USC

USC Viterbi
School of Engineering

Scalable High-Throughput SRAM-based Architecture for IP-lookup Using FPGA

Hoang Le, Weirong Jiang, Viktor K. Prasanna
Department of Electrical Engineering
University of Southern California, USA

FPL 2008 **Sep. 23, 2008**

Our solution

USC Viterbi School of Engineering

- **Dual Linear Pipeline Architecture for IP Lookup (DuPI)**
  - Linear pipeline
  - Memory sizes proportional to the sizes of supported routing tables
- **Complete binary tree structure (BST)**
  - Addresses of the child nodes eliminated
  - Amount of memory doubled in the next level
  - Number of pipeline stages determined by the size of the supported routing tables
- **Large routing tables with external SRAMs**
  - Last few stages moved onto external SRAMs
  - Number of external stages depends on number of I/O pins

Pre-processing

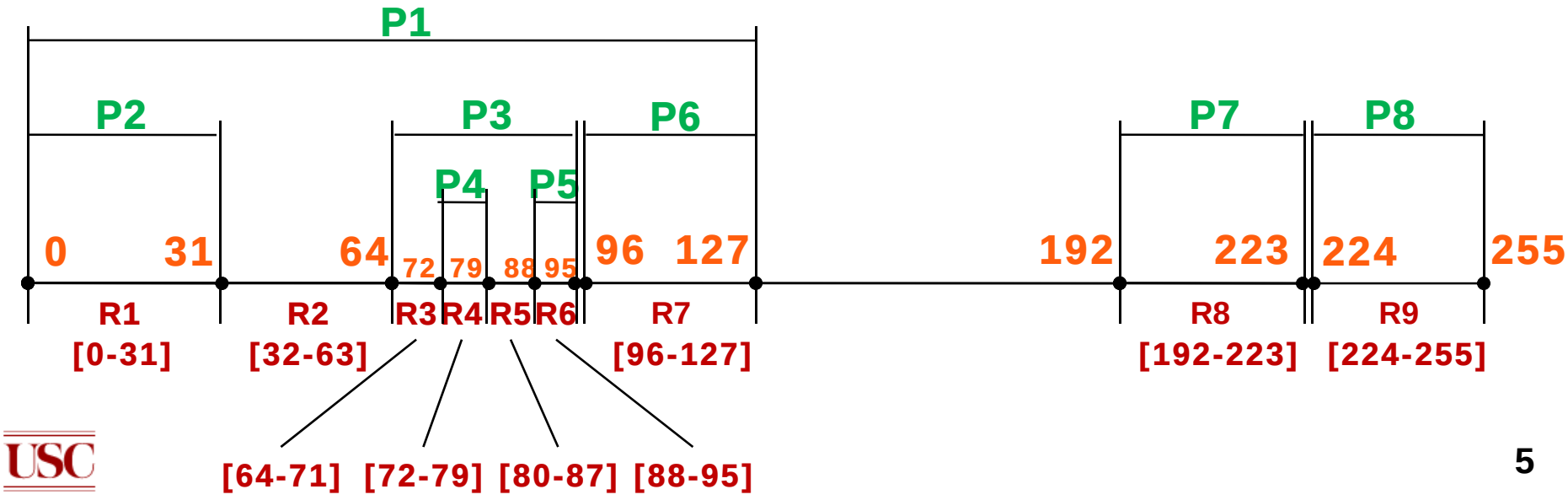
USC Viterbi School of Engineering

- **Convert given set of prefixes to a set of non-overlapping ranges:**
  - **Convert prefixes to ranges**
    - Pad '0' for lower bound, '1' for upper bound
  - **Map all ranges onto [0 – (2^{32}-1)]**
  - **Algorithm to cut overlapped ranges**
    - Scan from left to right, cut at any prefix's bound
    - Assign routing index of the longest prefix among the overlapping prefixes to the cut range
- **Set of *N* prefixes**
  - **At most *2N* distinct bounds**
  - ***N* non-overlapping ranges (best case)**
  - **(*2N* -1) non-overlapping ranges (worst case)**

