Data @ Bare Metal Speed
The Wisconsin Quickstep Project

Jignesh M. Patel

Blog: http://bigfastdata.blogspot.com
Good News: Big Data Opportunity

Databases are at the core of the big data revolution
Bad News: Big Data Software

What we have delivered looks like a pig with lipstick
Good News: Big Data Opportunity

People are focused on the massive and real growth in this area
Bad News: The game is up!

We can’t hide for much longer

a disruptive way
Disruptive hardware trends

Want

High Performance

Low Cost

Constraint

Power

Lower-powered, lower latency, higher bandwidth, persistent stores

CPU
DRAM
cache

Multicores, multi-packet, heterogeneous cores

Magnetic Hard Disk Drives

NVRAM (e.g. SSDs)
Moore’s Law continues, but ...

Oazizi, 2010
Future processor design?

1. Keep adding cores
   (~40% per generation)
2. Heterogeneous cores
3. Programmable functional units

But, systems must work within a power budget
Data growth continues unabated

Need to do more with less.
Scan: A Key Data Processing Kernel

What?
- Scan a column of a table applying some predicate

Why?
- A key primitive in database
- Often most invoked kernel in deep analytic processing

How?
- Conserve memory bandwidth: BitWeaving the data
- Use every bit of data that is brought to the processor efficiently using intra-cycle parallelism

© 2013, Jignesh M. Patel, University of Wisconsin
Focus on Column Scan (can be generalized)

Traditional Row Store

<table>
<thead>
<tr>
<th>shipdate</th>
<th>discount</th>
<th>quantity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mar-12-2013</td>
<td>5%</td>
<td>5</td>
</tr>
<tr>
<td>Jan-08-2013</td>
<td>2%</td>
<td>4</td>
</tr>
<tr>
<td>Apr-29-2013</td>
<td>10%</td>
<td>3</td>
</tr>
<tr>
<td>May-14-2013</td>
<td>0%</td>
<td>6</td>
</tr>
<tr>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>Feb-28-2013</td>
<td>5%</td>
<td>0</td>
</tr>
</tbody>
</table>

Column Store

<table>
<thead>
<tr>
<th>shipdate</th>
<th>discount</th>
<th>quantity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mar-12-2013</td>
<td>5%</td>
<td></td>
</tr>
<tr>
<td>Jan-08-2013</td>
<td>2%</td>
<td></td>
</tr>
<tr>
<td>Apr-29-2013</td>
<td>10%</td>
<td></td>
</tr>
<tr>
<td>May-14-2013</td>
<td>0%</td>
<td></td>
</tr>
<tr>
<td>...</td>
<td>...</td>
<td></td>
</tr>
<tr>
<td>Feb-28-2013</td>
<td>5%</td>
<td></td>
</tr>
</tbody>
</table>

File: n-1

File: n

Column Codes: 5 4 3 6 2 7 1 0

Blog: http://bigfastdata.blogspot.com

Li and Patel, SIGMOD'13

© 2013, Jignesh M. Patel, University of Wisconsin
First batch of Processor Words

(batch size = code size in bits)

- Extra bit
- Vertical code stacking

© 2013, Jignesh M. Patel, University of Wisconsin
SELECT SUM(l_discount * l_price) FROM lineitem
WHERE l_shipdate BETWEEN Date AND Date + 1 year
AND l_discount BETWEEN Discount - 0.01 AND Discount + 0.01
AND l_quantity < Quantity

Result bit vector
RID List: 9, 15

Aggregation

BitWeaving/H

l_price

l_discount

BitWeaving/H

Result bit vector

BitWeaving/V

Result bit vector

l_shipdate

BitWeaving/V

Result bit vector

l_discount

BitWeaving/H

Result bit vector

l_quantity
Disruptive hardware trends

- **CPU cache**
- **DRAM**
- **NVRAM (e.g. SSDs)**
- **Magnetic Hard Disk Drives**

- Multicores, multi-socket, heterogeneous cores
- Lower-powered, lower latency, higher bandwidth, persistent stores

© 2013, Jignesh M. Patel, University of Wisconsin
Long Term: Raw computing and storage costs tends to zero!

The cost is in moving data and powering the circuits/devices
There are similar ways of using hardware creatively, e.g. ASICs or GPUs.

Basically, need hardware and software synergy!
Transformative architectural changes at all levels (CPU, memory subsystem, I/O subsystem) is underway

Need to rethink data processing kernels; e.g. BitWeaving

Need to think of hardware software co-design
Thanks!

The Quickstep Team

Blog: http://bigfastdata.blogspot.com

© 2013, Jignesh M. Patel, University of Wisconsin