Big Data implementation: Role of Memory and SSD in Microsoft SQL Server Environment

Scenario Analysis of Decision Support System with Microsoft Windows Server 2012 OS & SQL Server 2012 and Samsung 20nm-class DRAM & SSD on Dell PowerEdge R910 Server

A cooperation of

Microsoft Technology Center (MTC)
& Samsung Semiconductor

Document Version:
1.0

Authors:
MTC Paris: Stéphane Goudeau, Laurent Marzouk
Samsung Semiconductor: Sanggeol Lee, Sungjun Kim

Contacts:
Alain Le Hegarat, MTC Paris, alainle@microsoft.com
Peyman Blumstengel, Samsung Semiconductor Europe, peyman.b@samsung.com
Dominique Leong, Global Solution Architect, Dell, dominique_leong@dell.com
Introduction

Today’s organizations face growing challenges extracting business value from their data. Firstly, the relentless growth of data continues, due to the proliferation of new devices and sensors, and rapidly declining hardware cost. More organizations now store terabytes and even petabytes of data. Secondly, data complexity is increasing as customers store both structured data in relational format and unstructured data such as Word or PDF files, images, videos and geo-spatial data.

"Big data" is a term applied to data sets whose size is beyond the capability of commonly used software tools to capture, manage, and process within a tolerable elapsed time. Big data sizes are a constantly moving target, with 2012 figures ranging from a few dozen terabytes to many petabytes in a single data set.¹

Microsoft began doing Big Data long before it became on-trend in the market. Microsoft’s RDBMS SQL Server offers a robust and scalable platform for storing and analyzing big volumes of data in a traditional data warehouse.

When applications need to tap data from a database, they run queries against the RDBMS engine, which in turn searches for these data in the most efficient way; firstly in memory, where they will be residing if they were accessed (loaded) recently, then on disk if they were not found in memory (first access). In this whitepaper, we will use the terms **Cold cache** when data are accessed from disk and **Warm cache** when they are accessed directly from memory.

Since main memory provides access to data in mere nanoseconds (ns), which is one million times faster than HDDs performing in milliseconds (ms), there is a huge drive in the IT industry to get the whole database loaded into main memory. This would enable fast access and analysis of data; the so-called “In-Memory” solutions. The new Flash-based storage devices also offer opportunities in this space. With data access times in the Micro Second (µs) range, SSDs provide a good compromise in terms of performance/cost ratio.

This whitepaper describes the results of a Proof of Concept (PoC) performed by Microsoft and Samsung semiconductor at the MTC Paris, aiming to illustrate the ability of Microsoft SQL Server 2012 and Windows Server 2008 R2 SP1 & Windows Server 2012 operating systems to deal with and benefit from large amounts of memory. It also measures the impact of Samsung’s brand new memory modules and SSDs on performance and energy savings, translating this into a measurable user experience at lower operating costs. All this have been performed on Dell’s brand new PowerEdge R910 servers.

For this purpose two key indicators were measured during this benchmark: queries duration and power consumption savings.

Two identical Dell PowerEdge R910 servers were used to run our tests in parallel, which avoided switching memory modules and HDD/SSD between tests, and helped reduce the PoC duration. Both Dell servers were installed in dual boot with Windows Server 2008 R2 and Windows Server 2012.

One Dell server was populated with main stream 50nm technology memory and Hard disk drives in the Windows Server 2008 R2 SP1 OS environment. The other system was setup with the Samsung leading edge 20nm Green main memory and SSDs working with Windows Server 2012 RC. A 1TB database was generated on both systems with MS SQL server 2012 setup.
We selected 4 types of queries reflecting the major tasks conducted in a typical data analysis within standard benchmarking software (TPC-H):

- Scenario 1: Minimum Cost Supplier Query
- Scenario 2: Forecasting Revenue Change Query
- Scenario 3: Promotion Effect Query
- Scenario 4: Potential Part Promotion Query

We measured the performance of both systems when running these queries and accessing the 1TB database both in Cold and Warm cache conditions.

**Summary of Results**

The system powered with Samsung Green 20nm-class DDR3, SSDs, Microsoft SQL server 2012 and Microsoft Windows Server 2012 OS (“High Performance & Green Server”) compared to an enterprise-grade traditional server with 50nm DDR3 and HDDs (“Standard & Non-Green Server”).

