DB2 Analytics Accelerator for z/OS
A deep dive into the real workings

Willie Favero, IBM
DB2 SME
Data Warehousing for System z Swat Team
IBM Silicon Valley Laboratory
Agenda

☑ DB2 Analytics Accelerator Refresher

☒ V3 New Function Overview
☒ V3 Incremental Update
☒ V3 High Performance Storage Saver
☒ PureData for Analytics N2001 Hardware Refresh
☒ What are Customers' Saying
☒ Conclusions
Introducing

**DB2 Analytics Accelerator**

*Now even faster!*

**Question:** What’s faster than a DB2 Analytics Accelerator with a N1001 server?

**Answer:** A DB2 Analytics Accelerator with a N2001 server!
Benefits of the N2001
The fastest performance of Netezza technology to date!

- **Accelerate Performance of Analytic Queries**
  - 3X faster performance\(^1\) for Big Data analytics
  - 128 GB/sec effective scan rate per rack\(^2\) to tackle Big Data faster

- **Increase Efficiency of your Data Center**
  - 50% greater data capacity per rack\(^3\) helps optimize data center efficiency
  - More capacity and less power per rack than both Oracle and Teradata

- **Simplicity and Ease of Administration**
  - Improved system management and resilience to spend less time managing and more time delivering value

---

\(^1\) Based on a comparison of the IBM PureData System for Analytics N2001 to the IBM PureData System for Analytics N1001. The performance speed refers to the query times on both macro-analytic and mixed workload tests as conducted in IBM engineering lab benchmarks. The N2001 query times were an average of 3x faster than those of the N1001. Individual results may vary.

\(^2\) 128 GB/sec scan rate assuming an average of 4x compression across the system. Individual results may vary.

\(^3\) Capacity of IBM PureData System for Analytics N2001 compared to previous generation IBM PureData System for Analytics N1001.
DB2 Analytics Accelerator V3

Further extending the features

More insight from your data

- Unprecedented response times for “right-time” analysis
- Complex queries in seconds rather than hours
- Transparent to the application
- Inherits all System z DB2 attributes
  - i.e. A secured environment
- No need to create or maintain indices
- Eliminate query tuning
- Fast deployment and time-to-value
- Minimize the need to create data marts for performance

Blending System z and Netezza technologies to deliver unparalleled, mixed workload performance for complex analytic business needs.
DB2 Analytics Accelerator

zEnterprise (zEC12)

Data Studio Foundation
DB2 Analytics Accelerator
Admin Plug-in

CLIENT

10 GbE

Users/ Applications

Netezza Technology

Private Service Network
Primary

10Gb
Backup

Netezza: Recognized leader in cost-effective high speed deep analytics

z/OS: Recognized leader in Security, Availability, & Recoverability for OLTP

DB2 Analytics Accelerator for z/OS
Creating the Hybrid Data Server
Combine DB2 for z/OS with Netezza to provide an industry exclusive

Best in OLTP and Transactional Analytics
Industry recognized leader in mission critical transaction systems

Best in Complex Analytics
Proven appliance leader in high speed analytic systems

Best in Consolidation
Unprecedented mixed workload flexibility and virtualization providing the most options for cost effective consolidation

Together:
Destroying the myth that transactional and decision support workloads have to be on separate platforms
Workload Assessment Query Results
International Securities Company

Customer Goals utilizing IBM DB2 Analytics Accelerator and Smart Analytic System 9700

- The customer is currently utilizing Sybase for their OLTP and Data Warehouse environments and moving the OLTP environment to DB2 for z/OS.
- The customer conducted a benchmark to determine if DB2 for z/OS with the accelerator is a viable platform for data warehousing.
- Customer would like to reduce MIPs, Software Costs and the Total Cost of Ownership with the 9700 and Accelerator

