UPC Ph.D. Course on Parallel Computer Architecture

Distributed Shared Memory Part 2 (Chapters 9 & 11)

Copyright 2003 Mark D. Hill University of Wisconsin-Madison

Slides are derived from work by Sarita Adve (Illinois), Babak Falsafi (CMU), Alvy Lebeck (Duke), Steve Reinhardt (Michigan), and J. P. Singh (Princeton). Thanks!



Outline

• Cache-Only Memory Architecture (COMA)

• Paged-Based Distributed Shared Memory

UPC Parallel Computer Architecture

- Simple-COMA (S-COMA)
- Hierarchical Coherence
- Latency Tolerance

(C) 2003 Mark D. Hill from Adve, Falsafi, Lebeck, Reinhardt, & Singh



## COMA E.g.: Data Diffusion Machine (DDM)

- All hardware COMA
- Attraction Memory => One giant hardware cache
- Maintains both address tags and state
- Data addressed, allocated, & kept coherent in blocks
- Directory info on a per cache-block basis
- Not Home Based:
  - data is migratory => AM attracts data
  - must find a home when replacing the data
    must find the directory entry before finding the data

```
(C) 2003 Mark D. Hill from Adve,
Falsafi, Lebeck, Reinhardt, & Singh UPC Parallel Computer Architecture
```























## Sun Wildfire

- [Hagersten/Koster HPCA99]
- Begin with up to four SMP nodes
- Add pseudo-processor board to each as proxy for rest of system Can run CC-NUMA directory protocol
- •
- Can selectively use S-COMA (called Coherent Memory Replication)
- Selects between with competitive algorithm [Falsafi/Wood ISCA97] •
- A hierarchical methods of building parallel machines

## (C) 2003 Mark D. Hill from Adve, Falsafi, Lebeck, Reinhardt, & Singh UPC Parallel Computer Architecture















- Bandwidth shared among nodes
- Bus increases latency to local memory
- With coherence, typically wait for local snoop results before sending remote requests
- Snoopy bus at remote node increases delays there too, increasing latency and reducing bandwidth
- Overall, may hurt performance if sharing patterns don't comply

22

(C) 2003 Mark D. Hill from Adve, Falsafi, Lebeck, Reinhardt, & Singh UPC Parallel Computer Architecture











## Outline

- Cache-Only Memory Architecture (COMA)
- Paged-Based Distributed Shared Memory
- Simple-COMA (S-COMA)
- Hierarchical Coherence
- Latency Tolerance

(C) 2003 Mark D. Hill from Adve, Falsafi, Lebeck, Reinhardt, & Singh UPC Parallel Computer Architecture

28