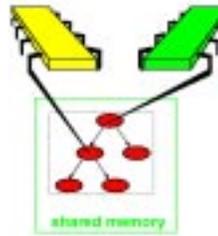


A Tale of Two Directories

Maurice Herlihy Michael Warres
Brown University Sun Microsystems

Shared Objects Rule!

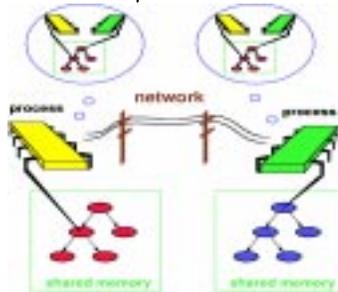


- Familiar model
- easier reasoning
- simplifies data movement and synchronization

18-Jun-99

2

Distributed Shared Object Systems



18-Jun-99

3

Distributed Directory Service



- locates cached copies
- coherence
- synchronization

18-Jun-99

4

This Talk

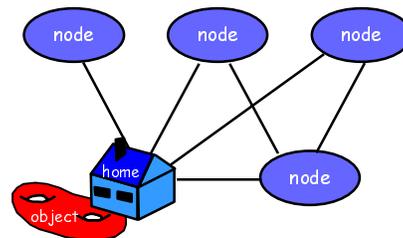
- Experience implementing new distributed directory protocol in Java
- Unexpected performance
- Detective story
 - Java artifact?
 - Something subtle?



18-Jun-99

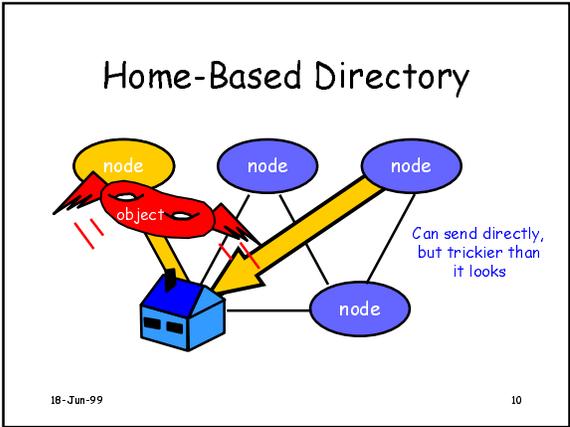
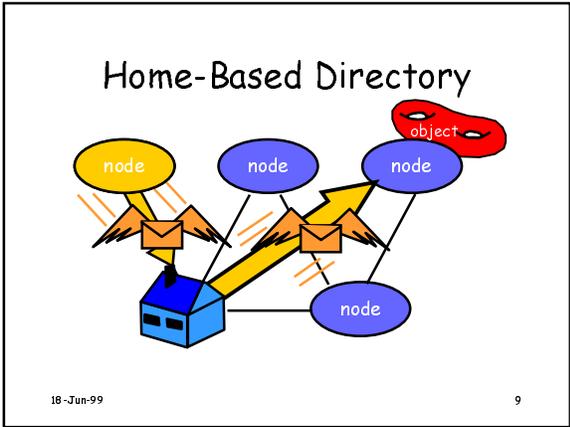
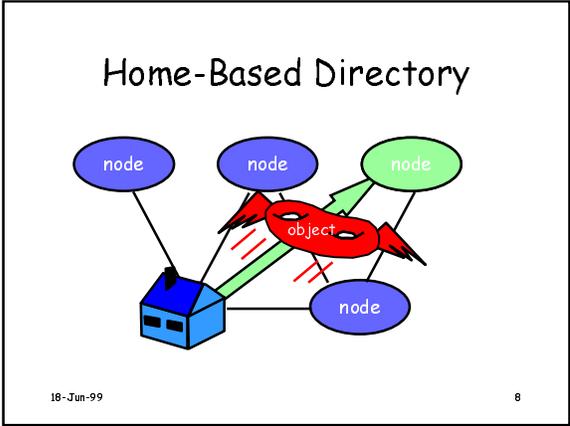
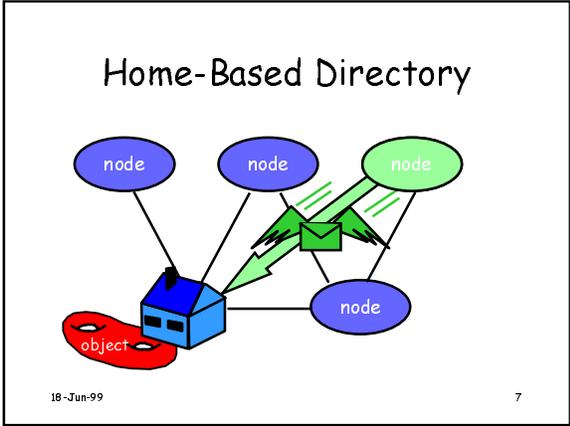
5

