ATOMIC-2

Production Use of a Gigabit LAN
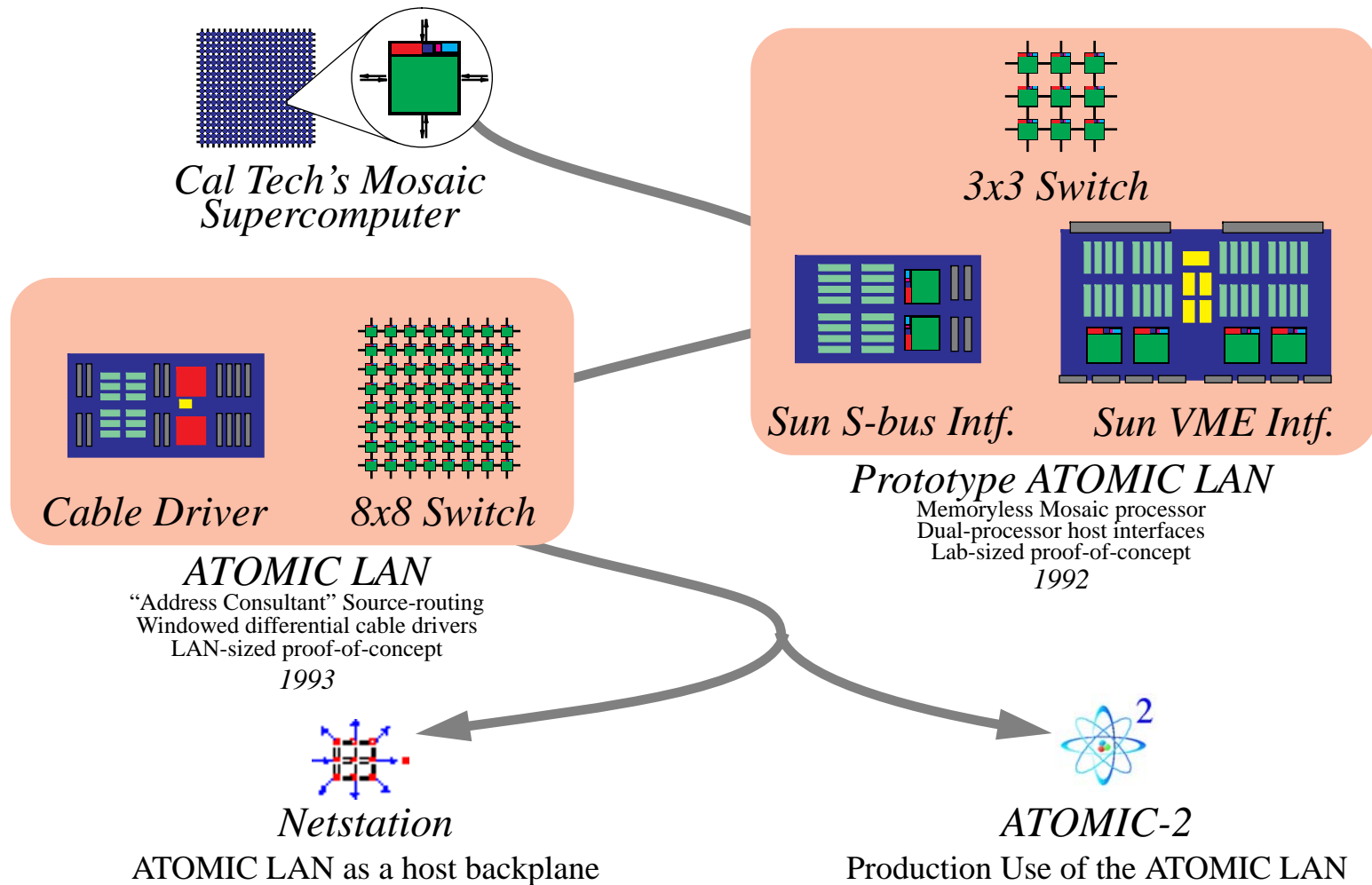
Joe Touch, Annette DeSchon, Hong Xu, Ted Faber, Tom Fisher, Avneesh Sachdev
USC/ISI HPCC Division

