

Get most out of IBM DB2 Analytics Accelerator: Customer and PoC Experiences

Patric Becker

IBM

Session Code: F15

17.10.2013 08:30 - 09:30 | Platform: DB2 for z/OS



Disclaimer

© Copyright IBM Corporation 2013. All rights reserved.

U.S. Government Users Restricted Rights - Use, duplication or disclosure restricted by GSA ADP Schedule Contract with IBM Corp.

IBM's statements regarding its plans, directions, and intent are subject to change or withdrawal without notice at IBM's sole discretion. Information regarding potential future products is intended to outline our general product direction and it should not be relied on in making a purchasing decision. The information mentioned regarding potential future products is not a commitment, promise, or legal obligation to deliver any material, code or functionality. Information about potential future products may not be incorporated into any contract. The development, release, and timing of any future features or functionality described for our products remains at our sole discretion.

*IBM, the IBM logo, ibm.com, DB2, and DB2 for z/OS are trademarks or registered trademarks of International Business Machines Corporation in the United States, other countries, or both. If these and other IBM trademarked terms are marked on their first occurrence in this information with a trademark symbol (® or ™), these symbols indicate U.S. registered or common law trademarks owned by IBM at the time this information was published. Such trademarks may also be registered or common law trademarks in other countries. A current list of IBM trademarks is available on the Web at "Copyright and trademark information" at www.ibm.com/legal/copytrade.shtml
Other company, product, or service names may be trademarks or service marks of others.*



Agenda

Finding queries to route to the accelerator

Considerations for using incremental update

Optimizing load strategy for maximum throughput

High Performance Storage Saver - Application integration

High Availability and Disaster Recovery thoughts

"Nice to know" pieces



Agenda

Finding queries to route to the accelerator

Considerations for using incremental update

Optimizing load strategy for maximum throughput

High Performance Storage Saver - Application integration

High Availability and Disaster Recovery thoughts

"Nice to know" pieces

Finding queries to route to the accelerator

- Finding appropriate workload
 - Reporting queries that can tolerate latency of queried data
 - Long-running, scanning large amounts of data
 - eg credit card reporting in online banking or financial reporting for banks or insurance companies
 - Involve DWH / BI folks for application context
 - Dynamic Statement Cache (IFCIDs 316 - 318)
 - Package Analysis (Performance DWH)
 - Watch out for ELT/ETL processing queries, these might not be the best candidates
 - Data transfer rates of IDAA
 - ROT: Calculate with 60 MB / sec for DRDA transfer to DB2 for z/OS
 - Watch out for streaming queries
 - Assuming ROT: 300 GB can take 5.000 seconds to return to DB2 for z/OS
 - Watch out for queries requiring transactional consistency of data

Transactional consistency - not a use case...

| | DB2 for z/OS | Accelerator |
|---------|---|---|
| Content | <i>Account 18, Balance 500 US\$</i> | <i>Account 18, Balance 500 US\$</i> |
| | Transaction 1: <i>Update Balance: + 300 US\$</i> | |
| Content | <i>Account 18, Balance 800 US\$</i> | <i>Account 18, Balance 500 US\$</i> |
| | Transaction 1 continues and reads data for account 18 to display the correct balance | |
| | Result of SELECT from DB2 for z/OS: | Result of SELECT from the accelerator: |
| Result | <i>Account 18, Balance 800 US\$</i> | <i>Account 18, Balance 500 US\$</i> |

Watch out for "read your own writes" semantics



Finding queries to route to the accelerator

- Verify eligibility of candidate queries **without** an accelerator
- Prerequisites (might be applied by maintenance already):

<http://www-01.ibm.com/support/docview.wss?uid=swg27035960>

- Configure subsystem for acceleration:
- DSNZPARMs
 - ACCEL COMMAND and QUERY_ACCELERATION NONE

- Installation of IBM DB2 Analytics Accelerator Studio (GUI):

<http://www.ibm.com/developerworks/downloads/im/data/>

- Installation documentation:

<http://www-01.ibm.com/support/docview.wss?uid=swg27036848>

Finding queries to route to the accelerator

- Add virtual accelerator:



- Add tables (virtually) to virtual accelerator:

Virtual Accelerator: TEST @ DWHDA11

Acceleration: Explain Only [Stop](#)

► About

▼ Tables (1 of 1 enabled for acceleration)

Add... Remove Acceleration ▼

Name like:

| Name | Acceleration |
|------------|--------------|
| TPORT2 | 1 of 1 |
| PLAN_TABLE | Enabled |

Finding queries to route to the accelerator

- Create DSN_QUERYINFO_TABLE (see DSNTESC in SDSNSAMP)

- Explain query:

```
SET CURRENT QUERY ACCELERATION = ENABLE;
```

- Or: ALL / ELIGIBLE to ignore heuristics, but you might want to start with ENABLE

```
EXPLAIN ALL SET QUERYNO = 1 FOR
```

```
SELECT * FROM schema.table
```

```
WHERE [...]
```

```
WITH UR;
```

- Check DSN_QUERYINFO_TABLE:

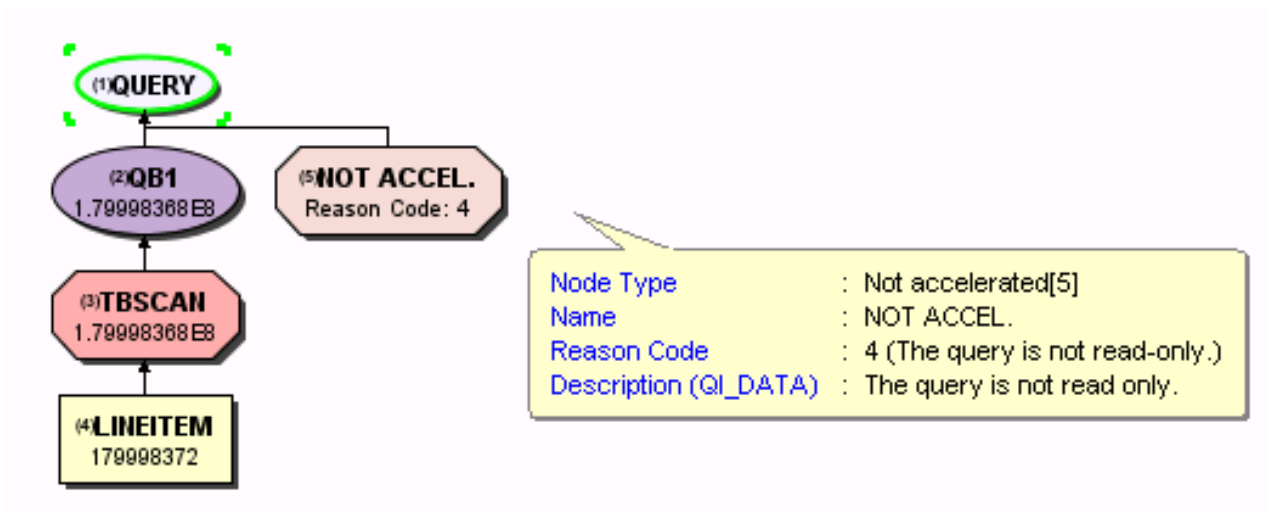
```
SELECT REASON_CODE, QI_DATA
```

```
FROM schema.DSN_QUERYINFO_TABLE
```

```
WHERE QUERYNO = 1;
```

Finding queries to route to the accelerator (c'ted)

- Visual Explain with V4.1



Finding queries to route to the accelerator

- APAR PM90886 (PTF UK97693) brings accelerator estimate for dynamic SQL (comparable to zIIP / zAAP offload estimates: IICPU in AccRep, parameter PROJECTPU in SYS1.PARMLIB(IEAOPT00))
 - DB2 10 for z/OS / DB2 11 for z/OS
 - Accumulated values for elapsed / CP time on PLAN level
 - Accelerator eligibility depends on application context
- APAR PM95035 for static SQL (only APARfix yet) - REBIND needed

| TIMES/EVENTS | APPL (CL.1) | DB2 (CL.2) | IFI (CL.5) |
|------------------|-------------|-------------------|------------|
| ----- | ----- | ----- | ----- |
| ELAPSED TIME | 23:41.8231 | 23:41.8118 | N/P |
| NONNESTED | 23:41.8231 | 23:41.8118 | N/A |
| [...] | | | |
| ELIG TIMES | | | |
| ACCEL ELA | N/A | 2:53:22.55 | N/A |
| [...] | | | |
| ELIG TIMES | | | |
| SECP CPU | N/A | N/A | N/A |
| ACCEL CPU | N/A | 3:32.17088 | N/A |
| [...] | | | |
| ELIG TIMES | | | |
| ACCEL SE | N/A | 4:40.90741 | N/A |

Static SQL and IBM DB2 Analytics Accelerator

- Default from DSNZPARM - Not taken into account for REBIND
- Explicit SET not possible due to code execution sequence
- No message is externalized to determine query routing at BIND time
 - Columns STATUS in SYSIBM.SYSPACKSTMT contains 'O' for a static query that has been bound for acceleration
 - Use DSN_QUERYINFO_TABLE and EXPLAIN(YES) or SYSIBM.SYSPACKSTMT