Pre-processing (example)

USC Viterbi School of Engineering

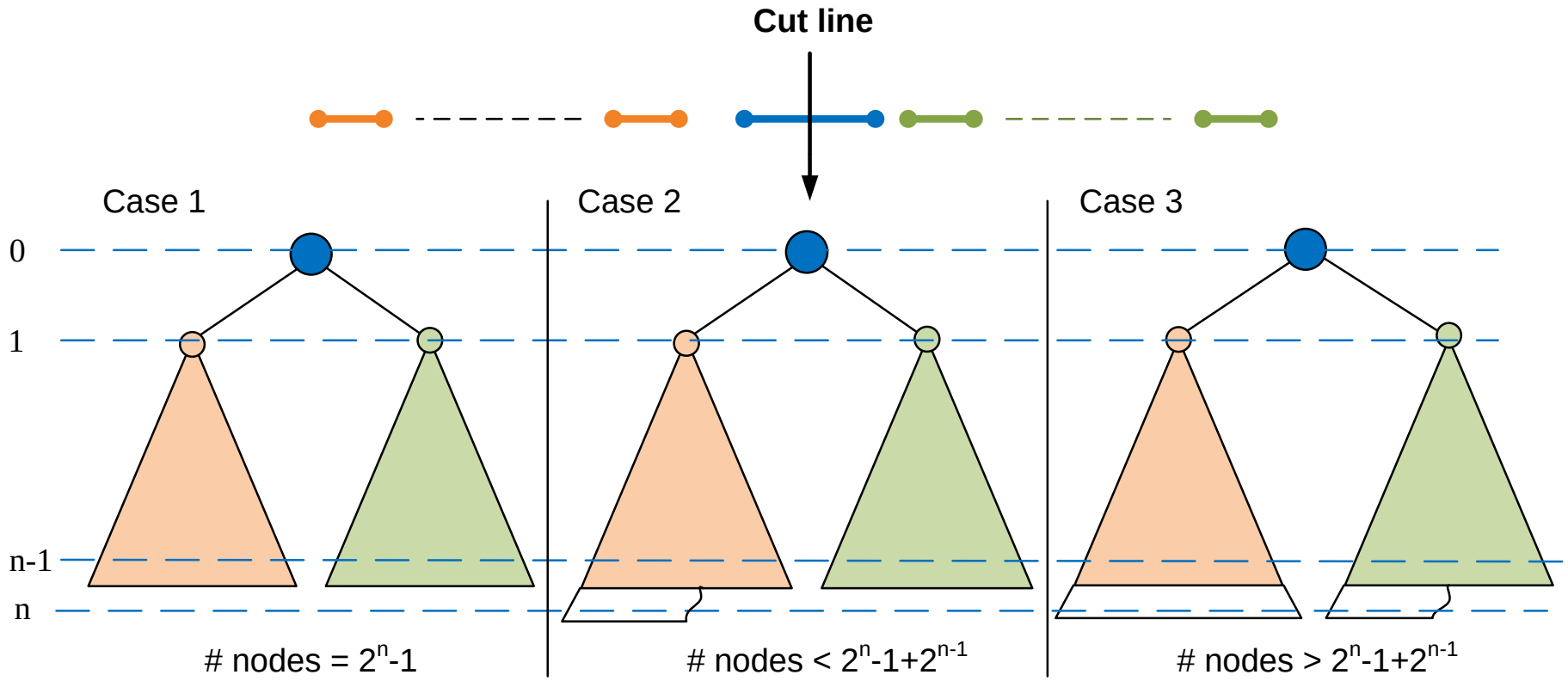
	Prefix	Lower bound		Upper bound		Range(s)
P1	0*	00000000	0	01111111	127	R2
P2	000*	00000000	0	00011111	31	R1
P3	010*	01000000	64	01011111	95	R3, R5
P4	01001*	01001000	72	01001111	79	R4
P5	01011*	01011000	88	01011111	95	R6
P6	011*	01100000	96	01111111	127	R7
P7	110*	11000000	192	11011111	223	R8
P8	111*	11100000	224	11111111	255	R9



