High-End Computing Systems

Fall 2011 EE380 State-of-the-Art Lecture

Hank Dietz
Professor & Hardymon Chair in Networking
Electrical & Computer Engineering Dept.
University of Kentucky
Lexington, KY 40506-0046
http://aggregate.org/hankd/

UK UNIVERSITY OF KENTUCKY

What Is A Supercomputer?

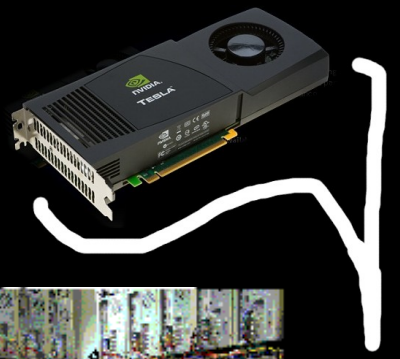
- One of the most expensive computers?
- A very fast computer?
- Really two key characteristics:
  - Computer that solves big problems...
    stuff that wouldn't fit on a PC
    stuff that would take too long to run
  - Performance can scale...
    more money buys a faster machine
- A supercomputer can be cheap!

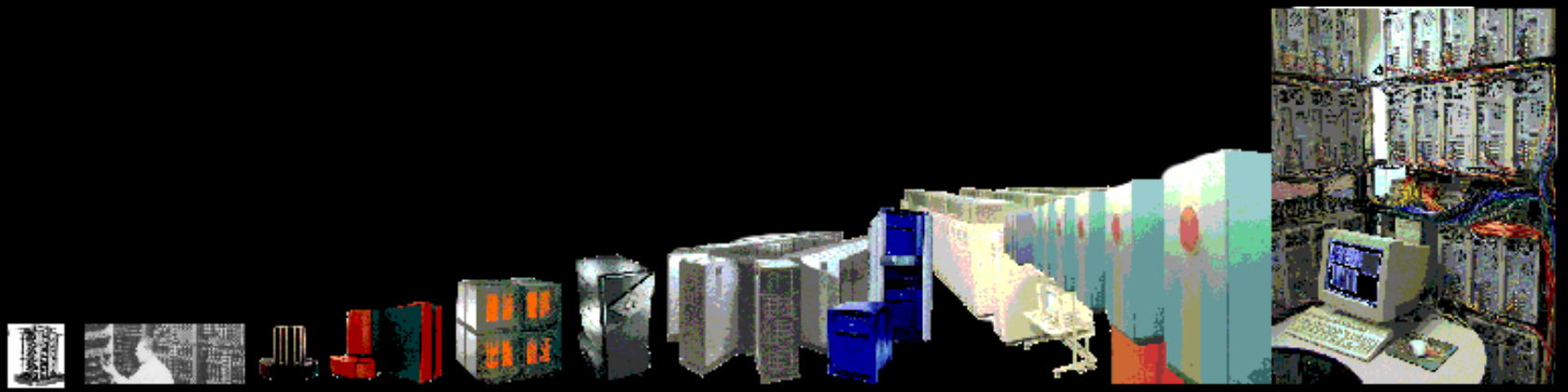
The Key Is Parallel Processing

- Process *N* "pieces" simultaneously, get up to factor of *N* speedup
- Modular hardware designs:
  - Relatively easy to scale – add modules
  - Higher availability (if not reliability)

The Evolution Of Supercomputers

- Most fit survives, even if it's ugly
- Rodents outlast dinosaurs...
  and bugs will outlast us all!





UNIVERSITY OF KENTUCKY

What Is A Cluster Supercomputer?

- Not a “traditional” supercomputer?
- Is a Cloud a cluster?
- Is The Grid a cluster?
- Is a Farm a cluster?
- A Beowulf?
- A supercomputer made from ***Interchangeable Parts*** (mostly from PCs)
- Some PC parts you don't need or want
- Often, Linux PC “nodes”

Types Of Hardware Parallelism

- Pipeline
- Superscalar, VLIW, EPIC
- SWAR (SIMD Within A Register)
- SMP (Symmetric MultiProcessor; multi-core)
- GPU (Graphics Processing Unit)
- Cluster
- Farm
- Grid
- Cloud

GPUs?

- The thing(s) on video cards
- Not really about graphics, but scalability... How many pixels do you have?
- SIMDish... but:
  - Lots of little SIMDs (low fanout)
  - Multithreading to hide memory latency
  - Various restrictions to simplify HW
- NVIDIA CUDA and OpenCL...
- GPU(s) will be on chip with SMP cores