<table>
<thead>
<tr>
<th>QUERY</th>
<th>Returned Rows</th>
<th>Class 1 Elapse Time</th>
<th>Class 1 CPU Time</th>
<th>Class 2 Elapse Time</th>
<th>Class 2 CPU Time</th>
<th>ISAS 9700 Only (sec.)</th>
<th>ISAS 9700 with IDAA (sec.)</th>
<th>IDAA Effects</th>
</tr>
</thead>
<tbody>
<tr>
<td>ME01210A</td>
<td>33138</td>
<td>13.52</td>
<td>87.88</td>
<td>13.35</td>
<td>87.71</td>
<td>2.10</td>
<td>0.49</td>
<td>1.92</td>
</tr>
<tr>
<td>ME01210B</td>
<td>33244</td>
<td>13.11</td>
<td>97.78</td>
<td>12.84</td>
<td>97.51</td>
<td>1.82</td>
<td>0.48</td>
<td>1.66</td>
</tr>
<tr>
<td>ME01210C</td>
<td>23240</td>
<td>10.57</td>
<td>87.77</td>
<td>10.43</td>
<td>87.53</td>
<td>1.95</td>
<td>0.39</td>
<td>1.81</td>
</tr>
<tr>
<td>ME01210D</td>
<td>23086</td>
<td>10.88</td>
<td>81.87</td>
<td>10.74</td>
<td>81.72</td>
<td>1.90</td>
<td>0.40</td>
<td>1.75</td>
</tr>
<tr>
<td>ME01210E</td>
<td>33973</td>
<td>11.29</td>
<td>58.29</td>
<td>11.12</td>
<td>58.12</td>
<td>1.92</td>
<td>0.47</td>
<td>1.78</td>
</tr>
<tr>
<td>MG26410A</td>
<td>372</td>
<td>12.32</td>
<td>133.51</td>
<td>12.30</td>
<td>133.51</td>
<td>1.60</td>
<td>0.01</td>
<td>1.49</td>
</tr>
<tr>
<td>MG26410B</td>
<td>372</td>
<td>13.48</td>
<td>146.84</td>
<td>13.48</td>
<td>146.83</td>
<td>1.35</td>
<td>0.01</td>
<td>1.35</td>
</tr>
<tr>
<td>MG26410C</td>
<td>372</td>
<td>12.36</td>
<td>127.59</td>
<td>12.35</td>
<td>127.56</td>
<td>1.35</td>
<td>0.01</td>
<td>1.37</td>
</tr>
<tr>
<td>MG26410D</td>
<td>338</td>
<td>12.20</td>
<td>131.73</td>
<td>12.19</td>
<td>131.73</td>
<td>1.22</td>
<td>0.01</td>
<td>1.19</td>
</tr>
<tr>
<td>MG26410E</td>
<td>372</td>
<td>14.53</td>
<td>172.12</td>
<td>14.82</td>
<td>172.11</td>
<td>1.49</td>
<td>0.01</td>
<td>1.48</td>
</tr>
<tr>
<td>AD01210A</td>
<td>13</td>
<td>22.45</td>
<td>211.57</td>
<td>22.44</td>
<td>211.55</td>
<td>2.20</td>
<td>0.01</td>
<td>2.19</td>
</tr>
<tr>
<td>AD01210B</td>
<td>111</td>
<td>440.63</td>
<td>2088.55</td>
<td>440.61</td>
<td>2088.41</td>
<td>11.69</td>
<td>0.03</td>
<td>11.67</td>
</tr>
<tr>
<td>AD01210C</td>
<td>111</td>
<td>641.42</td>
<td>3381.25</td>
<td>543.41</td>
<td>3381.25</td>
<td>30.94</td>
<td>0.02</td>
<td>30.83</td>
</tr>
<tr>
<td>AD01210D</td>
<td>110</td>
<td>671.43</td>
<td>4406.00</td>
<td>671.41</td>
<td>4405.89</td>
<td>38.85</td>
<td>0.02</td>
<td>38.64</td>
</tr>
</tbody>
</table>

Customer Statistics

Data Warehouse is approximately 12 TB accelerated into the Accelerator

100 Simple queries and 112 complex queries were tested in the benchmark

212 queries were executed sequentially (9700 only and then with the Accelerator)

IBM DB2 Analytics Accelerator (Netezza 1000-12)

Day 1-2: Hardware installation and configuration
Day 3: Created and loaded tables in the Accelerator and executed first queries
### Query Details

<table>
<thead>
<tr>
<th>EXEC_DATE</th>
<th>Without Acceleration (seconds rounded)</th>
<th>Without Acceleration CPU (seconds rounded)</th>
<th>Accelerated (seconds rounded)</th>
<th>X Factor</th>
<th>ROWS_RETURNED</th>
<th>RESULT_SIZE</th>
<th>STMT_TEXT</th>
</tr>
</thead>
<tbody>
<tr>
<td>5/9/12 12:27 PM</td>
<td>210</td>
<td>7</td>
<td>6</td>
<td>35</td>
<td>26787</td>
<td>628K</td>
<td>T3.CNTR_NMBR, T3.MBR_ID,</td>
</tr>
<tr>
<td>5/9/12 12:43 PM</td>
<td>2146</td>
<td>261</td>
<td>6</td>
<td>358</td>
<td>30688</td>
<td>60.4K</td>
<td>T1.PRFX_CD,</td>
</tr>
<tr>
<td>5/9/12 12:45 PM</td>
<td>755</td>
<td>33</td>
<td>20</td>
<td>38</td>
<td>624</td>
<td>30.9K</td>
<td>T1.PRFX_CD,</td>
</tr>
<tr>
<td>5/9/12 12:46 PM</td>
<td>40</td>
<td>14</td>
<td>7</td>
<td>6</td>
<td>72627</td>
<td>2.98M</td>
<td>T1.CNTR_NMBR, T1.MBR_ID,</td>
</tr>
<tr>
<td>5/9/12 12:47 PM</td>
<td>921</td>
<td>186</td>
<td>7</td>
<td>132</td>
<td>4799</td>
<td>136K</td>
<td>PRFX_CD, CNTR_NMBR,</td>
</tr>
<tr>
<td>5/9/12 12:48 PM</td>
<td>2543</td>
<td>1695</td>
<td>15</td>
<td>170</td>
<td>3626</td>
<td>299K</td>
<td>T1.MBR_TRNS_EFCTV_DT,</td>
</tr>
<tr>
<td>5/9/12 1:02 PM</td>
<td>1128</td>
<td>536</td>
<td>6</td>
<td>188</td>
<td>122948</td>
<td>2.81MB</td>
<td>CNTR_NMBR,</td>
</tr>
<tr>
<td>5/9/12 12:56 PM</td>
<td>126</td>
<td>43</td>
<td>1</td>
<td>126</td>
<td>0</td>
<td>0 B</td>
<td>S.CNTR_NMBR, S.MBR_ID,</td>
</tr>
<tr>
<td>5/16/12 12:00 AM</td>
<td>2</td>
<td>0</td>
<td>not accel (not expected to)</td>
<td></td>
<td>528701</td>
<td>9.38 MB</td>
<td>TBL1.PRFX_CD,</td>
</tr>
<tr>
<td>5/10/12 8:15 AM</td>
<td>34</td>
<td>13</td>
<td>11</td>
<td>3</td>
<td>189</td>
<td>131508</td>
<td>AS CL_MKT_SSEG_CD AS</td>
</tr>
<tr>
<td>5/10/12 8:37 AM</td>
<td>0</td>
<td>0</td>
<td>not accel (not expected to)</td>
<td></td>
<td>418</td>
<td>18904956</td>
<td>AS CL_MKT_SSEG_CD AS</td>
</tr>
<tr>
<td>5/10/12 9:31 AM</td>
<td>189</td>
<td>24</td>
<td>1</td>
<td>189</td>
<td>131508</td>
<td>1.63 MB</td>
<td>AS CL_MKT_SEG_CD,</td>
</tr>
<tr>
<td>5/10/12 9:47 AM</td>
<td>623</td>
<td>418</td>
<td>409</td>
<td>2</td>
<td>18904956</td>
<td>631 MB</td>
<td>K_AFLT_ID_AFILT_ID AS</td>
</tr>
<tr>
<td>5/10/12 9:51 AM</td>
<td>20</td>
<td>8</td>
<td>1</td>
<td>20</td>
<td>131508</td>
<td>1.38 MB</td>
<td>K_CNTR_ID, CNTR_NMBR</td>
</tr>
<tr>
<td>5/10/12 12:54 PM</td>
<td>0</td>
<td>0</td>
<td>not accel (not expected to)</td>
<td></td>
<td>299</td>
<td>33B</td>
<td>A.K_BS_RPT_UNT_ID AS</td>
</tr>
</tbody>
</table>