ATOMIC Evolution

A Mesh Supercomputer becomes an Gigabit LAN

Cal Tech's Mosaic Supercomputer

3x3 Switch

Sun S-bus Intf.   *Sun VME Intf.*

Cable Driver   *8x8 Switch*

Prototype ATOMIC LAN
Memoryless Mosaic processor
Dual-processor host interfaces
Lab-sized proof-of-concept
1992

ATOMIC LAN
"Address Consultant" Source-routing
Windowed differential cable drivers
LAN-sized proof-of-concept
1993

Netstation
ATOMIC LAN as a host backplane

ATOMIC-2
Production Use of the ATOMIC LAN

Goals

Given a 640-Mbps LAN... what challenges remain?

File Server
High-perf. NFS
RAM-disk interface

Application Demos
PVM benchmarks
Teleconferencing
Web performance

Gateway
ATM interface
Striping issues

Application BW
User-level TCP
Dual-stack arch.
Low-latency RPC

Security
IPv6 mechanisms
Perf. measures

Integrated Services
Reservation mech.
Traffic policing
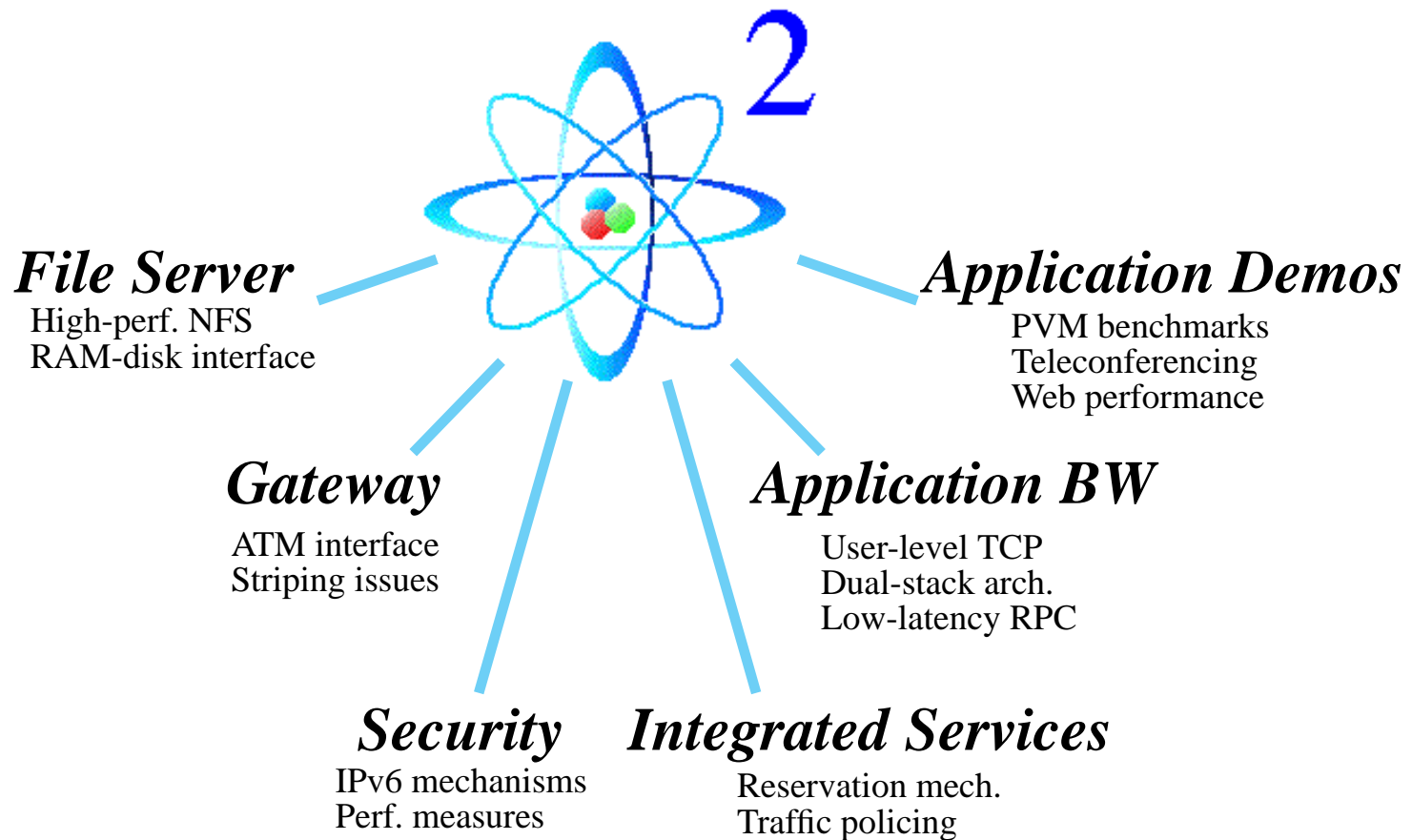
ISI has the only production stand-alone ATOMIC LAN
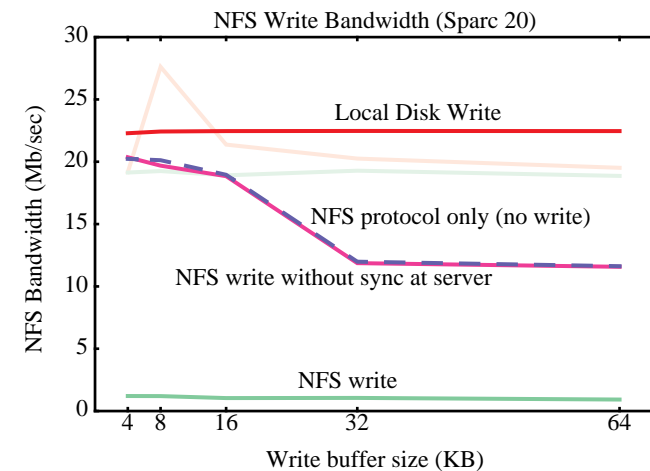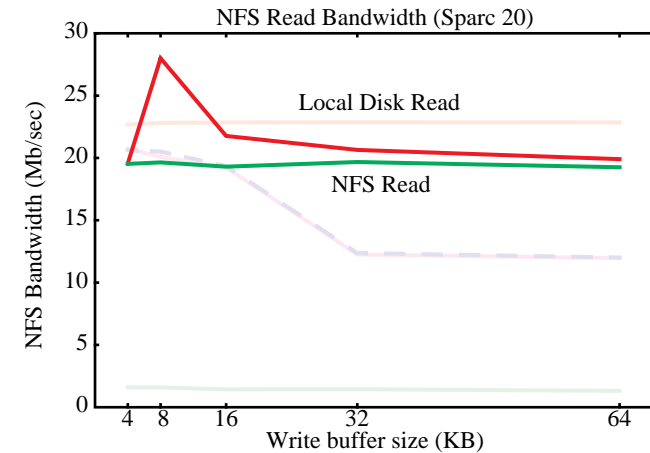
File Server

Goals

- *Files at full network BW*

- *Aggregate user requests fairly*

Software Issues

- *Change the protocol*
  - Pipeline RPCs
  - Dual-path to transport protocols

- *Measurements*

Hardware Issues

- *RAM-disk as support*
 - High-speed access to small files
  - Low-latency file operations

NFS Read Bandwidth (Sparc 20)

Local Disk Read

NFS Read

NFS Bandwidth (Mb/sec)

Write buffer size (KB)

NFS Write Bandwidth (Sparc 20)

Local Disk Write

NFS protocol only (no write)

NFS write without sync at server

NFS write

NFS Bandwidth (Mb/sec)

Write buffer size (KB)

