VISA: Netstation’s Virtual Internet SCSI Adapter

Rodney Van Meter

USC/Information Sciences Institute

rdv@isi.edu

http://www.isi.edu/netstation/

July 15, 1997

Talk Outline

- Netstation
- STORM & Derived Virtual Devices
- IP for NAPs
- VISA
- Conclusion

The Netstation Project

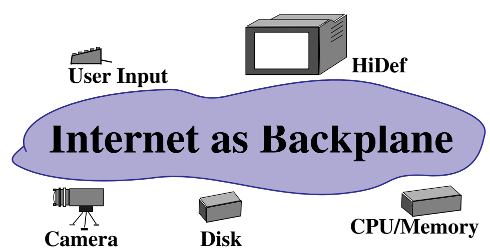
Gregory Finn (project leader), Rodney Van Meter, Steve Hotz, (Bruce Parham and Reza Rejaie)

Objective:

Overcome fixed bus-induced limitations by utilizing improved scaling properties offered by gigabit networking.

Netstation

Netstation is a system composed of **network-attached peripherals (NAPs)** created by replacing the system bus in a workstation with a gigabit network.



- Use Internet protocols for ubiquitous device access
- Based on ATOMIC 640 Mbps switched network

Why Netstation?

- Traditional buses don't scale in distance or bandwidth.
- Support efficient device-to-device transfer without consuming resources at main CPU.
  - ❏ e.g., incoming video data direct to display.
- Construct systems flexibly.

Netstation Problems Faced

Closed, bus-centric architecture allows simplifying assumptions about resource identification, security and sharing.

- Set of resources not constrained by architecture.
- Control of devices not limited to bus master.
- Non-dedicated network.
- Security now paramount.

Accomplishments

- DTP: 30,000 RPCs/sec
- Network-Attached Peripheral (NAP) Security Model: Derived Virtual Devices (DVDs)
- Netstation Display
- X on Netstation Display
- Zero-Pass Checksumming
- Netstation Keyboard
- ZCAV Disk Work

Derived Virtual Devices

A **derived virtual device** (DVD) is an *execution context* at a network virtual device (NVD); i.e. a set of resources and procedures to access them.

DVD concept provides a mechanism to support safe sharing of resources.

- Enforces resource bounds checking.
- Constrains operation functionality (e.g., read only).
- Checks authentication of **user**.

Who a request is from is **much** more important than *where*.

Current Work

- DVD Implementation w/ Kerberos
- IP Disk
- VISA: Virtual Internet SCSI Adapter
- STORM: A DVD File System
- Third Party Transfer
- Netstation Camera
- More Network Protocols (mostly TCP)

Related Work

- MIT Viewstation
- Cambridge Desk Area Network
- SGI Origin 2000?
- CMU Network-Attached Secure Disk (NASD)
- LLNL's Network-Attached Peripheral (NAP) RAID
- National Storage Industry Consortium's NASD Committee
- Fibre Channel Disk Drives
- Palladio at HP Labs

Talk Outline

- Netstation
- **IP for NAPs**
- VISA
- Conclusion

Networking Problems for NAPs

HiPPI-6400❋gigabit Ethernet❋Myrinet❋FC-AL

as I/O Nets Get Larger and More Complex:

- Media Bridging
  (Routing, Addressing)
- Congestion
- Flow Control
- Demultiplexing @ Endpoints
  (Destination Address Calculation, Control/Data Sifting, Upper Layer Protocols)
- Latency Variation
- Security
- Reliability
- Heterogeneity
  (Hosts, Traffic Types, Nets)

All Become Bigger Problems!
But...

1394❋HiPPI-800❋ATM❋SSA❋Fibre Channel

The Internet Community Has Solved Most of the Problems

- Strengths of IP: Issues of Scale and Heterogeneity
- Weakness: Performance

Advantages of IP

- Heterogeneous Interconnects
  - Intra-Machine Room
- Wide-Area Access
  - Enables Remote Mirroring and Backups
- Future Growth
  - Not Media-Specific
- Lower R&D Investment in Networking

Solving TCP/IP Performance Problems

Protocols

- Larger Packets (IPv6 MTU discovery)
- Zero-Pass Checksumming

Host Implementation

- Zero-Copy TCP (IPv6 Flow IDs)
- Early Demultiplexing

Device Controller

- Link & CPU Speeds Climbing Faster than Device Transfer Rate
- Implementation Can be Simple
- Scatter-Gather Real Memory Interface

Transport Layer Issues

- Want to Retain TCP's Reliability and Flow Control
- Need Application Framing

Application Layer Issues

- RPC Formatting
- App-Directed Out-of-Order Delivery

Conclusions

- IP Offers Significant Benefits with Little Cost
- Some Transport Issues are Still Open

Talk Outline

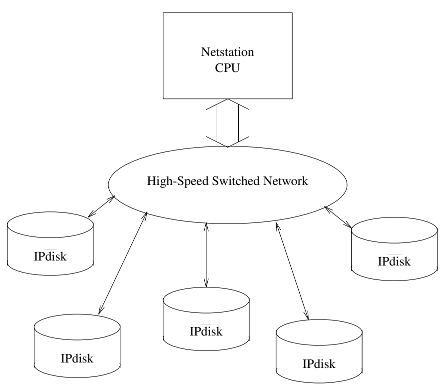
- Netstation
- IP for NAPs
- **VISA**
- Conclusion