BST Construction

USC Viterbi School of Engineering

- **Sort all non-overlapping ranges using lower bounds in increasing order (assume *N* ranges)**



IP look up

USC Viterbi School of Engineering

- **Start from root**

- **At each node**
  - **Compare incoming IP address with node's lower bound**
  - **If greater or equal**
    - Compare IP address with node's upper bound
    - **If** less than or equal, replace routing information with node's routing information

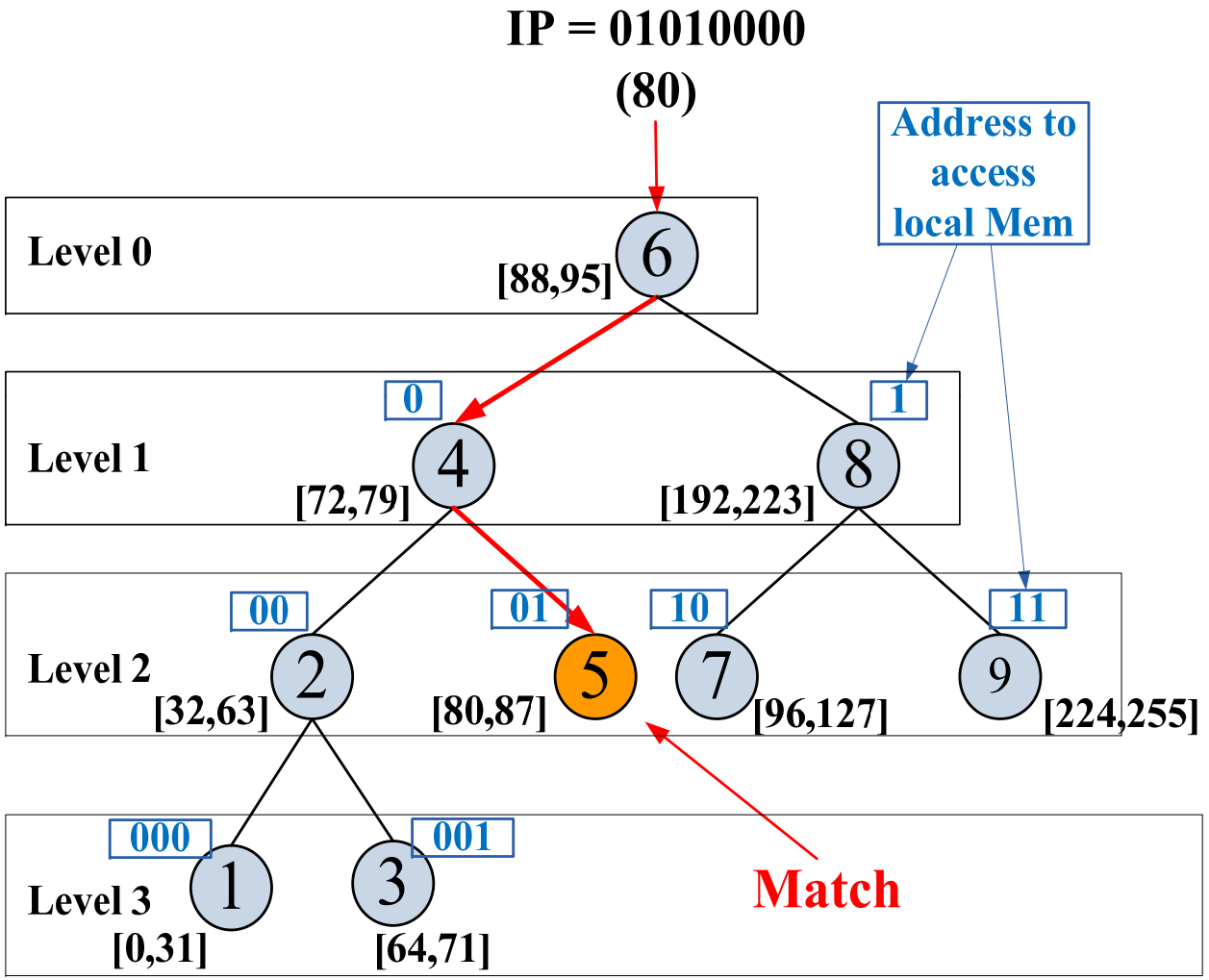
      Else Traverse to the right by appending 1 to the memory address