ATOMIC-ATM Gateway
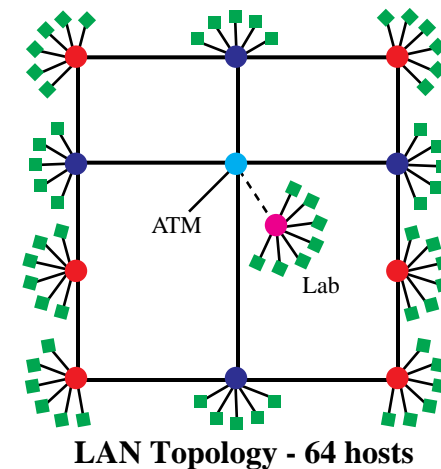
Supporting 64 hosts for everyday access

Host-based performance

- ## IP routing
  - BSD kernel-based
  - Direct inter-interface

- ## Striping issues

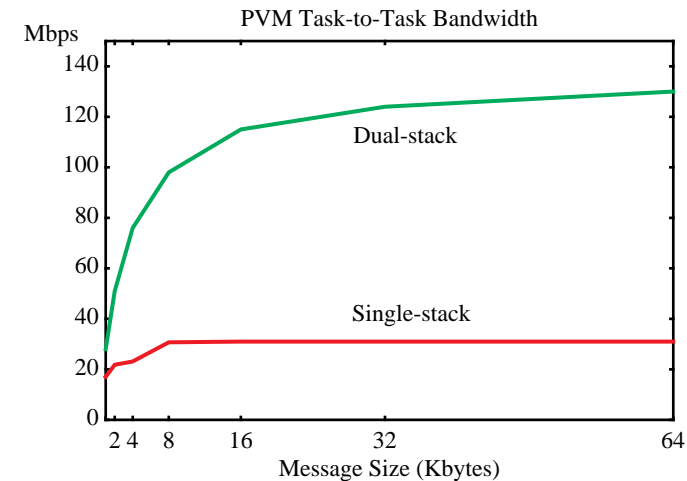- ## Hardware design issues
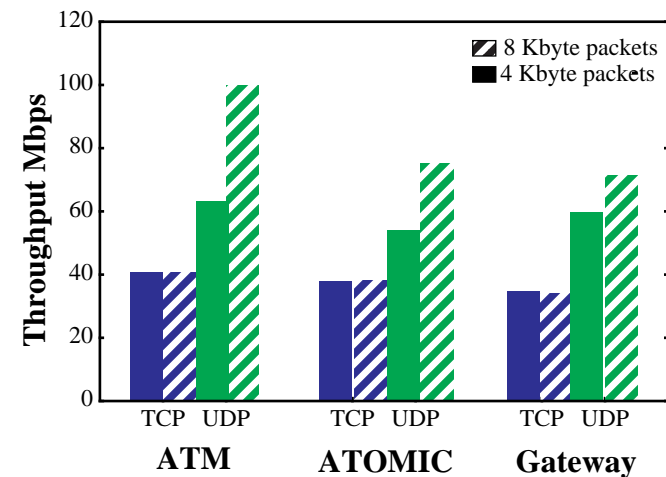
QoS / Integrated Services

- ## Source-route cut-through
  - Vary MTU
  - Pace

- ## Link-layer multicast

LAN Topology - 64 hosts

BW to the User

Dual-stack Protocols

- *PVM variant*

- Reliable user-level protocol for data
- Socket TCP/IP for control
- Dual-stack driver

- *User TCP/IP in progress*

Security

- *MD5 at 37 Mbps*

- vs. IPv4 at 60 Mbps
- 1.5 Mbps MD2, 20 Mbps SHA
- 7.5 Mbps DES

- *Seeking fast algorithm*

- MD5 is 45 opcodes per word
- 15 ops/wd OK / 2-4 for "low cost"

PVM Task-to-Task Bandwidth

Mbps — Dual-stack — Single-stack — Message Size (Kbytes)

Mbps — Raw / Optimized / External-Cache / Internal-Cache / UDP/IP — DEC 5K, DEC Alpha, HP 712, HP 730, IBM RS6k, Intel 486, Intel Pentium, SGI, Sun Sp2, Sun 10/51, Sun 20/61