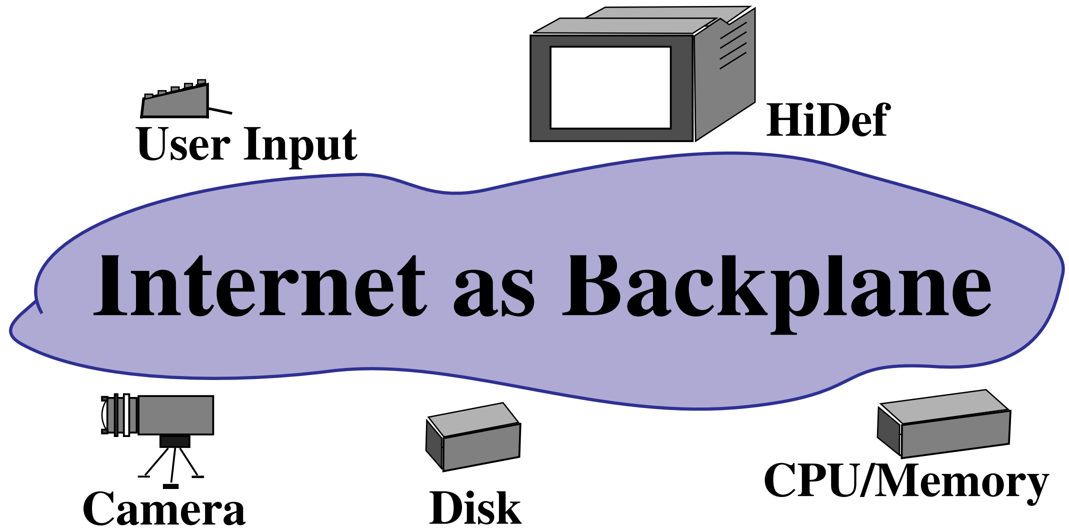
ATOMIC-2 & Netstation Projects

Gregory Finn, Joe Touch, Annette DeSchon, Ted Faber
Steve Hotz, Rodney Van Meter, Bruce Parham

Outline

1. ATOMIC-2 Project Overview
    2. Removing Protocol & OS Limits
    3. Accomplishments
4. Netstation Project Overview
    5. Derived Virtual Devices --- Why?
    6. Achieving Cut-through
    Performance for Internet protocols
    7. Multicomputer Node to Internet Node
    8. Planned Development

4. Netstation Project Overview



Tie workstation parts together via the Internet . . .

- Network-centered rather than bus-centered architecture

Areas and Issues

1. Presentation and control of network-attached peripherals

   RPC access program + RPC control programs --- one for each 'device'

   Derived virtual device (DVD) is what gets presented to the network.
   DVD is a synthetic device that fits between a client and the actual device.

   Each DVD has its own RPC program execution context

2. Performance --- Some devices require very fast RPCs

   Develop techniques that significantly improve RPC performance.
   --- RPC protocol - DTP - 35μs app-to-app, reliably across Myrinet

   --- Eliminate latency cost of embedded Internet checksumming

3. Security --- network-attached peripherals are Internet hosts

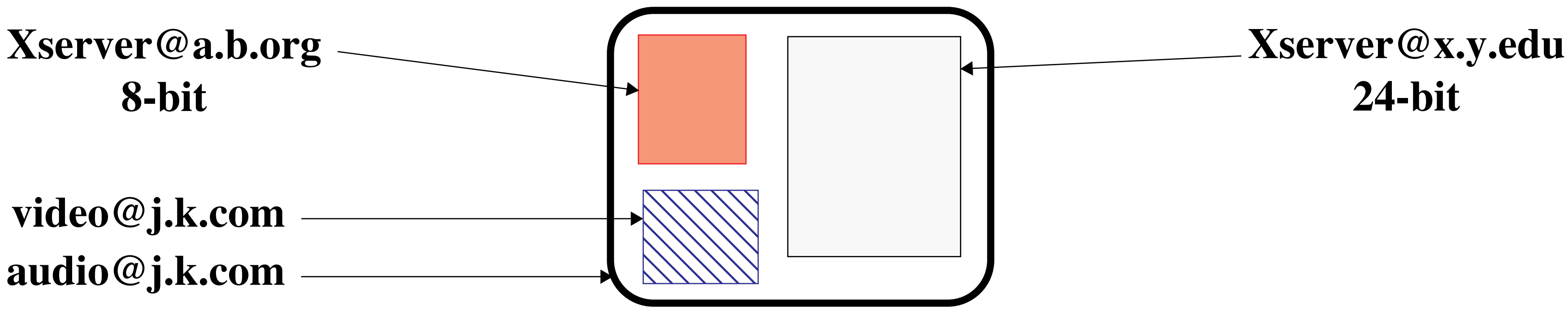
   Access-control lists + Kerberos authentication

