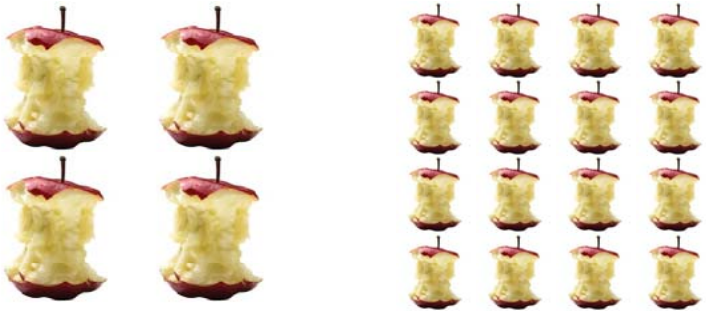
LABS hp

Corona project - solving the many core interconnect crisis

Cores Per Die

Past

Present

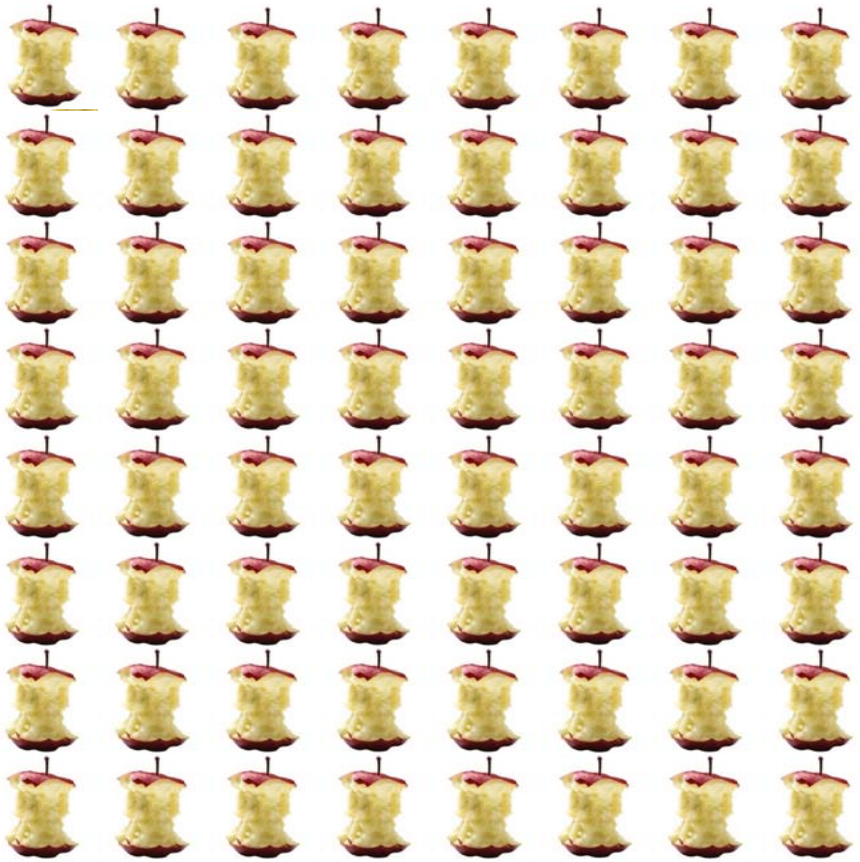


Future

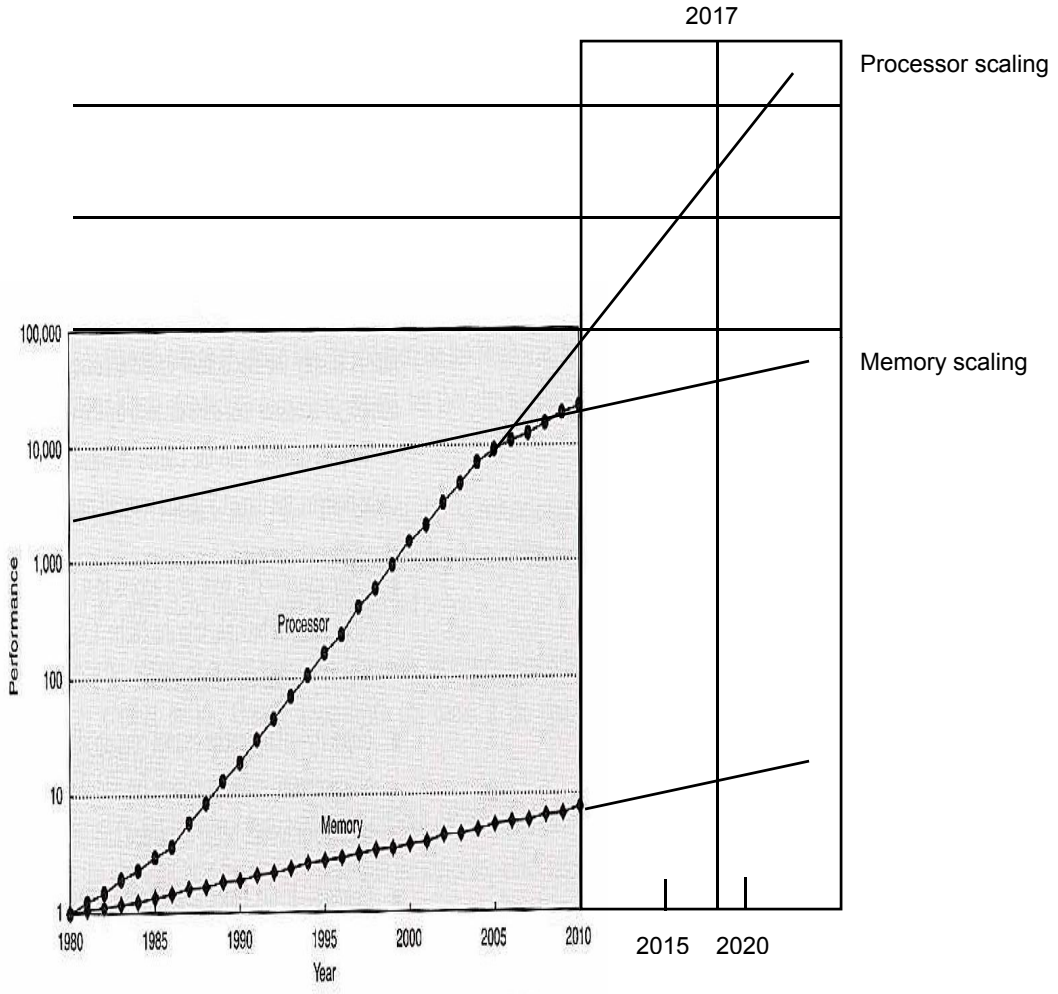
LABS hp

More cores, more flops, but

- External memory bandwidth limited by pincount and power
- Intrachip communications limited by electronic interconnect power