Agenda

Finding queries to route to the accelerator

Considerations for using incremental update

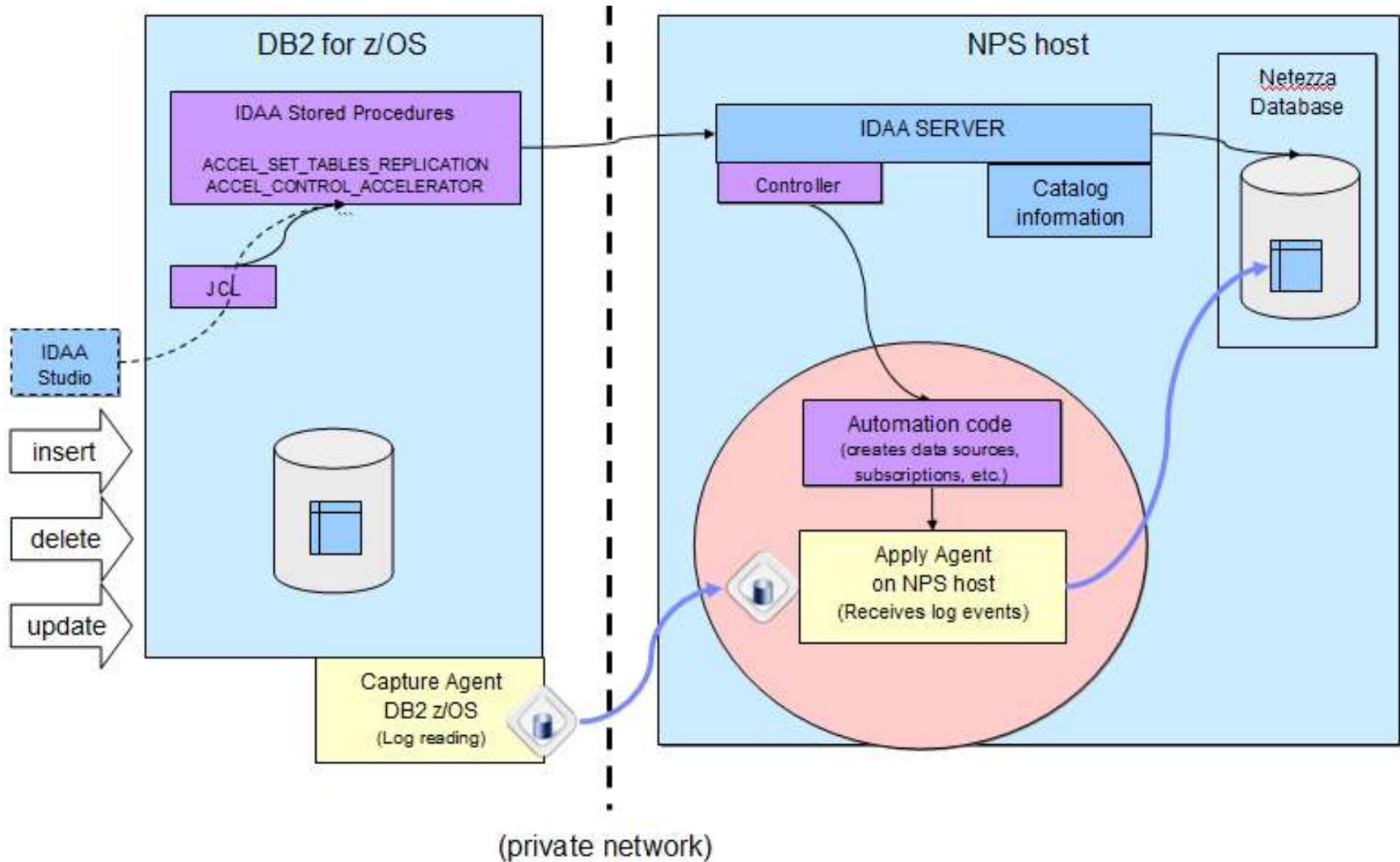
Optimizing load strategy for maximum throughput

High Performance Storage Saver - Application integration

High Availability and Disaster Recovery thoughts

"Nice to know" pieces

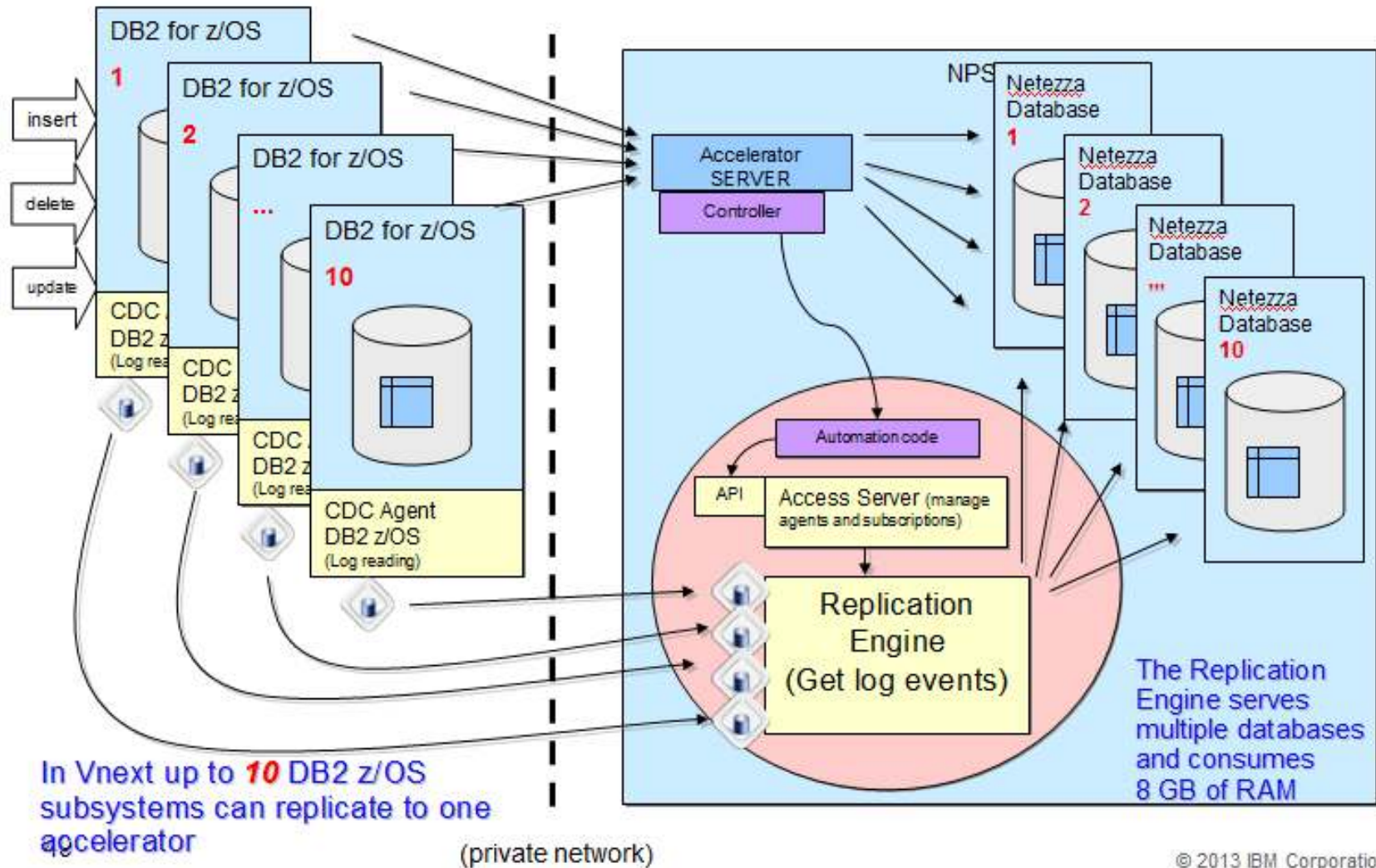
Incremental update solution architecture



Considerations for using incremental update

- Each DB2 for z/OS subsystem has its own database on the accelerator (don't worry about table naming conventions)
- One CDC apply agent per database on the accelerator
 - IDAA V3.1: Limited to 2 CDC apply agents per accelerator
 - IDAA V4.1: Up to 10 DBs per accelerator

Incremental update solution architecture



Considerations for using incremental update

- Throughput determination factors
 - **Incremental update comes with replication latency !!!**
 - Capture agent throughput
 - Analysis of heavily updated logs
 - Apply agent's handling of INSERTs, UPDATEs, DELETEs
 - INSERT => no scan
 - UPDATEs consist of DELETE and INSERT => 1 scan
 - DELETE => 1 scan
 - Scan performance improvement by using organizing keys
 - Unique constraints in DB2 for z/OS and SQL length
 - Data volume
 - Table size - replication performance better for smaller tables
 - Design goal: Maintain large tables that are slowly changing

Latency of incremental update

- Stored procedure ACCEL_CONTROL_ACCELERATOR
- XML input:

```
<?xml version="1.0" encoding="UTF-8" ?>  
<aqtables:controlCommand  
xmlns:aqtables="http://www.ibm.com/xmlns/prod/dwa/2011"  
version="1.2" >  
<getAcceleratorInfo/>  
</aqtables:controlCommand>
```


- Output:


```
<replicationInfo state="STARTED"  
lastChangeTimestamp="2013-05-09T20:35:04.296083Z"  
latencyInSeconds="11"
```

Latency of incremental update - GUI information

Accelerator: TESTIDAA @ TMCC037

Refresh:

Every minute 















Acceleration: Stopped [Start](#) Credentials valid since: 10.07.13 10:53 [Update](#)
 Status: Online Trace: DEFAULT / OFF [Configure](#) [Save](#) [Clear](#)
 Used space: N/A Active queries: 0 (0 queued)
 Replication: Started [Stop](#) Replication latency:  High [Show events](#)

► **Monitoring**

Replication latency is higher than expected or replication is stopped. Check the replication events to see why replication is delayed

