ECE/CS 552: Nanophotonics

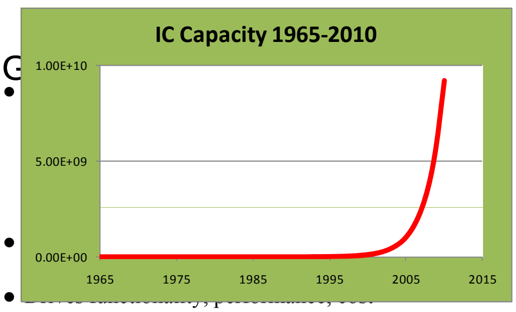
Instructor: Mikko H Lipasti

Fall 2010
University of Wisconsin-Madison

Nanophotonics

- ☑ Nanophotonics overview
- Sharing the nanophotonic channel
  – Light-speed arbitration [MICRO 09]
- Utilizing the nanophotonic channel
  – Atomic coherence [HPCA 11]

G

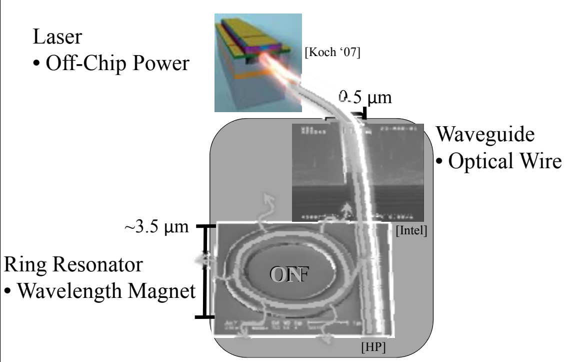


-
-
-
  – Exponential improvement for 40+ years
  – Built on Von Neumann model (fetch/execute)

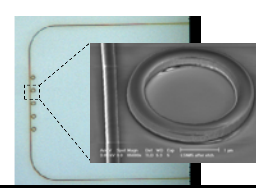
Distributed processing on chip

- Future chips rely on distributed processing
  – Many computation/cache/DRAM/IO nodes
  – Placement, topology, core uarch/strength, tbd
- Conventional interconnects may not suffice
  – Buses not viable
  – Crossbars are slow, power-hungry, expensive
  – NOCs impose latency, power overhead
- Nanophotonics to the rescue
  – Communicate with photons
  – Inherent bandwidth, latency, energy advantages
  – Silicon integration becoming a reality
- Challenges & opportunities remain

Si Photonics: How it works



Ring Resonators



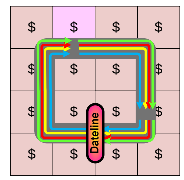
Key attributes of Si photonics

- Very low latency, very high bandwidth
- Up to 1000x energy efficiency gain
- Challenges
  - Resonator thermal tuning: heaters
  - Integration, fabrication, *is this real?*
- Opportunities
  - Static power dominant (laser, thermal)
  - Destructive reads: fast wired or

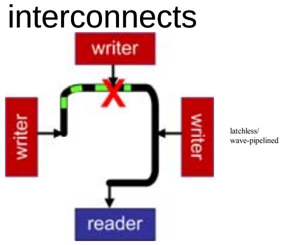
Corona substrate [ISCA08]

Targeting Year 2017

- Logically a ring topology
- One concentric ring per node
- 3D stacked: optical, analog, digital



Multiple writer single reader (MWSR) interconnects



Arbitration prevents corruption of in-flight data

Motivating an optical arbitration solution

MWSR Arbiter must be:

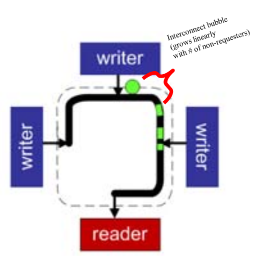
1. ***Global*** - Many writers requesting access
2. ***Very fast*** – Otherwise bottleneck

Optical arbiter avoids OEO conversion delays, provides light-speed arbitration

Proposed optical protocols

- Token-based protocols
  - Inspired by classic token ring
  - Token == transmission rights
  - Fits well with ring-shaped interconnect
  - Distributed, Scalable
  - (limited to ring)