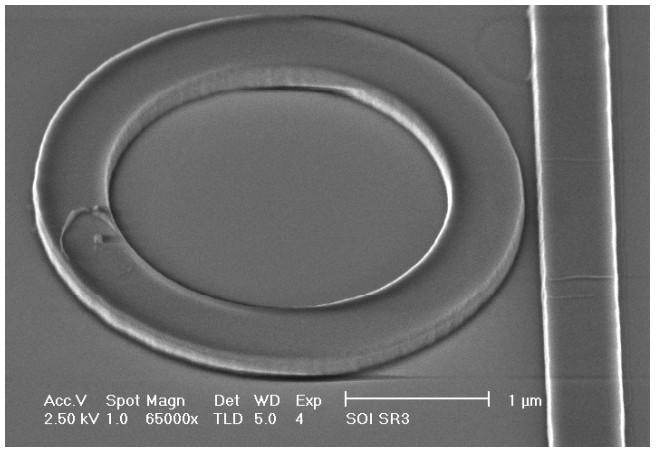


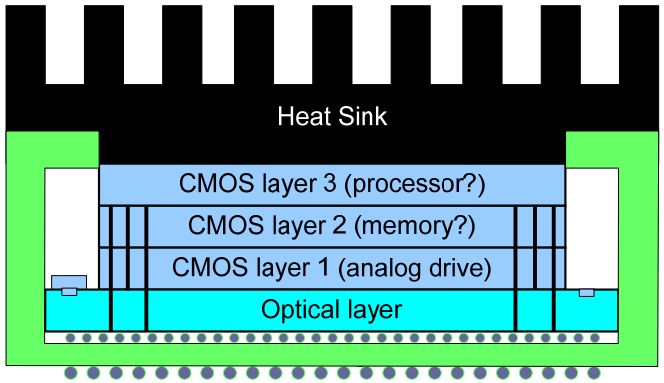
The Memory Wall in the Many Core Era



Solution - Integrated Photonics

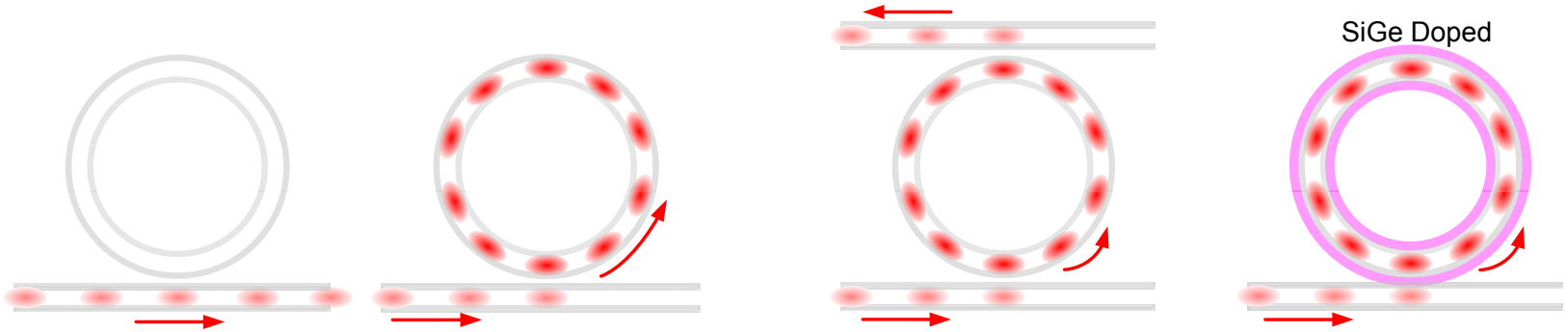
- Low power per bit – fJ range
- Requires modulation, detection and transmission all in an integrated process
- Laser is off-chip
- Compatible with standard CMOS processes
- Chip stacking allows integration with specialization