    Else
    - Traverse to the left by appending 0 to the memory address
  - **Memory address and routing information are forwarded**

Example

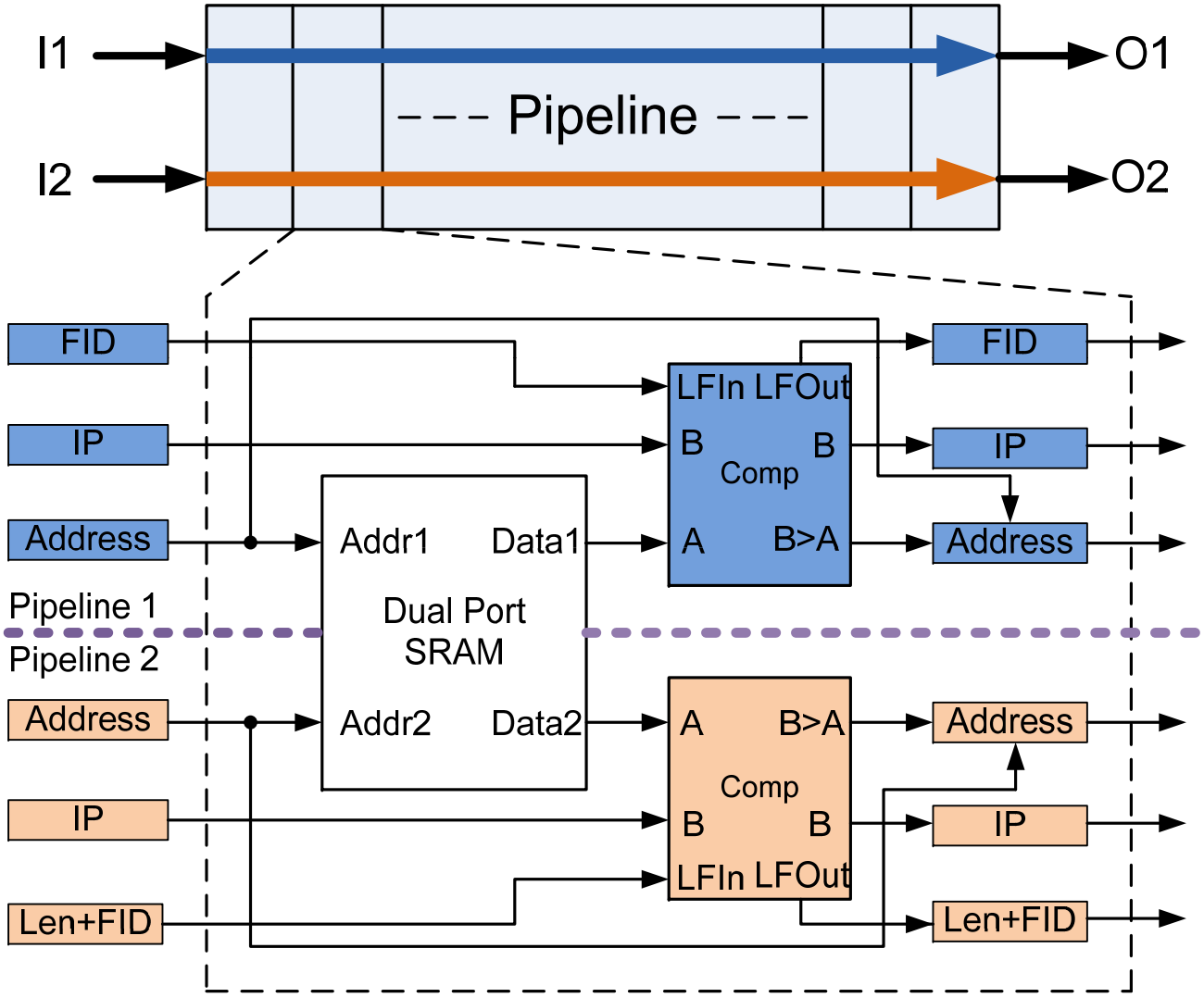
USC Viterbi School of Engineering

Range	Node
R1 [0,31]	1
R2 [32,63]	2
R3 [64,71]	3
R4 [72,79]	4
R5 [80,87]	5
R6 [88,95]	6
R7 [96,127]	7
R8 [192,223]	8
R9 [224,255]	9



Each node stores: 2 bounds, routing index

Architecture



FPGA Implementation (Non-optimized)

USC Viterbi School of Engineering

- **Each node needs approx. 71 bits**
  - **Lower bound: 32 bits**
  - **Upper bound: 32 bits**
  - **Routing index: 6 bits (arbitrarily chosen to represent ports, actions …)**
  - **Status: 1 bit**
- **Dual-ported Memory**
  - **71-bit data for each port**
  - **Address width equals stage index**
  - **Total of (142 + 2 x stage_index) pins**

- **Further memory optimizations are possible to achieve the compact design in FPL'08**

FPGA Implementation (non-optimized)

USC Viterbi School of Engineering

- **Virtex-4 FX140:**
  - **9936 Kb BRAM:**
    - Up to 17 stages
    - Up to 128K nodes
    - 92% BRAM utilization
  - **1517 I/O pins**
    - Interface up to 6 external dual-ported SRAM banks
    - Requires 969 pins coming into FPGA