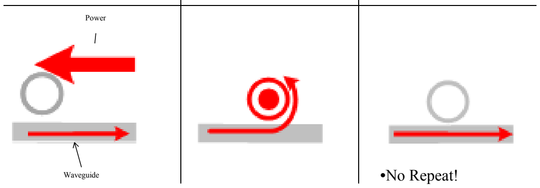
Baseline



- Based on traditional token protocols
- Repeat token at each node
  - But data is not repeated!
  - Poor utilization

Optical arbitration basics

Token - Inject | **Token - Seize** | **Token - Pass**

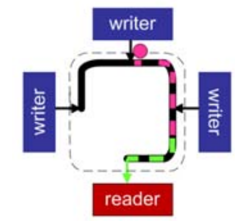


•No Repeat!
•Token latency bounded by the time of flight between requesters.

Arbitration solutions

Token Channel

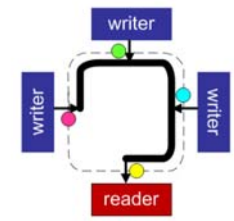
Single Token / Serial Writes



Token passing allows token to pace transmission tail (no bubbles)

Token Slot

Multiple Tokens / Simultaneous Writes



Token passing allows token to directly precede slot

Flow control and fairness

Flow Control:

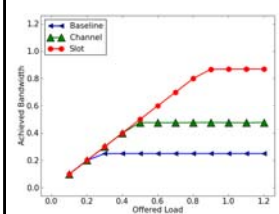
- Use token refresh as opportunity to encode flow control information (credits available)
- Arbitration winners decrement credit count

Fairness:

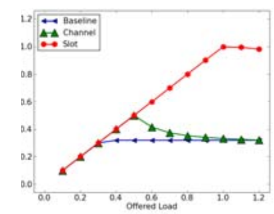
- Upstream nodes get first shot at tokens
- Need mechanism to prevent starvation of downstream nodes

Results - Performance

Uniform



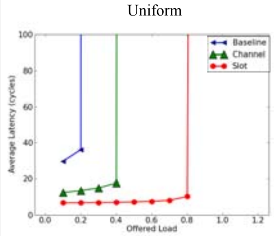
HotSpot



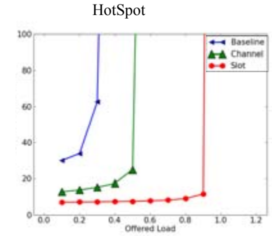
Token Slot benefits from
• the availability of multiple tokens (multiple writers)
• fast turn-around time of flow-control mechanism

Results - Latency

Uniform



HotSpot



Token Slot has the lowest latency and saturates at 80%+ load

Optical arbitration summary

- Arbitration speed has to match transfer speed for fine-grained communication
  - Arbiter has to be optical
- High throughput is achievable
  - 85+% for token slot
- Limited to simple topologies (MWSR)
- Implementation challenges
  - Opt-elec-logic-elec-opt in 200ps (@5GHz)

Nanophotonics

- ☑ Nanophotonics overview
- ☑ Sharing the nanophotonic channel
  - – Light-speed arbitration [MICRO 09]
- ● Utilizing the nanophotonic channel
  - – Atomic coherence [HPCA 11]

19

What makes coherence hard?

Unordered interconnects
- – split transaction buses, meshes, etc

Speculation
- – Sharer-prediction, speculative data use, etc.

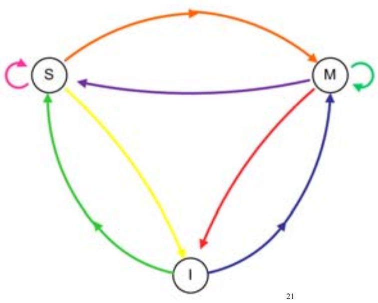
Multiple initiators of coherence requests
- – L1-to-L2, Directory Caches, Coherence Domains, etc

→State-event pair explosion

● →Verification headache

Example: MSI (SGI-Origin-like, directory, invalidate)

