ATM as the Interconnect for NOW

Who Cares - "It's the End-Station"

Gaurav Garg

SynOptics Communications, Inc.

October 4, 1994

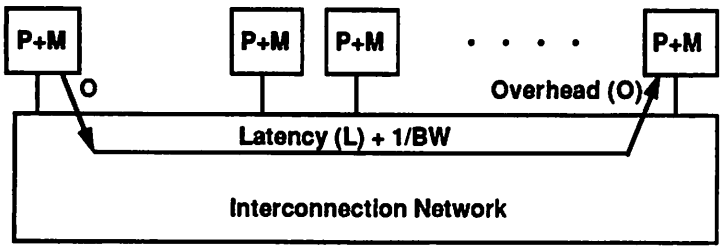
SynOptics

Outline

- NOW - Model & Assumptions
- ATM in NOW - Rationale & Issues
- NOW & Networks
- ATM Tomorrow
- NOW & End-Stations
- What is effective BW ?
- Conclusions

SynOptics

NOW Communication Model



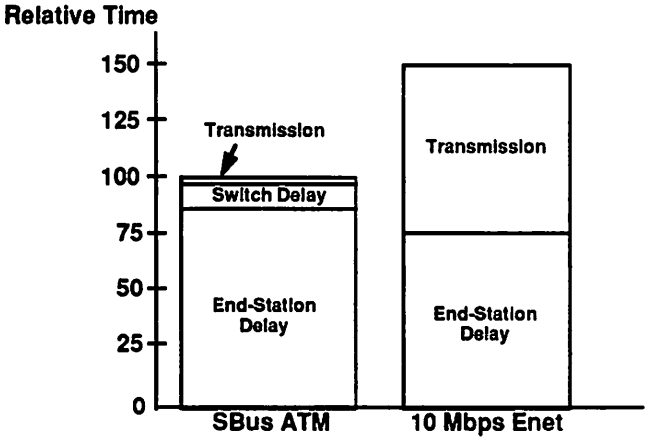
- Interconnect for NOW
  Ultimate Goal is a solution that maximizes effective bps/$, with an acceptable lower limit, given network application - NOW's AM paradigm
- NOW Transport Protocol designed for MPPs
  Request Response Protocol
  Message/Window Size About 20 Bytes
  Emphasizes Latency because :

  Effective BW Approx. = 160 Bits/RT Latency at Protocol Layer

SynOptics

Latency & Overhead in LANs

- Graphs below based on NFS traces over a week



SynOptics

NOW Commercial Assumptions

- NOW Justified based on cost assumptions of near-commodity Components
- Use near-commodity equipment for computing and network resources shipped in 100'000s, or Millions. Cost/Performance superior to Specialized products.
- Use of fixed individual resources of many in shared pool of computing resources impies use of desktop infrastructure (including building wiring).
- Goal is to not force individuals to join collective when using individual machine, but to harness idle resources.
- All of above imply standard HW and infrastructure, with SW converting between operating models.

Motivation for ATM (in NOW)

- MPP Networks Switched -> Go to most promising, most hyped, commercial switched network interconnection technology
- High Speed - 155, 622 Mbps etc.
- Predictable Response Time - Fixed Cell Size, BW allocation, Switching
- Cell size suited to Active Messages, unlike TCP/IP, where 48 byte cell size results in a loss of nearly 50% of the bandwidth for TCP Acks
- Scalable from low speed to high speed due to fixed cell size & switching; allows easy migration of 10 Mbps Ethernet ports also
- Simpler, cheaper switching model when using fixed size cells
- Full Duplex allows for Flow Control in a switched environment
- VCIs simple handle for switching, searching, forwarding, and VLANS

Commercial Status of LAN ATM

- **ATM Switches**
  Switches shipping at 155, 45, and 25 Mbps speeds
  Cost at 155 - Shipping $1500/Port, Announced $800/Port
  Cost at 25 Mbps - Shipping at $500/Port, Announced $300/Port
- **ATM Adapters**
  Shipping at 155 and 25 Mbps speeds
  Cost at 155 - Shipping $1000/Port, Announced $700/Port
  Cost at 25 Mbps - Shipping at $400/Port, Announced $300/Port
- **Internetworking**
  Very large number of Ethernet to ATM, Token Ring to ATM, Router Interface products announced and/or shipping
- **Industry Size**
  At least 100 companies including most large communications companies have announced and/or are already shipping ATM products
- **ATM Forum**
  700 Members currently - Largest standards group in the World dedicated to 1 topic

ATM Implementations Today

- **Latency : Switches & End-Stations**
  - **Switches not necessarily optimized for latency**
  - **SONET adds 12-15 us one-way**
  - **Adapter Card SW Overhead is PRIMARY system bottleneck**
- **Throughput/Loss**
  - **Today's switches don't do anything to address flow control; depending on TCP. Cell Loss in AAL5 magnifies Packet Loss.**
  - **End-Station with standard I/O Bus and protocol stack supports small number of messages/sec.**
- **Cost**
  - **Reasonably high compared to Ethernet today**