**Comments:**

- **Several Ran Sub-Second**: Some queries ran in sub-second time, indicating high performance.
- **Up to 358x Faster!**: The performance improvement is significant, with queries running much faster than expected.
- **Avoided CPU Consumption**: CPU usage was significantly reduced compared to previous measures.
- **Avoided Redirecting**: Redirection, which could lead to additional processing, was prevented, enhancing efficiency.
Deep DB2 Integration within zEnterprise

**Applications**
- Application Interfaces
  - (standard SQL dialects)

**DB2 for z/OS**
- Data Manager
- Buffer Manager
- IRLM
- Log Manager

**DBA Tools, z/OS Console, ...**
- Operational Interfaces
  - (e.g. DB2 Commands)

**z/OS on System z**
- Superior availability
  - reliability, security,
  - Workload management

**IBM DB2 Analytics Accelerator**
- Superior performance on analytic queries
Query Execution Process Flow

Application Interface → Optimizer → Accelerator DRDA Requestor → Accelerator

- Queries executed with Accelerator
- Queries executed without Accelerator
- Heartbeat (availability and performance indicators)

DB2 for z/OS

Accelerator
Agenda

- DB2 Analytics Accelerator Refresher
- V3 New Function Overview
- V3 Incremental Update
- V3 High Performance Storage Saver
- PureData for Analytics N2001 Hardware Refresh
- What are Customers' Saying
- Conclusions
### New Query Routing Options

**Values for CURRENT QUERY ACCELERATION**

<table>
<thead>
<tr>
<th>Value</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>NONE</td>
<td>No query is routed to the accelerator</td>
</tr>
<tr>
<td>ENABLE</td>
<td>A query is routed to the accelerator if it satisfies the acceleration criteria including the cost and heuristics criteria. Otherwise, it is executed in DB2. If there is an accelerator failure while running the query, or the accelerator returns an error, DB2 will return a negative SQL code to the application.</td>
</tr>
<tr>
<td>ENABLE WITH FAILBACK</td>
<td>A query is routed to the accelerator if it satisfies the acceleration criteria including the cost and heuristics criteria. Otherwise, it is executed in DB2. Under certain conditions the query will run on DB2 after it fails in the accelerator. In particular, any negative SQLCODE will cause a failback to DB2 during PREPARE or first OPEN. No failback is possible after a successful OPEN of a query.</td>
</tr>
<tr>
<td>ELIGIBLE</td>
<td>A query is routed to the accelerator if it satisfies the acceleration criteria irrespective of the cost and heuristics criteria. Otherwise, it is executed in DB2.</td>
</tr>
<tr>
<td>ALL</td>
<td>A query is routed to the accelerator. If it cannot be executed there, the query fails and a negative return code is passed back to the application.</td>
</tr>
</tbody>
</table>
Workload Manager Integration Usage scenarios

• Workload Isolation:
  Ensure that the workload of one DB2 subsystem doesn’t monopolize the resources of a shared accelerator. A development subsystem, attached to the same accelerator as a production subsystem, should not be able to drain all accelerator resources.

• Query Prioritization:
  More important queries should be executed before and faster than less important queries that are sent from the same DB2 subsystem against the accelerator.
Workload isolation

• The goal is to ensure that specific DB2 subsystems have a guaranteed amount of available resources on a shared accelerator.

• Netezza GRA (Guaranteed Resource Assignment) is used to ensure that workloads against a specific subsystem (identified by user group) get a guaranteed amount of resources.

• The system programmer can specify the required amount of resources on the customer configuration console of the accelerator after pairing.

• IDAA maps this request to GRA definitions against the Netezza system.
Workload isolation

Licensed Materials - Property of IBM
5697-SAO
(C) Copyright IBM Corp. 2009, 2012.
US Government Users Restricted Rights -
Use, duplication or disclosure restricted by GSA ADP Schedule
Contract with IBM Corporation

**********************************************************************
* Welcome to the IBM DB2 Analytics Accelerator Console
**********************************************************************

You have the following options:
(1) - Generate a pairing code and display IP-address and port.
(2) - Execute 'nzstart' on the Netezza host.
(3) - Execute 'nzstate' on the Netezza host.
(4) - Execute 'nzstop' on the Netezza host.
(5) - Change the configuration console password.
(6) - Change Resource Allocation for connected DB2 subsystems.
(x) - Exit the Configuration Console.
Workload isolation

- An minimum resource requirement can be specified per attached DB2 subsystem (or sharing group) in a range from 1%..100%.