Conventional Home-Based Directory



18-Jun-99

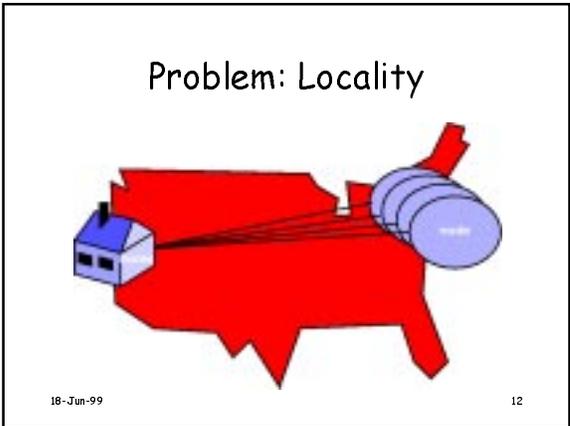
6



Home-Based Directories

- Simple
- Used in real DSO systems
- and in real multiprocessor systems (called **directory-based**)
- Nevertheless ...

18-Jun-99 11



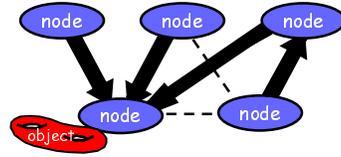
Problem: Hot-Spot Contention



18-Jun-99

13

Arrow Directory

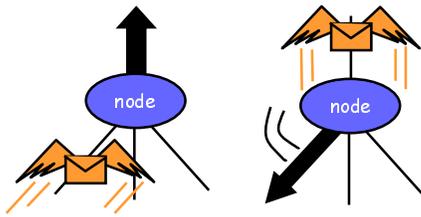


- Spanning tree
- edges directed toward object

18-Jun-99

14

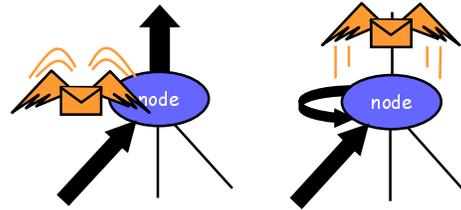
Path Reversal



18-Jun-99

15

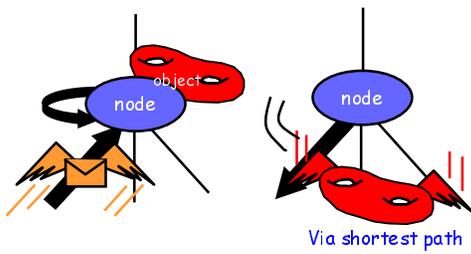
Request Source



18-Jun-99

16

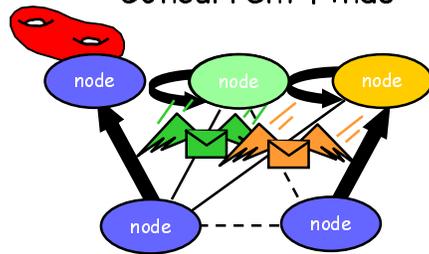
Request Destination



18-Jun-99

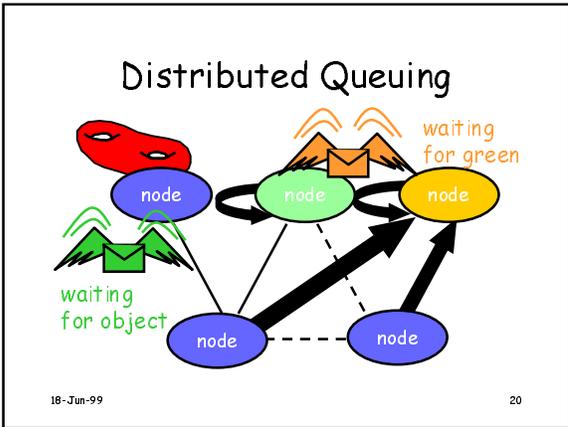
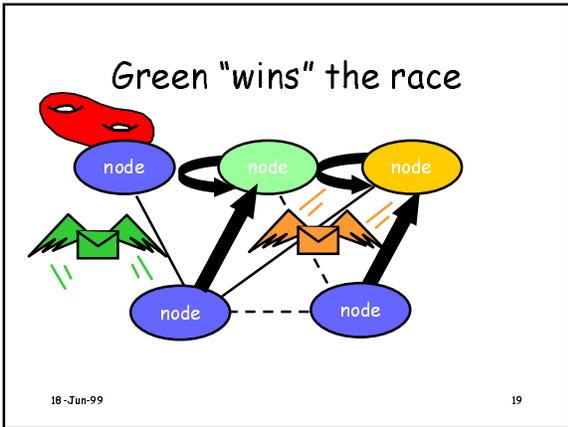
17

Concurrent Finds



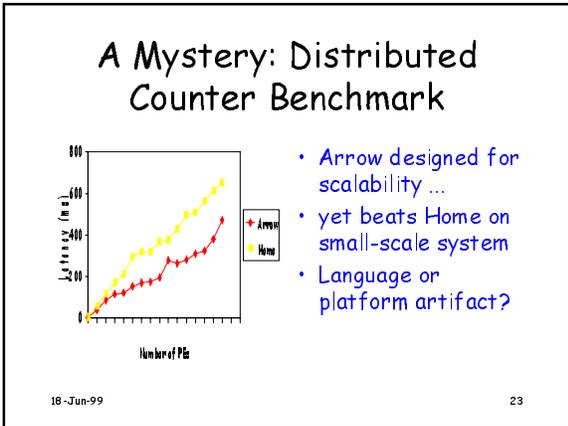
18-Jun-99

18



- ### Aleph Toolkit
- Java packages for DSO systems
 - Set of basic interfaces:
 - ``active'' messages (sort of)
 - reliable multicast
 - directory service
 - remote threads and method invocation
- 18-Jun-99 21

