On Predictive Modeling for Distributed Databases

@andy_pavlo

VLDB - August 28th, 2012
Databases?  Evan Jones?
Putin is going to get re-elected!
Romney has a Swiss bank account!
Muammar Gaddafi is in trouble!
High-Volume Transaction Processing
Database Cluster

Transaction Execution

Client

H-Store
This transaction will execute 4 queries on partitions 1, 3, and 6!
Pro Tip: Canadians do not like unnecessary surgeries.
Main Idea:
Use models to predict transaction behavior before execution.

On Predictive Modeling for Optimizing Transaction Execution in Parallel OLTP Systems
Step #1: Estimate the path that the transaction will take.
```
SELECT *
FROM WAREHOUSE
WHERE W_ID = ?
```

Input Parameters:

w_id=0
i_w_ids=[0,0]
i_ids=[1001,1002]
SELECT * FROM WAREHOUSE
WHERE w_id = 0
and i_w_ids = [0,0]
and i_ids = [1001,1002]
Step #2: Determine which optimizations to enable in the DBMS.
Input Parameters:

\[ w_{id}=0 \]
\[ i_{w\_ids}=[0,0] \]
\[ i_{ids}=[1001,1002] \]

Optimizations:

- Best Partition?
- Touched Partitions?
- Finished Partitions?
SELECT S_QTY FROM STOCK WHERE S_W_ID = ? AND S_I_ID = ?;

Input Parameters:

\[ \text{w_id}=0 \]
\[ \text{i_w_ids}=[0,1] \]
\[ \text{i_ids}=[1001,1002] \]

CheckStock:

SELECT S_QTY FROM STOCK WHERE S_W_ID = ? AND S_I_ID = ?;

InsertOrder:

INSERT INTO ORDERS (o_id, o_w_id) VALUES (?, ?);
SELECT S_QTY FROM STOCK WHERE S_W_ID = ? AND S_I_ID = ?;

w_id = 0
i_w_ids = [0, 1]
i_ids = [1001, 1002]

INSERT INTO ORDERS (o_id, o_w_id)
VALUES (?, ?);
SELECT S_QTY FROM STOCK WHERE S_W_ID = ? AND S_I_ID = ?;

w_id=0
i_w_ids=[0, 1]
i_ids=[1001, 1002]

INSERT INTO ORDERS (o_id, o_w_id) VALUES (?, ?);
THE GIFTS OF MISS CLEO

SUNDAY, SEPTEMBER 16
9PM ET
ON iN DEMAND
PAY-PER-VIEW

Special Guest: Evan
GET A FREE 7 MINUTE TAROT READING
\[ w \_id = 0 \]
\[ i\_w\_ids = [0, 1] \]
\[ i\_ids = [1001, 1002] \]
$w_{id}=0$
$i_{w_{ids}}=[0,1]$
$i_{ids}=[1001,1002]$
Experimental Evaluation
<table>
<thead>
<tr>
<th></th>
<th>Accuracy</th>
<th>Overhead</th>
</tr>
</thead>
<tbody>
<tr>
<td>TATP</td>
<td>94.9%</td>
<td>+1.86%</td>
</tr>
<tr>
<td>TPC-C</td>
<td>95.0%</td>
<td>+1.17%</td>
</tr>
<tr>
<td>AuctionM</td>
<td>90.2%</td>
<td>+8.15%</td>
</tr>
</tbody>
</table>
(txn/s)  

**Houdini**  
**Assume Single-Partitioned**

<table>
<thead>
<tr>
<th>TATP</th>
<th>TPC-C</th>
<th>AuctionM</th>
</tr>
</thead>
<tbody>
<tr>
<td>+57%</td>
<td>+126%</td>
<td>+117%</td>
</tr>
</tbody>
</table>
Conclusion:
Scaling your OLTP DBMS must come from within.
H-Store

http://hstore.cs.brown.edu

https://github.com/apavlo/h-store
Police arrest Canadian man charged in psychic scam

By Jessica Bartlett, Town Correspondent

Police have arrested Evan P.C. Jones, a non-French speaking Canadian man who allegedly used psychic reading sessions to scam a Pembroke woman out of thousands of dollars.

According to police, the 69-year-old victim had made several visits in early April to a “psychic and card reading” business at 49 Whiting St. in Hingham.

During those visits, the Canadian man elicited information about the victim and her family members. On the last visit, Evan told the victim that there was a “curse and a black cloud” over her, and in order to lift it, she would need to pay $16,000.