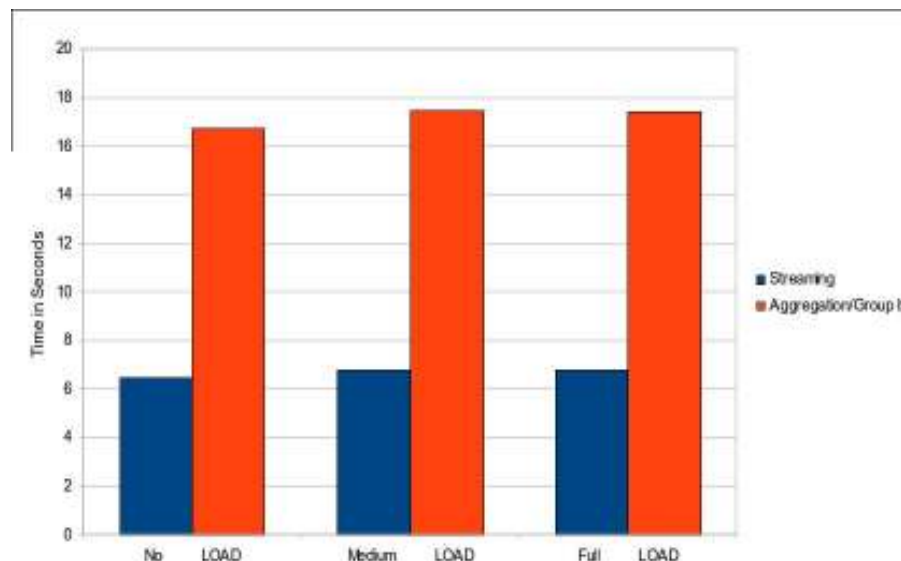
Replication Events: TESTIDAA @ TMCC037

View: Show:

| Severity | Message | Event ID | Originator | Time |
|---|--|----------|------------|---------------------|
|  Information | CHC0057I System 10.16.36.10 is unavailable. A restart of subscription TESTIDAA will... | 57 | DSC | 13.09.2013 13:47:05 |
|  Error | CHC7057E TCP/IP failure to connect to remote port 11301. The returned errno is ET... | 7.057 | DSC | 13.09.2013 13:47:05 |
|  Information | CHC0057I System 10.16.36.10 is unavailable. A restart of subscription TESTIDAA will... | 57 | DSC | 13.09.2013 13:42:00 |
|  Error | CHC7057E TCP/IP failure to connect to remote port 11301. The returned errno is ET... | 7.057 | DSC | 13.09.2013 13:42:00 |
|  Information | CHC0057I System 10.16.36.10 is unavailable. A restart of subscription TESTIDAA will... | 57 | DSC | 13.09.2013 13:36:55 |
|  Error | CHC7057E TCP/IP failure to connect to remote port 11301. The returned errno is ET... | 7.057 | DSC | 13.09.2013 13:36:55 |
|  Warning | CHC6436W Communications session 10.16.36.11:51786 in session slot 1 and state 8 ... | 6.436 | CMO | 13.09.2013 13:33:04 |
|  Warning | CHC6448W Communications link in link slot 1 has terminated abnormally | 6.448 | CMO | 13.09.2013 13:32:49 |
|  Error | CHC6503E Communications link to 10.16.36.11:0 abnormally terminated by remote... | 6.503 | CCO | 13.09.2013 13:32:49 |
|  Error | CHC7061E Failure in TCP/IP send. The returned errno is EPIPE for a value of 32 | 7.061 | CCO | 13.09.2013 13:32:49 |
|  Warning | CHC6436W Communications session 10.16.36.10:11301 in session slot 2 and state 8 ... | 6.436 | CMO | 13.09.2013 13:32:04 |
|  Information | CHC0055I Subscription TESTIDAA has been queued for restart | 55 | DSC | 13.09.2013 13:31:49 |
|  Information | CHC0221I The control task for subscription TESTIDAA is terminating | 221 | SCT | 13.09.2013 13:31:49 |
|  Information | CHC0223I The data task for subscription TESTIDAA is terminating | 223 | SDT | 13.09.2013 13:31:49 |

Incremental update and concurrent workload

- 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





Agenda

Finding queries to route to the accelerator

Considerations for using incremental update

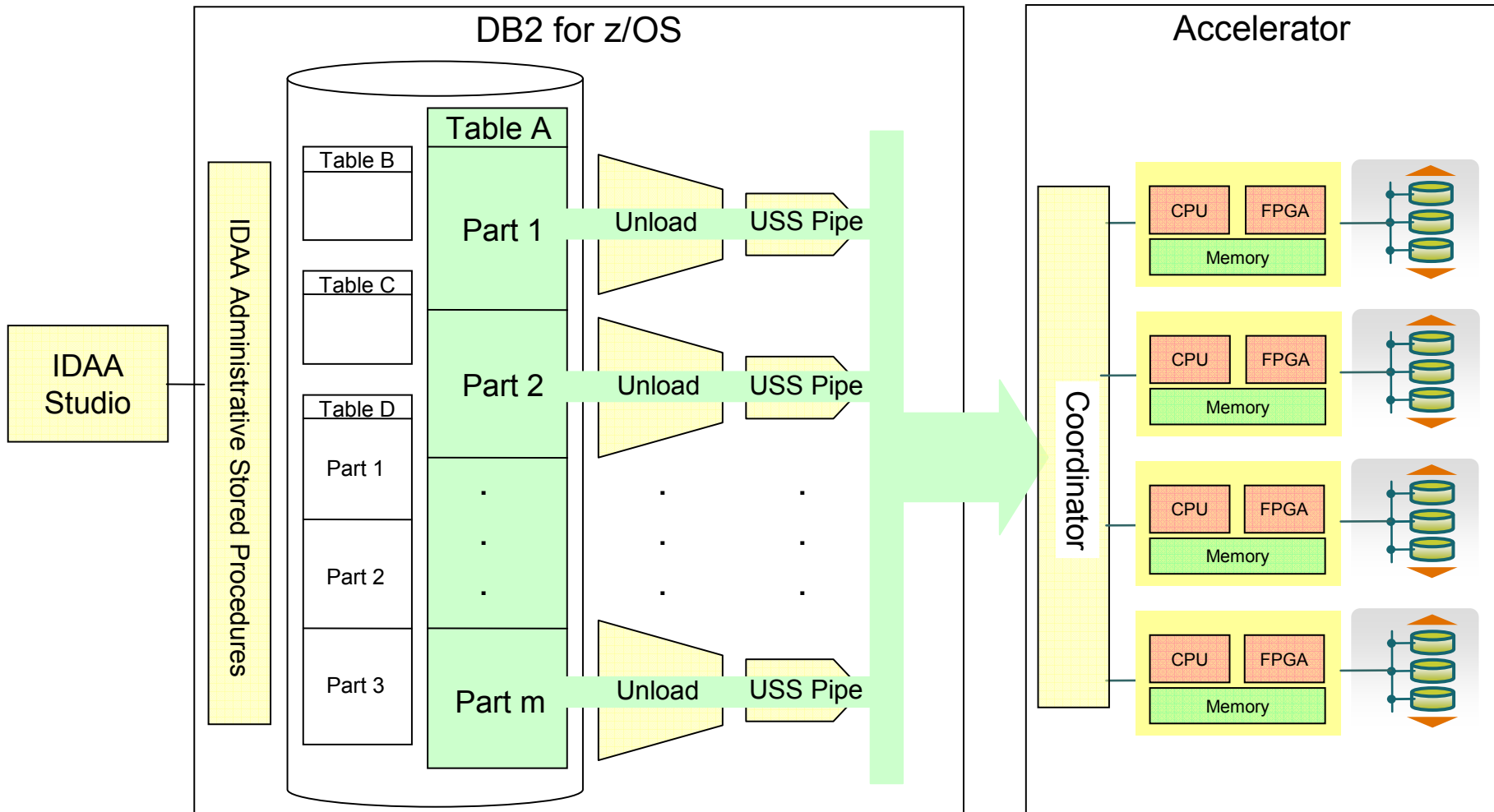
Optimizing load strategy for maximum throughput

High Performance Storage Saver - Application integration

High Availability and Disaster Recovery thoughts

"Nice to know" pieces

Loading strategies for maximum throughput



Loading strategies for maximum throughput

| Synchronization options | Use cases, characteristics and requirements |
|--|---|
| <p>Full table refresh</p> <p>The entire content of a database table is refreshed for accelerator processing</p> | <ul style="list-style-type: none"> Existing ETL process replaces entire table Multiple sources or complex transformations Smaller, un-partitioned tables Reporting based on consistent snapshot |
| <p>Table partition refresh</p> <p>For a partitioned database table, selected partitions can be refreshed for accelerator processing</p> | <ul style="list-style-type: none"> 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 |
| <p>Incremental update</p> <p>Log-based capturing of changes and propagation to IDAA with low latency (typically few minutes)</p> | <ul style="list-style-type: none"> 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 |

- Change detection: DB2 automatically determines if table / partition was changed, otherwise skips the table partition in the load request
- Uses DB2 real time statistics to determine data changes
 - Externalization of RTS in-memory counters every 30 mins (STATSINT) to SYSTABLESPACESTATS

Recommended maintenance for maximum throughput

- Recommendation: Use the latest IDAA PTF if possible
- Significant reduction of CP and accelerator load times (- 30%)
 - DFSMS APAR OA41706

| PTF No. | Rel Id | z/OS |
|----------------|---------------|-------------|
| UA68971 | HDZ1C10 | V1.12 |
| UA68972 | HDZ1D10 | V1.13 |
| UA68973 | HDZ2210 | V2.1 |

- Eliminate invocation overhead if multiple tables are specified:
 - Most elapsed time savings if loading a large number of small tables
 - IDAA V2.1 PTF7
 - IDAA V3.1 PTF3
 - IDAA V4.1 PTF1

Recommended maintenance for maximum throughput

- GUI Performance can be slow, depending on number of objects on IDAA
- **Recommendation:** Upgrade to IDAA Studio V3.1 PTF3
(see <http://www-01.ibm.com/support/docview.wss?uid=swg27035960>)
- Multi-row fetch
- Static versus dynamic SQL
- Bringing accelerator cache to DB2 for z/OS tables
- See Redbook Section 10.1.3 - "Monitoring the Accelerator using stored procedures"
 - 1. Calling ACCEL_GET_QUERY_DETAILS
 - 2. Parse output
 - 3. Insert runtime information from queries into a DB2 table

Loading strategies for maximum throughput

- Meaning of AQT_MAX_UNLOAD_IN_PARALLEL in AQTENV environment variables
 - Determines unload parallelism for partitioned table space
 - AQT_MAX_UNLOAD_IN_PARALLEL = 8 , unloading 10 partitioned table spaces in // to the accelerator => 80 invoked UNLOAD utilities in // ...
- Parallelize load streams to find your system's optimum
 - Customer case:
Elapsed time of sequential load > 500 tables: approx. 2.5 hours
 - Best throughput: Parallelize load streams
 - Recommendation: Start with max. 8 parallel load streams
 - Even distribution of data amounts among load streams
 - Customer case: 12 minutes with 7 parallel streams
 - Question: How to parallelize load streams ? See next slide...

