SILICON ACCESS™
NETWORKS

…Enabling the Future of the Internet

The iFlow Address Processor™

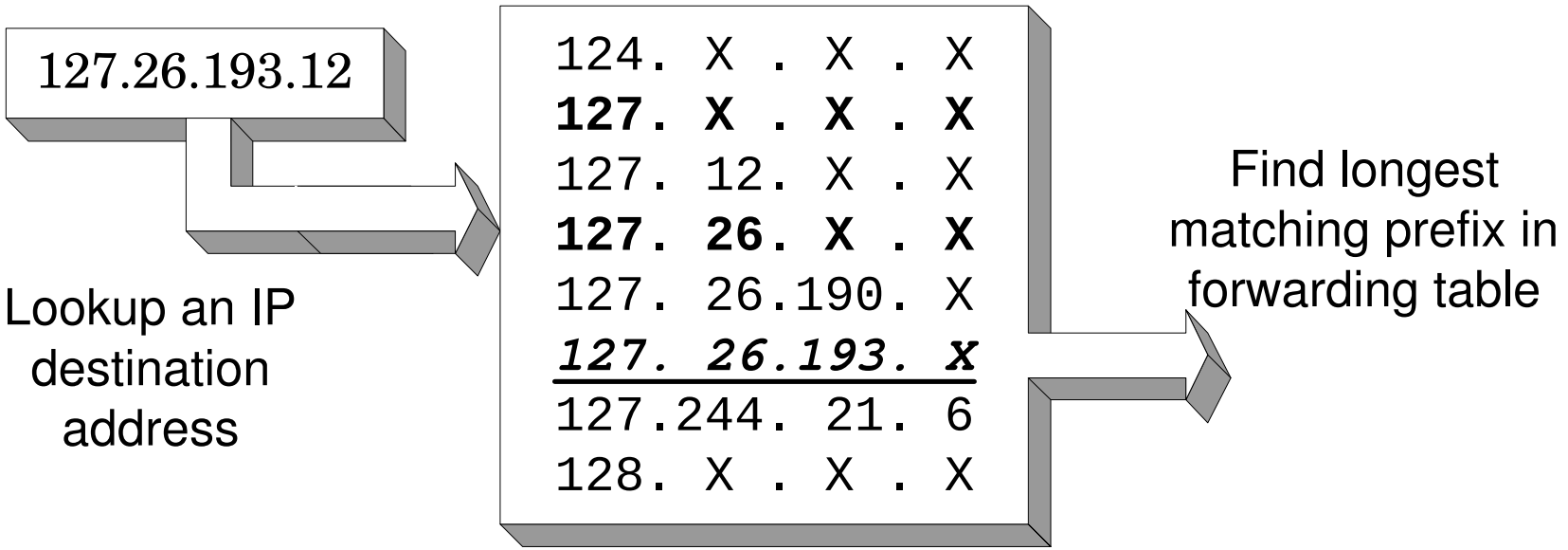
Forwarding Table Lookups using Fast, Wide Embedded DRAM

