CMOS Crossbar

Ting Wu, Chi-Ying Tsui, Mounir Hamdi

Hong Kong University of Science & Technology Hong Kong SAR, China



OUTLINE

- Motivations
- Problems of Designing Large Crossbar
- Our Approach Pipelined MUX Core
- Interface Link and Clocking Design
- Conclusions



Motivations

- Advances in fiber optic link technology and WDM have made raw bandwidth in abundance
- Switches/Routers are replacing the transmission link as the bottleneck of the network
- Switches/Routers with high speed (OC-192, 10Gb/s) and large number of I/O ports (128*128 or 256*256) are becoming a necessity
- Key issues for designing high speed router:
 - Switching Fabric Interconnects
 - Queuing Schemes
 - Arbiters
 - Multicast

IKUST

Fabric Interconnects: Crossbar

- Crossbar (Crosspoint) Fabric is becoming the preferable interconnect fabric for high speed and scalable switching
- It has been proven that crossbar (even input-queued) can have as high throughput as any switch.
- A crossbar inherently supports multicast efficiently.
- QoS can be implemented reasonably easy.

IKUST

- The key challenge is the scalability for high line rate and large number of ports
- CMOS technology can achieve high density and low cost

Architecture of the CMOS Crossbar Switch

- Crossbar Switch Core fulfills the switch/router function
- Controller configures the crossbar core switching
- High speed data link communicates between switch fabric and line card
- PLL provides on-chip precise clock





Two Approaches to Build the Core

• X-Y Based Crossbar



- Scalability : N²
- Speed : limited by Cap at input and output lines
- Control: N² bits





- Scalability : N²
- Speed : limited by Cap only at input line
- Control: N*Log₂N bits

Problems of Designing Large Crossbar Switch

- The switch core scales to N²
- Design complexity increases
- The performance requirement increases much faster than that can be achieved through CMOS technology scaling
- The throughput can be satisfied by using multiple bit-slices (e.g. 8) of the core, however, the core size increases by 8 times
- Wire delay is also substantial in high performance chip



Our Approach – Pipelined MUX Crossbar Digital Core

- Digital MUX tree based design technique guarantees the high performance as well as the low design complexity
- In order to integrate large crossbar switch, only 2 bit-slices are embedded in the digital core instead of 8 (60%+ area saving)
- 1GHz digital core is demanded for the 2 Gb/s interface, the MUX tree can be pipelined to fulfill the requirement
- Additional pipeline stage is added to drive long wire



SDFF embedded with MUX



- High performance Semi-Dynamic Flip-Flop (SDFF) is used [Klass98, Stojanovic et.al. 99]
 - One of the fastest Flip-Flops due to negative setup time
 - Little overhead for embedding with MUX function



Pipeline Stages Partition



The pipeline of the 256-to-1 MUX can be partitioned as

 Natural: 16-to-1 MUX in 1st stage; 16-to-1 MUX in 2nd stage
 Balanced: 8-to-1 MUX in 1st stage; 32-to-1 MUX 2nd stage

Driving Long Wire – Adding Repeater cannot Satisfy the 1GHz Requirement

- The 1st stage is critical due to the large capacitor at the input line CLK
- Distributed R-C wire model is employed

HKUST

- Repeater can be inserted to reduce the wire delay
- For 256 ports, even inserting the optimal size and number of repeater, the delay is still larger than 1ns



2

Repeater number

3

0

0

Adding One Pipeline Stage to Drive Long Wire

- Add one more stage for driving the long wire by inserting a Flip-Flop
- The whole 256*256 crossbar is divided into 4 128*128 -- sub-crossbar, so that the input line only need to drive 128 cells instead of 256
- For 128 ports, sub-ns delay time is achievable



3-stages Pipelined MUX Crossbar Floor-Planning

- The 256*256
 crossbar consists of
 4 sub-crossbars
 (128*128) running
 at 1GHz frequency
- 2 pipeline stages in each sub-crossbar
- 2 bit-slices are embedded matching with 2Gb/s data link



3-stages Pipelined MUX Crossbar Timing Diagram

- In sub-crossbar 0, inputs [0:127] are switching in the 1st and 2nd stages, while in sub-crossbar 3, inputs [128:255] are switching in the 2nd and 3rd stages
- Finally, the two groups of outputs are fed into SDFF_embedded with 2-to-1 MUX to complete the 256-to-1 MUX action



The Sub-Crossbar Circuits Simulation Results





Technology		TSMC 0.25µm SCN5M Deep	
Sub-crossbar No.		Sub-crossbar0	Sub-crossbar3
Post layout simulation	1st stage	959 ps	845 ps
	2 nd stage	866 ps	959 ps
	3 rd stage	236 ps	866 ps

Control Circuits

- Control bits are used to configure the corresponding MUX in the crossbar pipeline in the correct timing stage.
- For saving the pin counts, the control inputs are embedded within the data inputs, each incoming frame packet includes one byte control word and 64 bits of data
- The timing constraints can be satisfied by careful pipelining the control path



Control Circuits (cont'd)

 Bang-Bang PD: samples the 2Gb/s inputs, converts to 2bits, each at 1Gb/s

HKUST

 Re-synchronization: synchronizes each input signal to the main clock



- DMUX: demultiplexs signal to data & control bits
- Counter: counts 4/36 and controls the DMUX

Full Crossbar Core Layout and Specification



Full 256*256 crossbar core with 2 bit-slices

- Technology:
 - TSMC 0.25µm SCN5M Deep, 5 Layer Metal
- Layout size:

14 mm*8 mm

- **Transistor counts**: 2000k
- Supply voltage: 2.5v
- Clock frequency: 1GHz
- **Power**: 40W



Interface Link and Clocking Design

- The dual loop delay locked loop (DLL) design technique is adopted in the data link for data and clock recovery
- The main analog DLL generates multiple clock phases for the interpolation in the full digital periphery loop
- A half rate bang-bang phase detector is used in the periphery loop to sample the 2Gb/s incoming signal by using 1GHz clock
- A 3rd loop, an analog PLL, provides the 1GHz on-chip clock



Interface Link and Clocking Design

- The system clock is at 250MHz, PLL provides the precise 1GHz clock for the whole chip
- Several periphery loops share one analog DLL



Conclusions

- A 2Gb/s 256*256 CMOS Crossbar Switch Core is achievable with current process technology
- Significant area saving is obtained by using only 2 bit-slices in the crossbar switch core
- 3-stages pipelined MUX circuit is proposed to decrease the cycle time to less than 1ns
- Post layout simulation results show that each stage can run at a clock rate higher than 1GHz
- Full 256*256 crossbar core has been laid out to demonstrate the design
- PLL & dual DLL circuits have been designed for the clocking and high speed link in the whole chip