Crossbow
Network Virtualization & Resource Management

www.opensolaris.org/os/project/crossbow

www.opensolaris.org/os/project/vnm

Real Scenarios

Financial Services

- *Trading house starts offering free financial information to attract customers*
- *Brokerage customers start complaining that trading site slows down*
- *The paying customers start deserting*

Large ISP

- *ISP wants to deploy virtual systems on same physical machines*
- *ISP sells each virtual system at different price levels to its customers*
- *Any virtual instance can overwhelmed the shared networking resource*

Enterprise Computing

- *A large company uses a workgroup server for day to day as well as critical traffic*
- *IT Ops doing non critical stuff started a backup over the network*
- *Users doing time critical work can't get bandwidth to do their job*

What Happened?

- *Critical services are overwhelmed by non-critical services, traffic types, or virtual systems*
- *No usable mechanism available for fairness, priority and resource control for networking bandwidth*

Network Landscape Challenges

High Bandwidth Demand

> High volume of data sharing over network
> More applications require remote data access
> Scalability, network expands quickly

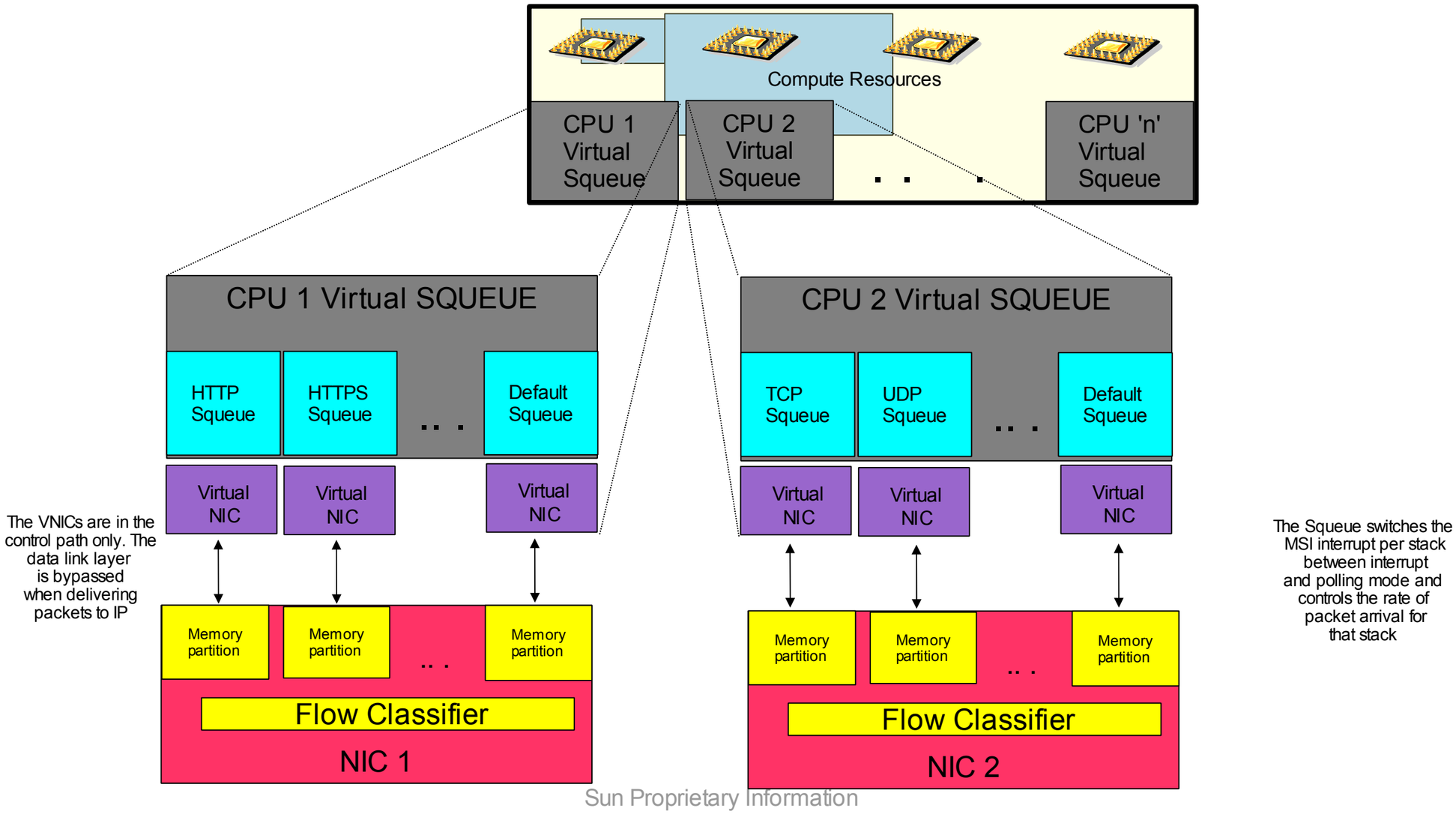
Network Management Complexity

> Too many different technologies: Ethernet, IB, FC, etc.
> Too many network cables and equipment