Stable States

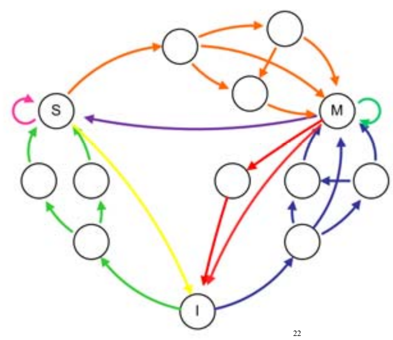


21

Example: MSI (SGI-Origin-like, directory, invalidate)

Stable States

Busy States



22

Example: MSI (SGI-Origin-like, directory, invalidate)

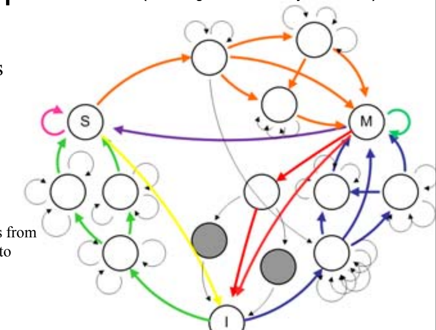
Stable States

Busy States

Races

↑

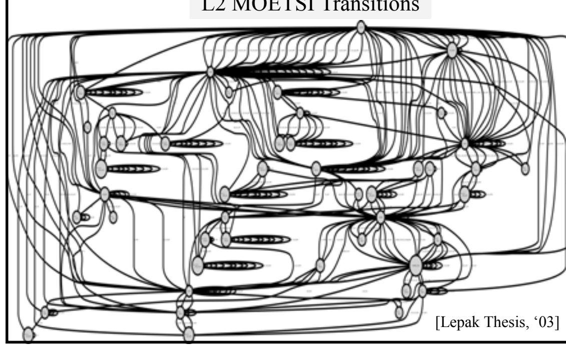
"unexpected" events from concurrent requests to same block