Mike O'Connor - Director, Advanced Architecture

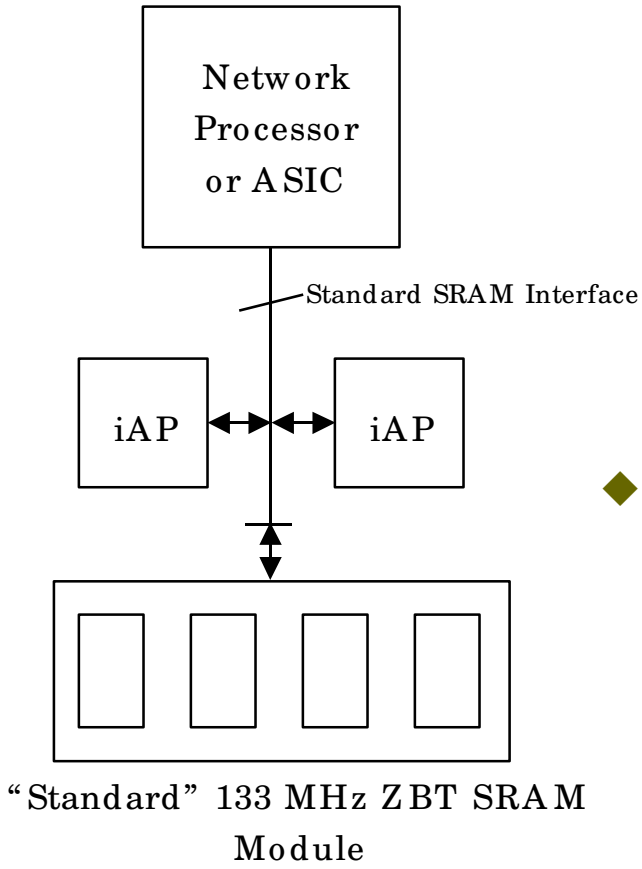
SILICON ACCESS NETWORKS

Forwarding Table Lookups

- **One key component of routing an IP packet is the forwarding table lookup of the destination address**
  - » Uses Classless Inter-domain Routing (CIDR) protocol
  - » Requires finding longest prefix match



iFlow Address Processor



- **Designed for layer 2 and layer 3 switching and routing applications requiring high table densities at up to OC-192 / 10 Gb/s Ethernet line speeds**
  - 10 Gb/s line rates translate into over 25 million packets per second
  - Enables at least 2 lookups per packet
- **Operates as a coprocessor to a Network Processor or ASIC on a router line card**
  - Sits on a standard SRAM bus
  - Configurable as 32, 64, 96, or 128-bits wide (+ parity)
  - Bus and chip operate up to 133 MHz
  - Cascade up to 4 iAP chips at full speed

iAP Applications

- **Longest Prefix Match/Exact Match Lookup Table**
  - Up to 256K 48-bit keys
    - IPv4 plus optional (up to 16-bit) VLAN or MPLS tag
    - IEEE 802.3 MAC Address
  - Up to 128K 96-bit keys
    - IPv4 <S,G> Lookups for multicast
    - MAC Address + VLAN tag
  - Up to 80K 144-bit keys
    - IPv6 plus optional VLAN or MPLS tags
    - Exact Flow Lookups (e.g. Cisco NetFlow)
  - Any combination of above subject to capacity
    - e.g. 128K 48-bit and 40K 144-bit

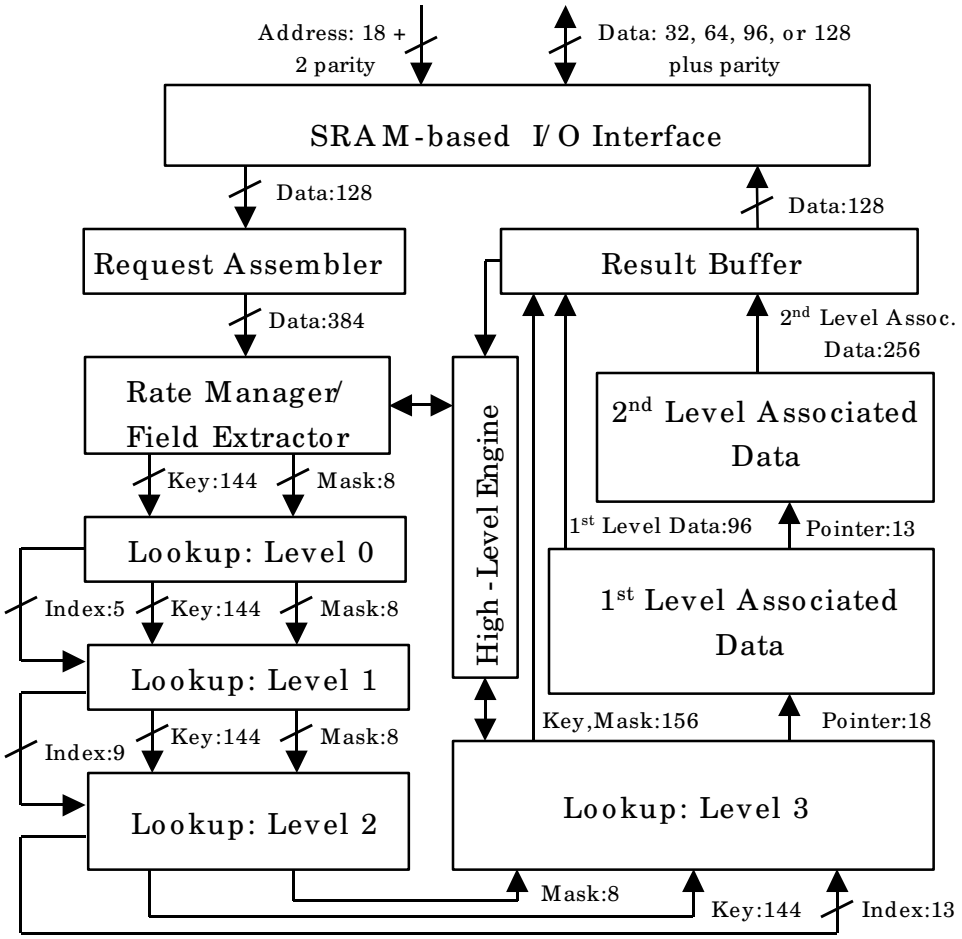
iAP Applications (cont.)