- The following example shows a test system that requires at least 30% resource allocation while the production requires 80% of all resources.

- If the \( \sum (\text{MinResources}) \) over all subsystems is >100%, the resource allocation is distributed proportional. In the below example, 'TEST' receives 3/11 (27%) and 'PRODUCTION' 8/11 (73%) as a real minimum.

- The implicit maximum resource allocation is 100%

<table>
<thead>
<tr>
<th>Index</th>
<th>Location Name</th>
<th>Minimum Resource Allocation (in percent)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>TEST</td>
<td>40</td>
</tr>
<tr>
<td>2</td>
<td>PRODUCTION</td>
<td>80</td>
</tr>
</tbody>
</table>

Select database system by location name or index (use empty string or 0 to go back): 1

Minimum resource allocation for system with location name 'TEST' (1-100, 0 or empty to abort): 30

Successfully updated minimum resource allocation for system with location name 'TEST' to 30.

Press <return>

<table>
<thead>
<tr>
<th>Index</th>
<th>Location Name</th>
<th>Minimum Resource Allocation (in percent)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>TEST</td>
<td>30</td>
</tr>
<tr>
<td>2</td>
<td>PRODUCTION</td>
<td>80</td>
</tr>
</tbody>
</table>
Query Prioritization

- DB2 sends the importance level to IDAA with each query.
- IDAA maps the importance level to a Netezza priority and alters the session prior to query execution, using the corresponding priority. Also threads scheduled will have their priorities adjusted.
  - The changes in prioritization after query start are not reflected
- Netezza supports 4 different priority levels, therefore multiple WLM importance levels have to be mapped against the same Netezza priority.

<table>
<thead>
<tr>
<th>WLM Importance Level</th>
<th>Netezza Priority</th>
</tr>
</thead>
<tbody>
<tr>
<td>System</td>
<td>Critical</td>
</tr>
<tr>
<td>Importance 1</td>
<td>Critical</td>
</tr>
<tr>
<td>Importance 2</td>
<td>High</td>
</tr>
<tr>
<td>Importance 3</td>
<td>Normal</td>
</tr>
<tr>
<td>Importance 4</td>
<td>Normal</td>
</tr>
<tr>
<td>Importance 5</td>
<td>Normal</td>
</tr>
<tr>
<td>Discretionary</td>
<td>Low</td>
</tr>
</tbody>
</table>
Maintenance tasks

* Welcome to the IBM DB2 Analytics Accelerator Console *

You have the following options:
(1) - Generate a pairing code and display IP-address and port.
(2) - Execute 'nzstart' on the Netezza host.
(3) - Execute 'nzstate' on the Netezza host.
(4) - Execute 'nzstop' on the Netezza host.
(5) - Change the configuration console password.
(6) - Change resource allocation for connected DB2 subsystems.
(7) - Clear the query history.
(8) - Specify the priority of maintenance tasks.
(x) - Exit the Configuration Console.

Configured database systems:
1: DWAD911
2: UTEC812

Select database system or enter 0 to go back: (Default 0) > 2

Queries are run with the priority that is set in the corresponding System z Workload Manager (WLM) environment. You can set the priority of maintenance operations, such as loading data, analogously for the 'UTEC812' subsystem:
1: SYSTEM
2: HIGHEST
3: HIGH
4: NORMAL
5: LOW
6: LOWEST
7: DISCRETIONARY

Enter the appropriate number to set the priority level.
To return to the main menu of the Configuration Console, just press Enter. (Default 0) > Canceled!
Detect Staleness of Data

**Goal:** Make data maintenance on the accelerator easier and reduce resource consumption

**Situation in IDAA V2:**
- Accelerator data can be refreshed from DB2 at table or partition granularity
- Admin must know explicitly which data needs to be refreshed (e.g. from ETL process)

**New feature:** find out if DB2 data for a table or partition has changed since last load on accelerator
- Avoid refreshing data that is already up-to-date without requiring explicit knowledge about data changes
- Based on DB2 real-time statistics about data updates

IDAA V3 also adds support for partition-wise reload of partition-by-growth tables
- Explicit partition reload is not suitable, because you cannot usually know which partitions are affected by a data change …
- … but the change detection mechanisms work the same as for range-partitioned tables
Interactive usage scenario

In the IDAA admin GUI

- When re-loading tables, show which tables or partitions need to be reloaded and how much data needs to be loaded
- Allow to manually correct “in-doubt” cases where no automated decision is possible
- Implemented as a new load dialog with additional features
- There is also a new stored procedure behind the dialog that can be called directly to retrieve the information
Batch usage scenario

From JCL or DB2 client

- Refresh one or more tables by loading only the smallest necessary amount of data (partitions or unpartitioned tables)
- If a table has not changed since last load, no data is transferred
- In-doubt cases will trigger a reload -- “false positives” are possible
- Stored procedure call can be scheduled in intervals (e.g. daily) as a simple synchronization mechanism
- Implemented as a new input flag for ACCEL_LOAD_TABLES stored procedure

Note:

Since V2 PTF3, IDAA ships a sample C program that can be used to call some IDAA stored procedures (especially ACCEL_LOAD_TABLES) easily from JCL

```plaintext
//LOADCHNG EXEC PGM=IKJEFT01
//AQTP1 DD *
MYACCEL
//AQTP2 DD *
PARTITIONS
//AQTP3 DD *
<?xml version="1.0" encoding="UTF-8" ?>
<table xmlns:aqt="..." version="1.1">
detectChanges="DATA"/>
</table>
//SYSTSIN DD *
DSN SYSTEM(!DSN!)
RUN PROGRAM(AQTSCALL) -
   PLAN(AQTSCALL) -
   LIB('AQT.V3.USRMOD') -
   PARMS('LOADTABLES')
END
```
Misc. IDAA V3 Functionality

