Bandwidth, Throughput and Goodput

- Bandwidth – how much data can pass through a channel.
- Throughput – how much data actually travels through a channel.
- Goodput is often referred to as application level throughput.

But bandwidth can be limited below link’s capacity and vary over time, throughput can be measured differently from bandwidth etc.....
• Higher bandwidth does not necessarily mean higher speed

• E.g., can mean the aggregation of links

• $100G = 2 \times 50G$ or $4 \times 25G$ or $10 \times 10G$

• A very common practice in interconnects
Packet Rate

- Throughput may change under different conditions, e.g. packet size
- Packet Rate: how many packets can be processed in a given amount of time
  - Also changes under different conditions
  - But often provides better insights
Switch Models

- A perfect fluid mental model:
Switch Models

• A single packet mental model:
Circuit Switches

- Input A is connected to output X
- Example: a crossbar
  - Not the only option
- Used mostly in optical switching
  - No header processing!
- But also in electrical switching
  - E.g., high frequency trading (HFT)
- Scheduling is a limiting factor
Packet Switches

• In a circuit switch:
  The path of a sample is determined at time of connection establishment
• In a packet switch, packets carry a destination field
• Need to look up destination port on-the-fly
  • Two sequential packets may head to different destinations
To achieve high throughput, packet switches are pipelined:
Store and Forward

- Wait for the entire packet to arrive
- Check the FCS, then start processing
  - FCS – frame check sequence, terminates the packet
- Once the packet is checked, it starts propagating through the pipeline
  - Not necessarily the entire packet
Cut Through

- Start processing the packet as soon as the first chunk arrives
  - Does not wait for the FCS
- If FCS error is detected, the packet is dropped somewhere along the pipeline
Measuring Performance

• Bandwidth: number of bits (or bytes) through the channel every unit of time

• One way to calculate: \textit{bus width} \times \textit{clock frequency}
Measuring Performance

Throughput = clock frequency x bus width?

<table>
<thead>
<tr>
<th>PACKET 512B</th>
<th>Data path Width e.g. 256B</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>256B</td>
</tr>
<tr>
<td></td>
<td>256B</td>
</tr>
</tbody>
</table>

CLOCK CYCLE1

CLOCK CYCLE2
Measuring Performance

Throughput ≠ clock frequency x bus width!
Example: NetFPGA SUME – Switch Throughput

NetFPGA SUME Switch Throughput

Throughput [Gbps]

Packet Size [B]

120MHz

150MHz

180MHz
Performance Profile

- An aggregation of the performance estimates of a device
- Considering both I/O and modules
- Helps identify bottlenecks
- Informs design decisions, e.g.:
  - Bus width
  - Clock frequency
Packet Rate “on the wire”

- Calculating the packet rate on (e.g.,) 10GE port:

- Line rate = 10.3125Gbps

- Line coding 64b/66b
  \[10.3125G \times \frac{64}{66} = 10Gbps\]

- Before each Ethernet packet there is 8B preamble and 12B Inter Packet Gap (IPG)
  \[\text{Data rate: } 10.3125G \times \frac{64}{66} \times \text{(packet size)} / \text{(packet size + 12B +8B)}\]

- Divide it by packet size and convert bytes to bits:
  \[\text{Packet rate: } 10.3125G \times \frac{64}{66} / \left(\text{(packet size + 12B +8B)} \times 8\right)\]
Packet rate within a device

• “Cycles per packet”
  – How many cycles are required to process a single packet within a module?

• Example 1: Cycles required to fit a packet into the data bus

• On NetFPGA SUME:
  – 10G Port: 64b (8B) wide
    – 64B packet: 8 clocks
    – 65B packet: 9 clocks
  – Data path: 256b (32B) wide
    – 64B packet: 2 clocks
    – 65B packet: 3 clocks
Packet rate within a device

• Why 64b bus in the 10G port?
  – $64b \times 156.25\text{MHz} \text{ clock} = 10\text{Gbps}$
  – Also, 64b fits the 64b/66b coding
    – No need for a gear box

• Why 256b bus in the data path?
  – Need to process data from $4 \times 10\text{G ports}$
    – $4 \times 64b = 256b$

• Important! Packet size may differ between modules
  – 10G port packet size includes FCS, in the data path FCS is not included
    – 4B difference
Packet rate within a device

• “Cycles per packet”
  – How many cycles are required to process a single packet within a module?

• Example 2: Packet processing
  – How many cycles are required to process the packet’s header?
    – E.g. Look up in the memory
    – Actions that are not pipelined

• Example 3: Scheduling / Arbitration
  – How many cycles are required to schedule a packet?
    – E.g. can a new packet enter the arbiter every clock cycle?
## Performance profile - example

- 10G Port clock – 156.25MHz
- Data path clock – 200MHz

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>64</td>
<td>14.88</td>
<td>8</td>
<td>19.53</td>
<td>59.52</td>
<td>2</td>
<td>100</td>
<td>1.68</td>
</tr>
<tr>
<td>65</td>
<td>14.71</td>
<td>9</td>
<td>17.36</td>
<td>58.82</td>
<td>3</td>
<td>66.66</td>
<td>1.14</td>
</tr>
<tr>
<td>66</td>
<td>14.53</td>
<td>9</td>
<td>17.36</td>
<td>58.14</td>
<td>3</td>
<td>66.66</td>
<td>1.15</td>
</tr>
<tr>
<td>…</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
### Performance profile - example

- **10G Port clock** – 156.25MHz
- **Data path clock** – 150MHz

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>64</td>
<td>14.88</td>
<td>8</td>
<td>19.53</td>
<td>59.52</td>
<td>2</td>
<td>75</td>
<td>1.26</td>
</tr>
<tr>
<td>65</td>
<td>14.71</td>
<td>9</td>
<td>17.36</td>
<td>58.82</td>
<td>3</td>
<td>50</td>
<td>0.85</td>
</tr>
<tr>
<td>66</td>
<td>14.53</td>
<td>9</td>
<td>17.36</td>
<td>58.14</td>
<td>3</td>
<td>50</td>
<td>0.86</td>
</tr>
<tr>
<td>...</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Example: NetFPGA SUME – Switch Throughput

NetFPGA SUME Switch Throughput

Throughput [Gbps]

Packet Size [B]

120MHz  150MHz  180MHz
Packet rate within a device

• Question:

• 10G Port supports full line rate at 156.25MHz

• Data path supports full line rate for 64B packets at 150MHz (measured!)

• But data path input is 4 x 10G port…

• How do we get 100% throughput at a lower data path frequency?

• And what did we neglect?
The Truth About Switch Silicon Design

12.8Tbps Switches!

Let's convert this to packet rate requirements:

5.8 Gpps @ 256B

19.2 Gpps @ 64B

But clock rate is only ~1GHz....
Multi-Core Switch Design

Barefoot Tofino

Broadcom Tomahawk 3

https://www.nextplatform.com/2018/01/20/flattening-networks-budgets-400g-ethernet/