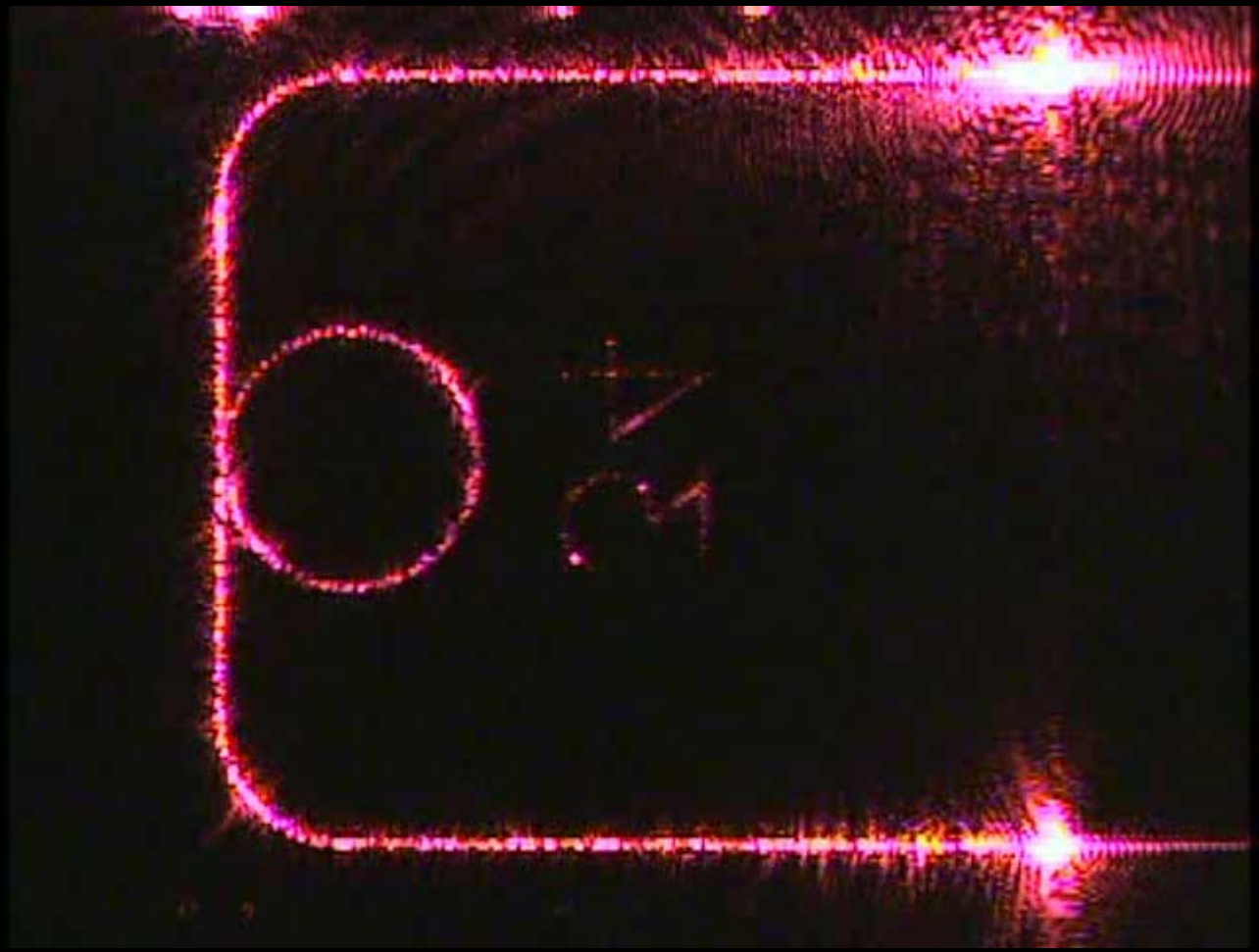
Ring Resonators

One basic structure, 3 applications



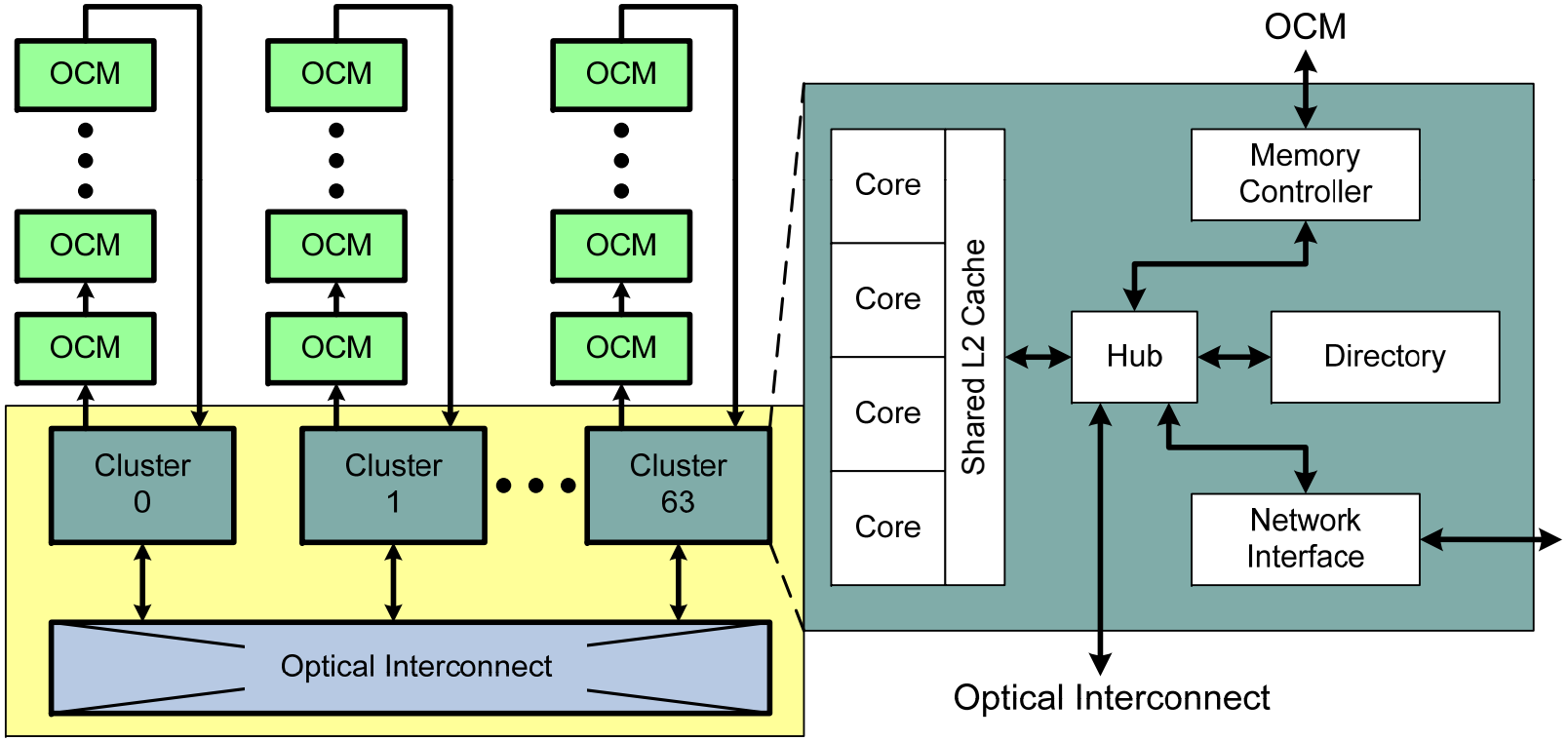
- A modulator – move in and out of resonance to modulate light on adjacent waveguide
- A switch – transfers light between waveguides only when the resonator is tuned
- A wavelength specific detector - add a doped junction to perform the receive function

