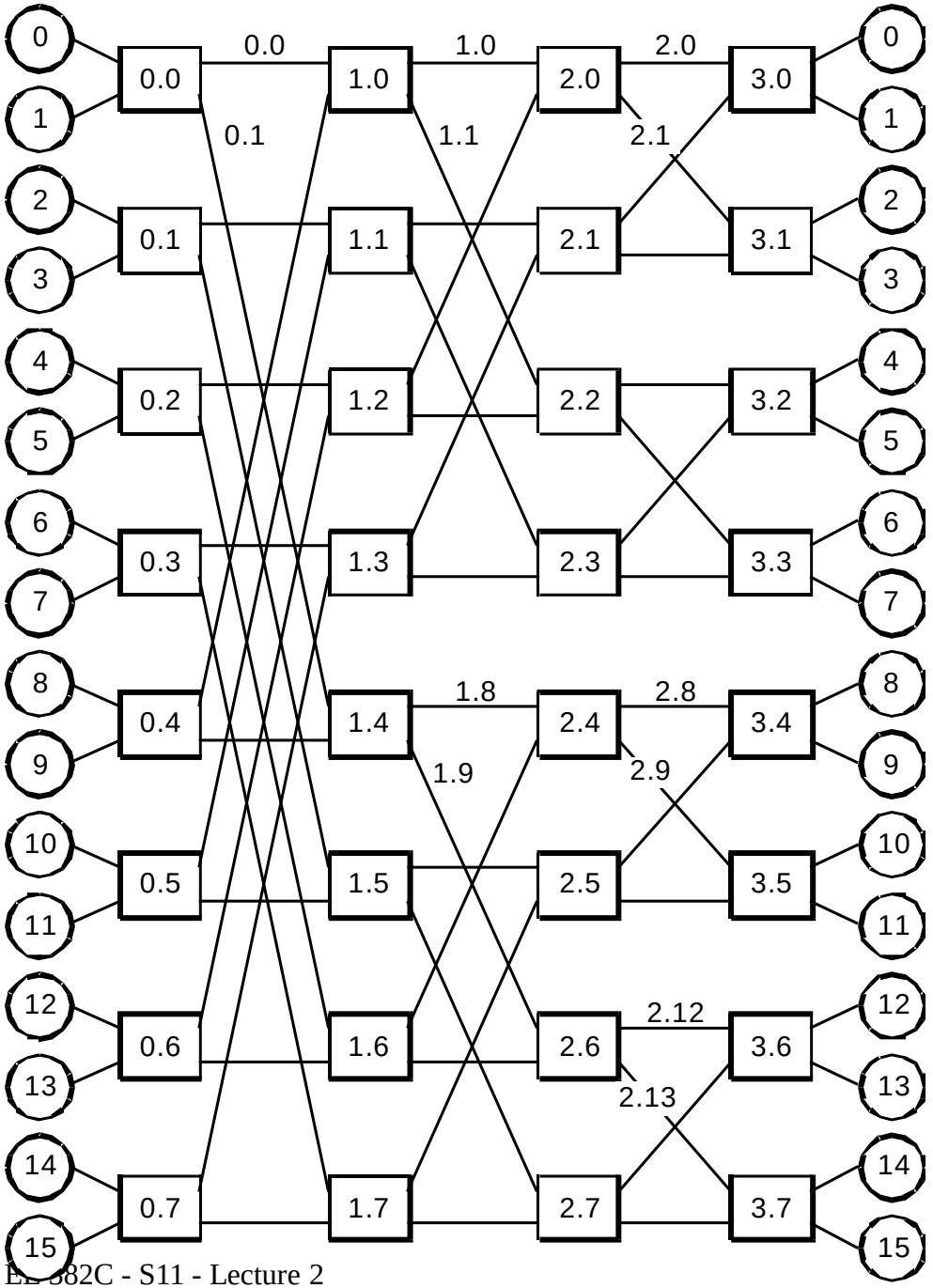


- Consider 4-ary 2-cube: δ=8, B_c=16
- Suppose B_n = 128 Gb/s and B_s = 1024 Gb/s
  - B_n : $b_c <= B_n/\delta$ = 16 Gb/s
  - B_s: $b_c <= B_s/B_C$ = 64 Gb/s
  - Network is pin-limited at the node