  - **Support up to 8M nodes**
    - 8M prefixes (best case)
    - 4M prefixes (worst case)

FPGA Implementation (Non-optimized)

USC Viterbi School of Engineering

- **Throughput: over 300 MLPS (150 MHz)**
- **Required external SRAM:**

External stage	1	2	3	4	5	6
Mem (Mb)	9	27	63	135	279	567
# nodes	256K	512K	1M	2M	4M	8M

- **Architecture with duplicated pipelines**

# duplications	External SRAMs (Mb)	# nodes	Throughput
2x	31.5 x 2	512M	**600 MLPS (192 Gbps)**
3x	13.5 x 3	256K	**900 MLPS (288 Gbps)**

- **State-of-the-art TCAM: 533 MLPS**

Power Dissipation: SRAM vs. TCAM

USC Viterbi School of Engineering

	TCAM (18 Mb)	SRAM (18 Mb)
Clock rate (MHz)	266	250
Cell size (# of transistors per bit)	16	6
Power consumption (W)	15	1

$$P_{18\text{Mb}-SRAM}^{f=150MHz} = 1 \times \frac{150}{250} = 0.6\ \text{W}$$

Source: Samsung Specification

Power Dissipation (Non-optimized)

USC Viterbi School of Engineering

	Amount of Memory	# prefixes	Power consumption (W)		
			Logic	Memory	**Total**
TCAM	36 Mb	**819.2K**	N/A	30	**30**
USC-4	135 Mb	**1M**	< 1	4.5	**< 5.5**
USC-2x	63 Mb	**256K**	< 1	2.1	**< 3.1**

USC-4: no duplication + 4 external SRAMs

USC-2×: 2× duplication + 3 external SRAMs

Worst case

Source: Xilinx