SPINE

An Operating System for Intelligent Network Adapters

Marc E. Fiuczynski
Brian N. Bershad
Department of CS&E
University of Washington

Richard P. Martin
David E. Culler
Department of EE&CS
University of California Berkeley

http://www.cs.washington.edu/research/spine

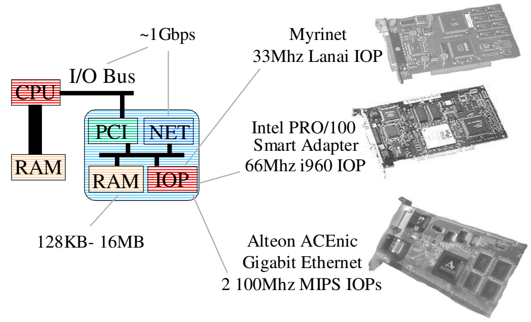
Observation

- Computer Systems are fancy communication devices
  - Multimedia clients
  - IP forwarding/routing
  - Firewalls
  - Archive and Retrieval (file-, web-, multimedia-servers)
- Unbalanced increase in system performance
  - CPU, Link bandwidth, and disk capacity steadily increasing
  - I/O subsystem lags in scalability and performance increase

SPINE -- Intelligent I/O Processing

- Problem: I/O performance is not tracking workstation performance
- Implication: Need to stay off the I/O bus as much as possible
- Opportunity: Many I/O controllers come with their own CPUs
- Solution: Migrate application-specific processing directly on to the I/O card itself
- Strategy: Develop an extensible runtime system for I/O controllers

Technology trends argue for smarter I/O devices



Strategy

- Migrate functionality to network adapter
  - enable application to act on data closer to media
  - reduce I/O related control & data transfers to host
- Techniques to improve performance
  - Device to device transfers
  - Host/Device protocol partitioning
  - Device level memory management

Three Key Questions

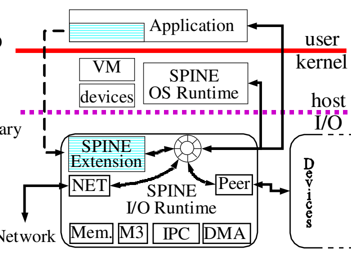
- What kind of system software support is necessary?
- Can we really partition applications?
- Are the slower I/O processors fast enough?

SPINE Properties

- Performance
  – single address space design
  – low overhead access to system services
  – exchange data with peer devices w/o host support
- Run-time adaptation
  – applications dynamically load extensions onto the intelligent adapter
- Safety and Fault Isolation
  – type-safe language & defensive interface design

System Structure

Dynamically load extensions onto I/O device to avoid control and data transfer across the host I/O bus boundary



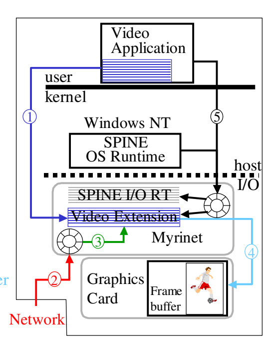
System Characteristics

- Message Driven Scheduling
- Active Message Communication Layer
- *"Terminate with prejudice"* scheduling
- Reference counting GC

Do Applications Partition?

Device-to-Device Transfers • Video Client • IP forwarding • 3D distributed RT rendering	Packet Manipulations • Fire walls • Filtering • Accounting
Remote Memory Systems • GMS & Network RAM • treadmarks, quarks (DSM)	Operating System Services • Quality of Service • Lazy Receive Processing • Scout Paths
High Perf. Networking • Checksum offloading • TCP/IP offloading • VIA	Management & Availability • SNMP - statistics/configuration • Crash detection

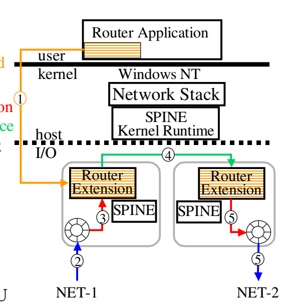
SPINE Video



1. Download extension to I/O card
2. Video data arrives from Network
3. Dispatch to the Video Extension
4. DMA video image to frame buffer
5. Window events sent to extension

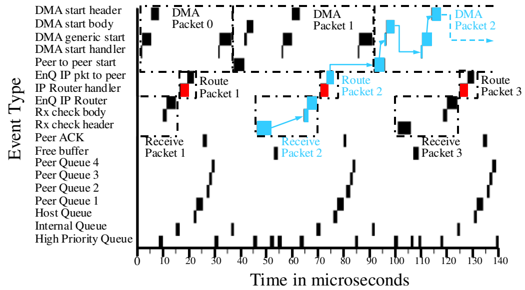
SPINE IP Forwarding

1. Download extension to I/O card
2. IP packet arrives from NET-1
3. Dispatch to the Router Extension
4. Transfer IP packet to peer device
5/6. Send IP packet out on NET-2



SPINE Performance:
11,300 PPS + zero host CPU
Host base implementation:
12,000 PPS + 100% host CPU

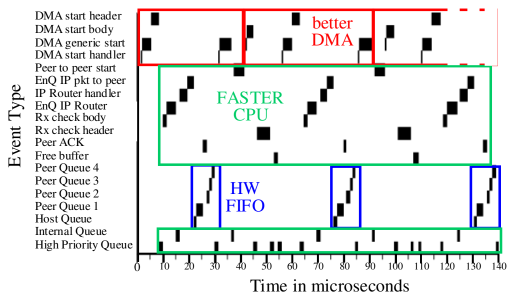
SPINE Event Processing for IP Forwarding Extension



Hardware Enhancements

- What we had to use: 33 MHz Lanai Processor
  - 1989 processor technology (think Sparc-1)
  - no I/D cache
  - highly contended shared memory resource
  - software emulation of msg FIFOs & DMA control
- What would help:
  - hardware FIFOs
  - chained & strided DMA
  - recent Processor technology
  - higher memory bandwidth

SPINE Event Processing with improved hardware support



Summary

- Intelligent Front-end I/O devices are nothing new
  - IBM Channel Controllers, Cray Y-MP, Auspex NFS server
- Processors used in most major system components
  - Disks: NASDs, IDISK, Active Disks (Database Machines)
  - Graphics: Phillips Trimedia, Equator VLIW, …
  - NICs: Alteon Tigon2, 3COM X?X?, …
- SPINE is an OS for Intelligent Network Adapters
  - example applications work even with *slow I/O processors*
  - faster I/O processor and hardware structures enable other interesting applications

http://www.cs.washington.edu/research/spine

Performance

- Video client supports up to 4 video streams at 40Mbps
  - 0% host processor utilization
- Router handles 11300 packets per second (PPS)
  - 0% processor utilization
  - Compare to ~12000 PPS for a host-centric router using the same hardware (100% CPU utilization for a 200 MHz Pentium)