Real Routing at Gigabit Speeds

Debra Deutsch
Product Line Manager, Layer 3 Switching
Prominet Corporation
Design Drivers

- Market Requirements
  - Integrated Layer 2 and Layer 3 product line
  - Scale from small to high capacity configurations
  - Low latency for selected traffic; easy to use
  - Deliver full wire speed performance

- Technology
  - Division of function among Cajun Switch core ASICS
  - Non-blocking crossbar switch
  - Dual-queue, weighted round-robin queuing
  - Cajun ASICS, Intelligent Cache, undersubscribed switch

*This presentation will focus on the areas highlighted in gray*
System Architecture

Packet Flow
Packet Routing ASIC and Address Cache

- Complete IP and IPX packet processing
  - Designed to a cycle budget
  - Some parallelism
  - Filtering, prioritization in silicon
- Address cache in SSRAM for speed, modularity, cost
  - 20,000 fine-grained entries manage RSVP flows
  - Proprietary fast look-up and management algorithms
    - chose against longest-prefix matching
Interconnects, Queues, and QOS

- Choice of interconnect and mechanisms for QOS are interdependent
- Prominet’s solutions
  - Switching fabric
  - Support 8 queues in architecture, implement 2
- How did we arrive at this design?
  - Technical alternatives
  - Market and product requirements
Shared Memory Interconnect

● A memory pool connects all ports
● Memory bandwidth at least 2x the sum of port bandwidths; controller speed is 2x the sum of port packet rates
● Disadvantages
  » Memory speed limits ability to scale aggregate bandwidth
  » Large-scale design not cost-effective in small unit implementation
● Advantages
  » Minimizes packet copying (low latency)
  » Controller can be a smart queue manager

Example: 16 ports @ 1 Gbps
Memory: 4 ns, 128 bits wide
Controller: 48 million packets per second
Crossbar Interconnect

- Crossbar provides network of paths between ports
- Use a combination of input and output queues, run fabric fast to counteract head-of-line blocking

Disadvantages
- Design must address head-of-line blocking

Advantages
- Scales to high aggregate bandwidths
- Relative simplicity
How Many Queues?

<table>
<thead>
<tr>
<th>Latency</th>
<th>Utilization</th>
</tr>
</thead>
<tbody>
<tr>
<td>System Minimum Latency</td>
<td></td>
</tr>
<tr>
<td>Best Effort Traffic One Queue</td>
<td></td>
</tr>
<tr>
<td>Best Effort Traffic Multiple Queues</td>
<td></td>
</tr>
<tr>
<td>Priority Traffic Two Queues</td>
<td></td>
</tr>
<tr>
<td>Priority Traffic Third Queue</td>
<td></td>
</tr>
<tr>
<td>Priority Traffic Fourth Queue</td>
<td></td>
</tr>
</tbody>
</table>

See Difference Between 1 and 2 Queues Here

See Difference Between 2 and 4 Queues Here
LAN Backbone Utilization

<table>
<thead>
<tr>
<th>Percent Loaded</th>
<th>Percent of Respondents</th>
</tr>
</thead>
<tbody>
<tr>
<td>up to 25%</td>
<td>41</td>
</tr>
<tr>
<td>26-50%</td>
<td>29</td>
</tr>
<tr>
<td>51-75%</td>
<td>17</td>
</tr>
<tr>
<td>76-100%</td>
<td>8</td>
</tr>
<tr>
<td>Don't know</td>
<td>5</td>
</tr>
</tbody>
</table>

source: User Plans for High Performance LANs

copyright 1997, Infonetics Research, Inc.
Prominet P550 Queuing Performance

Mean Delay for Low and High Priority Queues

Delay

Offered Load

Low (90/10)
Low (70/30)
Low (50/50)
High (90/10)
High (70/30)
High (50/50)

System Minimum Latency < 10 us
Going Even Faster

More ports
- Consider size of master routing table
- No change to packet path on media modules and supervisor
- Use a larger crossbar
  » datapath speed unchanged
  » more ports
  » faster controller to allocate more ports in same cycle time

Faster ports
- Need larger address tables at routing ASICs
- Parallelism can get us close to 10 Gbps today
- We are using “sweet spot” clock rates
- Semiconductor speeds double every 3 years
- Prominet would likely use an architecture very similar to what we have today!
Conclusions

Gigabit-scaled routing is simply a matter of great engineering!
- Many feasible design choices
  - Work around memory speed limitations
- Latitude to match technology to market requirements

Network managers should
- Plan to use routers where routing is the right answer
- Expect interoperable, cost-effective systems
- Take advantage of inexpensive capacity to simplify campus network management