**Cold Cache Test:**

- Cold cache test primarily highlights SSD performance. Memory usage is only 5%, which means all the data are located on disks before the benchmark starts:

  ![Task Manager](image)

  - The “High performance & green server” takes **11 minutes and 8 seconds** and consumes **126Wh** to run 4 queries of TPC-H for decision support (DSS), while standard & non-green server takes **2 hours 41 minutes and 33 seconds** and consumes **2,196Wh**.

  - “High performance & green server” is **14.5 times faster** and saves **94% of system power consumption**.

  - In this scenario we definitely see the leverage of SSD vs. HDD.
Warm Cache Test:

- Warm cache test highlights the benefits of in-memory computing and power saving of 20nm-class DDR3 modules. Memory usage is 95%, which means almost all the data is located in memory before the benchmark starts:

- "High performance & green server” takes **1 minute and 45 seconds** and consumes **23Wh** to run 4 queries of TPC-H for decision support (DSS), while standard & non-green server takes **1 minute 47 seconds** and consumes **32Wh**.
High performance & green server is only 2% faster than the standard and non green configuration since almost all the data is located in memory, but it still saves 28% of system power consumption.

Dual Boot Test:

- Windows server 2012 RC (Release Candidate) shows 6.5% faster performance in cold cache test and 7.1% faster performance in warm cache test compared to Windows Server 2008 R2 SP1, as illustrated by the picture below:
Test Environment

The table below provides the technical characteristics of the platform used to achieve this benchmark:

<table>
<thead>
<tr>
<th></th>
<th>High Performance &amp; Green Server Solution</th>
<th>Standard &amp; Non-Green Server Solution</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Server</strong></td>
<td>Dell PowerEdge R910 (Intel Westmere-EX, four sockets)</td>
<td>Dell PowerEdge R910 (Intel Westmere-EX, four sockets)</td>
</tr>
<tr>
<td><strong>BIOS version</strong></td>
<td>2.4.5</td>
<td>2.4.5</td>
</tr>
<tr>
<td><strong>CPU</strong></td>
<td>4x Intel Xeon E7-4870 10C/20T 2.40GHz 30MB (40 Cores)</td>
<td>4x Intel Xeon E7-4870 10C/20T 2.40GHz 30MB (40 Cores)</td>
</tr>
<tr>
<td><strong>DRAM</strong></td>
<td>1TB (20nm-class DDR3-1066) (based on 16GB modules, 4GB chips)</td>
<td>1TB (50nm-class DDR3-1066) (based on 16GB modules, 2GB chips)</td>
</tr>
<tr>
<td><strong>Storage</strong></td>
<td>4TB (Samsung 2.5” PM830 SSD) - OS : 2x 256GB (RAID1) - DB+TempDB : 4x 256GB + 6x 512GB (RAID0)</td>
<td>3.2TB (2.5” SAS HDD) - OS : 2x 600GB (RAID1) - DB+TempDB 6x 146GB + 4x 600GB (RAID0)</td>
</tr>
<tr>
<td><strong>OS</strong></td>
<td>Dual Boot (x64): - Win Server 2008 R2 SP1 - Win Server 2012 RC (Power profile = Max Performance)</td>
<td>Dual Boot (x64): - Win Server 2008 R2 SP1 - Win Server 2012 RC (Power profile = Max Performance)</td>
</tr>
<tr>
<td><strong>DB Engine</strong></td>
<td>SQL Server 2012 CU1 x64</td>
<td>SQL Server 2012 CU1 x64</td>
</tr>
</tbody>
</table>

Table 1. Summary of technical architectures tested

<table>
<thead>
<tr>
<th>Total Memory</th>
<th>20nm-class 1TB DDR3</th>
<th>50nm-class 1TB DDR3</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Memory Voltage</strong></td>
<td>1.5V</td>
<td>1.35V</td>
</tr>
<tr>
<td>1TB(2DPC) *</td>
<td>1066Mbps</td>
<td>-</td>
</tr>
</tbody>
</table>

* DPC : DIMM Per Channel

Table 2. Summary of memory configurations tested

- Storage: 6x Samsung 2.5” PM830 512GB SSD + 6x Samsung 2.5” PM830 256GB SSD ; SATA 6G, 20nm-class MLC NAND, Data Center version vs. 6x 2.5” 10Krpm 600GB HDD + 6x 2.5” 15Krpm 146GB HDD; SAS 6G
Big Data implementation: Role of Memory and SSD in Microsoft SQL Server Environment