- ### Directory Manager Interface
-
- Multiple implementations
 - Home (900 lines)
 - Arrow (400 lines)
 - Others ...
- 18-Jun-99 22



- ### Platform Artifact?
- Solaris, NT, Linux, Digital Unix
 - JDK 1.1.*, 1.2
- Actual numbers changed, but curves look the same
- 18-Jun-99 24

Thread Hypothesis

- Home directory
 - Hot spot = lots of threads
 - Request need to block? wait()
 - Releasing the object? notifyAll()
 - wakes up all waiting threads
 - "thundering herd" effect?

18-Jun-99

25

Thread-Free Home Directory

- Blocked request
 - doesn't wait()
 - appends list entry
 - and vanishes
- Release object
 - don't notifyAll()
 - scans list
 - replies directly

Result: nothing changed

18-Jun-99

26

Other Rejected Hypotheses

- Serialization overheads
 - tried Externalizable
 - 20% speedup for both
- Arrow tree ordering
 - randomization had no effect
- transport-level protocols
 - tried UDP, TCP, RMI

18-Jun-99

27

Worst-Case Message Traffic

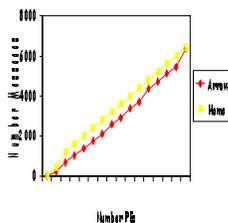
	request	response
home	2	2
arrow	$2 \log n$	1

Does Arrow really use more messages?

18-Jun-99

28

Actual Message Traffic



- Arrow uses slightly fewer messages
- most requests quickly "diverted" into queue
- worst-case doesn't happen

18-Jun-99

29

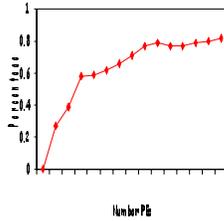
Latency Hiding Hypothesis

- If, when a PE releases an object,
- another request is already waiting,
- then computation time dominated by
 - local computation
 - one-hop transmission
- Request latency hidden by contention

18-Jun-99

30

Eureka!

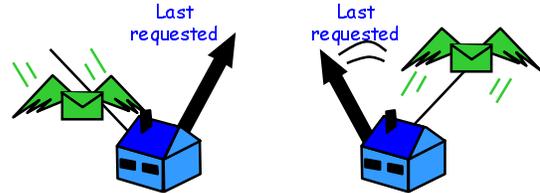


Percentage of time object is released to waiting request

18-Jun-99

31

Hybrid Directory



18-Jun-99

32

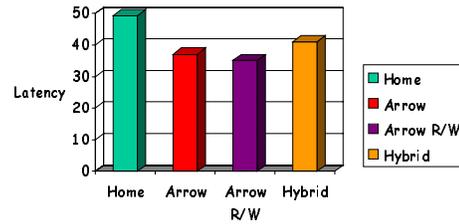
Compromise Position

- Hybrid reduces request latency
 - request: 2 messages
 - response: 1 message
- but reintroduces hot-spot
 - requests go through home
 - but don't queue there

18-Jun-99

33

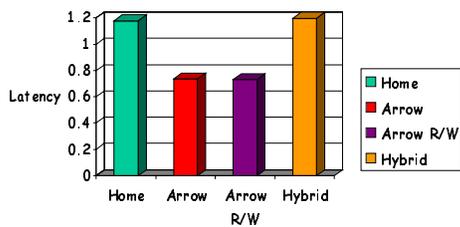
Cholesky



18-Jun-99

34

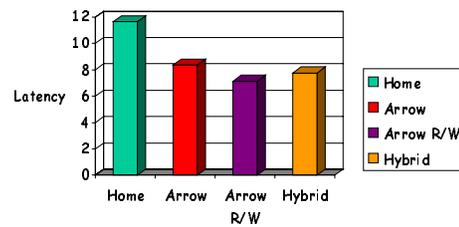
Traveling Salesperson



18-Jun-99

35

Ray Tracing



18-Jun-99

36

Other Issues

- Adding read-sharing modes
- internal synchronization
- applications to ordered multicast
- war stories ...

18-Jun-99

37

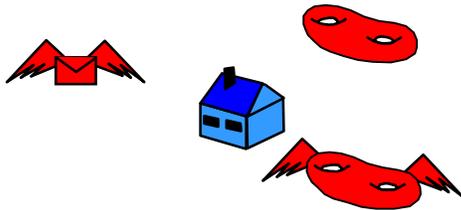
Java

- Mostly Java has been good to us
 - serialization
 - interface & abstract classes
 - rich class libraries
- some things could improve
 - performance debugging & introspection
 - serialization performance

18-Jun-99

38

Clip Art



18-Jun-99

39