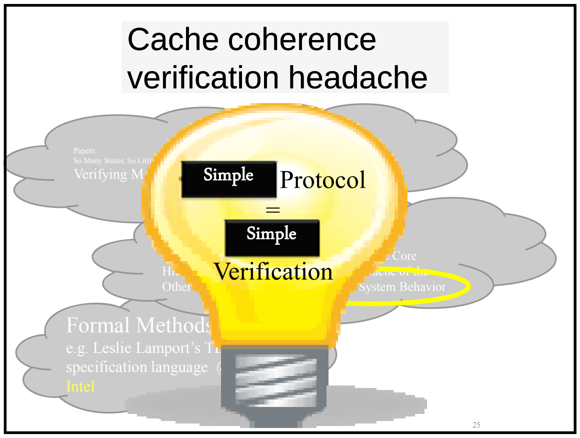
23

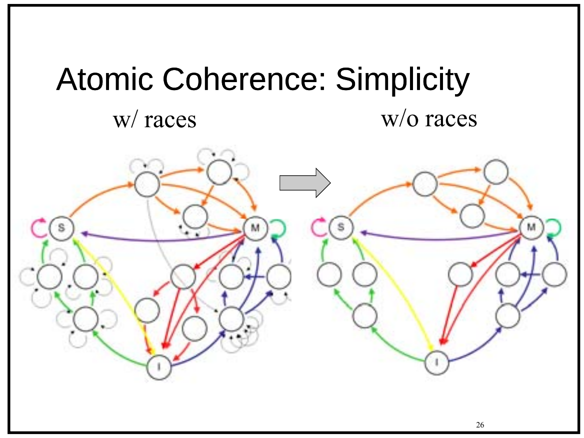
Cache coherence complexity

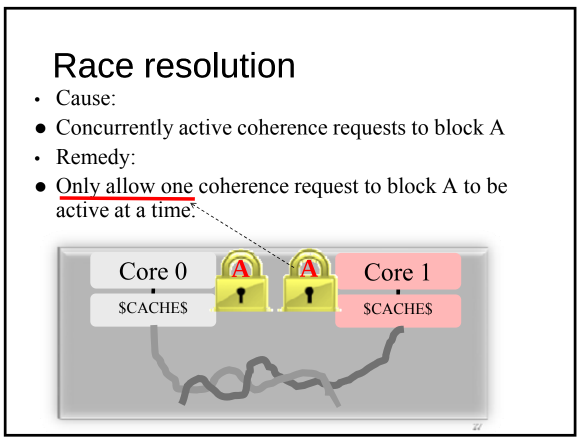
L2 MOETSI Transitions

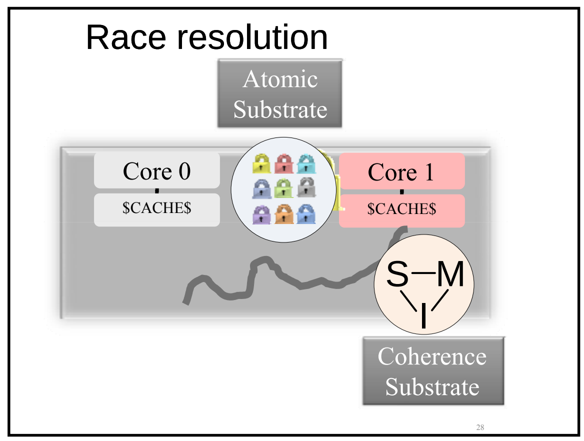


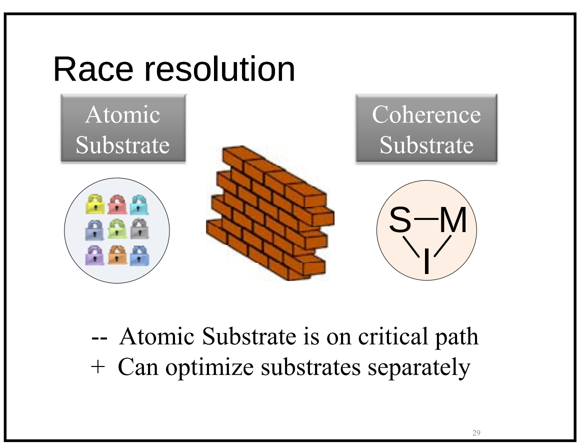
[Lepak Thesis, '03]

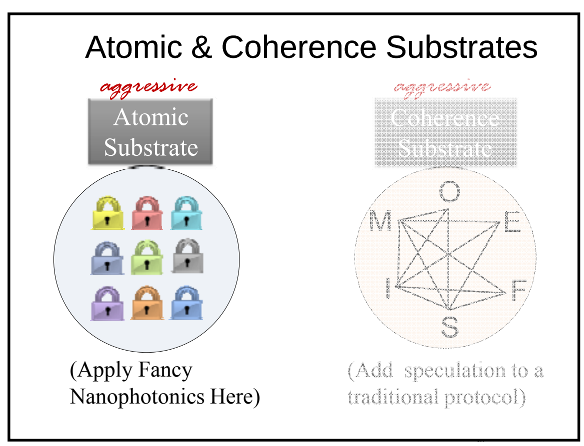

Cache coherence verification headache
Simple
Protocol
=
Simple
Verification
Formal Methods
System Behavior