4. Assembly --- DNS-like configuration database

5. Derived Virtual Devices - DVDs - Why?

Flexibility and Simplicity:

Single physical device can act as multiple virtual devices.
Multiple physical devices can act as single virtual device.



Integrity & Policy:

A DVD has a specific execution context that strictly limits client access to the underlying physical device(s).

Multiple Access:

DVDs let device owner share its device in a controlled manner.

Flexibility: Adapt to Client Needs

Actual frame buffer: **1200x1600x24 bit/pixel**

Client requests: **1024x1280x8 bit/pixel frame buffer**

Create an 8-bit frame buffer DVD with its RPC context.
<src, dst, port> is bound to that context and client is informed.

Simplicity: Present Only What's Needed

Frame buffer: RPCs are BitBlt(), DrawRegion(), ColorMap()

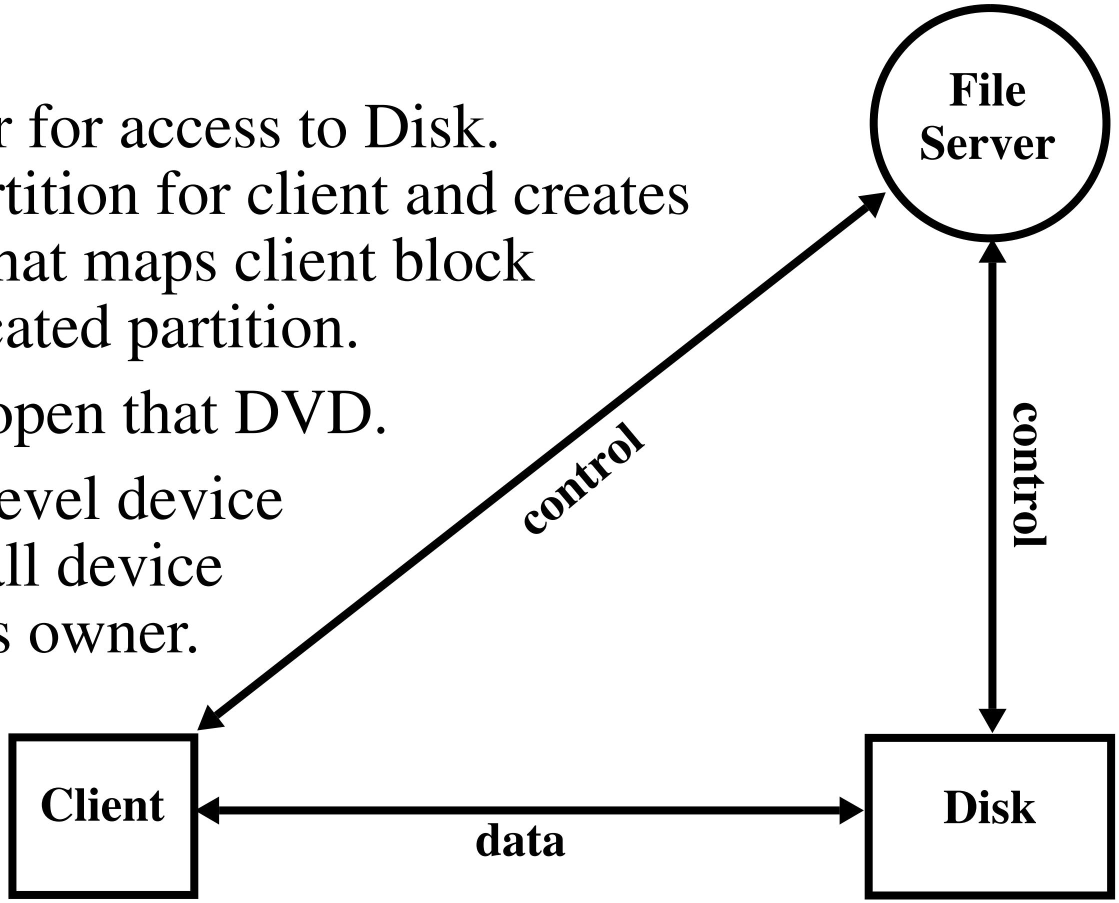
Netstation X Server as Client

Scroll (350 Kpixel) Window	**Mbone Video**	
Packet rate: 600 pkt/sec	**Packet rate: 875 pkt/sec**	
Data rate: 640 Kpixel/sec	**Data rate: 1.2 Mpixel/sec**	**~10 Mb/s @ 8 bit**