- Accelerating SELECT part of 'INSERT from SELECT'
- Further speeding-up IDAA content refresh process, and reducing associated resources consumption at IDAA
- Enhancing scope for query offloading
  - Additional functions
Agenda

- DB2 Analytics Accelerator Refresher
- V3 New Function Overview
- V3 Incremental Update
- V3 High Performance Storage Saver
- PureData for Analytics N2001 Hardware Refresh
- What are Customers' Saying
- Conclusions
Accelerator Data Load

- Rate can vary, depending on CPU resources, table partitioning, columns...
- Update on table partition level, concurrent queries allowed during load
- V2.1 & V3 unload in DB2 internal format, single translation by accelerator
Option 1: Full Table Refresh

- Changes in data warehouse tables typically driven by scheduled (nightly or more frequently) ETL process
- Data used for complex reporting based on consistent and validated content (e.g., weekly transaction reporting to the central bank)
- Multiple sources or complex transformations prevent propagation of incremental changes
- Full table refresh triggered through DB2 stored procedure (scheduled, integrated into ETL process or through GUI)
- Queries may continue during full table refresh for accelerator
Option 2: Table Partition Refresh

- Changes in data warehouse table typically driven by “delta” ETL process (considering only changes in source tables compared to previous runs) or by more frequent changes to most recent data

- Optimization of Option 1 when target data warehouse table is partitioned and most recent updates are only applied to the latest partition

- Table partition refresh triggered through DB2 stored procedure (scheduled, integrated into ETL process or through GUI)

- Maintains snapshot semantics for consistent reports

- Queries may continue during table partition refresh for accelerator
Option 3: Incremental Update

- Changes in data warehouse tables typically driven by replication or manual updates
  - Corrections after a bulk-ETL-load of a data warehouse table
  - Continuously changing data (e.g. trickle-feed updates from a transactional system to an ODS)

- Reporting and analysis based on most recent data

- May be combined with Option 1 & 2 (first table refresh and then continue with incremental updates)

- Incremental update can be configured per database table
Incremental Update Solution Architecture

DB2 for z/OS

IDAA Stored Procedures
ACCEL_SET_TABLES_REPLICATION
ACCEL_CONTROL_ACCELERATOR

JCL

Capture Agent
DB2 z/OS
(Log reading)

NPS host

IDAA SERVER

Controller

Catalog
information

Netezza Database

Automation code
(create data sources,
subscriptions, etc.)

Apply Agent
on NPS host
(Receive log events)

(private network)
User Interface

Incremental update UI elements only visible if it has been enabled on the DB2 subsystem via the accelerator configuration console

- Start / stop replication process (per subsystem-accelerator pair)
- Enable / disable replication (per table)
- Trace collection
- Information on replication latency and events
## Limitations of the Incremental Update Functionality

<table>
<thead>
<tr>
<th>Limitation</th>
<th>Details</th>
</tr>
</thead>
<tbody>
<tr>
<td>Incremental update support is limited to 2 subsystems per Accelerator.</td>
<td>Optimizations require considerable amounts of RAM per apply agent on the NPS host.</td>
</tr>
<tr>
<td>Only tables with (logical) unique key are supported.</td>
<td>Required to locate rows for delete/update operations (predicate). Our assumption is that most “real” warehouse tables do have a logical unique key. We either use the unique constraint (if defined in DB2) or default (“all columns” as key).</td>
</tr>
</tbody>
</table>
Incremental update considerations

- CDC apply agent requires 4 GB of main memory per NZ database
  - Each DB2 for z/OS subsystem has its own NZ database
  - Limit of two apply agents per IDAA

- Example
  - e0 byte row table… f0,000 inserts per second
  - Average of f0-ot0 CPU Secs per 100mb
    - Mostly influenced by:
      - Presence of Deletes and Updates
      - Total size of the tables (e.g. rowcount * rowsize) that are used for incremental update
      - Presence of unique index or a unique constraint on the tables to be replicated
Netezza Apply Tidbits

- INSERTs are always appended to the last extent, this is very efficient

- Apply throughput is dominated by the time it is required to apply the DELETEs:
  - The presence of a unique constraint or unique index on the replicated table (the predicate to DELETE a row is based on the values of the unique column). If there is no unique constraint or index, all values of the row to be DELETED are used for the predicate
  - Since CDC uses a predicate on the unique index columns, the table should be clustered (organizing key) on these columns so that zonemaps have maximum effect for the DELETE scan.
  - A DELETE causes a scan of the entire table, hence table size, distribution and machine size is an important factor

- UPDATEs are decomposed into <DELETE,INSERT> pairs, hence UPDATEs behave very similar to DELETEs

- All operations are batched in 1 minute batches, i.e. all changes collected during that time are applied to a table "at once", with queries similar to:
  - INSERT: insert into table xxx (select * from external ... /cdc/apply/pipe);
  - DELETE: delete from table xxx where name,zip in (select name,zip from external ... /cdc/apply/pipe);
Impact on concurrently running queries

- Validated that incremental update has only minor impact on query response time
  - "No" workload:
    - 10x parallel queries: 5 streaming, 5 aggregation / group by
  - "Medium" workload:
    - 10x parallel queries: 5 streaming, 5 aggregation / group by
    - Replication from 1 subsystem: 300,000 rows/minute / 5,000 rows/s
  - "Full" workload
    - 10x parallel queries: 5 streaming, 5 aggregation / group by
    - Replication from 2 subsystems: 2.0 mio rows/minute, 33.333 rows/s
## Synchronization Options with DB2 Analytics Accelerator