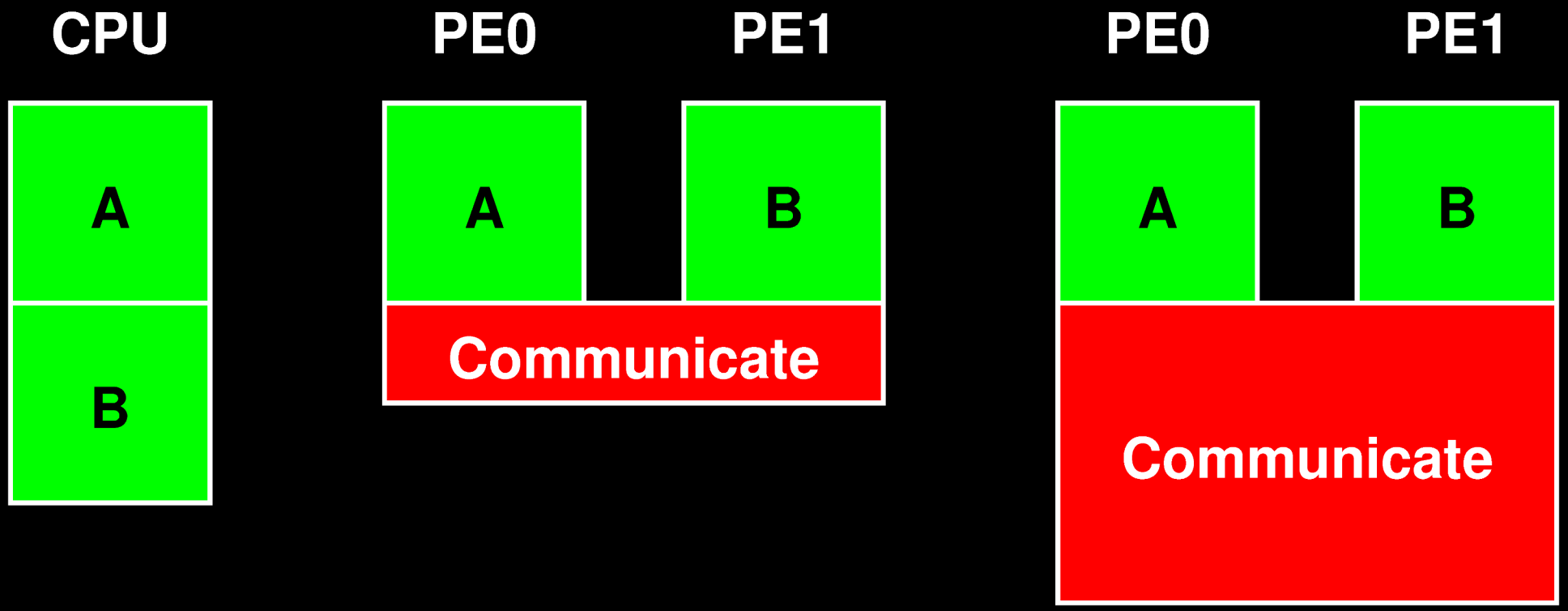
Engineering A Cluster

- This is a ***systems*** problem
- Optimize ***integrated effects*** of:
  - Computer architecture
  - Compiler optimization/parallelization
  - Operating system
  - Application program
- Payoff for good engineering ***can be*** **HUGE!**
  (penalty for bad engineering ***is*** **HUGE!**)

One Aspect: Interconnection Network

- Parallel supercomputer **nodes** interact
- **Bandwidth**
  - Bits transmitted per second
  - **Bisection Bandwidth** most important
- **Latency**
  - Time to send something here to there
  - Harder to improve than bandwidth....