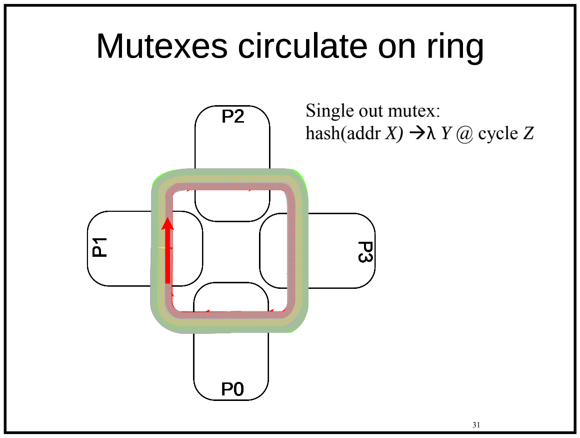

Atomic Coherence: Simplicity
w/ races
w/o races
S
M
I

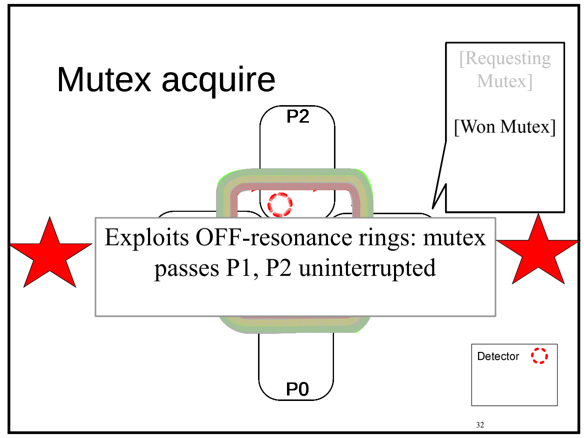

Race resolution
• Cause:
● Concurrently active coherence requests to block A
• Remedy:
● Only allow one coherence request to block A to be active at a time.
Core 0
$CACHE$
A
A
Core 1
$CACHE$

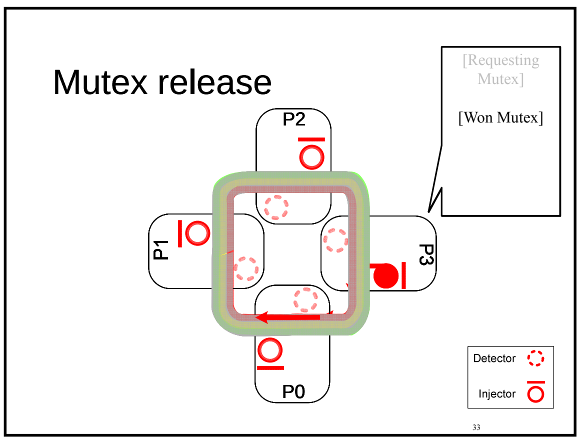

Race resolution
Atomic Substrate
Core 0
$CACHE$
Core 1
$CACHE$
S—M
I
Coherence Substrate

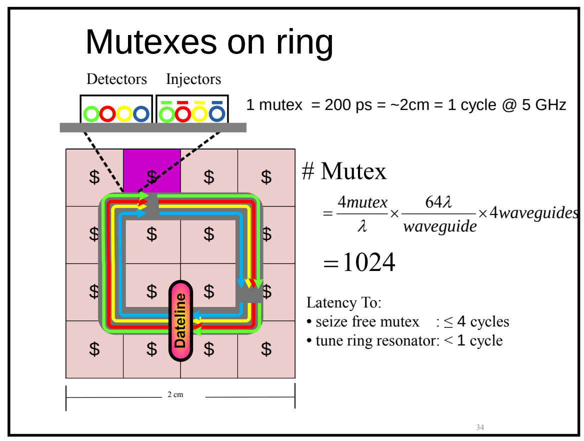

Race resolution
Atomic Substrate
Coherence Substrate
S—M
I
-- Atomic Substrate is on critical path
+ Can optimize substrates separately

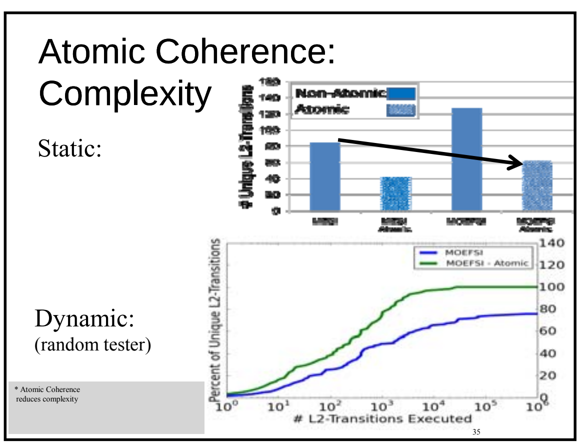

Atomic & Coherence Substrates
aggressive
Atomic Substrate
aggressive
Coherence Substrate
O
M
E
I
F
S
(Apply Fancy Nanophotonics Here)
(Add speculation to a traditional protocol)