<table>
<thead>
<tr>
<th>Synchronization options</th>
<th>Use cases, characteristics and requirements</th>
</tr>
</thead>
</table>
| **Full table refresh**        | ▪ Existing ETL process replaces entire table  
▪ Multiple sources or complex transformations  
▪ Smaller, un-partitioned tables  
▪ Reporting based on consistent snapshot |
| **Table partition refresh**   | ▪ Optimization for partitioned warehouse tables, typically appending changes “at the end”  
▪ More efficient than full table refresh for larger tables  
▪ Reporting based on consistent snapshot |
| **Incremental Update**        | ▪ Scattered updates after “bulk” load  
▪ Reporting on continuously updated data (e.g., an ODS), considering most recent changes  
▪ More efficient for smaller updates than full table refresh |
Agenda

- DB2 Analytics Accelerator Refresher
- V3 New Function Overview
- V3 Incremental Update
- V3 High Performance Storage Saver
- PureData for Analytics N2001 Hardware Refresh
- What are Customers' Saying
- Conclusions
High Performance Storage Saver
Reducing the cost of high speed storage

*Store historic data on the Accelerator only*

Tables can be resident on:
1. DB2 Only
2. DB2 and Accelerator
3. Archive to Accelerator

When data no longer requires updating, reclaim the DB2 storage

Managed by zPARMs
Controlled by Special Registers:
- CURRENT QUERY ACCELERATION
- CURRENT GET_ACCEL_ARCHIVE

- **Best for OLTP**
- **High Speed Indexed queries**
- **Mixed Workload**
- **Active Only**
- **Archive Only**
- **Active & Archive**
- **Mixed Workload**
Supplied Stored Procedure encapsulates Archiving Procedure

Application

DB2

IDAA

CALL storproc
ARCHIVE TABLE X
OLDER THAN date \(^{(1)}\)

(1) 'date' is specified in terms of the partitioning key values

in this particular example 'date' indicates that only last two partitions should remain in DB2

part n
part n-1
part n-2
part 2
part 1

backup part 1 backup part 2 \(\ldots\) backup part n-1 backup part n

part n
part n-1
part n-2
part 2
part 1

Partitions to be Archived are first backed up

CALL storproc
ARCHIVE TABLE X
OLDER THAN date

in this particular example 'date' indicates that only last two partitions should remain in DB2
Old Partitions are deleted from DB2 and Table X is split within IDAA

Application

DB2

IDAA

Old partitions are still present in the table, but they are empty and the disk space use is limited to the primary allocation quantity which can be made very small.

CALL storproc
ARCHIVE TABLE X
OLDER THAN date

backup part 1
backup part 2
backup part n-1
backup part n
Applications have transparent access (no SQL statement changes needed) to the Table

Application

SELECT FROM X

Set special register

SELECT FROM X

Routing?

DB2

yes

no

part n

part n-1

IDAA

UNION

TABLE X

part n

part n-1

part n-2

. .

part 2

part 1

backup part 1

backup part 2

. .

backup part n-1

backup part n
High Performance Storage Saver
Reducing the cost of high speed storage

- Time-partitioned tables where:
  - only the recent partitions are used in a transactional context (frequent data changes, short running queries)
  - the entire table is used for analytics (data intensive, complex queries).
- High Performance Storage Saver’s “Archive” Process:
  - Data is loaded into Accelerator if not already loaded
  - Automatically takes Image Copy of Each Partition to be Archived
  - Automatically Remove data from DB2 archived tablespace partitions
  - DBA starts archived partitions as read-only
Agenda

- DB2 Analytics Accelerator Refresher
- V3 New Function Overview
- V3 Incremental Update
- V3 High Performance Storage Saver
- PureData for Analytics N2001 Hardware Refresh
- What are Customers' Saying
- Conclusions
IBM PureData System for Analytics
Optimized exclusively for analytic data workloads

Speed
- 10-100x faster than traditional custom systems
- Patented, hardware accelerated MPP (Massively Parallel Processing)

Simplicity
- Data load ready in hours
- No database indexes
- No tuning
- No storage administration

Scalability
- Peta-scale data capacity

Smart
- Designed to run complex analytics in minutes, not hours

1 Based on IBM customers’ reported results. "Traditional custom systems" refers to systems that are not professionally pre-built, pre-tested and optimized. Individual results may vary.
2 Peta Scale capacity offered in the N1001 model
Accelerate Analytic Performance
With the fastest out of the box scan rates

Scan Rate limits how quickly data can be read and processed

- Big Data complex analytics requires fast access to big volumes of data
- Storage bottlenecks limit efficiency and mean long wait times
- Faster performance drives revenue for data driven businesses

PureData System for Analytics

- Increased parallelism
- Offloads decompression, filtering, and processing to FPGA for blazing fast performance!
- Industry leading out of the box effective scan rate of 128 GB/sec

1 Based on a scan rate of 38 GB/sec per rack for Teradata 2690 from http://www.teradata.com/News-Releases/2011/Fifth-Generation-Teradata-Data-Warehouse-Appliance-Delivers-Double-the-Performance-and-Triple-the-Data-Capacity/ versus IBM PureData System for Analytics N2001 scan rate of 32 GB/sec raw scan rate per rack, 128 GB/sec effective scan rate per rack. 1128 GB/sec scan rate per rack assuming an average of 4X compression across the system. Individual results may vary. Scan rates per rack are based on out of the box configurations.