VISA: Virtual Internet SCSI Adapter

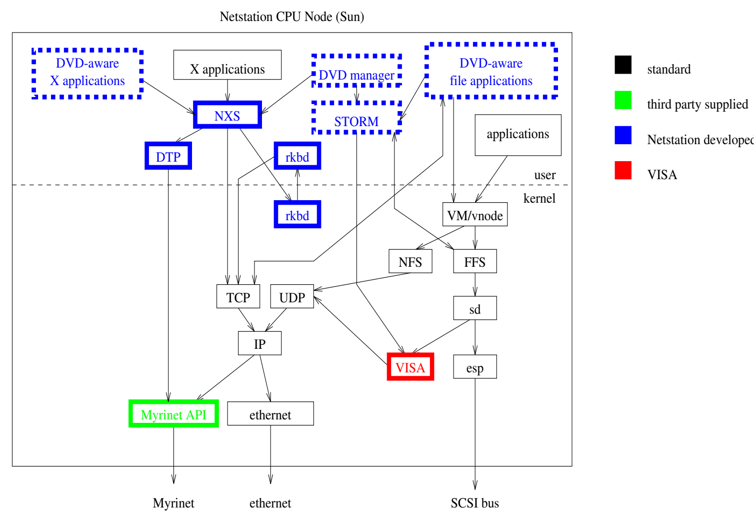
Goals:

- Demonstrate IP Acceptable for Peripherals
  - •Device Performance
  - •CPU Load
- Single Host to Device
- Platform for:
  - •Further NAP Protocol Research
  - •DVD/STORM
  - •Security
  - •Third Party Transfer
- Proof of Concept, not Production

Experimental Configuration



Netstation System Components



Architecture

What

- Sends SCSI RPCs, Receives Data via Network
- Accesses IPdisk
- SunOS 4.1.3 scsi_transport Layer
- Meshes with SCSI-3

How

- Simple Reliability over UDP
- Single-threaded Pseudo-process
- Prefetch/Consolidation Handled in FS Code Above
- Single Command per Target -- No Command Queueing

IPdisk

- Emulates Disk NAP as User Process
- Four Types of Store:
  - •RAM (done)
  - •File (coming soon)
  - •SCSI Disk (coming soon)
  - •SAM Solid State Disk (later)
- UDP (simple reliability) or TCP
- Third-party SCSI COPY w/ DVDs Planned

Transport Protocol

- Simplest Possible Reliability over UDP
- Fixed Size (Negotiated?):
  - •Packet Size (8KB)
  - •Window (48KB)
- Handles Errors, but not Efficiently
- Assumptions:
  - •Low Latency LAN
  - •In-Order Arrival
  - •Highly Reliable

Early Results

- 67 Mbps Write, 60 Mbps Read Through File System
  (Sparc 20/71, Myrinet, 8KB pkts, 48KB Window)
- Currently Limited by:
  - •CPU at IPdisk
  - •Brain-dead Reliability & Limited Buffering
- Compares to:
  - •60 Mbps NFS
  - •107 Mbps TCP Blast
  - •135 Mbps UDP Blast (8KB pkts)
- Requests up to 248KB Seen

Comparison

SCSI Bus

- Sun 4 ~1991: >75 Mbps SCSI Raw Device
- SCSI Coprocessors Very Effective
- Few Interrupts

VISA/IP for Disks

- 67 Mbps Through the File System
- Network Coprocessors not very Effective
- LOTS More Interrupts
- Lower Channel Efficiency not an Issue

Lessons Learned

- SunOS Layered/OO Modularity Made VISA Possible
- SCSI Configuration Happens EXTREMELY Early in Boot:
  - •No Timers, No Mbufs, No Networking
  - •Fake Device Config
  - •Kernel Rework Necessary to Correctly Identify Devices
- Packet Size Important (as Expected)
- CPU Load Significant Due to:
  - •Packet Overhead
  - •Extra Data Copies
  - •Underpowered/Underutilized Coprocessors

Future VISA Work

- Clean Up (Multi-Device Support, etc.)
- DVD Integration
- Transport Protocol:
  - •TCP
  - •Better Custom
  - •Preferred Framing/ACK Patterns
  - •Acceptable Assumptions
- Performance Measurement:
  - •CPU Utilization
  - •Macro FS Effects (File Create Time, Seeks, etc.)
  - •Paging & Raw Disk Performance
  - •Comparison of Same Disk Locally & via IP
  - •Host Saturation Point
- Test w/ Other Device Types (Tape Drive?)
- Fast Demultiplexing/Copy Reduction

Future Netstation Work

- STORM (STORage Manager)
  3rd-party capable FS w/ DVD mgmt
- Camera
- 3rd-Party Transfer
- Kerberos Integration

Conclusions

- Netstation: Exploring Space of Network-Based Architecture
- Virtual Internet SCSI Adapter (VISA)
  Working, Results Pending
- Assertion that IP for NAPs is:
  - Possible -- Done
  - Appropriate -- not yet Complete
- **`http://www.isi.edu/netstation/`**