Latency Determines Smallest Useful Parallel Grain Size

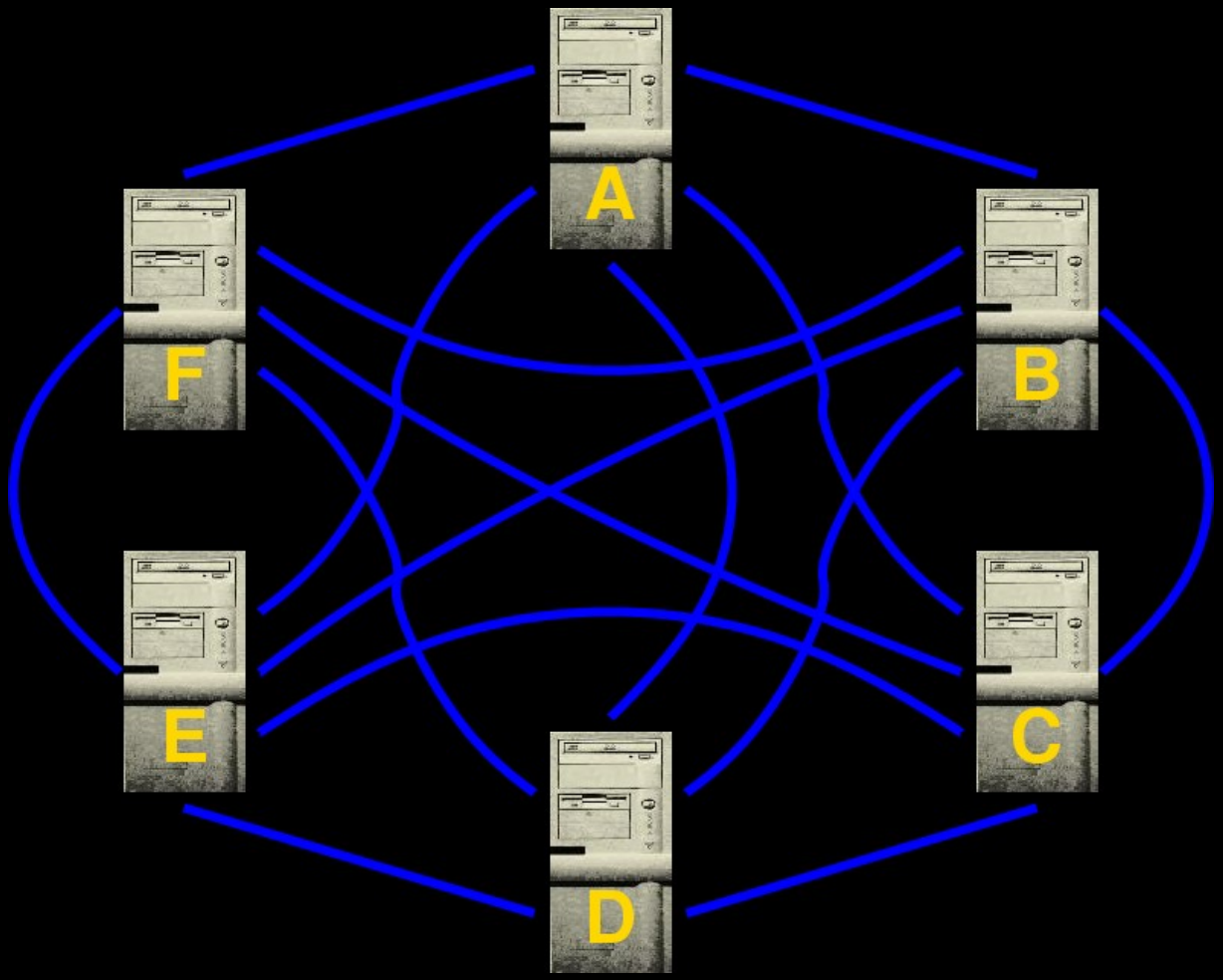


Network Design

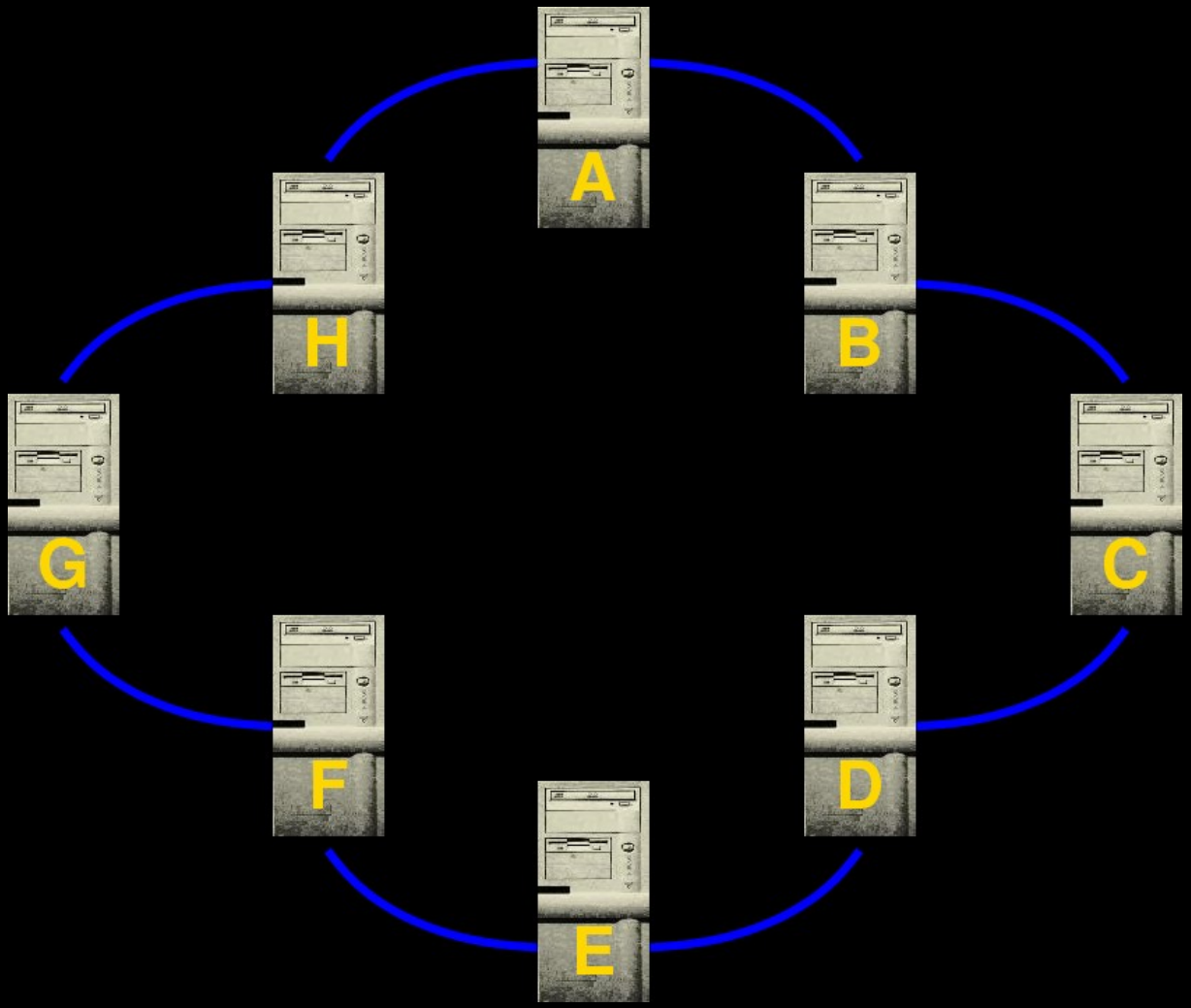
- Assumptions
  - Links are bidirectional
  - Bounded # of network interfaces/node
  - Point-to-point communications
- **Topology**
- Hardware
- Software

No Network

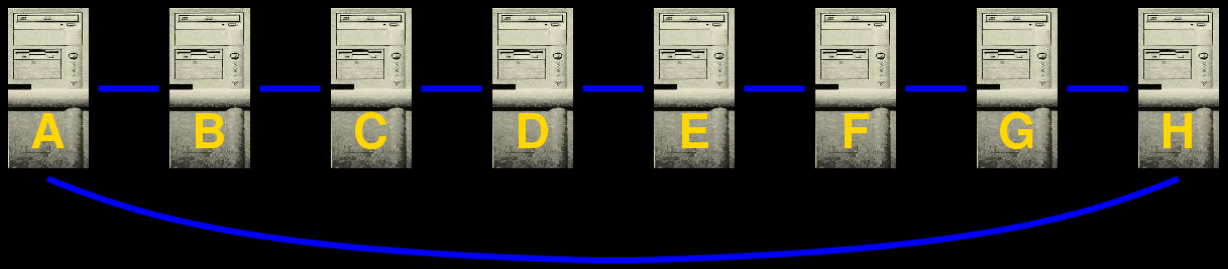
Direct Fully Connected

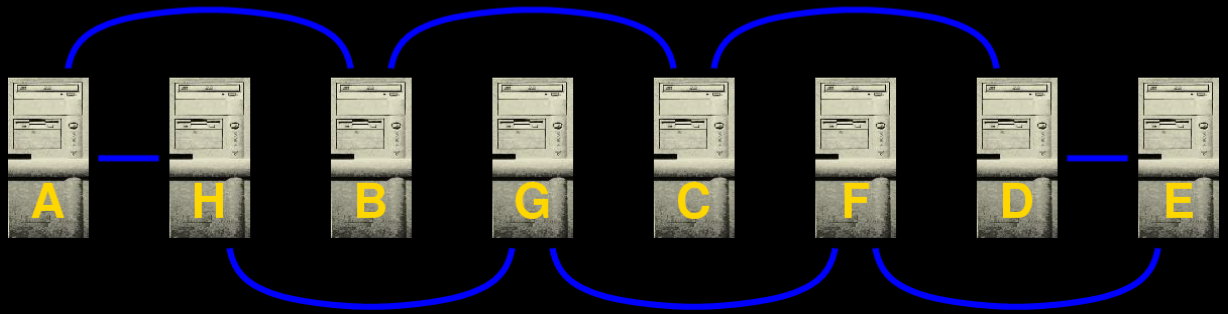


Toroidal 1D Mesh (Ring)