<table>
<thead>
<tr>
<th>Disk Configuration</th>
<th>2.5&quot; Samsung PM830 SSD</th>
<th>2.5&quot; SAS HDD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Interface</td>
<td>SATA 6G</td>
<td>SAS 6G</td>
</tr>
<tr>
<td>2x 256GB (RAID1)</td>
<td>2x 600GB (RAID1)</td>
<td></td>
</tr>
<tr>
<td>: Dual Boot OS</td>
<td>: Dual Boot OS</td>
<td></td>
</tr>
<tr>
<td>4x 256GB + 6x 512GB (RAID0)</td>
<td>4x 600GB + 6x 146GB (RAID0)</td>
<td></td>
</tr>
<tr>
<td>: DB + TempDB</td>
<td>: DB + TempDB</td>
<td></td>
</tr>
</tbody>
</table>

Table 3. Summary of storage configurations tested

- **Test Scenarios:**
  - MTC Paris and Samsung used the TPC-H benchmark from [www.tpc.org](http://www.tpc.org) to generate the load, which simulates decision support system (DSS) activities against a database. To test the impact of Samsung 20nm-class DDR3 memory and PM830 SSDs (Datacenter version) with Windows Server 2012 OS on system performance and power saving, MTC Paris generated a 1TB database based on the TPC-H scheme and ran a set of 4 predefined queries against the database.
  - **Scenario 1 (Minimum Cost Supplier Query)**
    : This query finds which supplier should be selected to place an order for a given part in a given region.
  - **Scenario 2 (Forecasting Revenue Change Query)**
    : This query quantifies the amount of revenue increase that would have resulted from eliminating certain company-wide discounts in a given percentage range in a given year. Asking this type of “what if” query can be used to look for ways to increase revenues.
  - **Scenario 3 (Promotion Effect Query)**
    : This query monitors the market response to a promotion such as TV advertisements or a special campaign.
  - **Scenario 4 (Potential Part Promotion Query)**
    : This query identifies suppliers in a particular nation having selected parts that may be candidates for a promotional offer.

```
<table>
<thead>
<tr>
<th>Table Name</th>
<th># Records</th>
<th>Reserved (KB)</th>
<th>Data (KB)</th>
<th>Indexes (KB)</th>
<th>Unused (KB)</th>
</tr>
</thead>
<tbody>
<tr>
<td>dbo.LINEITEM</td>
<td>3,960,004,561</td>
<td>688,720,232</td>
<td>590,416,000</td>
<td>98,296,584</td>
<td>7,648</td>
</tr>
<tr>
<td>dbo.PARTSUPP</td>
<td>1,600,000,000</td>
<td>294,507,808</td>
<td>253,239,016</td>
<td>41,266,456</td>
<td>2,336</td>
</tr>
<tr>
<td>dbo.ORDERS</td>
<td>1,500,000,000</td>
<td>183,360,360</td>
<td>182,971,608</td>
<td>386,664</td>
<td>2,088</td>
</tr>
<tr>
<td>dbo.PART</td>
<td>400,000,000</td>
<td>77,759,592</td>
<td>59,011,192</td>
<td>18,744,208</td>
<td>4,192</td>
</tr>
<tr>
<td>dbo.CUSTOMER</td>
<td>300,000,000</td>
<td>52,606,624</td>
<td>52,606,136</td>
<td>160</td>
<td>328</td>
</tr>
<tr>
<td>dbo.SUPPLIER</td>
<td>20,000,000</td>
<td>3,577,360</td>
<td>3,204,200</td>
<td>364,936</td>
<td>8,224</td>
</tr>
<tr>
<td>dbo.NATION</td>
<td>25</td>
<td>160</td>
<td>8</td>
<td>96</td>
<td>56</td>
</tr>
<tr>
<td>dbo.REGION</td>
<td>5</td>
<td>160</td>
<td>8</td>
<td>96</td>
<td>56</td>
</tr>
<tr>
<td>dbo.Log</td>
<td>178</td>
<td>16</td>
<td>8</td>
<td>8</td>
<td>0</td>
</tr>
</tbody>
</table>
```

Table 4. Table layout and tables sizes of the 1TB database used
**Test Protocol:**

