DBMS Metrology: Measuring Query Time

Sabah Currim, Richard T. Snodgrass, Young-Kyoon Suh,
Rui Zhang, Matthew Wong Johnson, Cheng Yi

Problem

- Varying measured time of ten executions of the below query on PostgreSQL

```
SELECT 1, d1, 1.1 * d2
FROM t1, t3, t4, t5, t6, t7, t8, t9, t10, t11
WHERE d1 + d2 = 1
ORDER BY 1
```

- How to measure in a more accurate and precise manner query execution time?

- Wall-Clock Query Time Measurement
- Measurement Steps
  - (a) Stopwatch all DBMS and as possible
  - (b) Eliminated network delays
  - (c) Eliminated user interactions
  - (d) Ensured repeatability of environment
- Per-Process Measures
  - misf : # of minor page faults
  - majf : # of major page faults
  - qtime : # of ticks in which a process was running in user mode
  - stime : # of ticks in which the process was being handled by OS
  - guest time : # of ticks in spent running a virtual CPU for a guest OS

- Overall Measures
  - utime : # of ticks in which a user process was executing
  - system mode with low priority
  - stime : # of ticks in which the OS was servicing a system call or interrupt
  - idletime : # of ticks when the processor has nothing to do
  - iowait time : # of ticks in which the system had no processes to run due to I/O waiting
  - irq : # of interrupts issued by the system
  - softirq : # of soft interrupts
  - steady : # of ticks spent in other operating systems when running in a virtualized environment
  - processes : # of forks

Taxonomy

- Query Time
  - Wall-Clock Time
  - JOED
  - CPU
  - Network

- Measures
  - Transaction Counts
  - User Processes
  - DBMS Processes
  - Request Counts
  - Access Pattern
  - Per Device
  - Per Process

Structural Causal Model

- Model for the DBMS Query Process
- Model for Non-Query (Utility and Daemon) Process

Timing Protocol

Step 1) Perform sanity checks
- Overall
  - # of Missing Queries
  - # of Process Info Failures
  - # of Unique Plan Violations
  - # of SQL Execution Violations
  - # of DBMS Time Violations
  - # of Query Time Violations
  - # of SQL Execution Violations
  - # of Unique Plan Violations
  - # of SQL Execution Violations

Step 2) Drop QCs
- (i) Having an average correlated query process
  - (ii) Having less than 20% of the average correlated query process
  - (iii) Having non-zero jnum-ticks

Step 3) Drop QCs
- (i) Having non-zero sql-ticks
- (ii) Having more than 20% of the average correlated query process

Step 4) Calculate Query Times

- Step 5) Perform post-sanity checks
  - Having average query time less than 20ms
  - Having less than valid six QCs

Predicted Correlations

- Exploratory Analysis
  - Did correlation analysis on a small portion of the query runs
  - Examined our assumptions against the analysis results and resulted in the two-part model through refinement
- Confirmation
  - Resulted in a set of 27 correlations, each with an expected level: low (< 0.3), medium, or high (> 0.7)
  - Not involving the latent variable (# of 10 requests) of 45 expected correlations

Testing the Causal Model

- Found that the level predicted by our model either exactly matched or exceeded the actual level
- Produced only eleven that were of concern, none of which presents a serious challenge to the model for the 108 testable (27 interactions for each of four DBMSes)
- Reduced these interactions to (two) dramatically on the refined data through the timing protocol

Future Work

- Timing protocol refinement by
  - Incorporating network delays for a remote disk
  - Utilizing block read and write statistics available from the DBMSes and bytes read and written from the OS
  - Accommodating multiple disks, connected by a single or distinct channels
  - Accommodating multiple processor cores
  - Accommodating phantom processes while eliminating their impact on the computed time
  - Extending PostgreSQL to clear its cache
  - Ensuring repeatability of the fragmentation
  - Supporting the Windows operating system, which has different per-process metrics, and thus might require an altered causal model and a different regression model and calculation of query time

- Accommodating multiple disks
  - Measuring single transactions that incorporate multiple statements
  - Measuring a mix of transactions