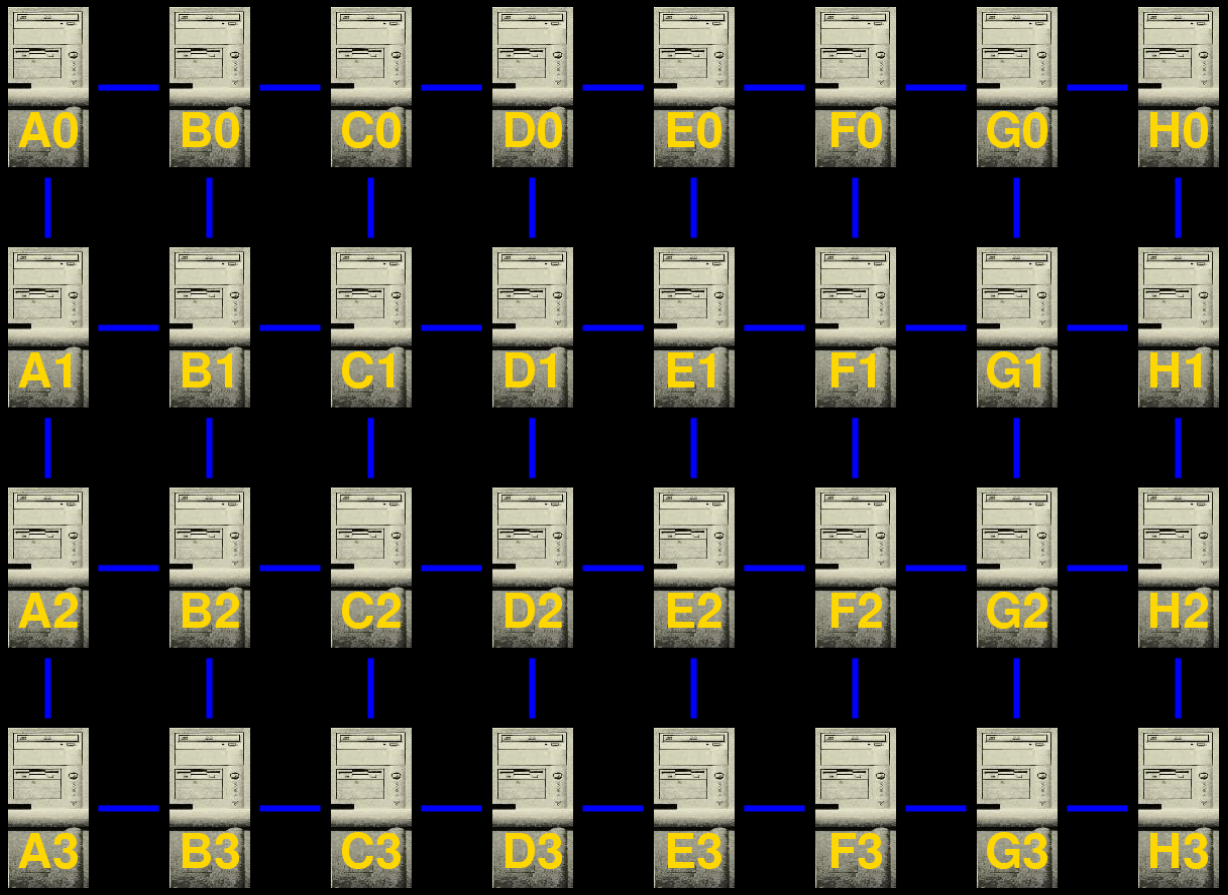


Physical Layout Of Ring

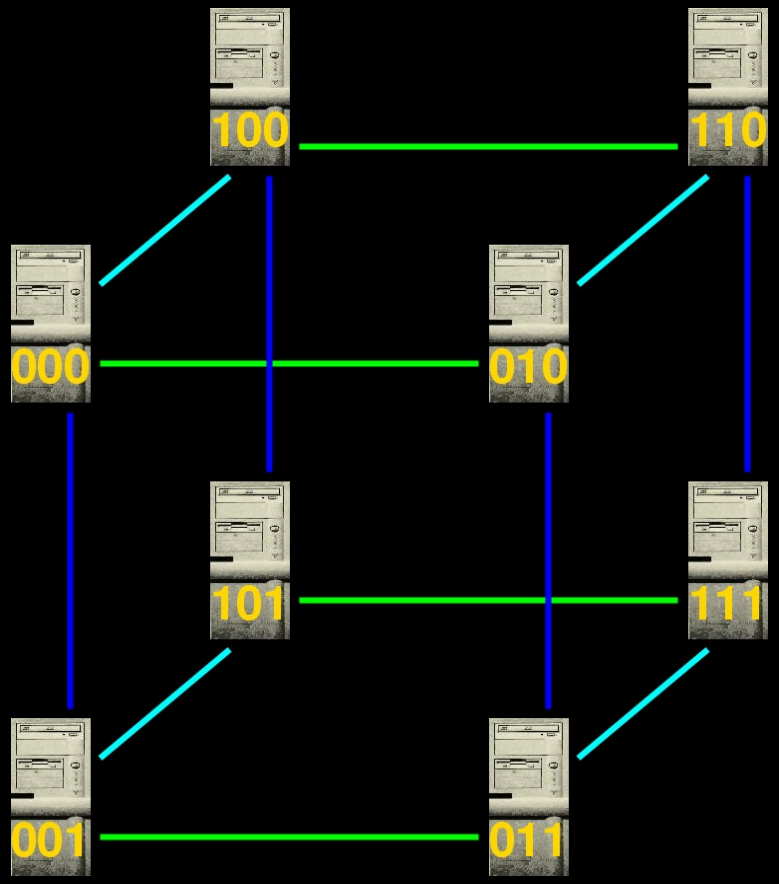




Non-Toroidal 2D Mesh



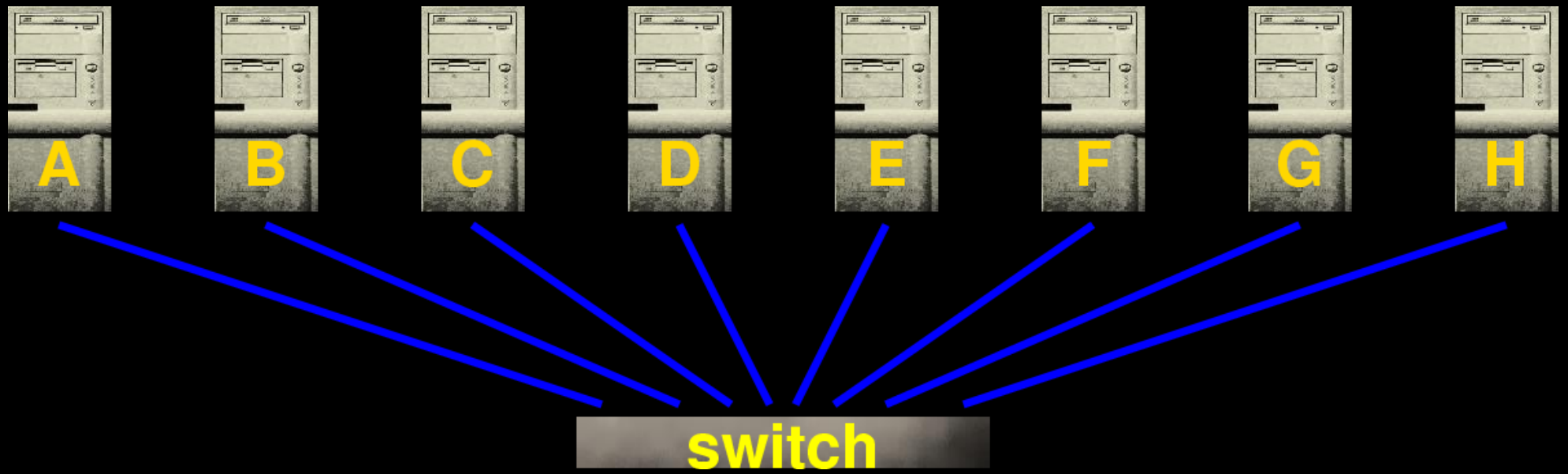
3-Cube (AKA 3D Mesh)