Check raw data table sizes for even distribution

```

SELECT
  AT.CREATOR    as "Table_Schema"
, AT.NAME      as "Table_Name"
, ST.RECLENGTH as "Table_RecordLength"
, ST.AVGROWLEN as "Table_AvgRowLength"
, ( SELECT COUNT(SC.COLNO)
    FROM SYSIBM.SYSCOLUMNS SC
    WHERE SC.TBCREATOR = AT.CREATOR
          AND SC.TBNAME = AT.NAME
          AND SC.COLTYPE IN
              ('VARCHAR', 'LONGVAR', 'VARG', 'LONGVARG', 'CLOB', 'BLOB', 'DECLOB', 'VARBIN', 'XML' )
    )          as "Table_NumberVarLengthColumns"
, TS.DBNAME    as "TableSpace_Database"
, TS.NAME      as "TableSpace_Name"
, TP.PARTITION as "Partition_Number"
, TP.AVGROWLEN as "Partition_AvgRowLength"
, TP.PAGESAVE  as "Partition_PageSave"
, TP.CARD      as "Partition_NumberRows"
, TSS.TOTALROWS as "TableSpaceStats_TotalRows"
, (CASE WHEN TP.CARD < 0
    THEN 'Nan'
    ELSE
      (CASE WHEN
        (
          SELECT COUNT(SC.COLNO)
          FROM SYSIBM.SYSCOLUMNS SC
          WHERE
            SC.TBCREATOR = AT.CREATOR
            AND SC.TBNAME = AT.NAME
            AND SC.COLTYPE IN
                ('VARCHAR', 'LONGVAR', 'VARG', 'LONGVARG', 'CLOB', 'BLOB', 'DECLOB', 'VARBIN', 'XML' )
        ) > 0
      )
    )

```

Check raw data table sizes for even distribution (c'ted)

```

THEN
    ( CASE WHEN TP.PAGESAVE > 0
      THEN
        (
          QUANTIZE( ROUND(
            (
              CAST(TP.AVGROWLEN AS DECFLOAT(34)) *
              (
                CAST(100E0 AS DECFLOAT(34)) /
                (
                  CAST(100E0 AS DECFLOAT(34)) -
                  CAST(TP.PAGESAVE as DECFLOAT(34))
                )
              )
            ) * CAST(TP.CARD AS DECFLOAT(34))
            ,0), DECFLOAT('1E+0'))
          )
        ELSE
          (
            CASE WHEN (TP.AVGROWLEN <> 0 AND TP.CARD <> 0)
              THEN
                ( CAST(TP.AVGROWLEN as DECFLOAT(34)) *
                  CAST(TP.CARD as DECFLOAT(34))
                )
              ELSE
                ( CAST(ST.RECLENGTH as DECFLOAT(34)) *
                  CAST(TSS.TOTALROWS as DECFLOAT(34))
                )
            )
          )
        END
      )
    END
  )

```

Check raw data table sizes for even distribution (c'ted)

```

ELSE ( CASE WHEN (TP.CARD = 0)
        THEN
          ( CAST(ST.RECLENGTH as DECFLOAT(34)) *
            CAST(TSS.TOTALROWS as DECFLOAT(34))
          )
        ELSE
          ( CAST(ST.RECLENGTH as DECFLOAT(34)) *
            CAST(TP.CARD as DECFLOAT(34)) )
        END
      )
    )
  END
)
)
END
) as "CalculatedUncompressedSize" -- <= ORDER BY column

FROM          SYSACCEL.SYSACCELERATEDTABLES  AT
LEFT OUTER JOIN SYSIBM.SYSTABLES           ST
  ON  AT.CREATOR = ST.CREATOR
  AND AT.NAME    = ST.NAME
LEFT OUTER JOIN SYSIBM.SYSTABLESPACE       TS
  ON  ST.DBNAME  = TS.DBNAME
  AND ST.TSNAME  = TS.NAME
LEFT OUTER JOIN SYSIBM.SYSTABLEPART        TP
  ON  TS.NAME    = TP.TSNAME
  AND TS.DBNAME  = TP.DBNAME
LEFT OUTER JOIN SYSIBM.SYSTABLESPACESTATS  TSS
  ON  TP.TSNAME  = TSS.NAME
  AND TP.DBNAME  = TSS.DBNAME
  AND TP.PARTITION = TSS.PARTITION

WHERE AT.ACCELERATORNAME = '<ACCELNAME>'
ORDER BY 13 DESC

```

Loading strategies for maximum throughput

- Sample output:

| | |
|---------|--------|
| Table A | 125 GB |
| Table B | 75 GB |
| Table C | 60 GB |
| Table D | 25 GB |

- A good way to start:

- Assumption: all tables unload with same througput
- Stream 1 to load table A
- Stream 2 to load table B
- Stream 3 to load table C + D
- Pre-IDAA LOAD performance testing:

```
TEMPLATE UD PATH /tmp/AQT.SZA1.AQT005A00010000 FILEDATA BINARY  
RECFM VB LRECL 32756 UNLOAD TABLESPACE "dbname". "tsname" FROM  
TABLE "schema". "table" UNLDDN UD FORMAT INTERNAL SHRLEVEL CHANGE  
ISOLATION CS SKIP LOCKED DATA
```



Loading strategies for maximum throughput

- Data availability during loading
 - Cost / benefit consideration
 - TRUNCATE TABLE (high-performing mass-delete) doesn't exist yet...
 - DELETE processing updates all rows in a table
 - LOAD adds new rows to the table
 - Consider impact when loading large number of tables within single stored procedure call (ROLLBACK if one table fails)
 - Consider REMOVE / ADD TABLE
 - Availability impact for dynamic queries
 - Invalidation of accelerated packages

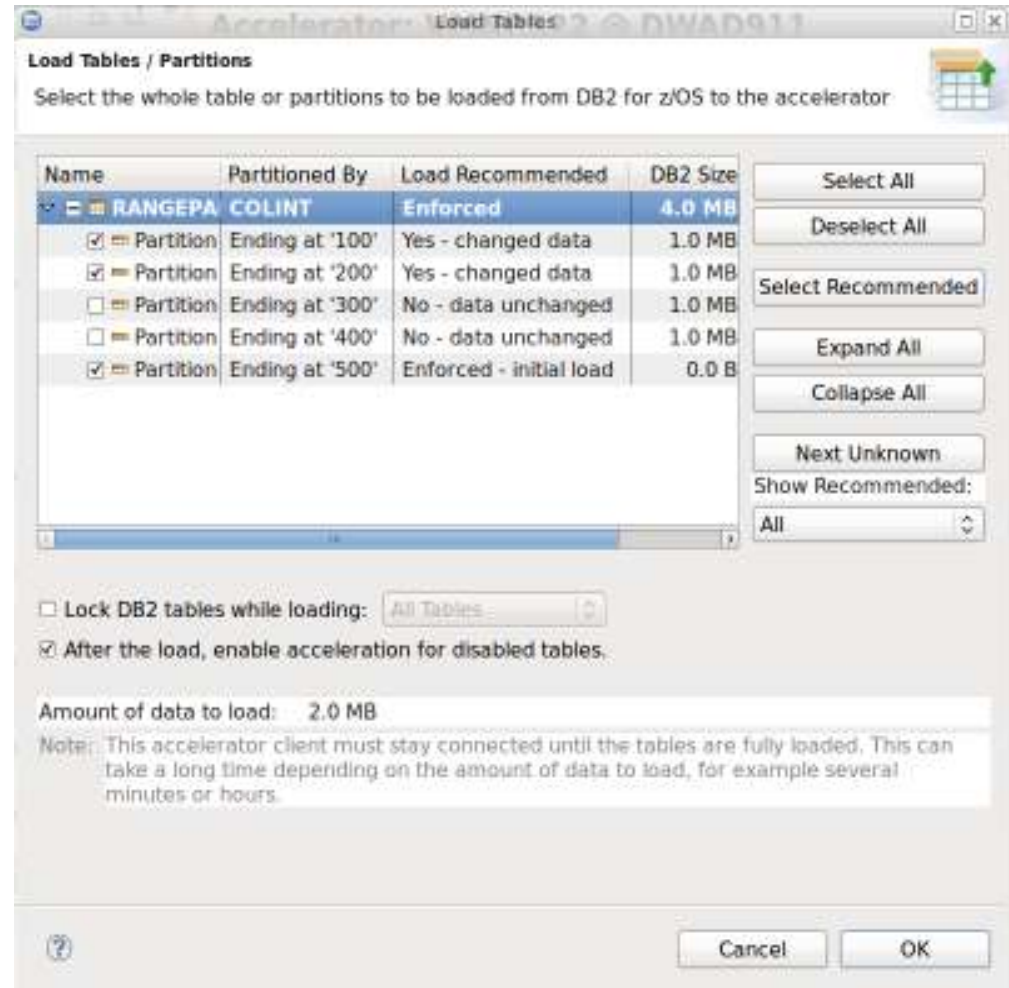
Loading strategies for maximum throughput

- Passing multiple objects within XML to stored procedure triggers sequential execution (eg ACCEL_LOAD_TABLES)
 - Parallel execution requires parallel invocation of stored procedures
 - Distribute tables among load jobs ,appropriately' and consider incremental updates where applicable
 - When partitioned tables are loaded into IDAA, AQT_MAX_UNLOAD_IN_PARALLEL determines parallelism of UNLOAD utility per table
- Determine data changes
 - Usage of RTS to detect whether a partition has changed or not
 - IDAA stores the current state of counters and timestamps for every loaded table/partition
 - In-memory RTS counters are regularly externalized to SYSIBM.SYSTABLESPACESTATS (STATSINT, default is 30 minutes)
 - RTS can be flushed using
ACCESS DB (dbname) SP (tsname) MODE (STATS)

Loading strategies for maximum throughput

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
(ACCEL_GET_TABLE_DETAILS)



Loading strategies for maximum throughput

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
- 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