- **1st Level Associated Data (1:1 per Flow)**
  - 256K 96-bit data words
  - Treated as up to 4 fields, including
    - 2 counters (up to 48-bits and 35-bits)
      - Optionally saturating
      - Typically byte and packet counters
    - 13-bit pointer to further 2nd level associated data
- **2nd Level Associated Data**
  - 8K 256-bit data words
  - Many-to-1 per flow, or 1-to-1 per Next-Hop
  - Two counter fields (up to 64 bits and 48 bits),
    - Optionally saturating
  - Typically contains next hop information
    - Next Hop IP
    - Destination Switch Fabric ID
    - Billing information, etc.

iAP Features

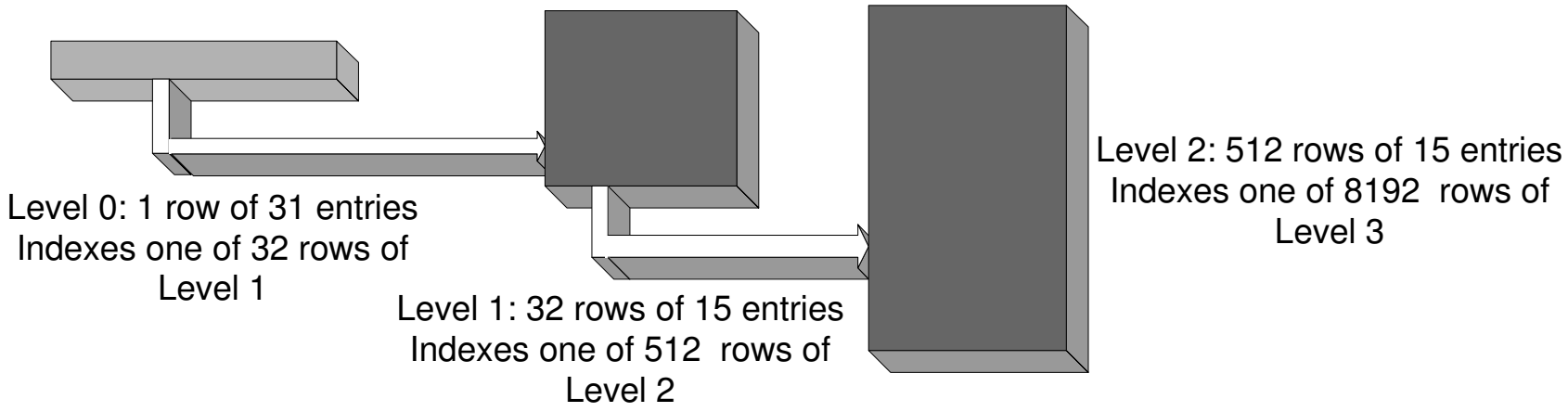
- **Automated incremental table maintenance ops**
  - Inserts and Deletes processed “in background” while lookups continue
  - Supports over 1M inserts/deletes per second using less than 10% of the lookup capacity
- **Designed for High-Availability Telco Applications**
  - Goal is to prevent “silent failures”
  - Parity on all internal memories larger than 1Kbyte
  - Parity on external address and data buses