LABS hp

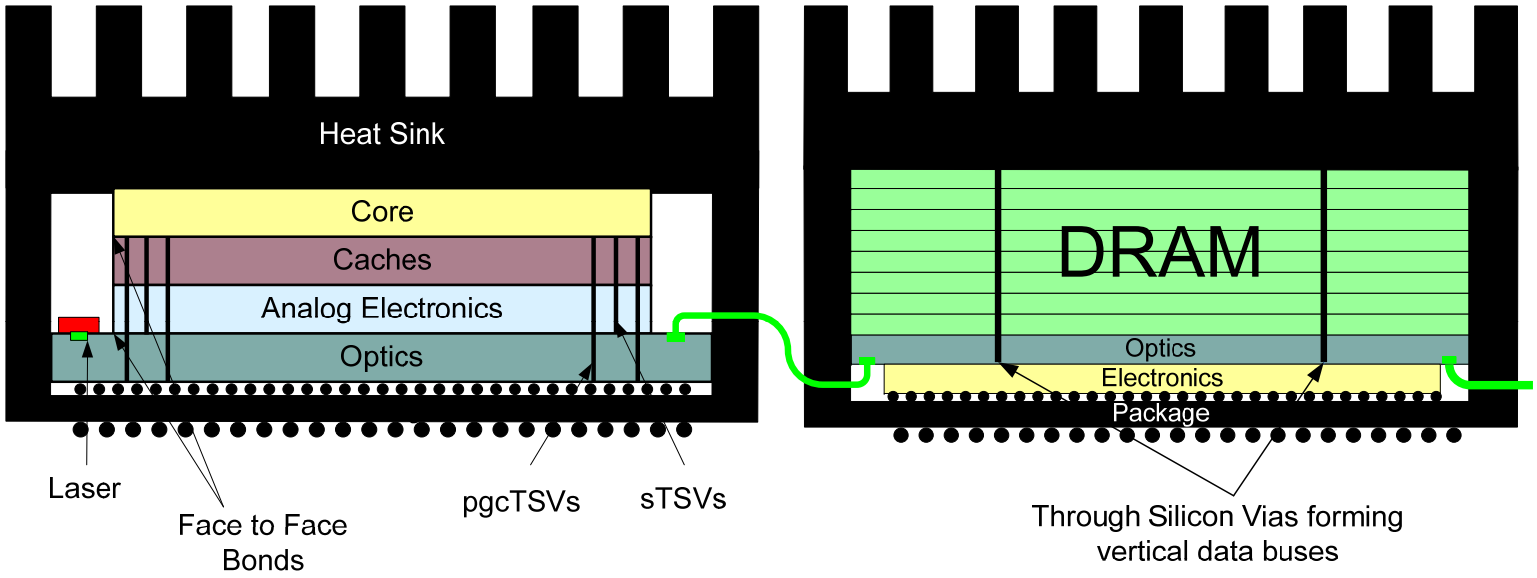


Corona many core architecture

Optically enabled 256 core processor

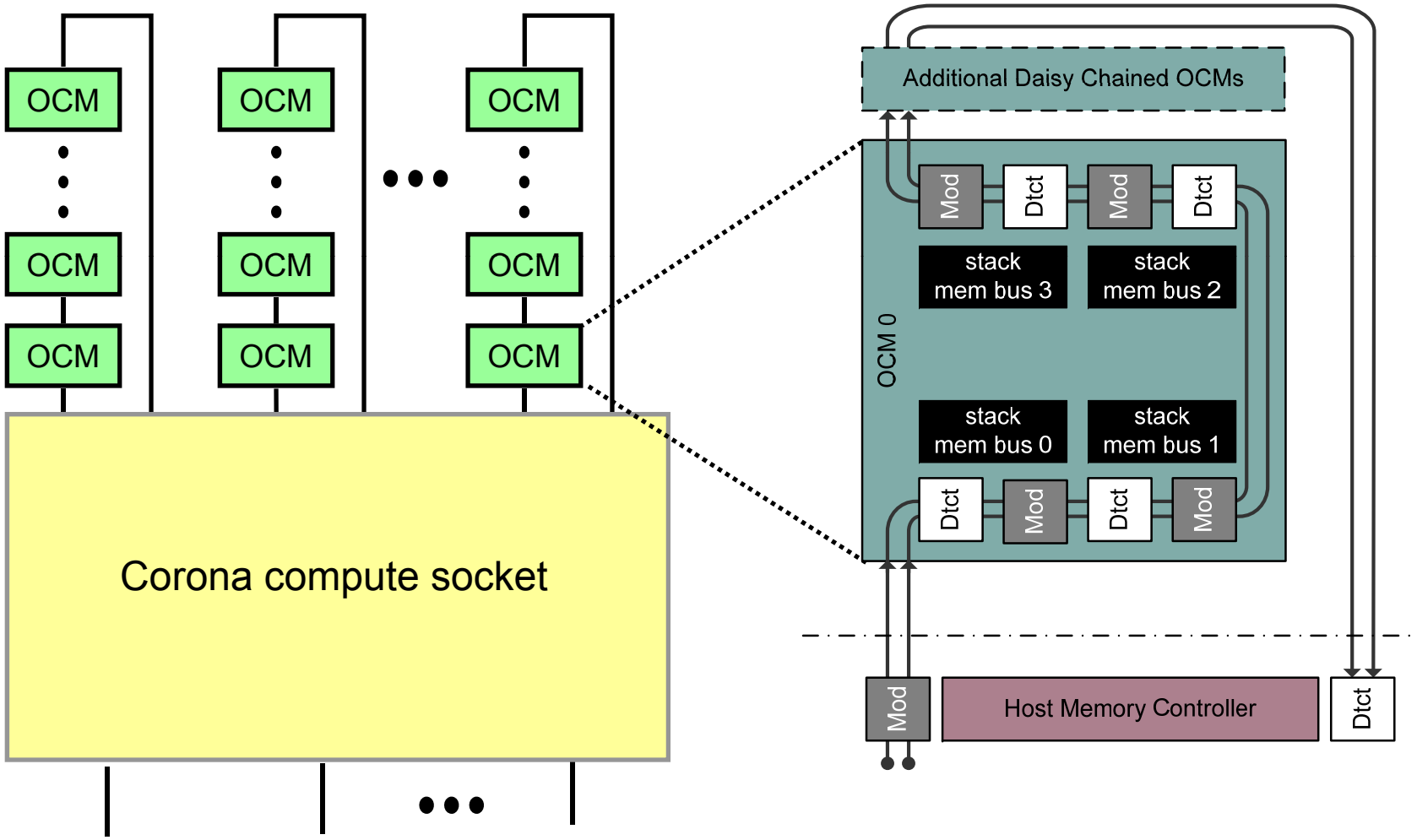


Corona Chip Stack