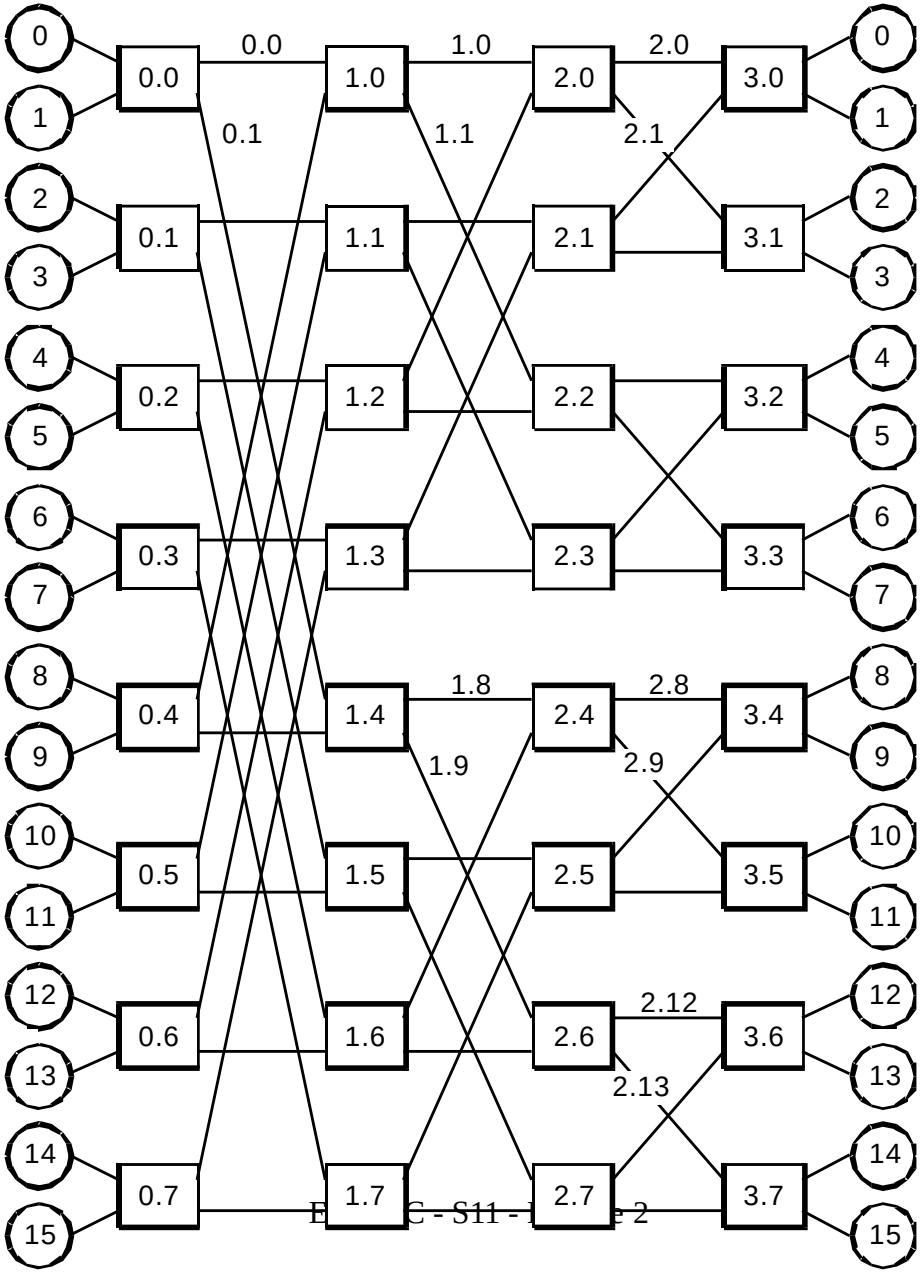
Butterfly network

- k-ary n-fly
  - k=2, n=4
  - $N = k^n$
  - $\delta = 2k$
  - $H_{max} = n+1$
  - $B_c = N/2$

- Wire from
  - $c_3c_2c_1c_0$ in stage 0 to $c_0c_2c_1c_3$ in stage 1
    - e.g., 1000 -> 0001
  - $c_3c_2c_1c_0$ in stage 1 to $c_3c_0c_1c_2$ in stage 2
    - e.g., x100->x001
  - Etc.