2 Based on a scan rate of 100GB/sec for flash cache operations from Oracle X3-2 datasheet versus IBM PureData System for Analytics N2001 scan rate of 32 GB/sec raw scan rate, 128 GB/sec effective scan rate. 1128 GB/sec scan rate assuming an average of 4X compression across the system. Individual results may vary. Based on per rack and out of the box configurations.

3 128 GB/sec scan rate assuming an average of 4X compression across the system. Individual results may vary.
How we did it, conceptually

More Drives with Faster Scan Rates

Faster FPGA Cores, Driving Higher Performance

Leading to Faster Performance

2.5 drives @ 130 MB/sec each
1 drive @ 120 MB/sec

FPGA Core
- Decompress
- Project
- Filter

1000 MB/sec

1000 MB/sec +

1000 MB/sec +

1000 MB/sec +

CPU Core
- Analyze

Balanced Performance
Outstanding Mixed Workload Performance

Benchmark of a mixture of light and power user queries

Queries Per Minute

- Single rack of N1001 – 13 QPM for mixed workload.
- Single rack of N2001 – 45 QPM for mixed workload.
N2001 Hardware Overview

**User Data Capacity:** 192 TB*
**Data Scan Speed:** 478 TB/hr*
**Load Speed (per system):** 5+ TB/hr

**Power Requirements:** 7.5 kW
**Cooling Requirements:** 27,000 BTU/hr

12 Disk Enclosures
- 288 600 GB SAS2 Drives
  - 240 User Data, 14 S-Blade
  - 34 Spare
  - RAID 1 Mirroring

2 Hosts (Active-Passive)
- 2 6-Core Intel 3.46 GHz CPUs
- 7x300 GB SAS Drives
- Red Hat Linux 6 64-bit

7 PureData for Analytics S-Blades™
- 2 Intel 8 Core 2+ GHz CPUs
- 2 8-Engine Xilinx Virtex-6 FPGAs
- 128 GB RAM + 8 GB slice buffer
- Linux 64-bit Kernel

Scales from ½ Rack to 4 Racks
The PureData System for Analytics AMPP Architecture

Field Programmable Gate Array = a blank canvas until it's configured

“Lite” Host (IBM xSeries, Red Hat Linux)

Disk Enclosures

S-Blades

Network Fabric

PureData System for Analytics Appliance

Applications

FPGA

CPU

Memory

Advanced Analytics

BI

ETL

Loaders

“Lite” Host

Applications

Disk Enclosures

S-Blades

Network Fabric

PureData System for Analytics Appliance

Field Programmable Gate Array = a blank canvas until it’s configured

“Lite” Host (IBM xSeries, Red Hat Linux)
Speed Through Hardware Acceleration

325 MB/sec
(2.5 drives / core)

N2001

FPGA Core

1600 MB/sec

CPU Core

800 MB/sec

Decompress
Project
Restrict Visibility
SQL & Advanced Analytics

From
Select
Where
Group by

120 MB/sec
500 MB/sec
800 MB/sec
480 MB/sec
1300 MB/sec
130 MB/sec

(2.5 drives / core)
Select State, Age, Gender, count(*) From MultiBillionRowCustomerTable Where BirthDate < '01/01/1960' And State in ('FL', 'GA', 'SC', 'NC') Group by State, Age, Gender Order by State, Age, Gender

S-Blade Data Stream Processing
## PureData System for Analytics Models

<table>
<thead>
<tr>
<th></th>
<th>Pure Data System for Analytics N1001</th>
<th>Pure Data System for Analytics N2001</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Blade Type</strong></td>
<td>HS22</td>
<td>HX-5</td>
</tr>
<tr>
<td><strong>CPU Cores / Blade</strong></td>
<td>2 x 4 Core Intel CPUs</td>
<td>2 x 8 Core Intel CPUs</td>
</tr>
<tr>
<td><strong># Disks</strong></td>
<td>96 x 3.5” / 1 TB SAS (92 Active)</td>
<td>288 x 2.5” / 600GB SAS2 (240 Active)</td>
</tr>
<tr>
<td><strong>Raw Capacity</strong></td>
<td>96 TB</td>
<td>172.8 TB</td>
</tr>
<tr>
<td><strong>Total Disk Bandwidth</strong></td>
<td>~11 GB/s</td>
<td>~32 GB/s</td>
</tr>
<tr>
<td><strong>S-Blades per Rack (cores)</strong></td>
<td>14 (112)</td>
<td>7 (112)</td>
</tr>
<tr>
<td><strong>S-Blade Memory</strong></td>
<td>24 MB</td>
<td>128 MB</td>
</tr>
<tr>
<td><strong>Rack Configurations</strong></td>
<td>¼, ½, 1, 1 ½, 2 – 10</td>
<td>½, 1, 2, 4</td>
</tr>
<tr>
<td><strong>FPGA Cores / Blade</strong></td>
<td>8 (2 x 4 Engine Xilinx FPGA)</td>
<td>16 (2 x 8 Engine Xilinx Virtex 6 FPGA)</td>
</tr>
<tr>
<td><strong>User Data / Rack</strong></td>
<td>128 TB</td>
<td>192 TB</td>
</tr>
</tbody>
</table>