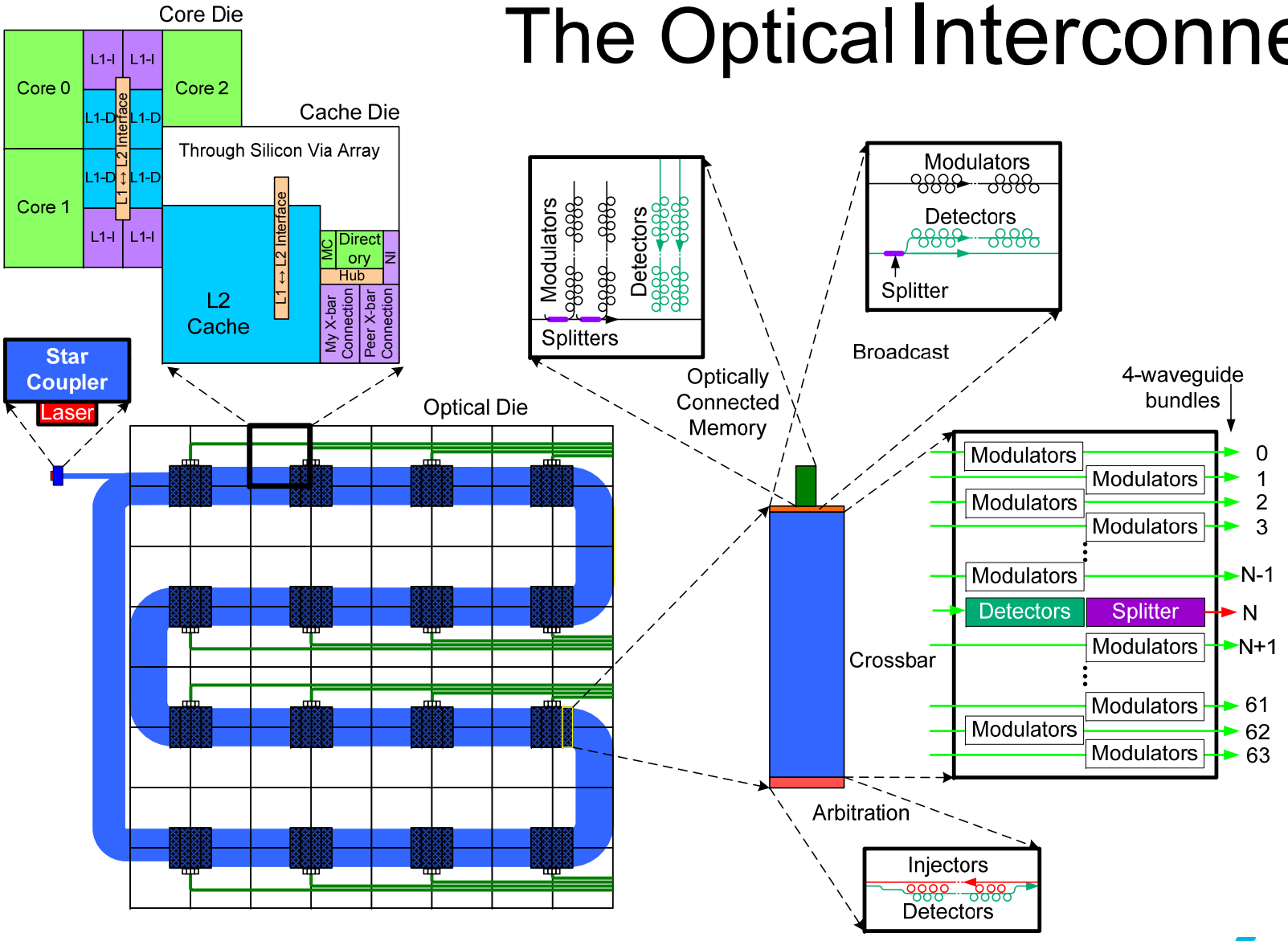


- Stacking technology minimizes electrical path lengths
- Compatible layers, each tailored to their function
- Chip to chip and intrachip communications > 5mm are optical

