Multi-gigabit Switching and Routing

Gignet ’97 — Europe: June 12, 1997.

Nick McKeown
Assistant Professor of Electrical Engineering and Computer Science
nickm@ee.stanford.edu
http://ee.stanford.edu/~nickm

Multigigabit and Terabits

1. How Routers have Evolved
2. Multigigabit Routing
3. Terabit Switching and Routing
The Evolution of Routers

The first shared memory routers
The Evolution of Routers

Reducing the number of bus copies

Routing CPU

Buffer Memory

DMA

Route Cache

Buffer Memory

MAC

Line Card

Routing CPU

Buffer Memory

DMA

Route Cache

Buffer Memory

MAC

Line Card

updates

Multigigabit and Terabit Switching and Routing  Page 5 of 19

Multigigabit and Terabit Switching and Routing  Page 6 of 19
The Evolution of Routers

Avoiding bus contention

- Memory
- Buffer
- ROUTE
- CPU

Advantage:
- Non-blocking backplane—high throughput

Disadvantage:
- Difficult to provide QoS

1. How Routers have Evolved
2. Multigigabit Routing
3. Terabit Switching and Routing
Multigigabit Routing

BBN’s Multigigabit Router

Routing Lookups

<table>
<thead>
<tr>
<th>Class A</th>
<th>Class B</th>
<th>Class C</th>
<th>D</th>
</tr>
</thead>
</table>

212.17.9.4

Routing Table:

- 212.17.9.0 Port 4
CIDR uses “longest matching prefix” routing:

CIDR uses “longest matching prefix” routing:

```
212.0.0.0/8
212.17.0.0/16
212.17.9.0/24
```

Hashing, caching and pipelining are hard!

Solution 1: Label Swapping

IP Switching, Tag Switching, ARIS, Cell-switched Router,....
Solution 2:
Perform Lookups Faster!

Solution 2 (cont):
20 million lookups per second

<table>
<thead>
<tr>
<th>Prefix length</th>
<th>Number in routing table</th>
</tr>
</thead>
<tbody>
<tr>
<td>24</td>
<td>0 2 32 -1 256</td>
</tr>
<tr>
<td>212.17.9.0/24</td>
<td>0 2 32 -1 256</td>
</tr>
<tr>
<td>212.17.9.4</td>
<td>0 2 32 -1 256</td>
</tr>
<tr>
<td>232-1</td>
<td>0 2 32 -1 256</td>
</tr>
</tbody>
</table>
1. How Routers have Evolved
2. Multigigabit Routing
3. Terabit Switching and Routing

The Tiny Tera

“Soda can” switch core
32x32 switch, ~16Gbps per port
Aggregate bandwidth: 0.5Tbps

Requires high speed chip-to-chip links.
Schedulers must be fast, fair and efficient.
High Bandwidth Parallel Datapath

The Tiny Tera Port Architecture
The Tiny Tera

http://tiny-tera.stanford.edu/tiny-tera/

32 ports, 16 Gb/s per port.
Input-queued architecture.
High bandwidth parallel datapath.
Efficient unicast and multicast.
Four priority levels.
Fixed and variable length packets.
Tag switching.
Asymmetric high-speed serial links.