*Assuming 4x Compression*
### N1001 Systems and Sizes

*PureData System for Analytics N1001*

<table>
<thead>
<tr>
<th></th>
<th>002</th>
<th>005</th>
<th>010</th>
<th>015</th>
<th>020</th>
<th>030</th>
<th>040</th>
<th>060</th>
<th>080</th>
<th>100</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cabinets</td>
<td>1/4</td>
<td>1/2</td>
<td>1</td>
<td>1</td>
<td>1/2</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>6</td>
<td>8</td>
</tr>
<tr>
<td>S-Blades</td>
<td>3</td>
<td>6</td>
<td>12</td>
<td>18</td>
<td>24</td>
<td>36</td>
<td>48</td>
<td>72</td>
<td>96</td>
<td>120</td>
</tr>
<tr>
<td>Processing Units</td>
<td>24</td>
<td>48</td>
<td>96</td>
<td>144</td>
<td>192</td>
<td>288</td>
<td>384</td>
<td>576</td>
<td>768</td>
<td>960</td>
</tr>
<tr>
<td>Capacity (TB)</td>
<td>8</td>
<td>16</td>
<td>32</td>
<td>48</td>
<td>64</td>
<td>96</td>
<td>128</td>
<td>192</td>
<td>256</td>
<td>320</td>
</tr>
<tr>
<td>Effective Capacity</td>
<td>32</td>
<td>64</td>
<td>128</td>
<td>192</td>
<td>256</td>
<td>384</td>
<td>512</td>
<td>768</td>
<td>1024</td>
<td>1280</td>
</tr>
</tbody>
</table>

*Predictable, Linear Scalability throughout entire family*

- Capacity = User Data space
- Effective Capacity = User Data Space with compression

*: 4X compression assumed*
# N2001 Systems and Sizes

*PureData System for Analytics N2001*

<table>
<thead>
<tr>
<th></th>
<th>005</th>
<th>010</th>
<th>020</th>
<th>040</th>
<th>Watch this space</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cabinets</td>
<td>1/2</td>
<td>1</td>
<td>2</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>S-Blades</td>
<td>4</td>
<td>7</td>
<td>14</td>
<td>28</td>
<td></td>
</tr>
<tr>
<td>Processing Units</td>
<td>56</td>
<td>112</td>
<td>224</td>
<td>448</td>
<td></td>
</tr>
<tr>
<td>Capacity (TB)</td>
<td>24</td>
<td>48</td>
<td>96</td>
<td>192</td>
<td></td>
</tr>
<tr>
<td>Effective Capacity</td>
<td>96</td>
<td>192</td>
<td>384</td>
<td>768</td>
<td></td>
</tr>
</tbody>
</table>

*Predictable, Linear Scalability throughout entire family*

- Capacity = User Data space
- Effective Capacity = User Data Space with compression

*: 4X compression assumed
What’s New: Summary

- 3x scan rate vs N1001 series
- 50% Greater Storage Capacity per rack
- 3x mixed workload throughput
- Improved Resiliency and Fault Tolerance
  - More spare drives per cabinet
  - Faster drive regeneration
  - Online Firmware upgrades
The Speed and Capacity for Data Mart Consolidation

Utilize virtualization to optimize the use of resources while reducing costs and gaining new agility

Benefits

- Consolidate isolated islands of data on one secure Hub
- Simplify management of costly/complex data copies while retaining the isolation benefits of individual platforms
- Provide consistency to informational data
- Increase time to value to deploy new Data Marts
- Enable application queries which would you prefer to run with more real-time data on System z

Gain control over the hundreds of data marts
Agenda

- DB2 Analytics Accelerator Refresher
- V3 New Function Overview
- V3 Incremental Update
- V3 High Performance Storage Saver
- PureData for Analytics N2001 Hardware Refresh
- What are Customers' Saying
- Conclusions
Business Innovation with IBM DB2 Analytics Accelerator

Banco do Brasil delivers IT at the speed of business by eliminating critical reporting latency

Banco do Brasil is using DB2 Analytics Accelerator to drive customer insight from operational data. Processes that previously took 11 hours to run now complete in 26 seconds!
Business Innovation with IBM DB2 Analytics Accelerator

Banca Carige is doing things they could never do before, changing the way they service their customers!

“DB2 Analytics Accelerator helps over 1,000 business users to get fast access to vital insights – informing the development of new products, services and strategies to grow the business.”

Daniele Cericola, CIO, Banca Carige
Business Innovation with IBM DB2 Analytics Accelerator

Aetna is now focused on business needs not technical constraints, positioned to expand their membership and provide insight faster without impacting existing applications and infrastructure.

“...it means our queries run dramatically faster”

“With the aging population, we expect a huge influx of data, so the cost of storing data is significant. By keeping data in the appliance, we expect substantial storage cost savings.”

Jeff Kohan, Systems Engineering Manager, Aetna
Agenda

- DB2 Analytics Accelerator Refresher
- V3 New Function Overview
- V3 Incremental Update
- V3 High Performance Storage Saver
- PureData for Analytics N2001 Hardware Refresh
- What are Customers' Saying
- Conclusions
IBM zEnterprise Analytics System 9700

DB2 Analytics Accelerator is now a standard feature

The next generation of System z analytics; an integrated solution of hardware, software and services that enables customers to rapidly deploy cost effective game changing analytics across their business.

Preselected
All the necessary components are identified and integrated into an end-to-end solution

Pretested
Over 20 different customer typical configurations are presized and tested

Solution Priced
Aggressively priced for a cost-effective add-on or new deployment for customers with critical data operations

Making every decision on facts at the point of impact
Visit the zAnalytics website
Shameless Self Promotion

Please Visit My DB2 for z/OS Blog

http://it.toolbox.com/blogs/db2zos/

QR code points directly to my blog
Thank You For Attending!
Willie Favero

DB2 SME
Data Warehousing for System z Swat Team
IBM Silicon Valley Laboratory

My DB2 Blog
www.it.toolbox.com/blogs/db2zos/

http://www.WillieFavero.com