Optically Connected Memory (OCM)



The Optical Interconnect



Corona Performance Projections

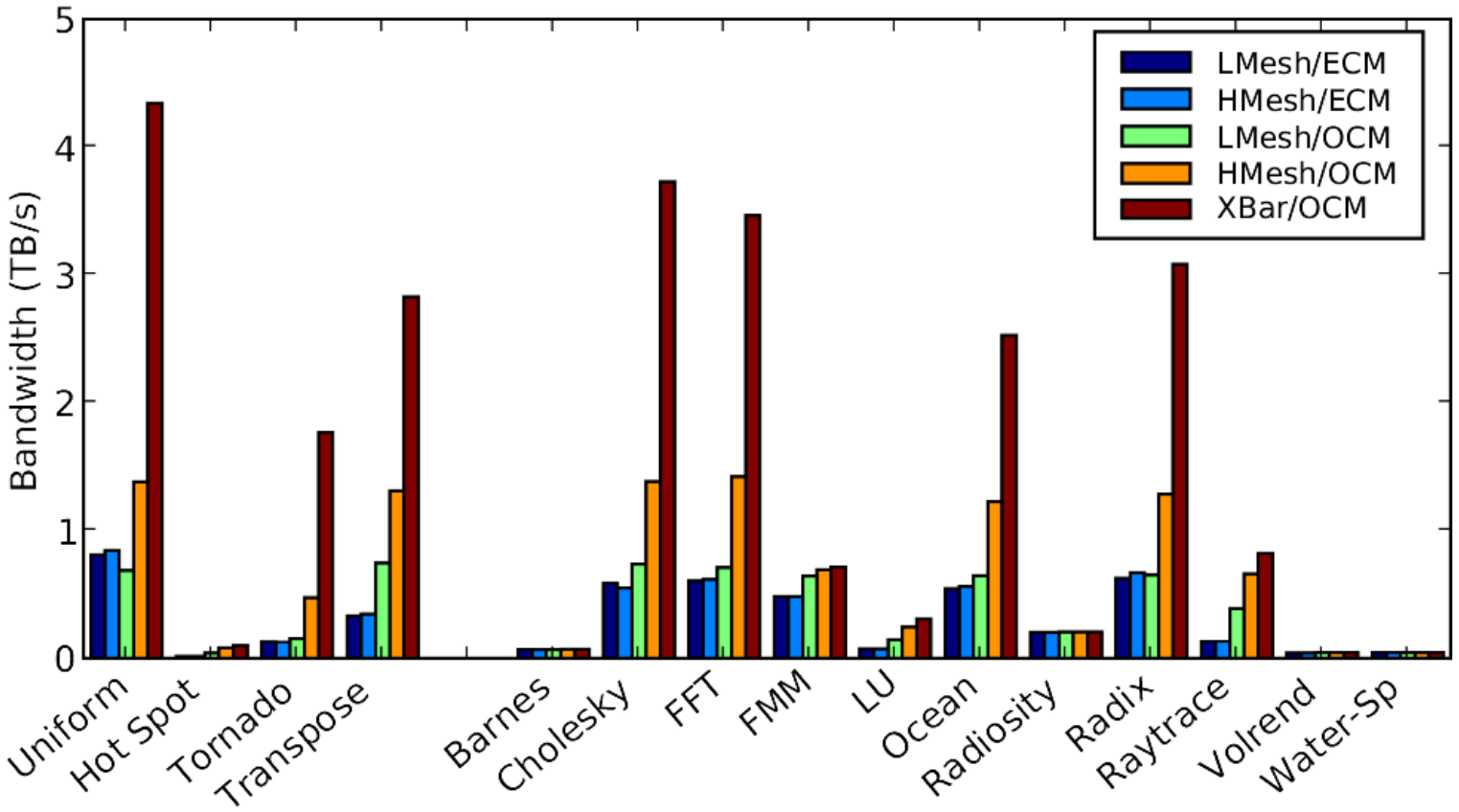
System Performance Simulation

- Compare 5 systems using:
  - Three different on-chip interconnects
    - Electrical 2D on-chip mesh, 0.64 TB/s and 5 cycle hops (LMesh)
    - Electrical 2D on-chip mesh, 1.28 TB/s and 5 cycle hops (HMesh)
    - Optical crossbar, 20.48 TB/s and 8 cycles total
  - Two different memory interconnects
    - Electrical 0.96 TB/s, 1536 signal pins, memory latency is 20 ns
    - Optical 10.24 TB/s, 256 fibers, memory latency is 20 ns