		~30 Mb/s @ 24 bit

Integrity and Multiple-Access: Third-Party I/O

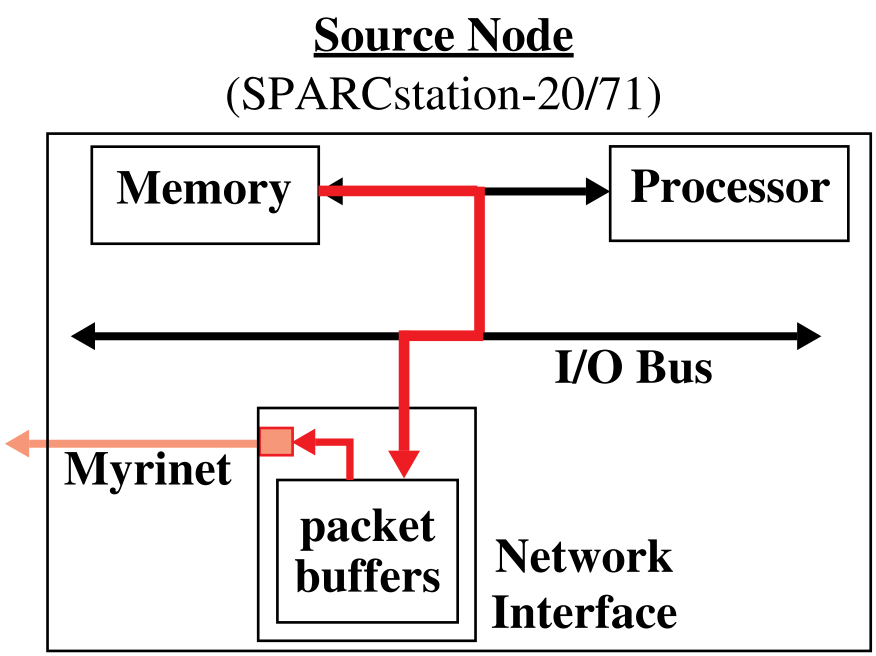
Provide efficient, block-level access to disks for multiple clients while ensuring integrity for all users.

1. Client asks File Server for access to Disk. File Server creates partition for client and creates a Disk DVD context that maps client block requests onto the allocated partition.
2. File server lets client open that DVD.
3. Client ‘sees’ a block-level device but cannot harm overall device integrity for the Disk’s owner.



6. Achieving Cut-through Performance for Internet Protocols

Cut-through transfer significantly reduces latency.



For 1024 byte packet:

Copy	Xmit	Prop	Total	
22 µs	13 µs	2 µs	37 µs	LAN
--	13 µs	2 µs	15 µs	1 km
22 µs	13 µs	20 µs	55 µs	Campus
--	13 µs	20 µs	33 µs	2 mi
22 µs	13 µs	200 µs	235 µs	MAN
--	13 µs	200 µs	213 µs	20 mi

Special device memory ⇒ copy packets into buffers.
Embedded transport-layer checksum ⇒ scan pass or copy.

Both prevent the application from using cut-through.