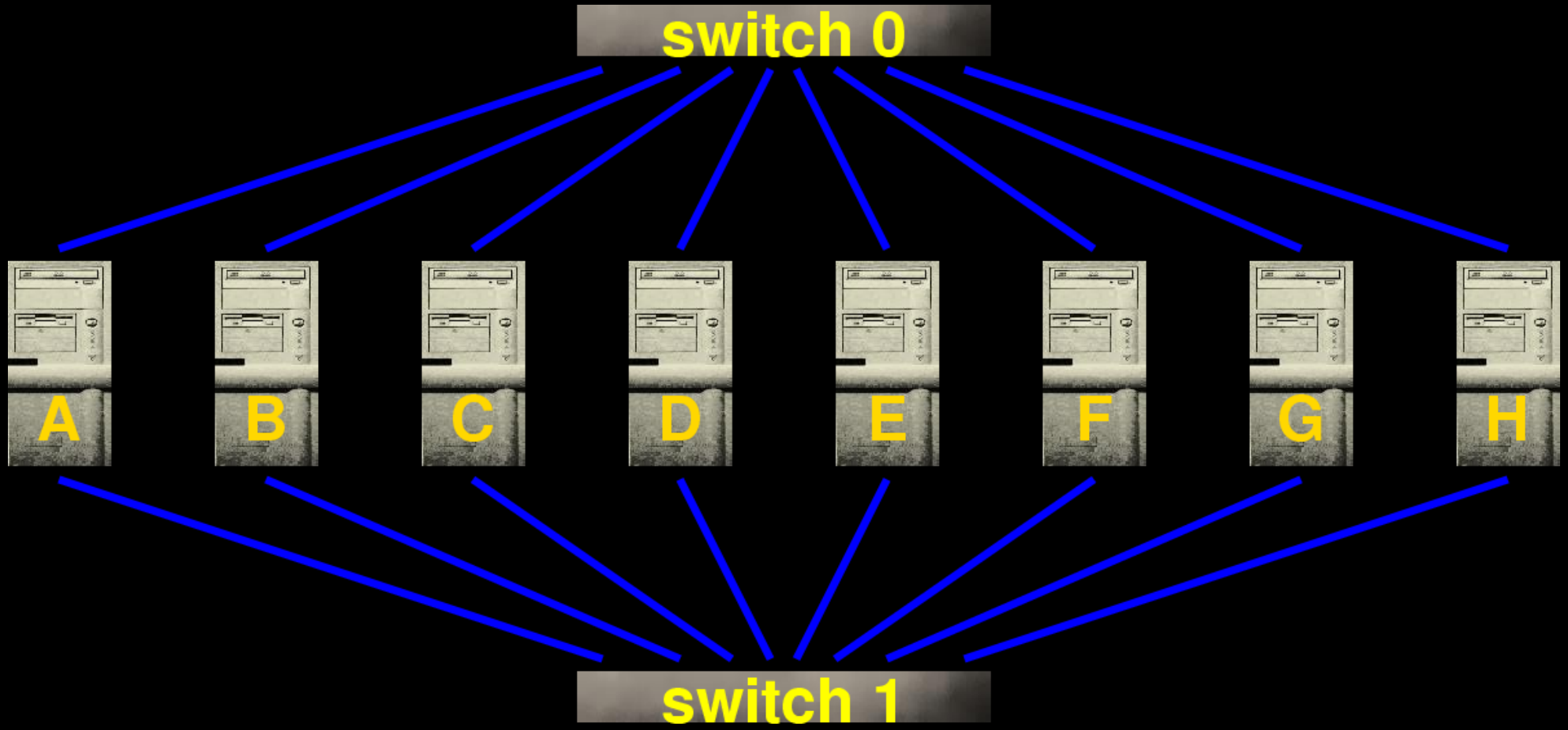
Switch Networks

- Ideal switch connects *N* things such that:
  - Bisection bandwidth = # ports
  - Latency is low (~30us for Ethernet)
- Other switch-like units:
  - Hubs, FDRs (Full Duplex Repeaters)
  - Managed Switches, Routers
- Not enough ports, build a Switch Fabric

Simple Switch (8-Port)