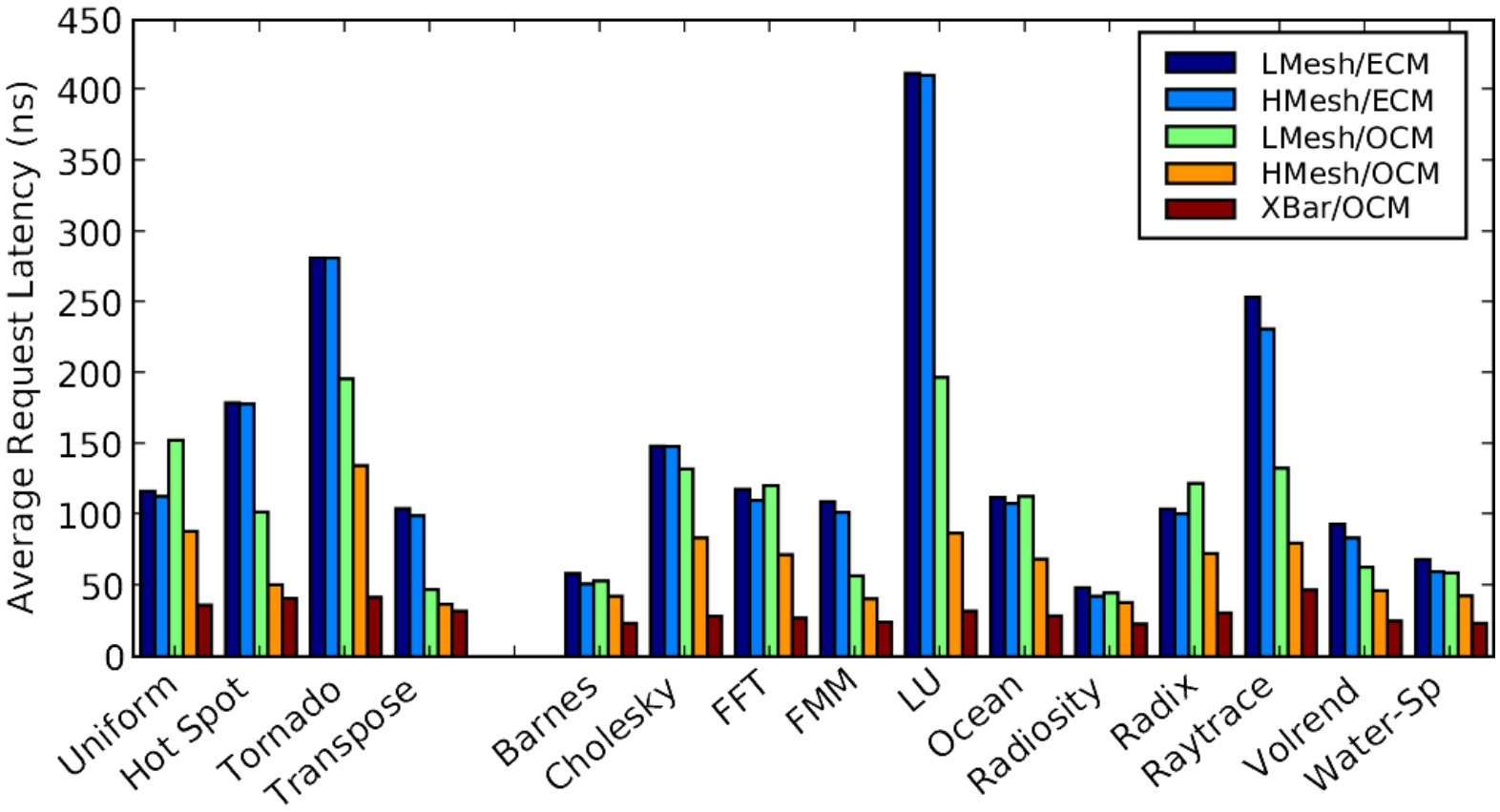
Methodology

- Simulate using COTSon + M5
- Workloads:
  - 5 synthetic benchmarks
  - SPLASH-2