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

Loading strategies for maximum throughput

- Batch operations via
 - AQTSCALL sample module (SAQTSAMP)
 - DSNREXX (samples in Redbook SG24-8005 Optimizing DB2 Queries with IDAA)
 - Cross-loader in DB2 10 for z/OS

```
EXEC SQL
  SET CURRENT QUERY ACCELERATION = ENABLE;
ENDEXEC

EXEC SQL
  DECLARE C1 CURSOR FOR
  SELECT Col1 FROM CREATOR.TABLE
  WHERE Col1 < 100 FOR FETCH ONLY
ENDEXEC

LOAD DATA INCURSOR C1 INTO TABLE CREATOR.TABLE1;
```

- Invocation from distributed platform via DSNUTILU

```
CALL SYSPROC.DSNUTILU
('UTILID'
, 'NO'
, 'EXEC SQL
DECLARE C1 CURSOR FOR SELECT * FROM CREATOR.TABLE
ENDEXEC
LOAD DATA INCURSOR C1
RESUME YES
INTO TABLE CREATOR.TABLE_1'
, ?)
```



Agenda

Finding queries to route to the accelerator

Considerations for using incremental update

Optimizing load strategy for maximum throughput

High Performance Storage Saver - Application integration

High Availability and Disaster Recovery thoughts

"Nice to know" pieces

High Performance Storage Saver

- HPSS moves single partitions from DB2 for z/OS to the accelerator
 - Step 1: Full Image Copy
 - Step 2: Loading data into the accelerator
 - Step 3: LOAD REPLACE for archived partition to regain PRIQTY
 - But it had a few pitfalls...
 - Write protection of partitions
 - Only HLQ could be specified for image copies
 - No automated recovery process - if needed...
- HPSS has improved with IDAA V4.1
 - Step 1: Partition 'Persistent Read-Only' status after data was archived
 - Step 2: Template name for image copies
 - Step 3: Automated recovery now possible



Stored Procedure ACCEL_ARCHIVE_TABLE encapsulates Archiving Procedure Initial situation

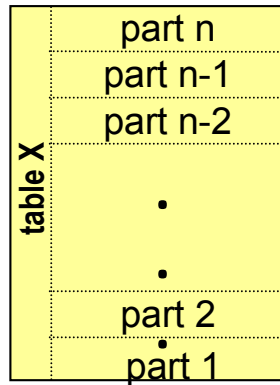
Application

CALL storproc
 ARCHIVE TABLE
 X
 OLDER THAN
 date⁽¹⁾

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

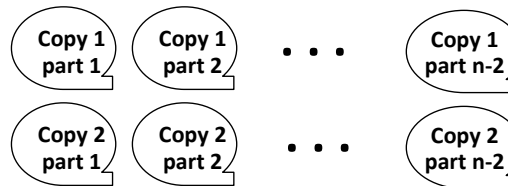
in this particular
 example 'date'
 indicates that only
 last two partitions
 should remain in DB2

DB2

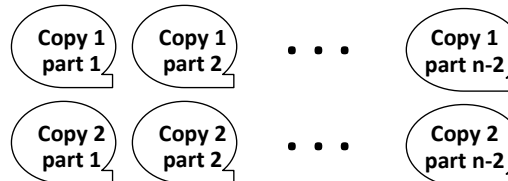


As of V4, the archiving
 process can generate
 multiple image copies
 according to specification
 in the template

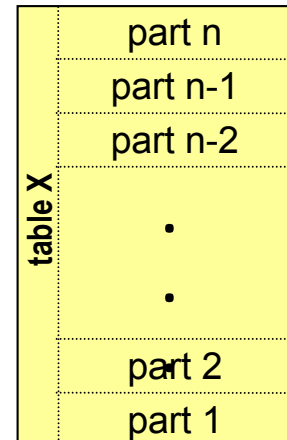
Image copy Definitions:



DB2 recovery site



Accelerator

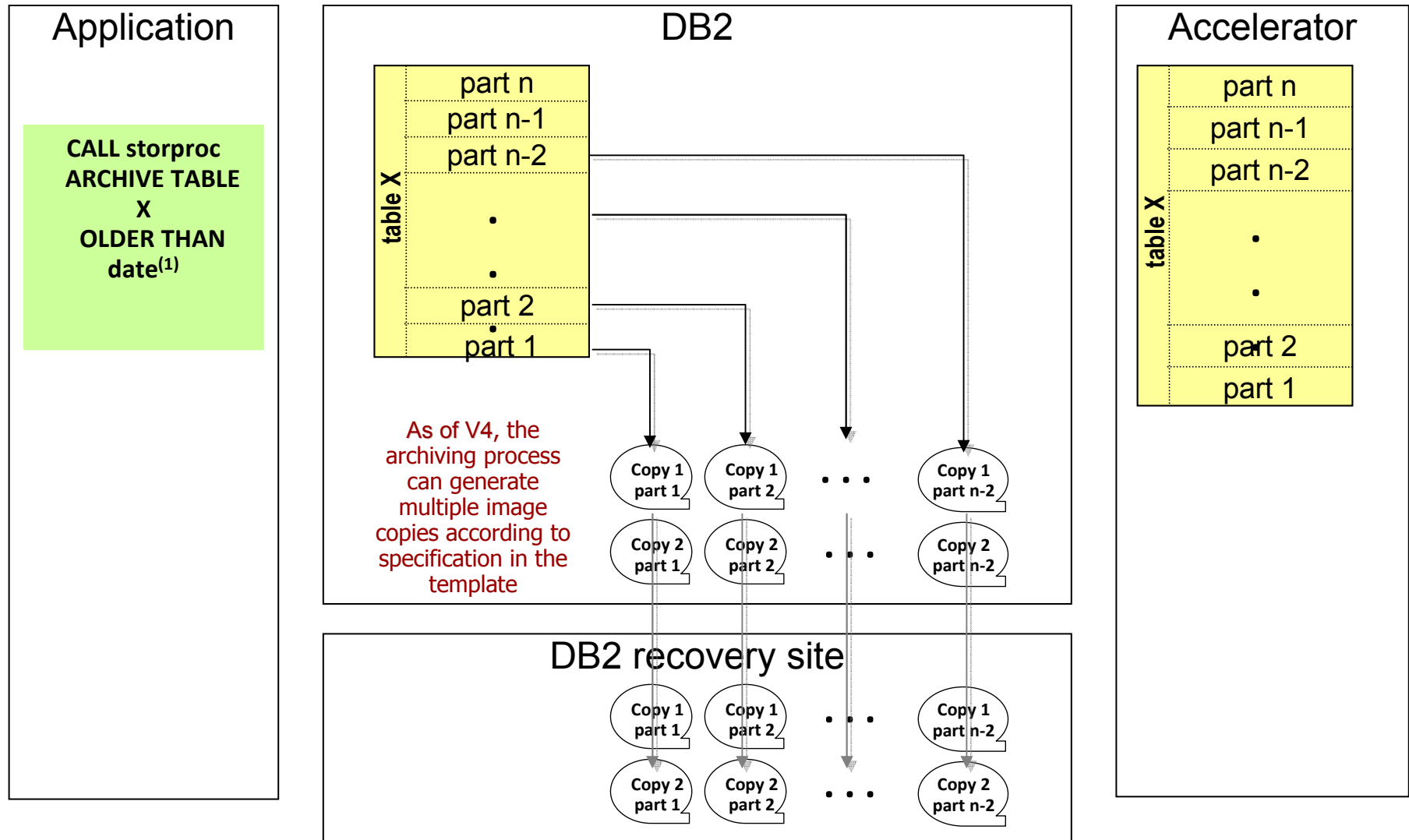


In this particular
 example the table is
 already added and
 loaded to the
 accelerator. It is also
 possible to archive
 partitions first and
 then load only the
 remaining partitions.
 The result is the same



Partitions to be Archived are first backed up

Create image copy step



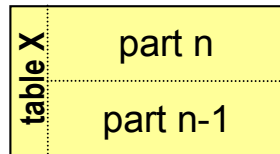


Old Partitions are deleted from DB2 and Table X is split on the accelerator. Old partitions and image copies are set to Read-only status

Application

CALL storproc
 ARCHIVE TABLE
 X
 OLDER THAN
 date⁽¹⁾

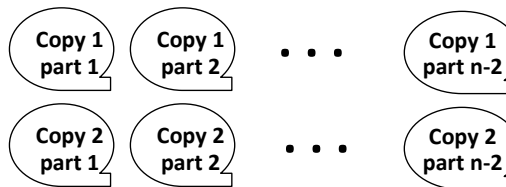
DB2



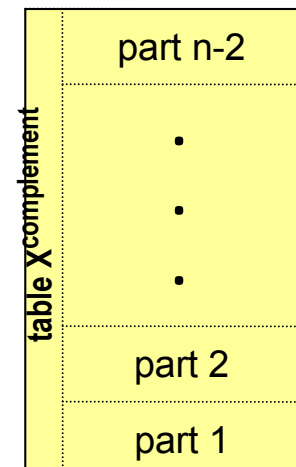
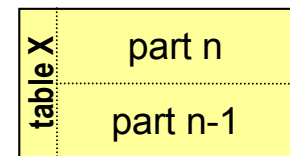
Old partitions are still present in the table, but they are empty and the disk space use is limited to the primary allocation quantity which can be made very small

As of V4, these partitions are set to the PRO status (PERSISTENT READ ONLY) which prevents data modifying operations. Offending applications receive -904, 00C90635

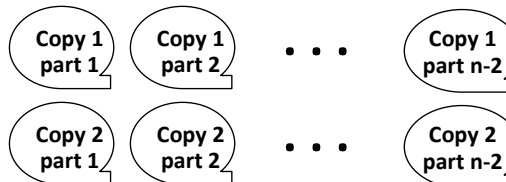
As of V4, the PRO status implicitly protects image copies, i.e. no further image copies can be taken, MODIFY doesn't clean-up those image copies



