SIMULATING OF QUERY PROCESSING ON MULTIPROCESSOR DATABASE SYSTEMS WITH MODERN COPROCESSORS

Pavel S. Kostenetskiy, PhD
Konstantin Y. Besedin,

South Ural State University, Russia
BACKGROUND

- GPUs and Intel Xeon Phi appeared to be very well-suited for general-purpose computations.
- There's an interest on using GPUs and manycore coprocessors in parallel database systems.
The Goal

Compare performance of different coprocessors in parallel database system.
WE IMPLEMENTED THE PARALLEL DBMS SIMULATOR

- Simple software that acts like real parallel DBMS.
- It can simulate SELECT and JOIN query processing in parallel DBMS.
- It can use CPUs, NVIDIA GPUs and Intel Xeon Phi coprocessors.
- It can use multiple computing nodes of a cluster.
## The Hardware

<table>
<thead>
<tr>
<th>Node</th>
<th>&quot;Tornado SUSU&quot; supercomputer</th>
<th>UNN cluster</th>
</tr>
</thead>
<tbody>
<tr>
<td>CPU</td>
<td>Intel Xeon 5680 3.33 GHz</td>
<td>Intel Xeon L5630 2.13 GHz</td>
</tr>
<tr>
<td>RAM</td>
<td>24 / 48 GB</td>
<td>24 GB</td>
</tr>
<tr>
<td>Number of CPUs</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>Coprocessor/GPU</td>
<td>Intel Xeon Phi 7110X</td>
<td>NVIDIA Tesla X2070</td>
</tr>
<tr>
<td>Number of nodes</td>
<td>480</td>
<td>16</td>
</tr>
</tbody>
</table>

Experiments with CPUs and Intel Xeon Phi were done on supercomputer in South Ural State University in Chelyabinsk. Experiments with GPUs were done on computing cluster in Lobachevsky State University in Nizhniy Novgorod.
EXPERIMENTS DESCRIPTIONS

At this stage, the simulator supports SELECT and JOIN queries.

SELECT
- Basic form of SELECT query on unindexed columns
- 370M tuples (5.5 GB) in relation.
- Each tuple consists of two long integer attributes.

JOIN
- INNER JOIN query using nested loops algorithm
- 33M tuples in outer relation
- 33K tuples in inner relation
SELECT EXPERIMENT

Time

![Graph showing time trends for different numbers of nodes for CPU, GPU, and MIC.]

Speedup

![Graph showing speedup trends for different numbers of nodes for CPU, GPU, and MIC.]

Nodes

1 2 3 4 5 6 7 8

Time (sec)

0 20 40 60 80

Speedup

0,8 1 1,2 1,4 1,6 1,8 2 2,2 2,4 2,6

CPU
GPU
MIC
JOIN EXPERIMENT

Time

Speedup

Nodes

Time (sec)

CPU
GPU
MIC

Nodes

CPU
GPU
MIC

1 2 3 4 5 6 7 8

1 2 3 4 5 6 7 8

1 3 5 7

0 50 100 150 200 250 300

0 1 3 5 7

1 2 3 4 5 6 7 8
SUMMARY

- Manycore coprocessors and GPUs are less efficient than CPUs for SELECT queries,
- They can be used effectively for processing INNER JOIN queries.
Thanks for the attention! 
Any questions?