The Good News

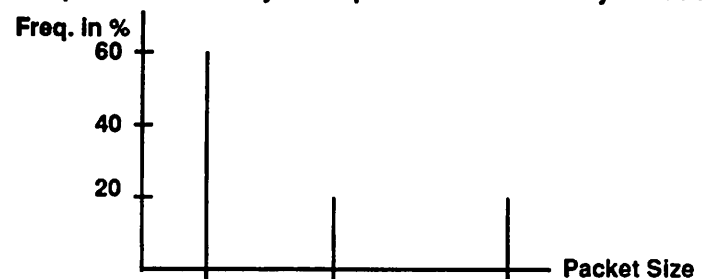
- Latency
  - 5-7 us ATM Processing latency is achievable
  - SONET delay to be eliminated via proposals at the ATM Forum
  - Protected systems such as NOW can blow away all those expensive (in overhead) protocol layers (as in MPP environments)
- Throughput/Loss
  - Rate based congestion/flow control approach for ABR (LAN) Traffic Management Adopted by ATM Forum 104-9 last week in Ottawa. Approach shows high utilization and fairness in strenuous simulation environments
- Cost
  - Innovation (industry R&D) driving cost down much faster than technology curve already - when will NOW be commercially viable ?
  - Believe that massive R&D investment will drive down cost much faster than most other technologies
  - Nothing in ATM to make it comparatively expensive

It's the End Station !

- End Station Latency, over standard I/O Bus and protocol stack is the biggest bottleneck in the system.
- Many specialized network interconnects not only address SW issues, but also move end station interconnect to graphics or even Processor-Memory bus. Why not try this with standard LANs ?
- Since MPPs use standard processors and memory, it should be possible to optimize NOW interconnect paradigm at end station in a manner similar to that used in MPPs.

LAN Traffic Patterns

- Strongly Bimodal Packet Sizes
- Small Packets typically
  - Protocol Acks
  - Slow keystroke driven applications
- Sometimes a Natural Application Unit (NAU) affects average packet size.
  Example : Patient Query in Hospital stored as 1000 Byte Unit by Database.

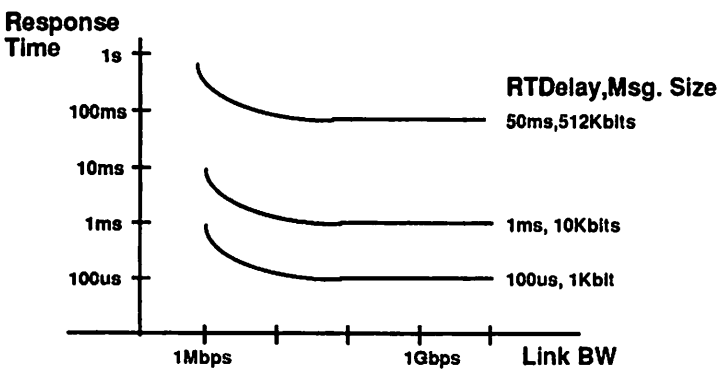


Ethernet	64 Bytes	1500 Bytes
NFS	64 Bytes	8K Bytes

SynOptics

Latency/BW Tradeoff

- User/Application/Protocol only care about response time to command
- Used Scaled Avg. Packet Size from previous page here.
  Examples :
  - 1Kbit (avg. packet in LANs)
  - 10Kbits (Large Enet Packet)
  - 512Kbits (64 KB from TCP Window)



SynOptics

Latency/BW Tradeoff (Con't)

- Effective BW = $\frac{\text{Avg. Message Size or Protocol Window Size}}{\text{Round Trip Delay at Protocol Stack}}$

- Typical Round-Trip Latency/Overhead distribution :
  - Network : 1-20 us
  - Adapter & HW transfers : 20-30 us
  - Protocol Stack including overhead : 300-800 us

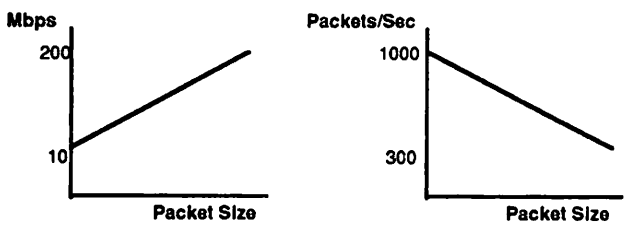
- Lesson : Dimension Control Loop at Protocol Layer appropriately;
  - Set Protocol Window size to Round-Trip Delay BW product.
  - TCP adjusts this, whereas many other protocols maintain a static size.

- Is this a show-stopper for NOW ?
  - NOW messaging model based upon specialized networks
  - Can NOW messaging model work on standard high volume LAN HW equipment with SW changes only ?
  - May affect NOW cost/performance assumptions in the future
  - NFS is an example of a protocol optimized for 1985

LAN Performance Measures

- LOG (latency,serial overhead, Bus BW) in End-Station limit LAN BW
- Serial Overhead is biggest factor in End-Station

- Two Views of End-Station limited BW for high-speed networks



- Packets/Second has higher predictability than Mbps

- Packet/Second is more relevant :
  - Protocol Acks, and Keystroke Apps are low BW
  - Bulk Transfers or NAUs need high BW
  - NAU is equal to Packet Size - thus PPS more appropriate

Conclusions

- **NOW must use widely available commercial computer and interconnect (including infrastructure) equipment to make a significant contribution.**
- **If switching is appropriate for NOW, ATM is the right answer for a variety of reasons - innovation, scalability, flow control etc.**
- **Address most frequent problems first in both network & end station**
- **May have contradiction between NOW messaging model and standard LAN interconnection HW. Hope that SW changes are adequate.**
- **When will NOW be viable commercially ? What technologies will be at the right place in the cost/performance space at that time ?**