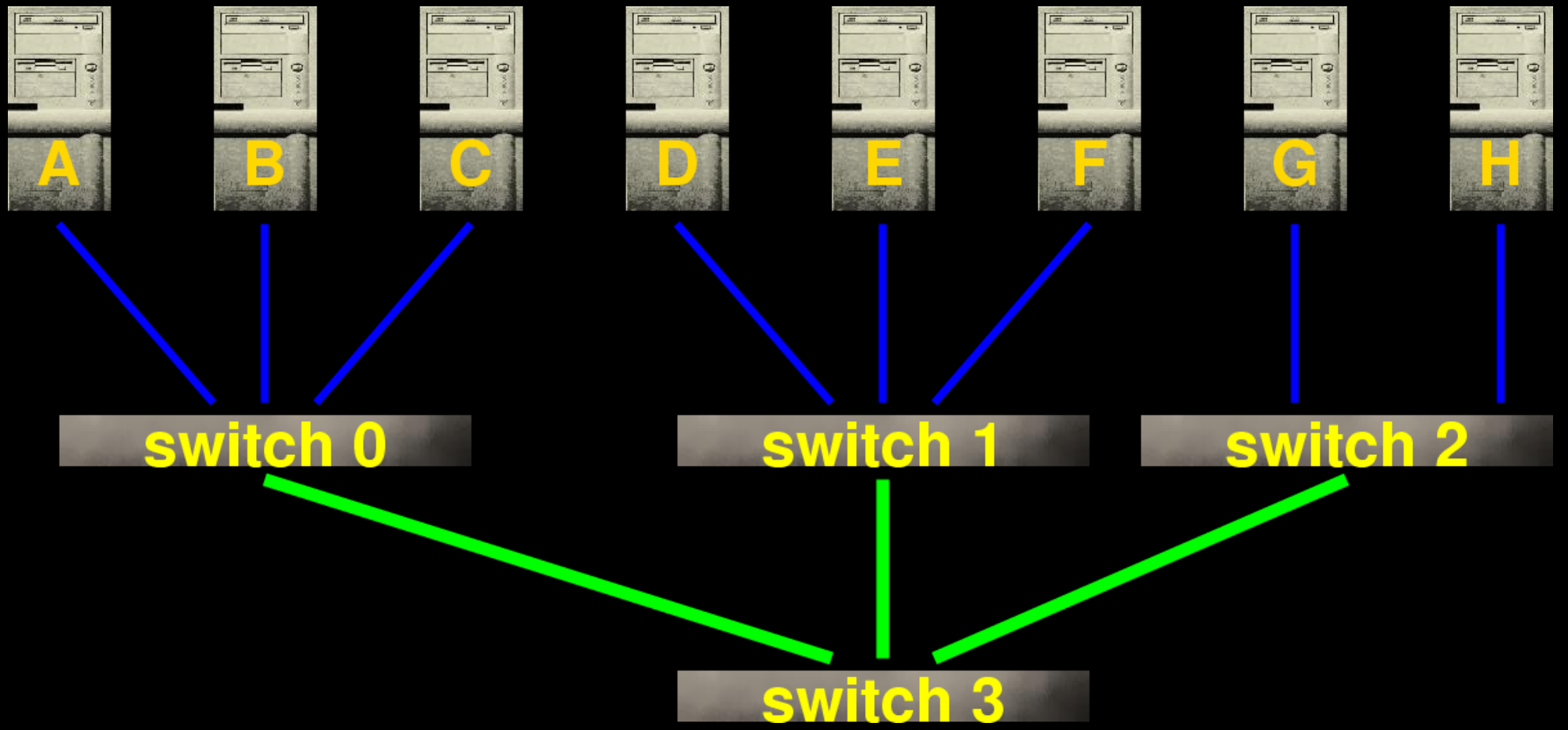


2-Way Channel Bonding

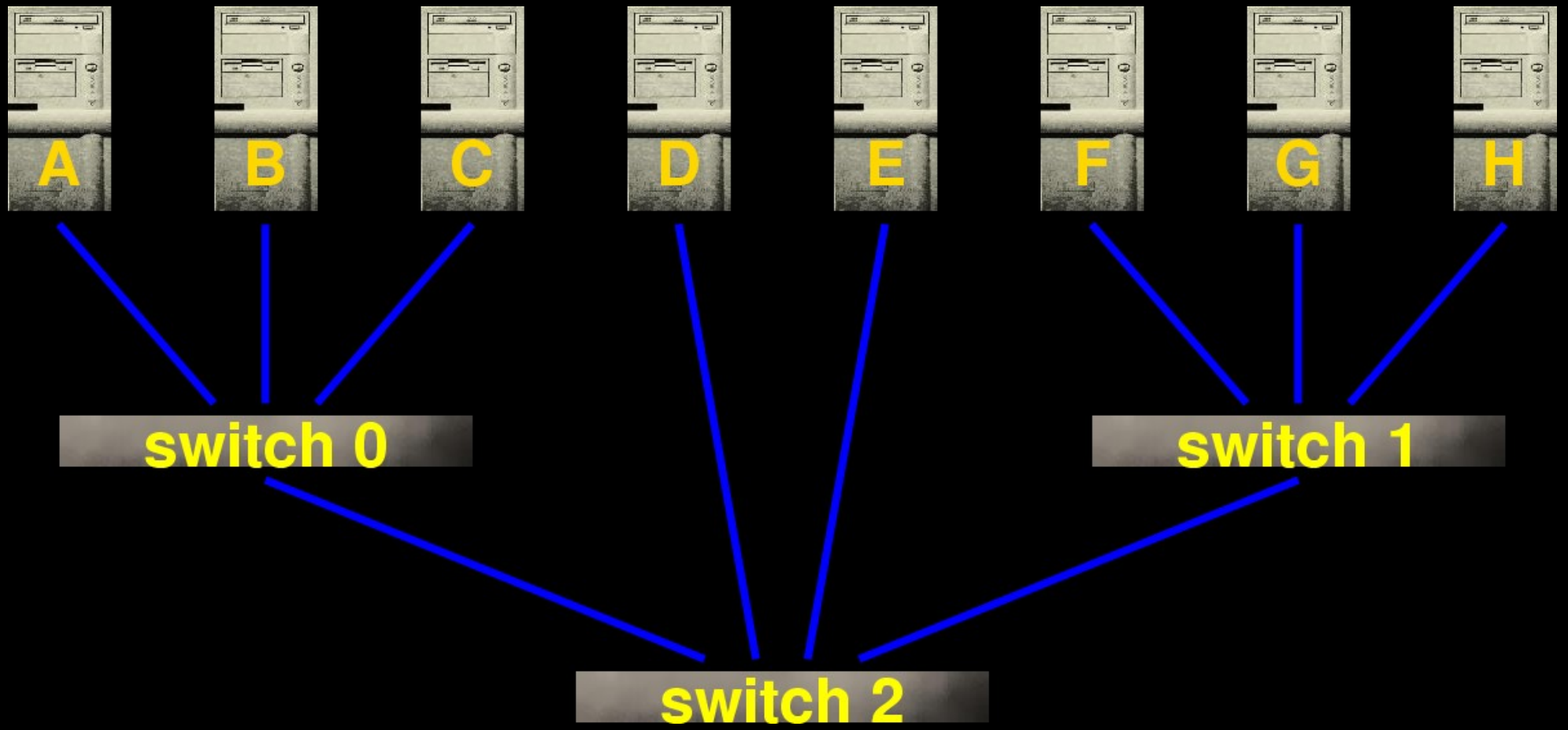


UK UNIVERSITY OF KENTUCKY

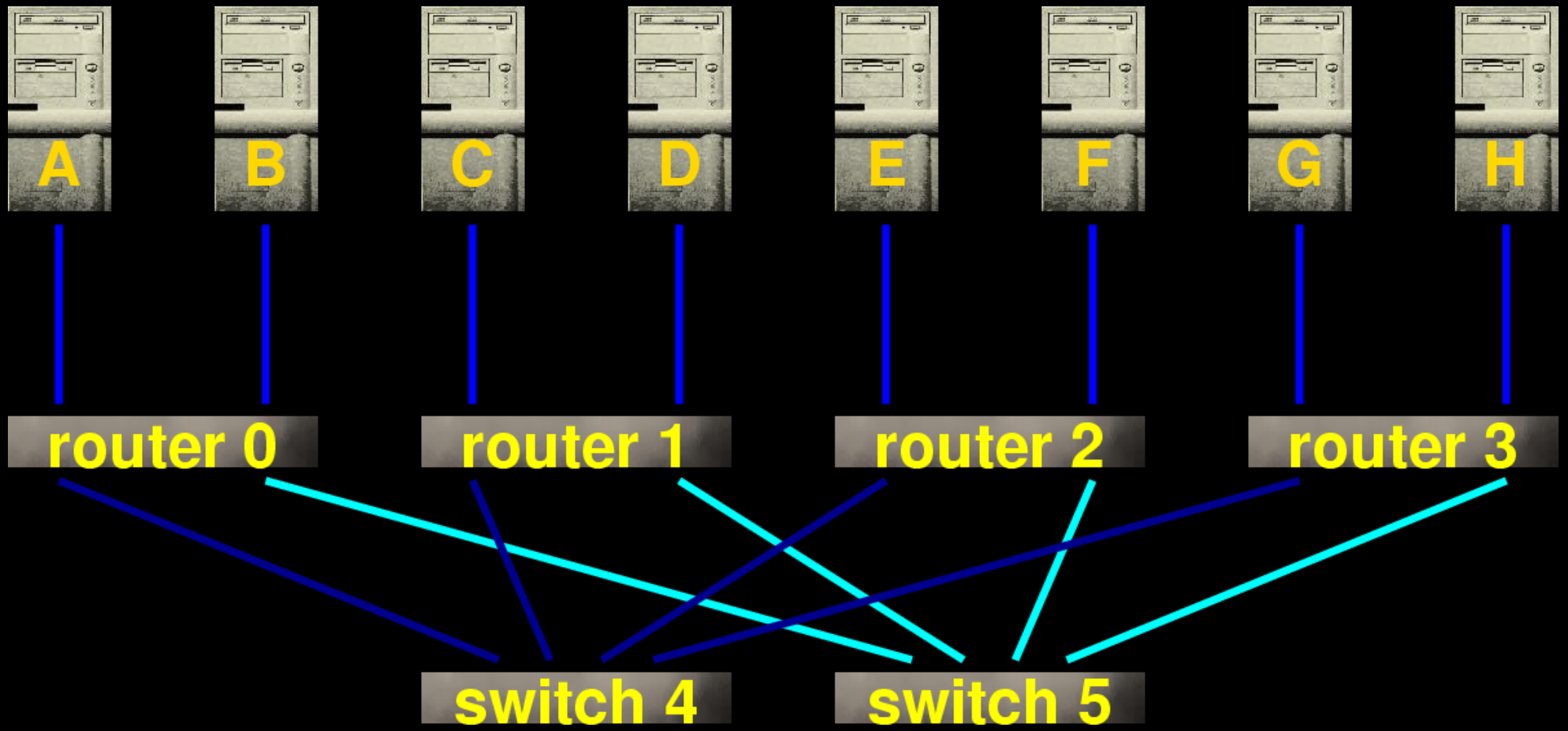
Tree (4-Port Switches)