Accelerator

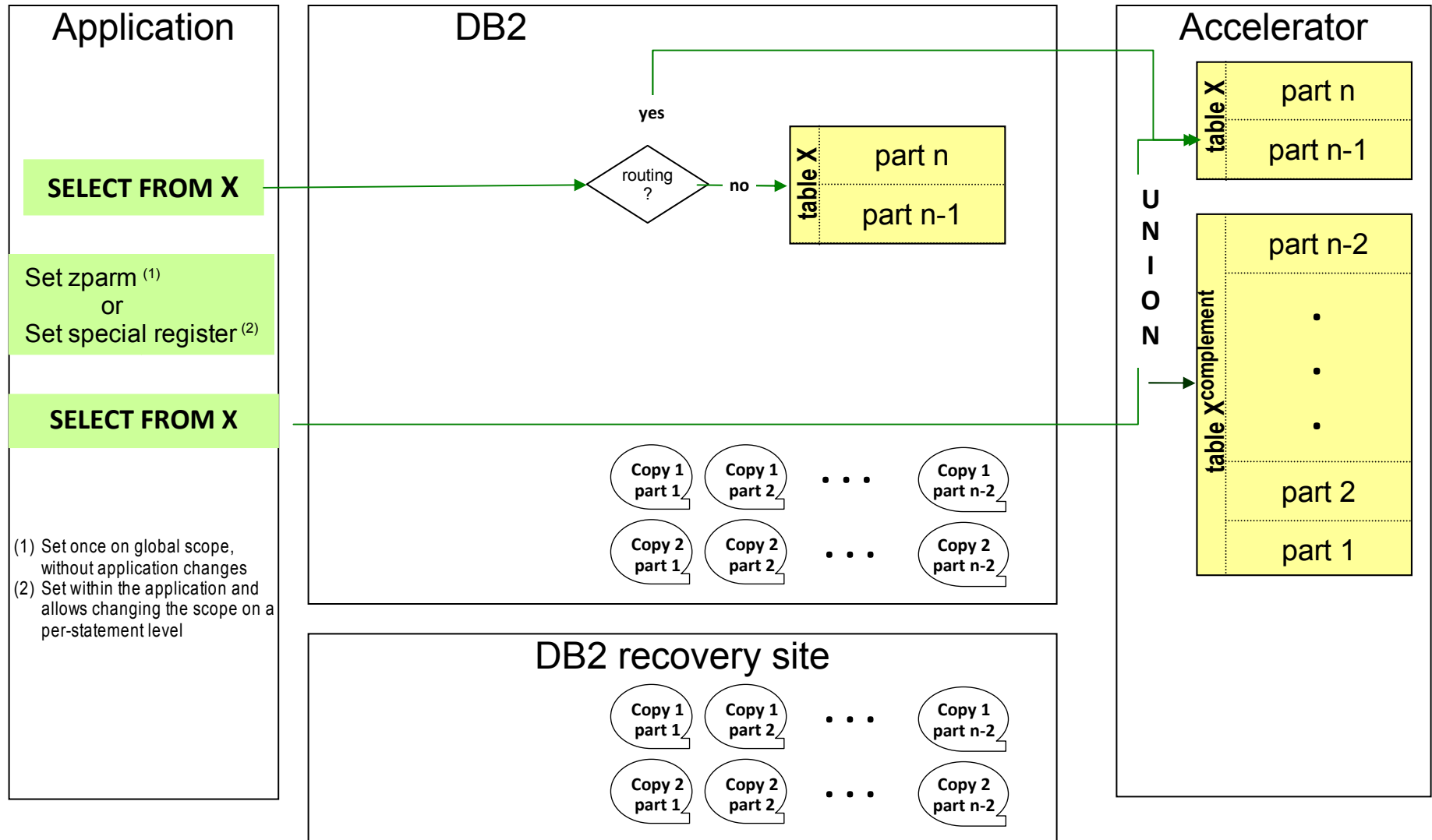


DB2 recovery site





Applications have transparent access (no SQL statement changes needed) to the table





High Performance Storage Saver Performance

- Current data is stored in table A
- Archived data is stored in table B
- When GET_CURRENT_ACCEL_ARCHIVE = YES, both tables are UNION ALL'ed
- UNION ALL processing in the accelerator
 - Today: Internal sequential processing per table

High Performance Storage Saver

- Data can't be modified once archived on the accelerator
- Identify candidate table space partitions by application context
- Check for regular 'batch cleanup processing' for old image copies
- Check SYSTABLESPACESTATS for modifications

```
SELECT TS.STATSLASTTIME, TS.STATSINSERTS, TS.STATSDELETES,
       TS.STATSUPDATES, TS.STATSMASSDELETE,
       TS.LOADRLASTTIME, TS.REORGLASTTIME
FROM   SYSIBM.SYSTABLESPACESTATS TS, SYSIBM.SYSTABLES T
WHERE  T.CREATOR = 'schema' AND T.NAME='table'
       AND TS.PARTITION = part#
       AND T.TSNAME= TS.NAME    AND T.DBNAME = TS.DBNAME
```

- Example: Two INSERT statements after Jan 18th

| L | STATSLASTTIME | STATSINSERTS | STATSDELETES | STATSUPDATES |
|-------|----------------------------|--------------|--------------|--------------|
| * | | * | * | * |
| ----- | ----- | ----- | ----- | ----- |
| | 2013-01-18-12.27.03.778123 | 2 | 1 | 0 |

High Performance Storage Saver

- Check SYSPACKDEP for dependencies of packages

```
SELECT DISTINCT DNAME, DCOLLID, DOWNER  
FROM SYSIBM.SYSPACKDEP  
WHERE BQUALIFIER = 'schema' AND BNAME='table' AND BTYPE='T'
```

- If applicable: check SYSPLANDEP for dependencies of plans

```
SELECT DNAME  
FROM SYSIBM.SYSPLANDEP  
WHERE BCREATOR = 'schema' AND BNAME='table' AND BTYPE='T'
```

- Check SYSPACKSTMT and Dynamic Statement Cache for
 - for DML statements that prevent HPSS usage
 - for non-eligible SELECT statements

High Performance Storage Saver

- Before table is archived, special register setting needs to be added to applications accessing archived data
 - `SET CURRENT GET_ACCEL_ARCHIVE = YES`
- SYSCOPY entries are persistent now throughout MODIFY RECOVERY
- Test and put recovery procedure in place before moving to production



Agenda

Finding queries to route to the accelerator

Considerations for using incremental update

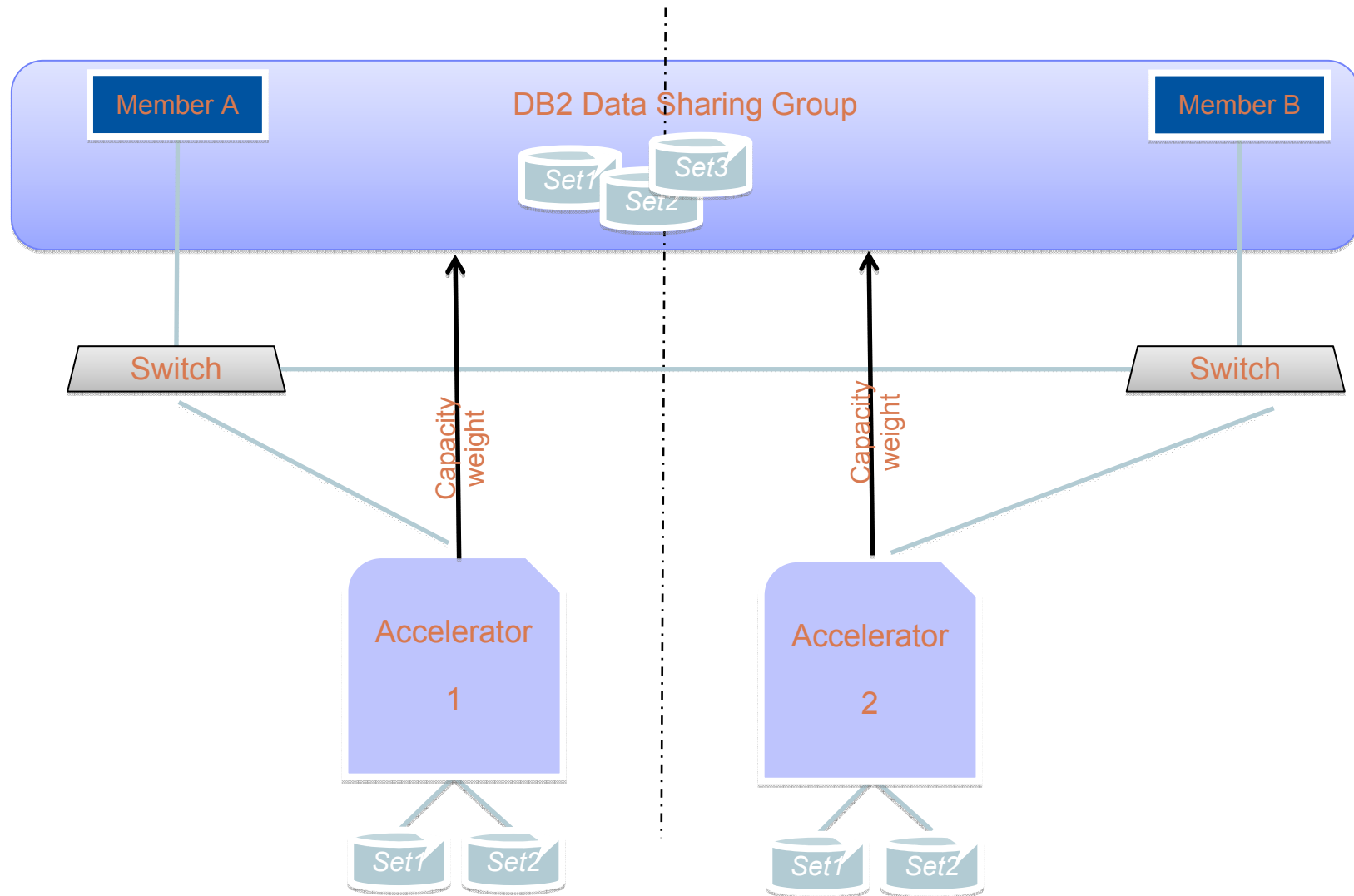
Optimizing load strategy for maximum throughput