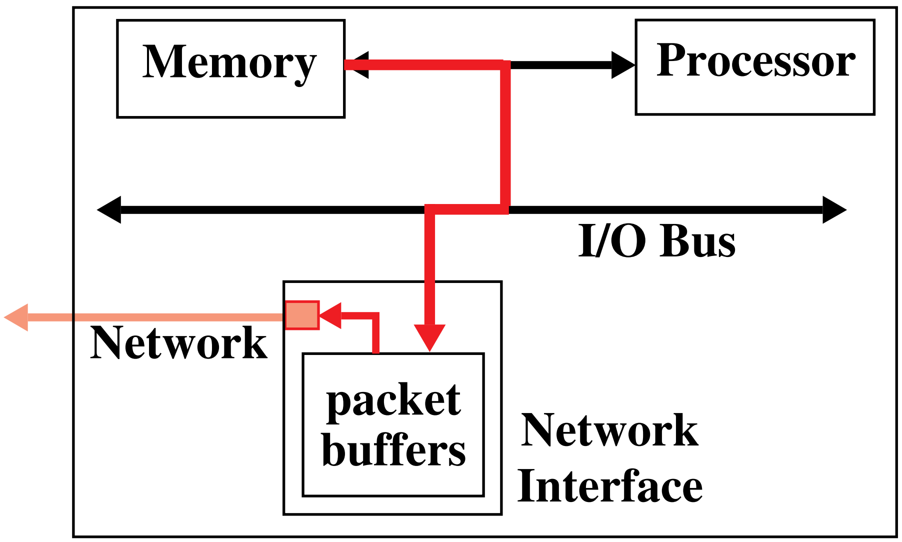
First Cut Solution

Have applications use the special device memory directly.

Buffer memory can be quite limited.

Don’t use complete TCP/UDP mechanisms within LAN.

DTP uses link CRC to avoid checksumming latency.

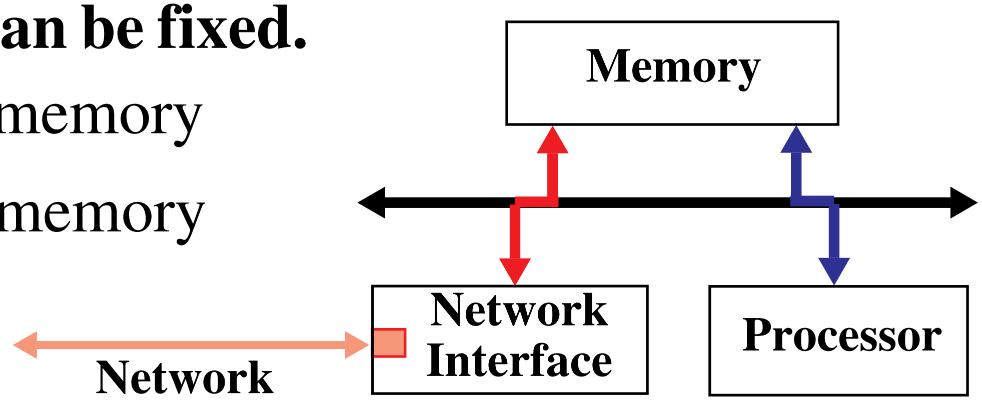


Result: You can achieve cut-through, but it may be a Pyrrhic victory.

What Next?

The interface architecture can be fixed.

1. Greatly increase device memory
2. Interface can use system memory



Embedded Checksum Still Stands in the Way

Switch to trailing end-to-end checksums?

Requires redefinition of TCP, UDP, . . . not likely.

Desired Solution

Keep embedded end-to-end checksum and remove the need to scan the payload prior to transmission.

Characteristic:

Must be transparent to the Internet.
$\Rightarrow$ No redefinition of TCP/UDP protocols

Observation

Transparency requires correct packet formats over the Internet. But within a source network we can be more flexible.