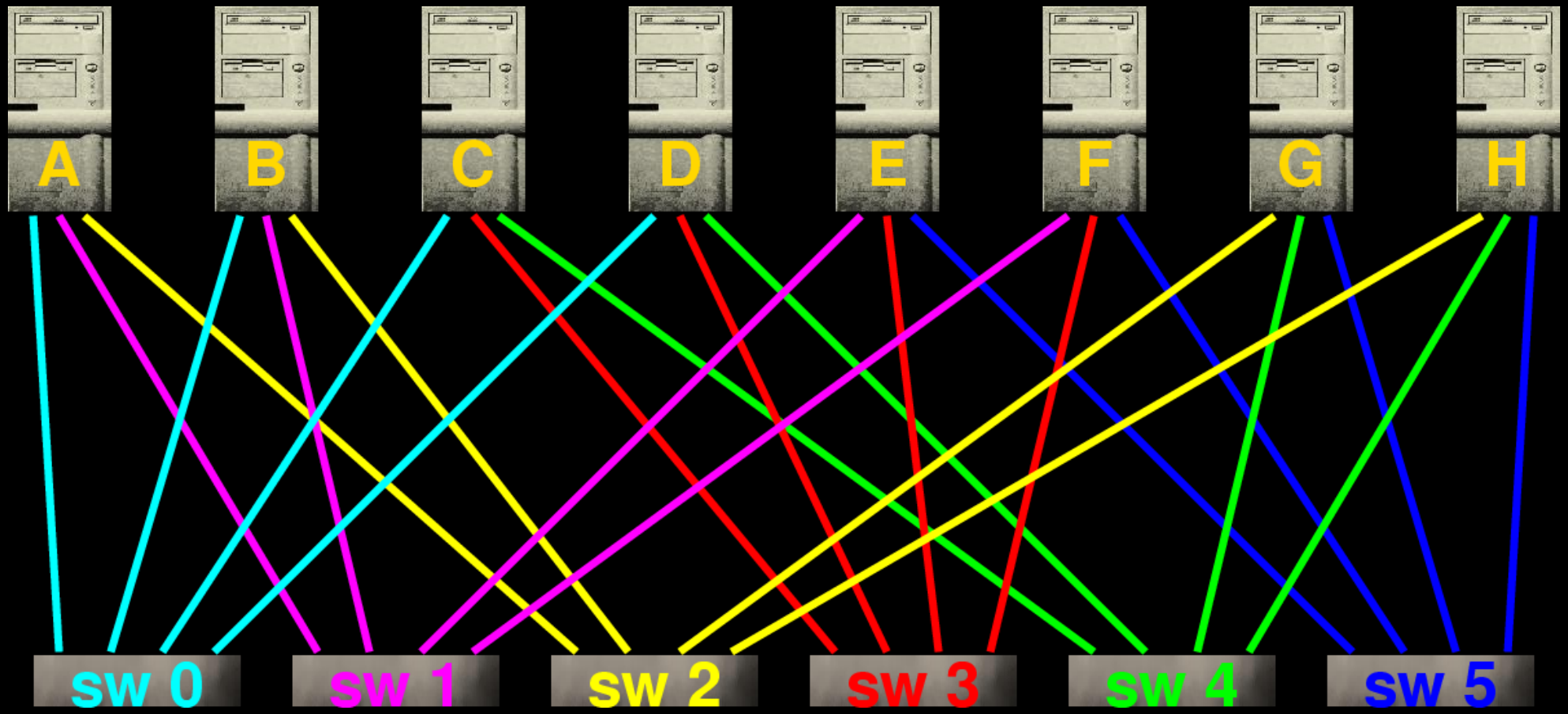
A Better Tree



Fat Tree


A
B
C
D
E
F
G
H
router 0
router 1
router 2
router 3
switch 4
switch 5

Flat Neighborhood Network… from UK!