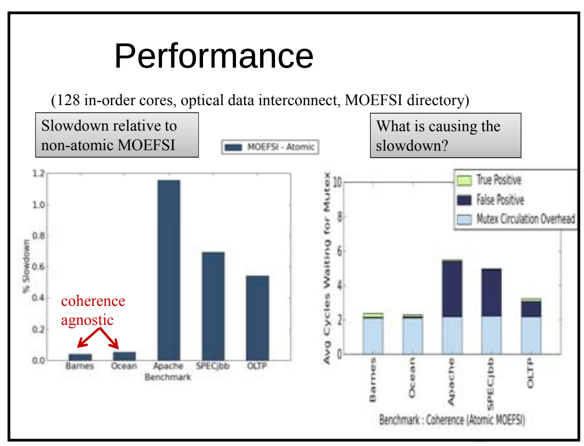

Mutexes circulate on ring
Single out mutex:
hash(addr X) →λ Y @ cycle Z
P2
P1
P3
P0


Mutex acquire
[Requesting Mutex]
[Won Mutex]
P2
Exploits OFF-resonance rings: mutex passes P1, P2 uninterrupted
P0
Detector


Mutex release
[Requesting Mutex]
[Won Mutex]
P2
P1
P3
P0
Detector
Injector


Mutexes on ring
Detectors
Injectors
1 mutex = 200 ps = ~2cm = 1 cycle @ 5 GHz
Mutex
= 4mutex/λ × 64λ/waveguide × 4waveguides
= 1024
Latency To:
• seize free mutex : ≤ 4 cycles
• tune ring resonator: < 1 cycle
Dateline
2 cm


Atomic Coherence:
Complexity
Static:
Non-Atomic
Atomic
Unique L2-Transitions
Dynamic:
(random tester)
MOEFSI
MOEFSI - Atomic
Percent of Unique L2-Transitions
L2-Transitions Executed
* Atomic Coherence reduces complexity


Performance
(128 in-order cores, optical data interconnect, MOEFSI directory)
Slowdown relative to non-atomic MOEFSI
MOEFSI - Atomic
What is causing the slowdown?
% Slowdown
coherence agnostic
Barnes
Ocean
Apache
SPECjbb
OLTP
Benchmark
True Positive
False Positive
Mutex Circulation Overhead
Avg Cycles Waiting for Mutex
Benchmark : Coherence (Atomic MOEFSI)

Optimizing coherence

Observation:

Holding Block B's mutex gives holder free reign over coherence activity related to block B

O.wned and **F.orward** State:

- Responsible for satisfying on-chip read misses

Opportunity:

- Try to keep O/F alive
- If O (or F) block evicted:
  While mutex is held, 'shift' O/F state to sharer
  (or hand-off responsibility)

37

Optimizing coherence

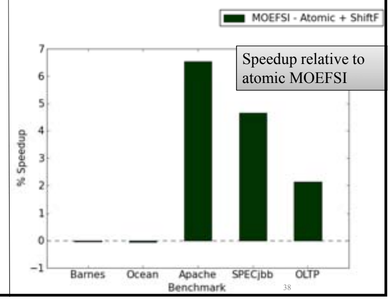
- If O (or F) block evicted: 'Shift' O/F state to sharer

Complexity:

L2 transitions ⇩

(b/c less variety in sharing possibilities)

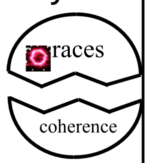
Performance:



38

Atomic Coherence Summary

- Nanophotonics as enabler
  - Very fast chip-wide consensus
- Atomic Protocols are simpler protocols
  - And can have minimal cost to performance (w/ nanophotonics)
  - Opportunity for straightforward protocol enhancements: ShiftF
- More details in HPCA-11 paper
  - Push protocol (update-like)



39

Nanophotonics

- ☑ Nanophotonics overview
- ☑ Sharing the nanophotonic channel
  - Light-speed arbitration [MICRO 09]
- ☑ Utilizing the nanophotonic channel
  - Atomic coherence [HPCA 11]

40