Example: Postpone embedding the transport checksum until reaching (1) initial gateway or (2) local destination

Zero-Pass Checksumming

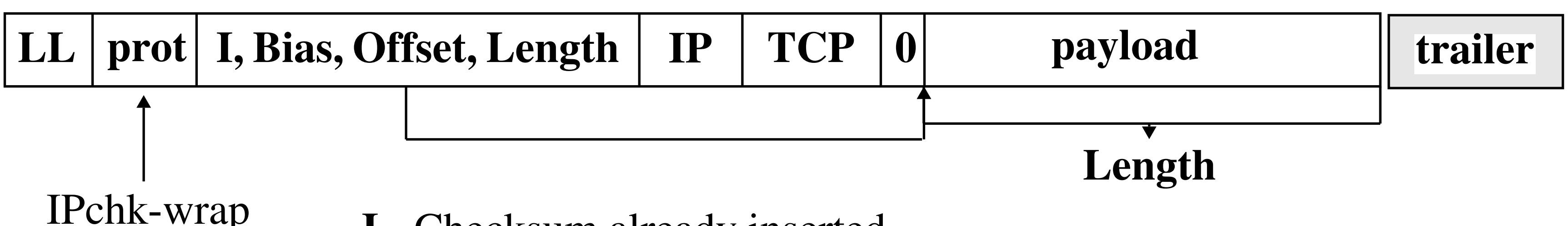
Place Internet checksum logic into the network interface

+

Postpone embedding of the end-to-end checksum

Example:

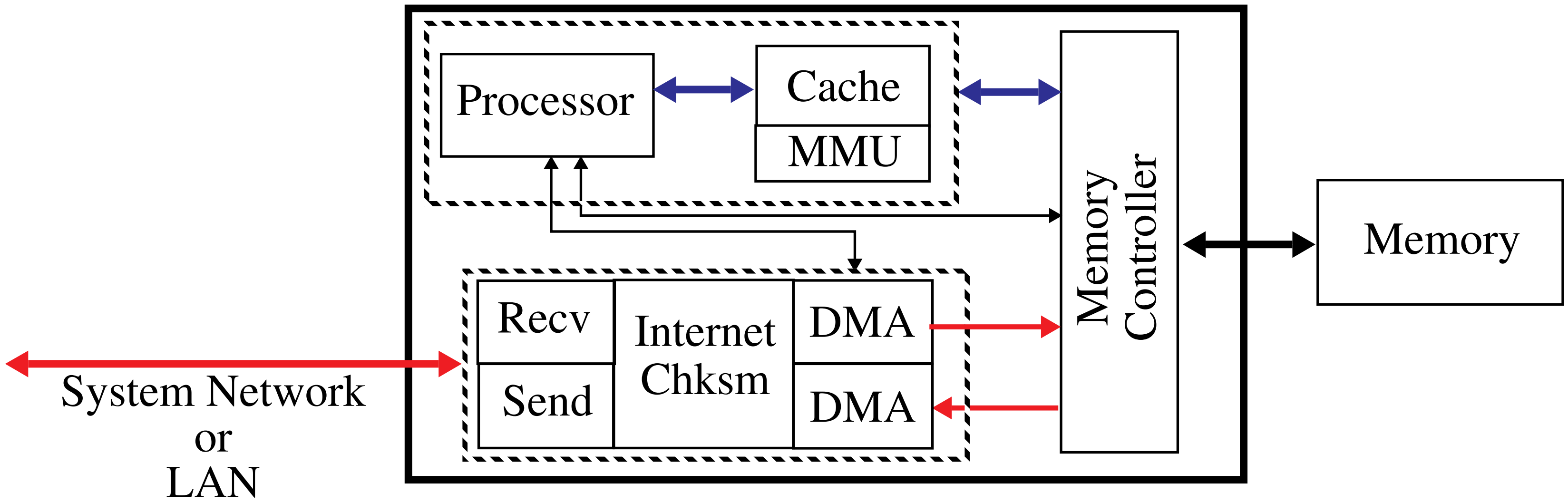
trailer + embedding at initial gateway



I - Checksum already inserted

Bias - Checksum initialization value (pseudo-header)

7. Multicomputer Node to Internet Node



Cut-through and Encryption

One can envision encryption logic in future network interfaces.

Payload checksumming and encryption can occur in parallel.

Encapsulated Security Payload (ESP) encrypts transport-layer header as well as payload. This may preclude cut-through.

8. Planned Development This Year

Security

Incorporate Kerberos as part of DVD access.

User-Input Node

Desktop hub around which to cluster low and high-speed devices.

- Myrinet switch + Myrinet port slots + PCMCIA slots
- Keyboard, Mouse, PCMCIA support

Camera Node

Full rate video, raw frame and JPEG compression.

Disk Node

Study efficiency and safety of third-party I/O using DVD model.