High Performance Storage Saver - Application integration

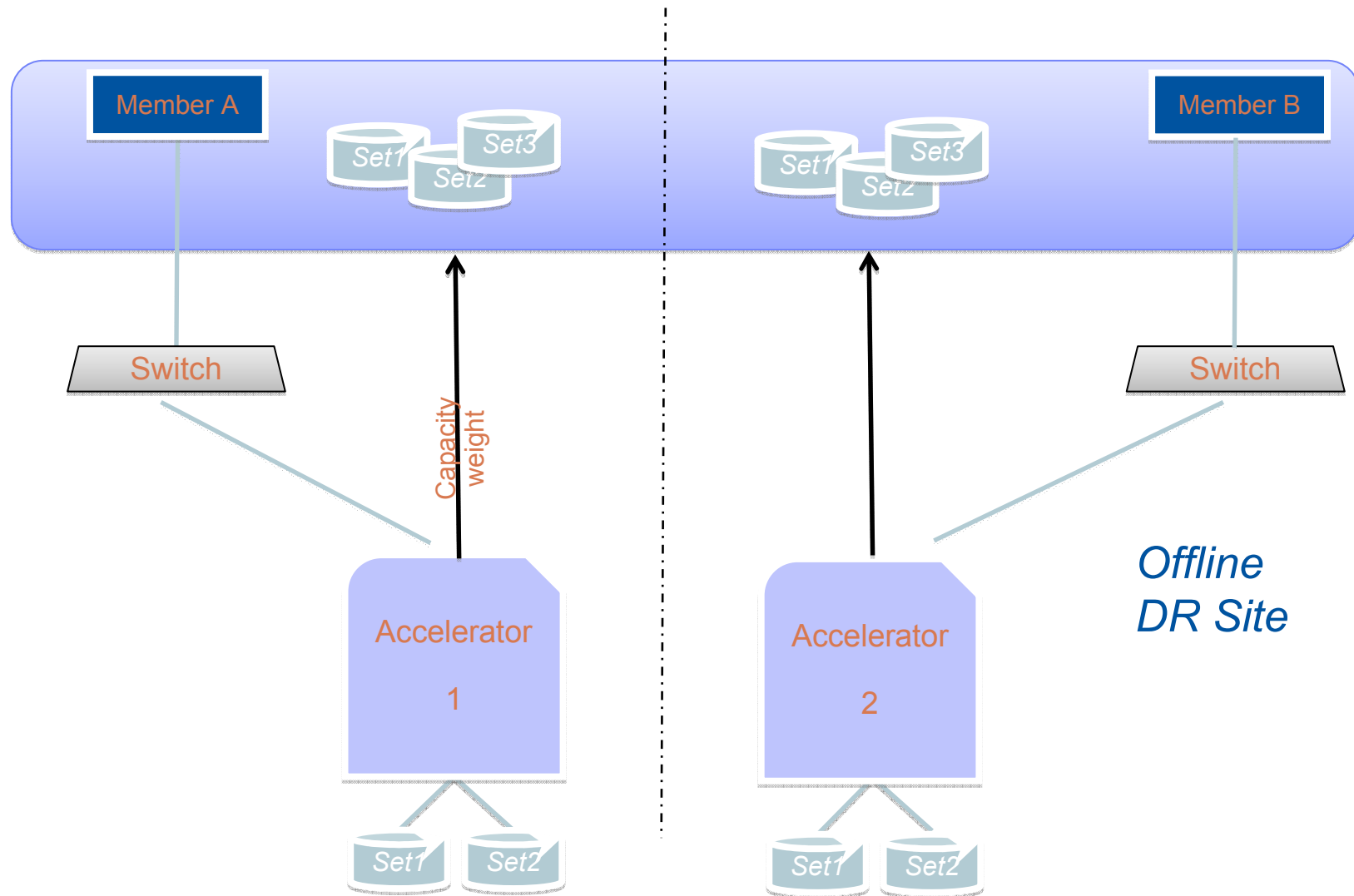
High Availability and Disaster Recovery thoughts

"Nice to know" pieces

High Availability and Disaster Recovery - Active/Active



High Availability and Disaster Recovery - Active/Passive





High Availability and Disaster Recovery

- DB2 for z/OS Active-Active scenario
 - When loading data into the accelerator at the primary site, data can be loaded into the accelerator at the DR site, too
- DB2 for z/OS Active-Passive scenario
 - In case of a disaster
 - DR DB2 for z/OS needs to be recovered
 - data needs to be reloaded from the DR DB2 for z/OS subsystem into the accelerator at the DR site

High Availability and Disaster Recovery - Incremental Update

- DB2 for z/OS Active-Active scenario
 - Accelerator at the DR site can be maintained via incremental update processing in the same way as the accelerator at the primary site
- DB2 for z/OS Active-Passive scenario
 - In case of a disaster
 - Recover DR DB2 for z/OS subsystem
 - Add relevant tables to the accelerator at the DR site
 - Activate incremental update for relevant tables
 - Load tables that are replicated from the DR DB2 for z/OS subsystem into the accelerator at the DR site



High Availability and Disaster Recovery - High Performance Storage Saver

- DB2 for z/OS Active-Active scenario
 - IDAA V4.1 archives data to > 1 accelerator inside the same network
- DB2 for z/OS Active-Passive scenario
 - In case of a disaster
 - Recover DR DB2 for z/OS subsystem
 - Recover archived tables based on full image copies
 - Add relevant tables to the accelerator at the DR site
 - Load tables from the DR DB2 for z/OS subsystem into the accelerator at the DR site
 - Archive data again into accelerator



Agenda

Finding queries to route to the accelerator

Considerations for using incremental update

Optimizing load strategy for maximum throughput

High Performance Storage Saver - Application integration

High Availability and Disaster Recovery thoughts

"Nice to know" pieces

Nice to know

- **SYSLOG messages of IDAA (DSNX881I)**

Status Change (ID 1, Info)
Hardware Service Requested (ID 2, Error)
Hardware Restarted (ID 3, Info)
Disk Space Threshold (ID 4, Warning)
Runaway Query (ID 5, Info)
System Stuck in State (ID 6, Error)
Predictive Failure (ID 7, Warning)
ECC Errors (ID 8, Warning)
Regeneration Error (ID 9, Error)
Disk Error (ID 10, Error)
Hardware Temperature (ID 11, Warning)
System Temperature (ID 12, Error)
SPU Coredumps (ID 13, Warning)
Voltage Fault (ID 14, Error)
Network Interface State Change (ID 15, Information)
S-Blade CPU Core Change (ID 16, Information)
IMM Event (ID 17, Warning/Error)
Hardware Path Down (ID 18, Error)
[Hardware needs attention \(ID 19, Warning\)](#)

Example:

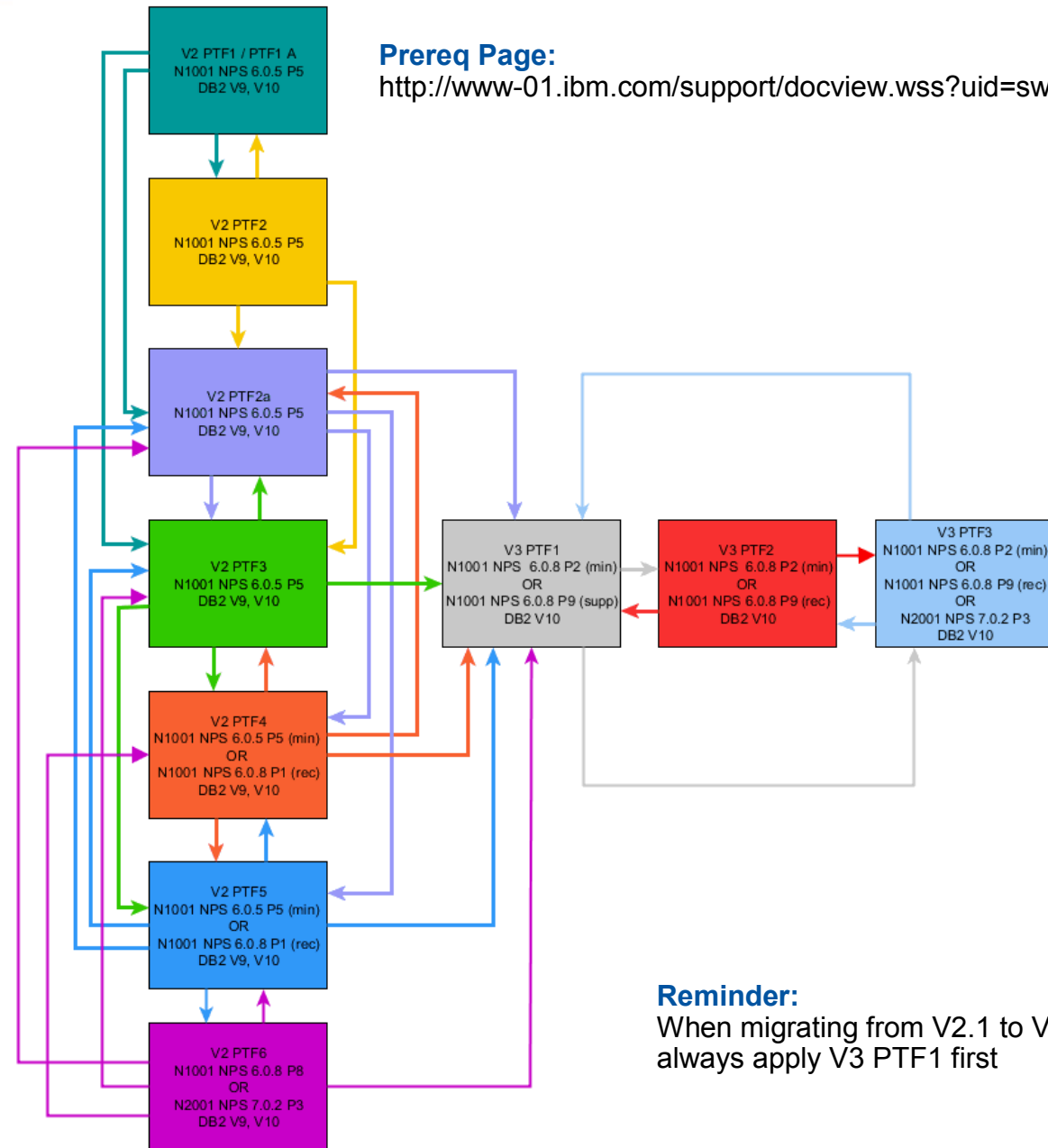
```
0010 DSNX881I -D916 19 W 314751 (10-Dec-12, 18:24:17 EST) 596
0010 Hardware needs attention. Host = dwatf0312ma.boeblingen.de.ibm.com
0010 Hardware type = SPA Hardware ID = 1003 Location = 1st Rack, 2nd SPA
0010 -msg
```