- To avoid variance and random effects on measurement between tests, the data buffers of SQL Server were emptied before each query was run (cold cache – this test primarily highlight the SSD throughput), and the query was run again (warm cache) to measure the raw performance provided by the memory (the data being read directly from the cache in this case)

1) Clean SQL Server’s buffer cache
   DBCC DROPCLEANBUFFERS
   GO

2) Cold cache test: Run a query and write down the performance metrics captured with the Profiler (Duration, #CPU cycles, #Reads, #Writes)

3) Warm cache test: Run a query again and write down the performance metrics captured with the Profiler (Duration, #CPU cycles, #Reads, #Writes)

4) Do this on both HW configurations and both O/S to get comparative result sets
Tests Results

1) Scenario 1 (Minimum Cost Supplier Query)

A. Query Text

```sql
SELECT s_acctbal, s_name, n_name, p_partkey, p_mfgr, s_address, s_phone,
       s_comment
FROM PART P, SUPPLIER S, PARTsupp PS, NATION, REGION
WHERE p_partkey = ps_partkey
  AND s_suppkey = ps_suppkey
  AND p_size = 7
  AND p_type LIKE '%STEEL'
  AND s_nationkey = n_nationkey
  AND n_regionkey = r_regionkey
  AND r_name = 'ASIA'
AND ps_supplycost = (SELECT MIN(ps_supplycost)
  FROM PARTsupp PS1, SUPPLIER S1, NATION N1, REGION R1
  WHERE P.p_partkey = PS1.ps_partkey
    AND S1.s_suppkey = PS1.ps_suppkey
    AND S1.s_nationkey = N1.n_nationkey
    AND N1.n_regionkey = R1.r_regionkey
    AND R1.r_name = 'ASIA')
ORDER BY s_acctbal DESC, n_name, s_name, p_partkey
```

B. Cold Cache Test

: High performance & green server solution is **13.4 times faster** than standard & non-green server solution and consumes **83% less power** to analyze minimum cost supplier query against 1TB database.

C. Warm Cache Test

: High performance & green server solution consumes **25% less power** than standard & non-green server solution to analyze minimum cost supplier query.
D. OS Performance Improvement

The running duration of the minimum cost supplier query on Windows server 2012 RC is **4.3% faster** in cold cache test than on Windows server 2008 R2 (no differences observed in the Warm cache scenario):

![System Performance Improvement w/ Windows Server 2012 OS](image)

**2) Scenario 2 (Forecasting Revenue Change Query)**

A. Query Text

```sql
SELECT SUM(l_extendedprice * l_discount) AS REVENUE
FROM LINEITEM
WHERE l_shipdate >= '1997-01-01'
  AND l_shipdate < DATEADD(yy,1,cast('1997-01-01' as SMALLDATETIME))
  AND l_discount BETWEEN 0.03 - 0.01 AND 0.03 + 0.01
  AND l_quantity < 24
```
B. Cold Cache Test

: High performance & green server solution is **14.8 times faster** than standard & non-green server solution and consumes **94% less power** to analyze forecasting revenue change query against 1TB database.

![System Performance and Power (Forecasting Revenue Change Query)](image)

**High Perf. & Green Server: 3 min 33 sec @ 40 Wh**

C. Warm Cache Test

: High performance & green server solution consumes **27% less power** than standard & non-green server solution to analyze minimum cost supplier query.

![System Performance and Power (Forecasting Revenue Change Query)](image)

**High Perf. & Green Server: 34 sec @ 8 Wh**

D. OS Performance Improvement

: The running duration of the forecasting revenue change query on Windows server 2012 RC is **5.3% faster** in cold cache test than on Windows server 2008 R2 (no significant difference observed in the Warm cache scenario):
3) Scenario 3 (Promotion Effect Query)

A. Query Text

```sql
SELECT 100.00 * SUM(CASE
    WHEN p_type LIKE 'PROMO%'
    THEN l_extendedprice * (1 - l_discount)
    ELSE 0
    END) / SUM(l_extendedprice * (1 - l_discount)) AS PROMO_REVENUE
FROM Lineitem, PART
WHERE l_partkey = p_partkey
    AND l_shipdate >= '1994-09-01'
    AND l_shipdate < DATEADD(MM,1,cast('1994-09-01' as SMALLDATETIME))
```

B. Cold Cache Test

: High performance & green server solution is **13.3 times faster** than standard & non-green server solution and consumes **94% less** power to analyze promotion effect query against 1TB database.
C. Warm Cache Test

High performance & green server solution consumes 23% less power than standard & non-green server solution to analyze promotion effect query.