Server and Network Consolidation

> Too many servers
> Consolidate multiple servers into one
> Consolidate multiple network connections to one link

Virtual Stacks



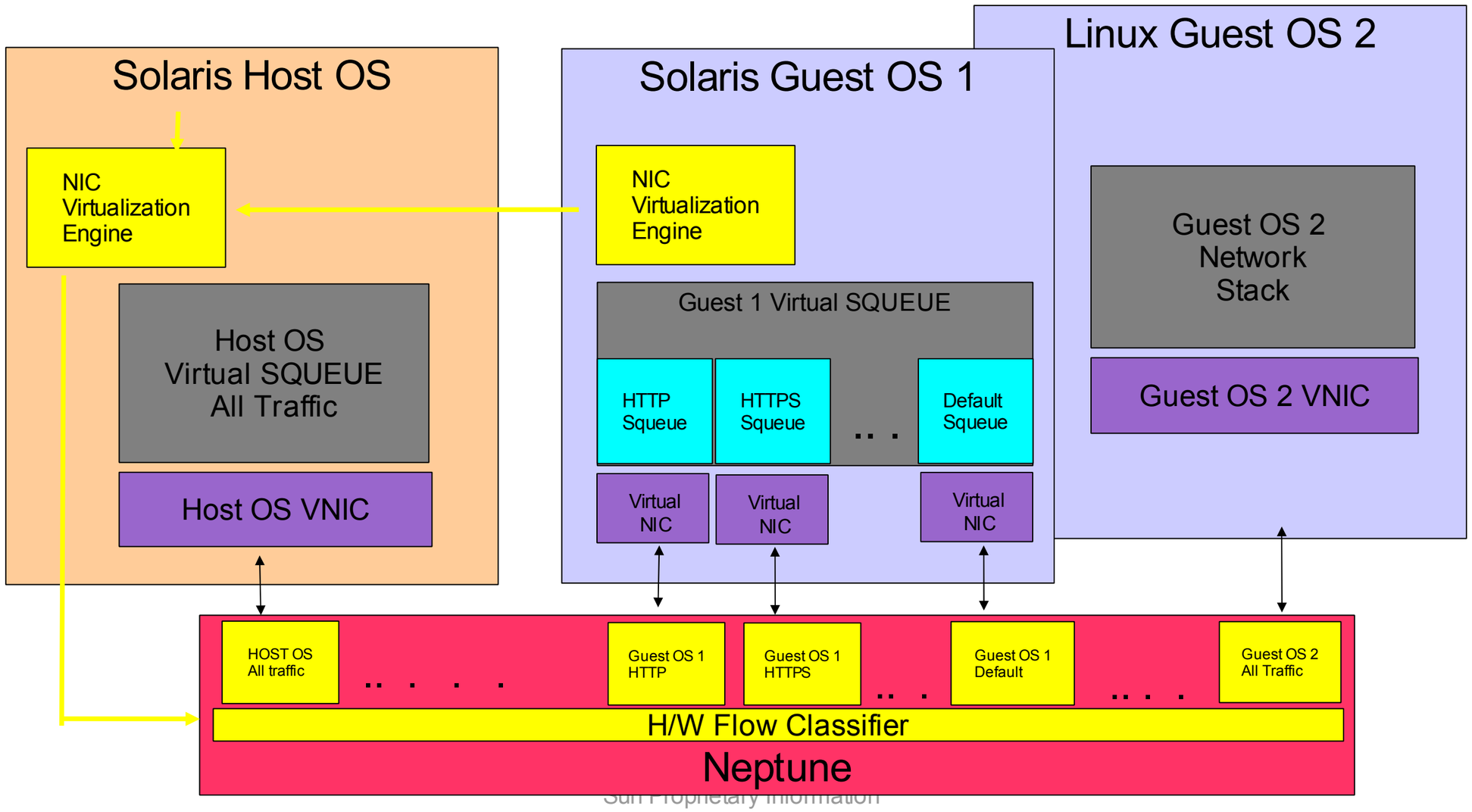
The VNICs are in the control path only. The data link layer is bypassed when delivering packets to IP

The Squeue switches the MSI interrupt per stack between interrupt and polling mode and controls the rate of packet arrival for that stack

The Crossbow Architecture

- Divide NIC memory, DMA channels, etc and use a flow classifier to build a virtual stack on each H/W partition
- Each Virtual NIC is owned by the FireEngine Squeue's which independently switch the VNIC between interrupt & polling mode
- Rate of packet arrival from a VNIC is independently controlled by the Squeue owning the VNIC