iAP Organization



- SRAM Write transactions go to Request Assembler
- Lookup Pipeline performs Longest-Prefix Match lookup
- High-Level Engine processes requests (like Inserts)
- 1st and 2nd Level Associated Data are indexed/updated
- Result Buffer holds results accessed via SRAM read transactions

iAP Lookup Pipeline

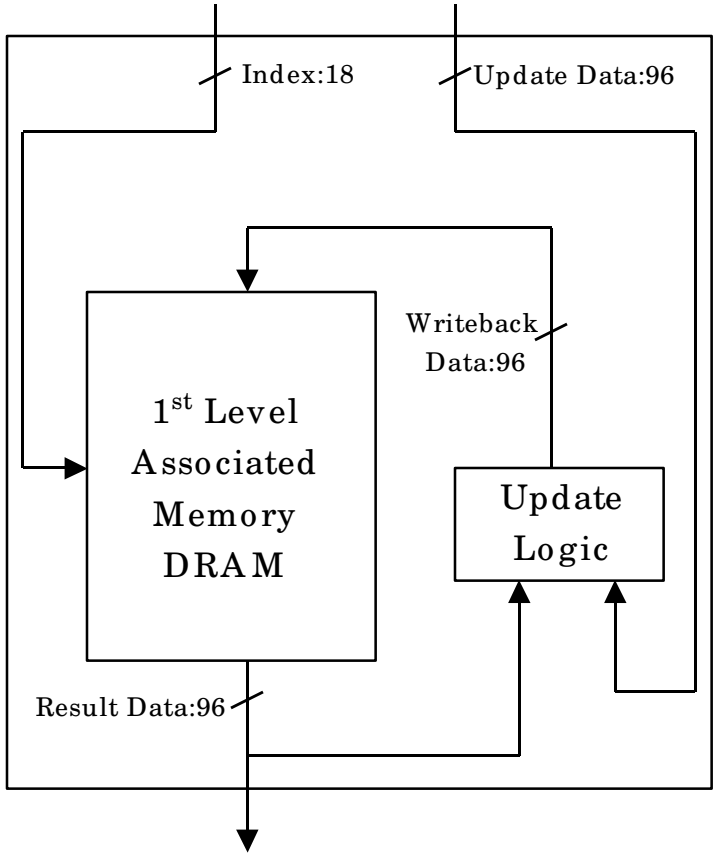


- **Organized as a pipelined tree-like structure**
  - First three levels are implemented in SRAM and total approximately 1.2 Mbits in size
  - First three levels are organized as a B-Tree index of the fourth level
  - Each SRAM Memory is >2000 bits wide

iAP Lookup Pipeline

- **Final level of lookup pipeline holds entire set of keys**
  - 25 Mbits of embedded DRAM with an access rate of 66 MHz
  - Organized as 8K rows of ***3,200*** bits wide
  - A single row contains enough information to determine the longest prefix match for a search key which indexes that row

Associated Memory



- **1st Level Associated Memory**
  - Each word can contain two counters
    - one up to 48-bits and one up to 35-bits
  - Supports reading counters, adding increments, and writing the values back every lookup
  - 25 Mbits of 256K x 100-bit fast embedded DRAM
    - Supports 133 MHz random-access rate
- **2nd Level Associated Memory**
  - Supports on-the-fly statistics updates
  - Implemented as 2 Mbits of 8K by 256-bit fast embedded DRAM

Table Maintenance

- **Embedded state machines control operations of high-level table updates**
- **Inserts and deletes into the table are processed in the background, while lookups take place.**
- **Appropriate bypassing allows updates of the table to be "invisible" to the lookups.**
  - » Lookups get either the original result or the new result
    - – never an inconsistent table state
- **Massive internal bandwidth makes this practical**