D. OS Performance Improvement

The running duration of the promotion effect query on Windows server 2012 RC is 4.8% faster in cold cache test and 17.3% faster in warm cache test than on Windows server 2008 R2.
4) Scenario 4 (Potential Part Promotion Query)

A. Query Text

```sql
SELECT s_name, s_address
FROM SUPPLIER, NATION
WHERE s_suppkey in
    (SELECT ps_suppkey
     FROM PARTsupp
     WHERE ps_partkey in
         (SELECT p_partkey
          FROM PART
          WHERE p_name like 'lace%')
     AND ps_availqty > (SELECT 0.5 * sum(l_quantity)
                         FROM LINEITEM
                         WHERE l_partkey = ps_partkey
                         AND l_suppkey = ps_suppkey
                         AND l_shipdate >= '1997-01-01'
                         AND l_shipdate <
                         DATEADD(YY,1,cast('1997-01-01' as SMALLDATETIME)))
     AND s_nationkey = n_nationkey AND n_name = 'JORDAN'
ORDER BY s_name
```
B. Cold Cache Test

: High performance & green server solution is **17.1 times faster** than standard & non-green server solution and consumes **95% less power** to analyze potential part promotion query against 1TB database.

![System Performance and Power (Potential Part Promotion Query)](image)

C. Warm Cache Test

: High performance & green server solution consumes **33% less power** than standard & non-green server solution to analyze potential part promotion query.

![System Performance and Power (Potential Part Promotion Query)](image)
D. OS Performance Improvement

The running duration of the potential part promotion effect query on Windows server 2012 RC is **12.4% faster** in cold cache test than on Windows server 2008 R2 (no differences observed on the Warm cache scenario):
Abbreviations and references used in this document:

1: Source: Wikipedia definition of BIG DATA

**DPC:** DIMM per Channel, number DRAM DIMMs modules populated into each available CPU Channel. In this set up the CPUs each have three channels for memory.

**TPC-H:** The TPC Benchmark™H (TPC-H) is a decision support benchmark. It consists of a suite of business oriented ad-hoc queries and concurrent data modifications. The queries and the data populating the database have been chosen to have broad industry-wide relevance. This benchmark illustrates decision support systems that examine large volumes of data, execute queries with a high degree of complexity, and give answers to critical business questions. [http://www.tpc.org/tpch/](http://www.tpc.org/tpch/)

**RDBMS:** Relational Database Management System
History of changes

Rev. 03: Final DRAFT from Samsung side. Passed to MTC team for review. the Q9 query is not included. 2TB configuration was not tested
Rev. 06: including feedback from Stephan, Gerd and SK
Rev. 07: including feedback from Laurent
REV. 10: candidate or release, agreed in con call with MTC team to keep the title from marketing perspective, although there are aspects that are not exactly fit.
About companies

**Microsoft Technology Center (MTC)**

Microsoft Technology Centers (MTCs) are collaborative environments that provide access to innovative technologies and world-class expertise, enabling you to envision, design, and deploy solutions to meet your exact needs.  [http://www.microsoft.com/mtc/](http://www.microsoft.com/mtc/)

**Samsung Electronics Co., Ltd.**

Samsung Electronics Co., Ltd. is a global leader in semiconductor, telecommunication, digital media and digital convergence technologies with 2011 consolidated sales of US$143.1 billion. Employing approximately 206,000 people in 197 offices across 72 countries, the company operates two separate organizations to coordinate its nine independent business units: Digital Media & Communications, comprising Visual Display, Mobile Communications, Telecommunication Systems, Digital Appliances, IT Solutions, and Digital Imaging; and Device Solutions, consisting of Memory, System LSI and LED. Recognized for its industry-leading performance across a range of economic, environmental and social criteria, Samsung Electronics was named the world’s most sustainable technology company in the 2011 Dow Jones Sustainability Index. For more information, please visit [www.samsung.com/greenmemory](http://www.samsung.com/greenmemory)

**Dell**

Dell Inc. (NASDAQ: DELL) listens to customers and delivers innovative technology and services that give them the power to do more. For more information, visit [www.dell.com](http://www.dell.com)