Destination Tag Routing



Packaging A Butterfly

- Wiring constraints
  - W_n per node
  - W_s across the system
- $W = \min(W_n/\delta, W_s/B_C) = \min(W_n/2k, 2W_s/N)$
- Would like to satisfy both at same time*
- e.g., W_n=128, W_s=1024, N=256

k	$W_n/2k$	$2W_s/N$	W	H
2	32	8	**8**	**9**
4	16	8	**8**	**5**
8	8	8	**8**	**4**
64	1	8	**1**	**3**

*Channel slicing changes this analysis

Packaging A Butterfly

- Wiring constraints
  - W_n per node
  - W_s across the system
- $W = \min(W_n/\delta, W_s/B_C) = \min(W_n/2k, 2W_s/N)$
- Would like to satisfy both at same time
- e.g., W_n=128, W_s=1024, N=256

k	$W_n/2k$	$2W_s/N$	W	H
2	32	8	**8**	**9**
4	16	8	**8**	**5**
8	8	8	**8**	**4**
64	1	8	**1**	**3**

Rotate traffic pattern

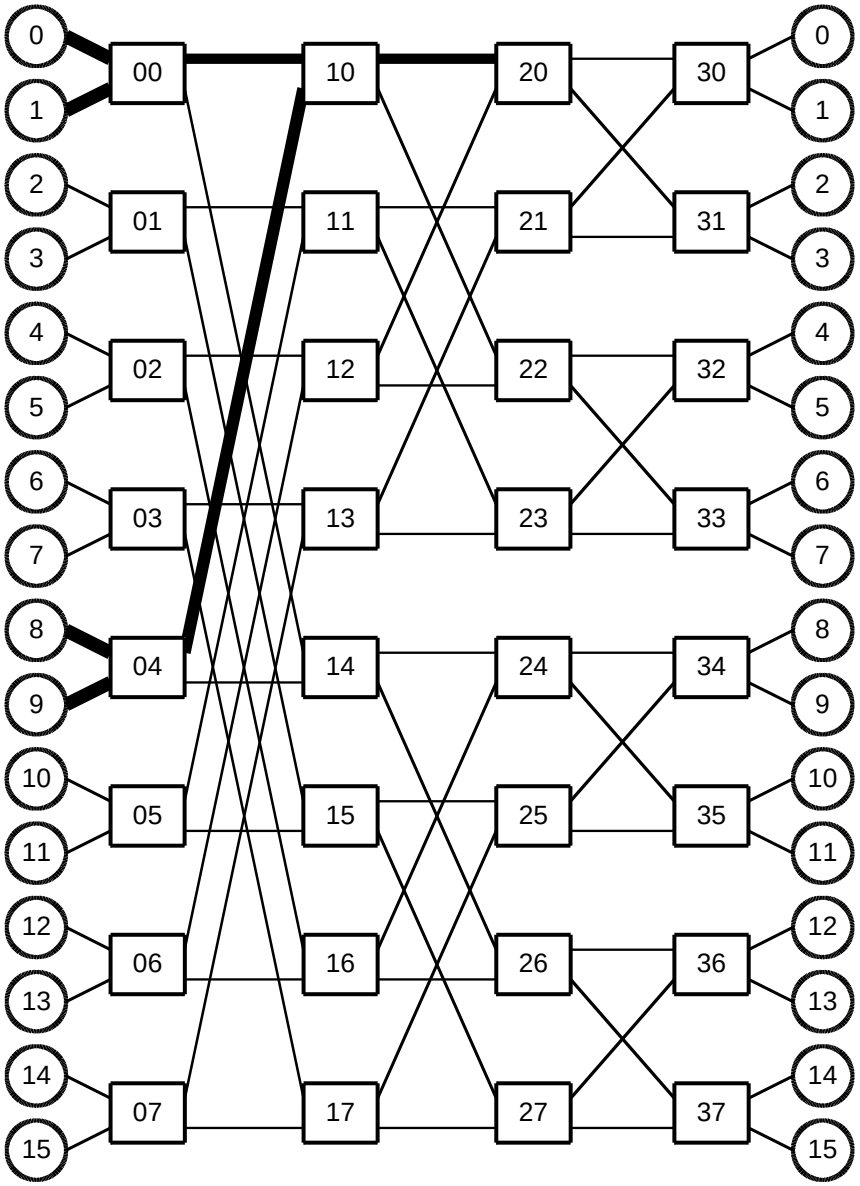
$a_3a_2a_1a_0 \rightarrow a_2a_1a_0a_3$

$0 \rightarrow 0$

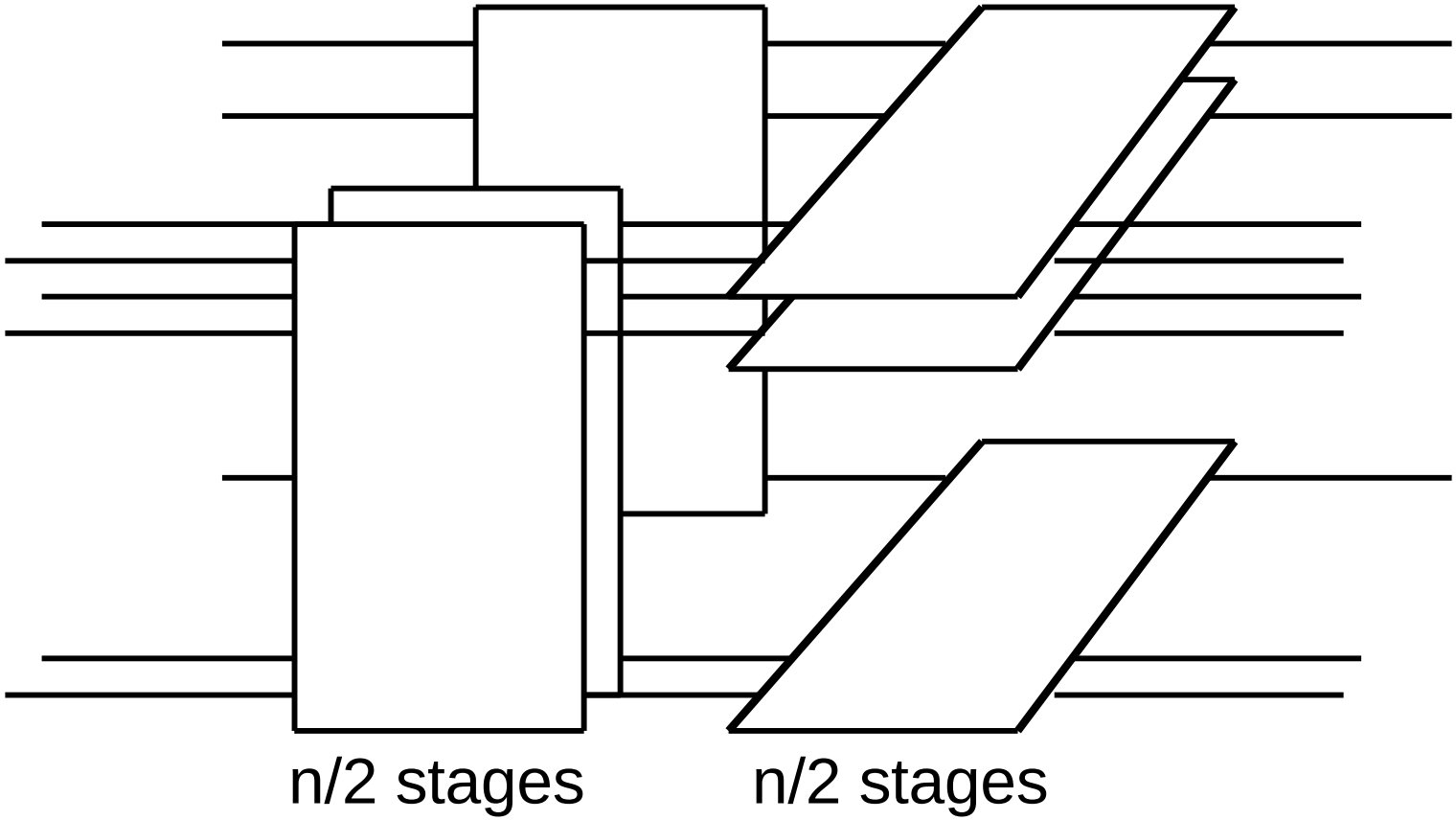
$1 \rightarrow 2$

$8 \rightarrow 1$

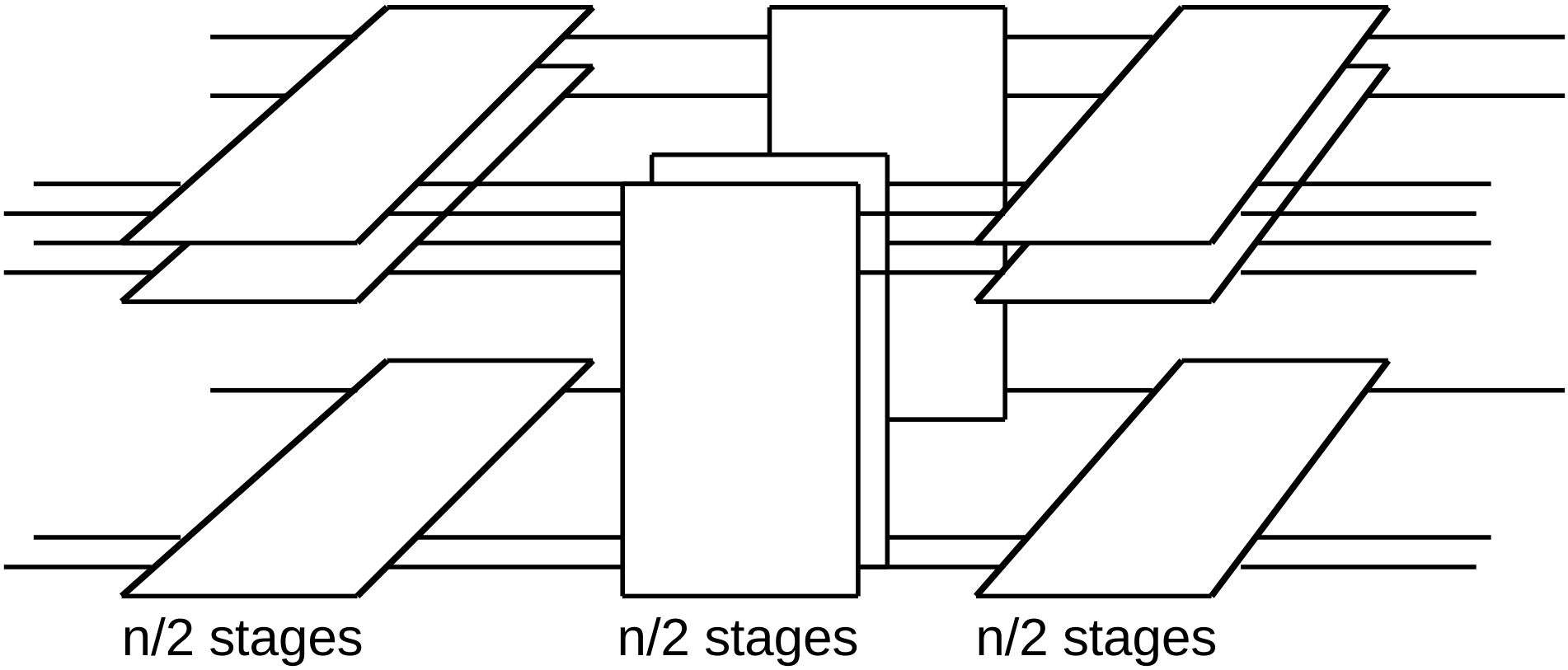
$9 \rightarrow 3$



Path diversity



Extra stages add diversity



Why different topologies for these two machines?

Cray T3D, 1995

Cray Black Widow, 2007

Summary

- Topology I(C,N*)
- Basics δ, b_c, B_c,, λ
- Channel load, γ_c
  - Definition
  - Bounds
- Traffic patterns, λ_{sd}, Λ
- Latency, $T_0 = Ht_r + D/v + L/b$
  - Trade diameter against channel width
- Packaging constraints – node and system
- Butterfly networks
  - Topology
  - Packaging
  - Load imbalance
- Next time: Torus networks, concentration and slicing