UK UNIVERSITY OF KENTUCKY

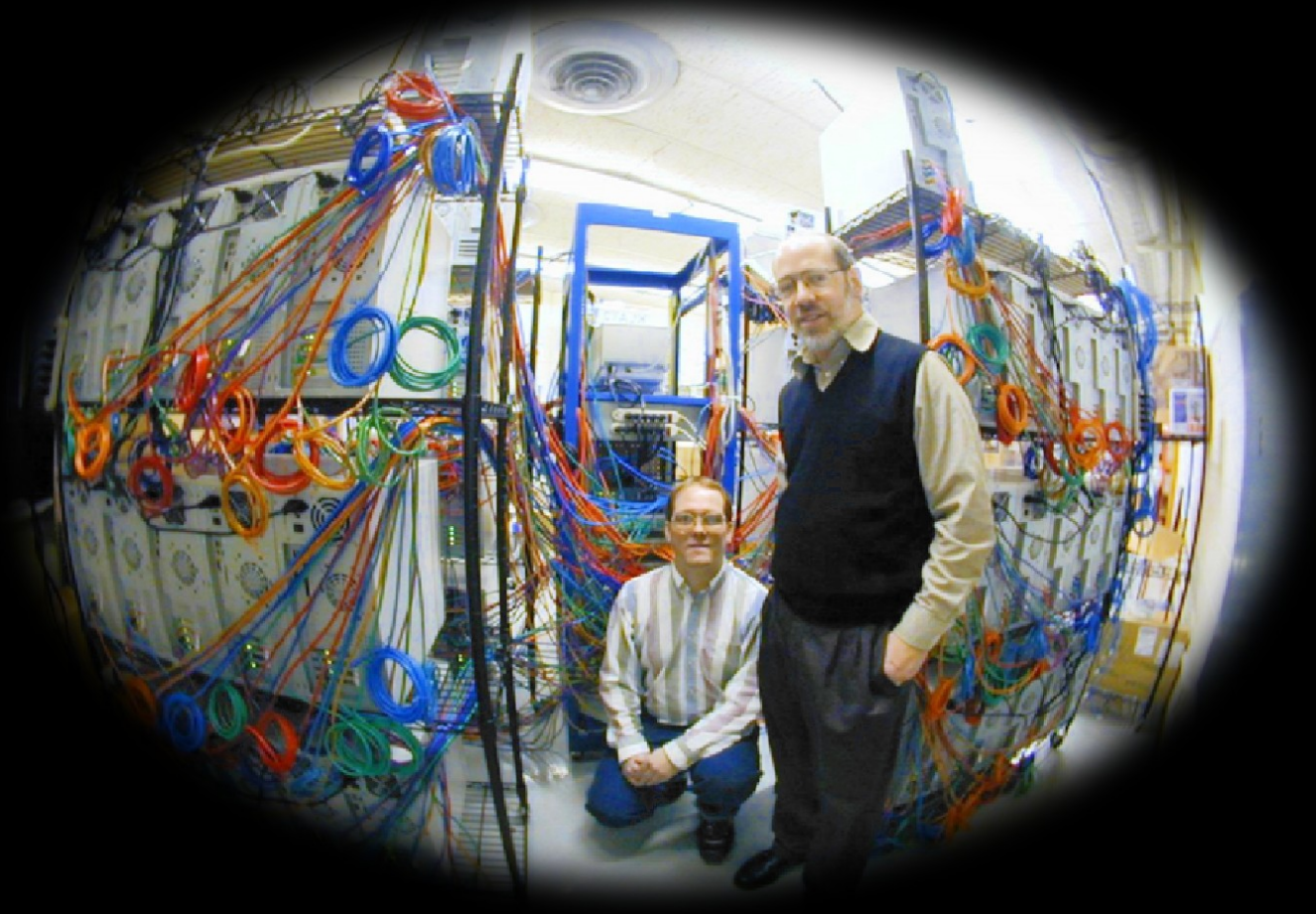
Flat Vs. Fat

- Latency:
  - 8 node, 4 port: 1.0 vs. 2.7 switch delays
  - 64 node, 32 port: 1.0 vs. 2.5
- Pairwise bisection bandwidth:
  - 8 node, 4port: 1.29 vs. 1.0 units
  - 64 node, 32 port: 1.48 vs. 1.0
- Cost: more interfaces vs. smart routers
- Summary: **Flat Neighborhood wins!**

KLAT2, Gort, & Klaatu

UK UNIVERSITY OF KENTUCKY

Behind KLAT2



UK UNIVERSITY OF KENTUCKY

KLAT2 Changed Everything

- KLAT2 (Kentucky Linux Athlon Testbed 2):
  - 1st network designed by computer
  - 1st network deliberately asymmetric
  - 1st supercomputer under $1K/GFLOPS
- 160+ news stories about KLAT2
- Various awards:
  - 2000 Gordon Bell (price/performance)
  - 2001 Computerworld Smithsonian, among 6 Its most advancing science

UK UNIVERSITY OF KENTUCKY

June 2003, KASY0

UNIVERSITY OF KENTUCKY

KASY0

- 128-node system using 24-port switches!
- KASY0 (Kentucky ASYmmetric zero):
  - 1st Sparse FNN
  - 1st physical layout optimized by GA
  - 1st TFLOPS-capable computer in KY
  - 1st under $100/GFLOPS
  - World record fastest POVRay 3.5

April 2010, NAK

UK UNIVERSITY OF KENTUCKY

NAK

- 9TFLOPS GPU-centric FNN cluster
- Appears as 64 shared memory nodes, each containing between 512-2,048 PEs
- NAK (NVIDIA Athlon cluster in Kentucky):
  - 1st *"nothing but power & ground"*
  - Relatively energy efficient
  - Arguably around $0.65/GFLOPS
  - Various SW environment claims...

MIMD On GPU (MOG)

- NVIDIA CUDA and OpenCL...
  not portable, no code base
- Our idea:
  Execute MIMD code (written for SMP, cluster) on SIMDish GPUs
  - MIMD Interpreter
  - Meta-State Conversion to SIMD code
- Alpha release of MOG was Nov. 2010

Supercomputers R Us

- We make supercomputing cheap!
- You can help...
  - Build parties
  - Weekly research group meetings
  - Projects
- Everything's at...

UK UNIVERSITY OF KENTUCKY


Aggregate.Org
UNBRIDLED COMPUTING