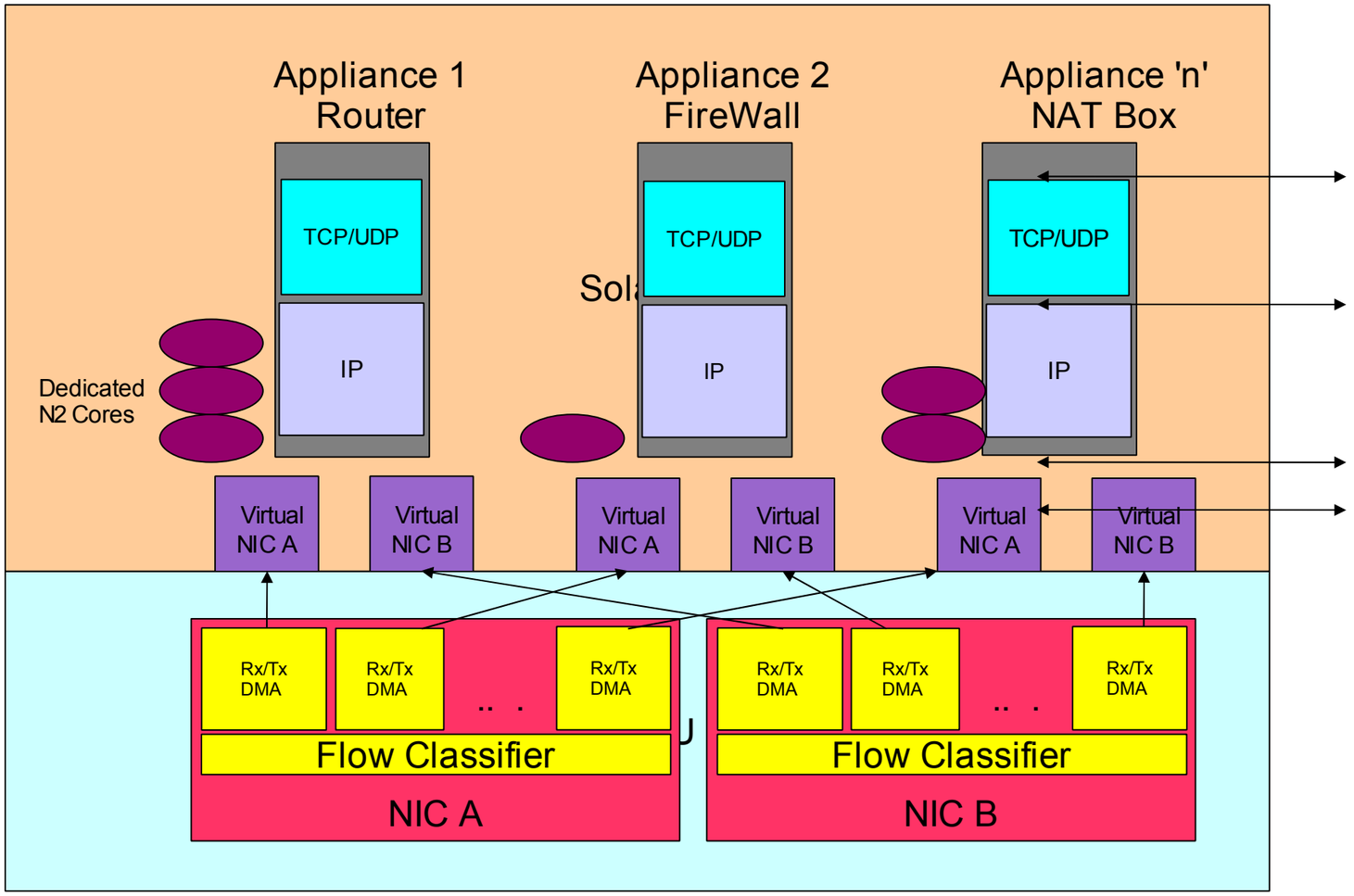
Crossbow:
Virtual Machines



Crossbow Class of Service

- **Glues parts of NIC (Rx/Tx Rings, DMA channels, interrupts) with Niagara strands or cores and assigns dedicated kernel threads, queues and other resources**
  - Spread packets from H/W and provide the “lane” discipline with dedicated H/W & S/W resources all the way from recv to xmit
  - Class of service per “lane” without performance overheads where critical traffic like VOIP can be tied to a high performance “lane” with more resources

Virtual Appliances



- Consolidate Various Network Appliances
- Dedicated B/W, kernel and H/W resources per appliances
- Developer Entry Points via Packet Event Framework

Defense against DOS/DDOS

- DDOS have the ability to cripple the entire grids and all services offered by them
- Only the impacted services or virtual machine takes the hit instead of the entire grid
- Under attack, impacted services start all new connections under lower priority (limited resource) stack
- Connections transition to appropriate priority stacks after application authentication

Crossbow and Lightweight Solaris Containers

Virtualization

- Virtual Stacks & VNICs for Zones
- Virtual Stacks & VNICs are totally independent of each other

Resource Control

- Without resource control, TCP stream between xb2 and zone xb1-z1 gets clobbered by UDP stream from xb3 to Zone xb1-z2

Observability

- Watch real time usage for each VNIC