Bandwidth

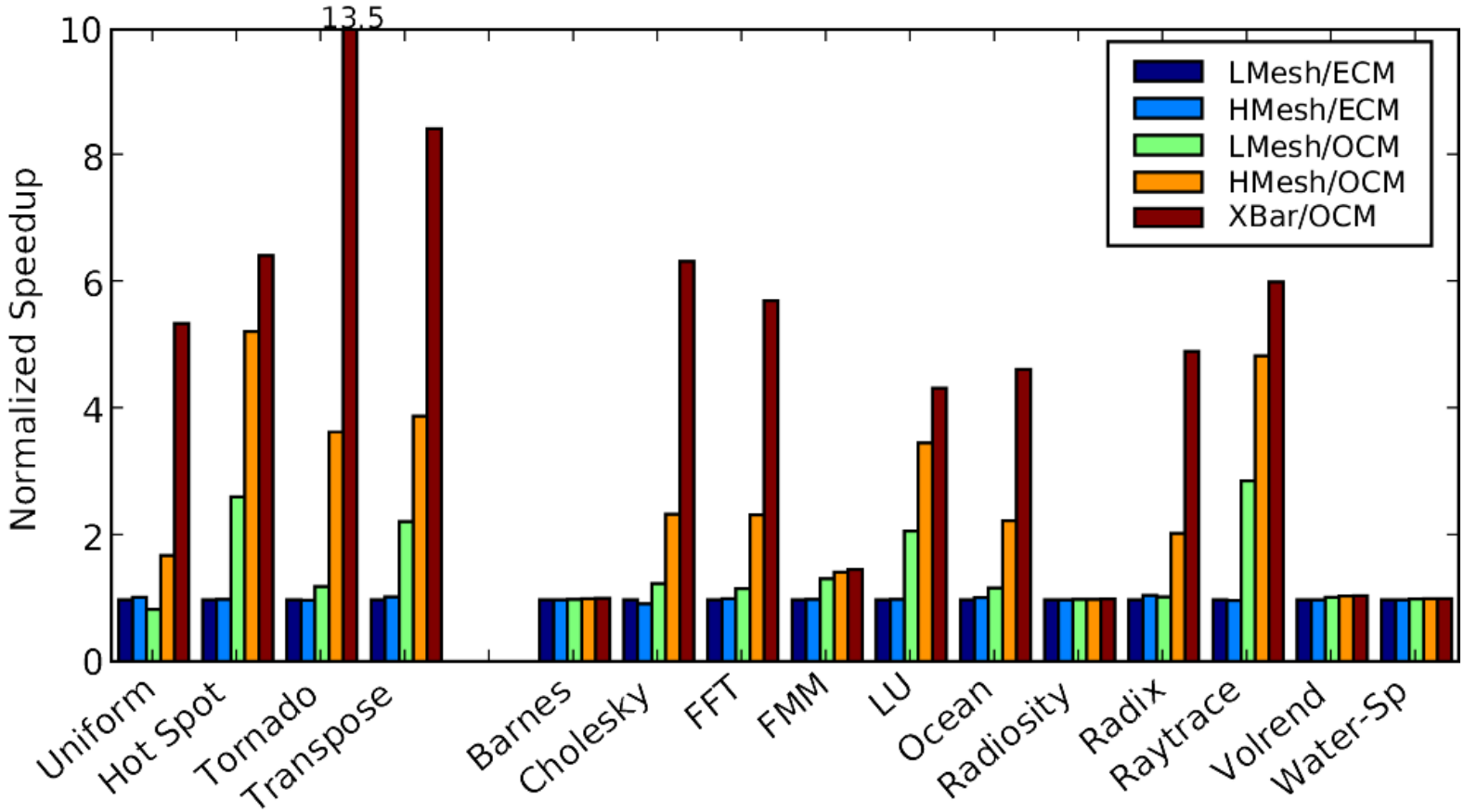

Bandwidth (TB/s)
0
1
2
3
4
5
LMesh/ECM
HMesh/ECM
LMesh/OCM
HMesh/OCM
XBar/OCM
Uniform
Hot Spot
Tornado
Transpose
Barnes
Cholesky
FFT
FMM
LU
Ocean
Radiosity
Radix
Raytrace
Volrend
Water-Sp

Latency



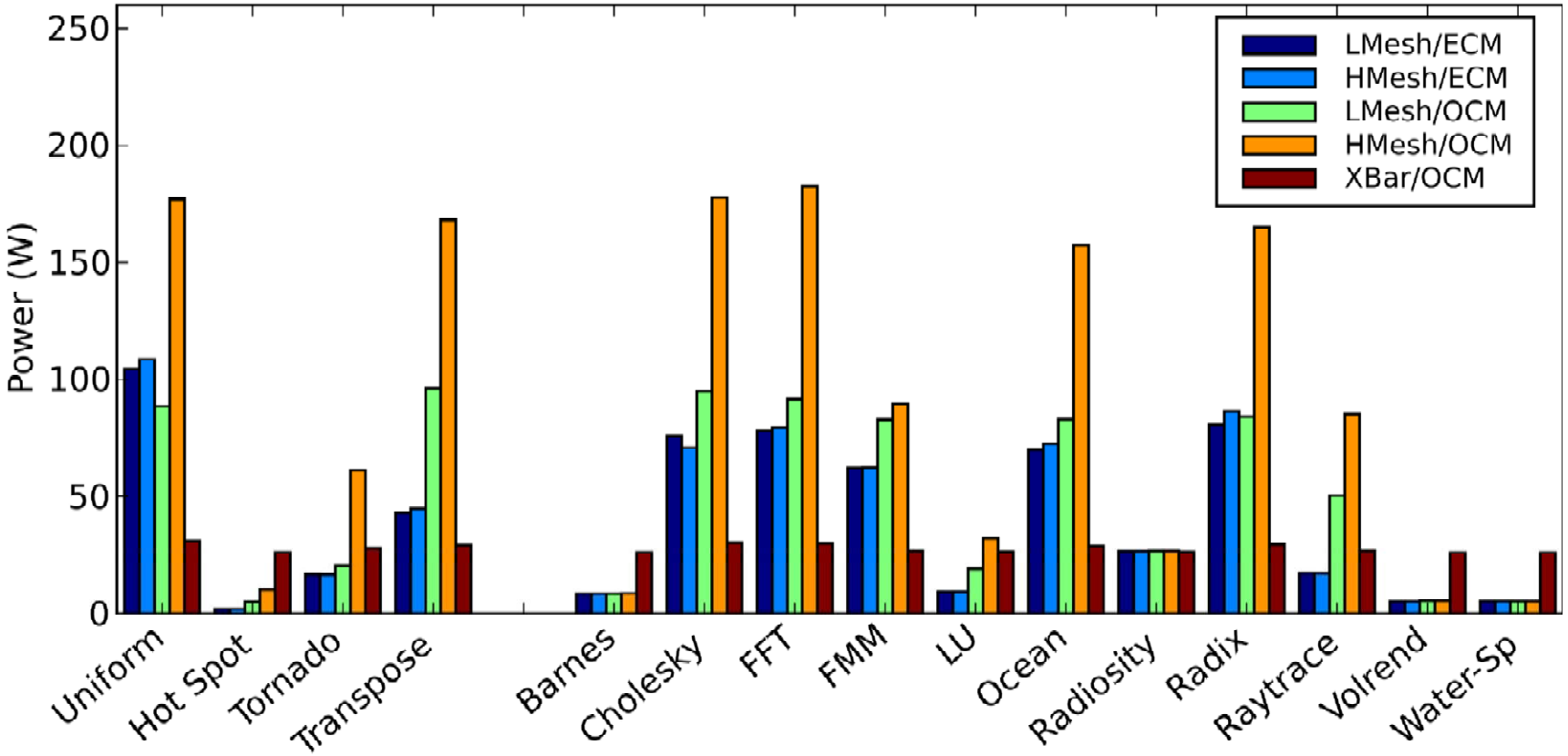
LABS hp

Relative Performance



LABS hp

Power comparison



Corona Benefits from Optics

- Bandwidth scales to 1,000 threads
  - 10 TB/s off-chip bandwidth
  - 20 TB/s bandwidth between cores
  - Modest power requirements
- Low, uniform latencies between cores & memory
  - Optical crossbar
- Coherent shared memory

LABS hp

Acknowledgements

- Ray Beausoleil, Marco Fiorentino, David Fattal
- Jung Ho Ahn, Nate Binkert, Al Davis, Norm Jouppi, Matteo Monchiero, Dana Vantrease

LABS hp

Some interesting programming issues

- When do large scale shared memory codes win over MPI codes?
- What factors limit the scalability of large scale shared memory codes?
- If you can have 256 cache coherent cores in one socket what should the programming model be between sockets?

LABS hp