Nice to know

- Stalled IBM DB2 Analytics Accelerator upgrades
 - Follow the required order:
 - 1. Update IDAA Studio
 - 2. Update IDAA Server code
 - **Always transfer new IDAA Server code with "old" Stored Procedures !!!**
 - 3. Update DB2 for z/OS incl. Stored Procedures
 - Most common issue:
Stored Procedures have been upgraded before IDAA Server
 - **How to resolve - Recommended way (if possible):**
 - If no other LPAR with "old" version is available:
downgrade Stored Procedures to "old" version
 - Use IDAA Studio to transfer & apply new accelerator server code
 - Upgrade Stored Procedures again

Upgrade paths

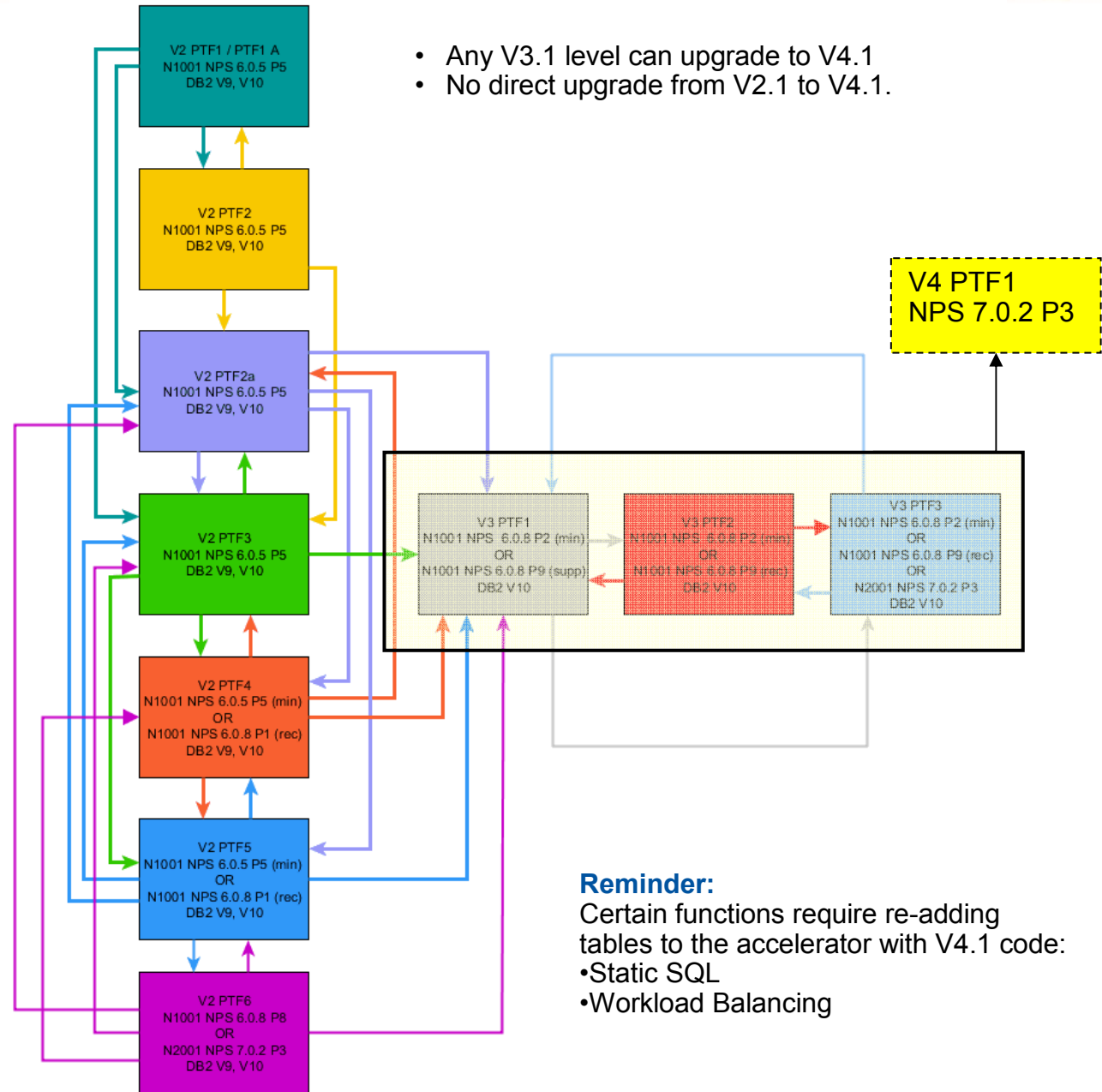
Prereq Page:
<http://www-01.ibm.com/support/docview.wss?uid=swg27035960>



Reminder:
 When migrating from V2.1 to V3.1,
 always apply V3 PTF1 first

Upgrade paths

- Any V3.1 level can upgrade to V4.1
- No direct upgrade from V2.1 to V4.1.



Reminder:

Certain functions require re-adding tables to the accelerator with V4.1 code:

- Static SQL
- Workload Balancing

Nice to know - Changing profile tables

- DSN_PROFILE_TABLE and DSN_PROFILE_ATTRIBUTES
 - REASON_CODE = 300 in DSN_QUERYINFO_TABLE
 - SUM(CARDF) of all tables accessed < 1,000,000

```
INSERT INTO SYSIBM.DSN_PROFILE_TABLE (PROFILEID)  
VALUES (4713);
```

```
INSERT INTO SYSIBM.DSN_PROFILE_ATTRIBUTES  
(PROFILEID,KEYWORDS,ATTRIBUTE1,ATTRIBUTE2,ATTRIBUTE3)  
VALUES  
(4713, 'ACCEL_TABLE_THRESHOLD', NULL, -1, NULL);
```

-1 disables check

Nice to know - Changing profile tables

- DSN_PROFILE_TABLE and DSN_PROFILE_ATTRIBUTES
(continued)
 - REASON_CODE = 301 in DSN_QUERYINFO_TABLE
 - TOTAL_COST in DSN_STATEMNT_TABLE < 5000

```
INSERT INTO SYSIBM.DSN_PROFILE_TABLE (PROFILEID)  
VALUES (4714);
```

```
INSERT INTO SYSIBM.DSN_PROFILE_ATTRIBUTES  
(PROFILEID,KEYWORDS,ATTRIBUTE1,ATTRIBUTE2,ATTRIBUTE3)  
VALUES  
(4714, 'ACCEL_TOTALCOST_THRESHOLD', NULL, NULL, 2500);
```

Not recommended: -1 to disable

Nice to know - Changing profile tables

- DSN_PROFILE_TABLE and DSN_PROFILE_ATTRIBUTES
(continued)
 - Estimated resultset size

```
INSERT INTO SYSIBM.DSN_PROFILE_TABLE (PROFILEID)  
VALUES (4715);
```

```
INSERT INTO SYSIBM.DSN_PROFILE_ATTRIBUTES  
(PROFILEID,KEYWORDS,ATTRIBUTE1,ATTRIBUTE2,ATTRIBUTE3)  
VALUES  
(4714, 'ACCEL_RESULTSIZE_THRESHOLD', NULL, -1, NULL);
```

Nice to know

- Acceleration benefits seen in the field:
 - Financial report for online banking customers
 - New reports that couldn't been created before
 - Insurance reporting and analysis - accelerating business processes
 - Most impressive customer testimonial:

"DB2 Analytics Accelerator makes it possible for us to execute 90 percent of our queries 25 times faster, and 50 percent of them 100 times faster"
 - Read the full story here:

http://www-01.ibm.com/software/success/cssdb.nsf/CS/STRD-9A3JU2?OpenDocument&Site=default&cty=en_us
- Something completely different: How the system is manufactured:
 - <https://www.youtube.com/watch?v=N0PibiG73CE>




Nice to know

- Synchronizing data in IBM DB2 Analytics Accelerator
<http://www-01.ibm.com/support/docview.wss?uid=swg27038501>
- Enabling query acceleration via ODBC/JDBC
<http://www-01.ibm.com/support/docview.wss?uid=swg27038078>
- Structure of DSNX881I messages
<http://www-01.ibm.com/support/docview.wss?uid=swg27037905>
- How IBM DB2 Analytics Accelerator handles correlated subqueries
<http://www-01.ibm.com/support/docview.wss?uid=swg27037928>




Much more information is available:

IBM, Information Management Software 

Draft Document for Review July 12, 2013 6:09 pm SG24-8151-00

Hybrid Analytics Solution Using IBM DB2 Analytics Accelerator for z/OS V3.1

- Leverage your investment in IBM System z for business analytics
- Transparently accelerate DB2 for z/OS complex queries
- Implement high performance and available analytics



Paolo Bruni
Willie Favero
James Guo
Ravi Kumar
Ruiping Li
Cristian Molano
Andy Perkins
Theresa Tai
Dino Tonelli

ibm.com/redbooks **Redbooks**



IDUG

Leading the DB2 User
Community since 1988

IDUG DB2 Tech Conference

Barcelona, Spain - October 2013



Thank You





IDUG
Leading the DB2 User
Community since 1988

IDUG DB2 Tech Conference

Barcelona, Spain - October 2013

Patric Becker

IBM

Patric.Becker@de.ibm.com

Session F15

Get most out of IBM DB2 Analytics Accelerator:
Customer and PoC Experiences