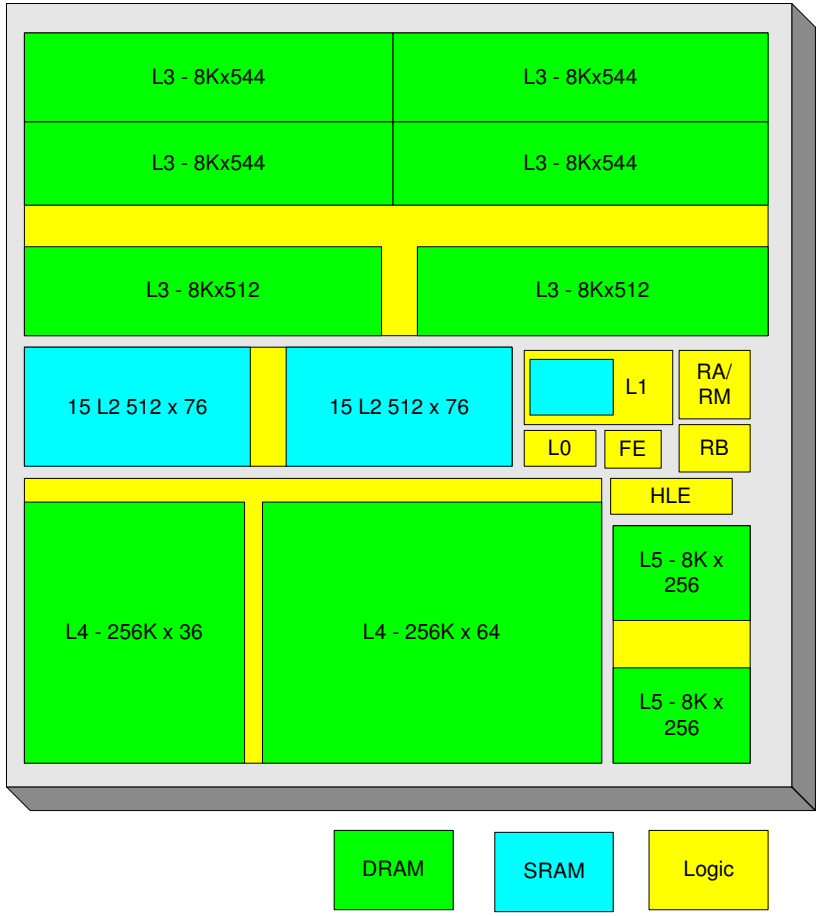
Embedded DRAM

- **The iAP contains a total 52 Mbits of custom embedded DRAM memories**
  - 25 Mb, 3200 bits wide at 66 MHz
  - 25 Mb, 100 bits wide at 133 MHz
  - 2 Mb, 256 bits wide at 133 MHz
- **252 Gb/s aggregate on chip DRAM bandwidth**
  - Equivalent bandwidth from external 133 MHz ZBT SRAMs would require approximately 1900 data pins
- **70% of the die area is DRAM**

Embedded DRAM challenges

- **Even Embedded DRAMs must be refreshed**
  - » Refresh cycles flow down the pipeline at regular intervals
    - – If user prefers completely deterministic lookup timing, the user can generate refresh commands externally
  - » Reduces effective maximum lookup rate to 64.6M/sec
- **Large, wide memory blocks create "interesting" floorplanning issues**
  - » 3200-bit wide memory spans almost entire width of die
- **Custom DRAMs require significant resources to develop**
  - » Memory design team is ~20 engineers

Physical Details



- iAP is implemented in a 0.18 micron TSMC embedded DRAM process.
- Die is 13.5 by 13.5 mm
- 476 pins
- ~750,000 logic gates
- 52 Mbits EDRAM
- 1.2 Mbits SRAM
- Power is ~5 W at 133 MHz

Conclusions

- **Embedded DRAM enables networking applications requiring high bandwidth to large tables**
- **Alternative solutions require more:**
  - Parts
  - Pins
  - Power
  - Cost
- **Silicon Access Networks is working on other chips too…**