WebSphere Data Interchange

# 2005 B2B Customer Conference

*Pioneering New Horizons – Solutions that Evolve*

## IBM Improves WDI Performance of z/OS

David Hixon
Product Architect, WDI and WPG

# Performance

- **Variables affecting performance**
- **Measures of performance**
- **Improvements for each measure**

# Performance Variables

- **Server Platform**
- **Number of concurrent executions**
- **Number of independent MQ messages**
- **Number of documents in an Interchange**
- **Number of Interchanges in a MQ Message**
- **Number of Documents in a file**
- **Type of map used - Send/Receive vs Data Transform**
- **Map complexity - Source / Target, number of mappings**
- **Data size**
- **Use of Auditing features - MR, OPTRECs, TS**
- **Persistent MQ messaging**
- **DB2 binding - Syncpoint TWOPHASE**
- **Functional Acknowledgments generated**

# Performance Measures

- **CPU Time**
- **Elapsed time**
- **Thread locking - stretch time**
- **File Activity**
- **Database activity - updates, reads, locks**
- **Number of transactions processed - Throughput**

# Types of Improvements

- **CPU usage**
  - Table searches and search techniques
  - Paradigm changes
- **Memory usage**
  - Reuse of AMM nodes
  - Restricted buffer size
  - Progressive "reallocation" of output buffers
- **Data Access**
  - Use of DB2 indexes
  - Reuse of frequently called objects, caching
  - Use of "RAM" files for work files
- **Throughput**
  - Elimination of "lock" contention

# Improvements Made

- **Pageable AMM**
  - 8010477/PK06404 - enhancement, PAMM, increased capacity for very large messages / transactions
- **Caching of frequently used objects**
  - Trading Partner Profiles, Rules, Control Strings
  - AMM node pool
- **Other PTRs**
  - 8010406/PK02319 - raw CPU improvement with a large number of transactions
  - 8010510/PK08194 - memory usage improvement, Internal - Storage usage, freeing data area instead of reusing to reduce storage used upper limit
  - 8010432/PK04612 - raw CPU improvement, large messages
  - Internal - CPU improvement with very large messages - buffer search

# Problem Statement

- **WDI 3.2 processes 1 unit of work at a time. For EDI, it is an Interchange, for XML it is one document, for Data Formats it is one Transaction. With Data Transformation maps, all output created is not presented to the Message Broker until a return from the translation process occurs. In the instance of an EDI Interchange with 200 transactions being translated to 200 XML documents, the components keep all the XML documents in memory until the Interchange processing is completed. The only reason to do so, since prior documents are not permitted to be referenced, is for Envelope recovery. In addition, the input interchange data is retained in memory until all transactions are processed - even though only one transaction is translated at a time.**
- **In the Health Care Industry is not uncommon to have 10,000 transactions in an interchange - e.g. EOB, Explanation of Benefits. In the Financial Industry, an Interchange might have one transaction for each employee in a company stock plan.**
- **This leads to large amounts of memory being required. A 50MB EDI Interchange may generate 100MB of XML data.**

# Current Processing

- **The WDI Utility reads the input file. The Logical Message Adapter identifies the input unit of work, e.g. an EDI Interchange, and passes it to the Message Broker. With the Aetna changes, the Message Broker does a "lazy parsing" of the Interchange to get Envelope Information and the ST/SE data. This is passed to the Deenvelope component and subsequent components for processing. The output from translation, the audit file / print file messages, and transaction store data are all stored in link lists. Control is then returned back to the Message Broker, and the process is repeated for the next ST/SE, until all transactions are processed.**
- **The output link lists are then returned to the Utility. The output files are updated, the messages are written, and the Transaction Store tables are updated.**
- **The Logical Message Adapter then gets the next input unit of work, and repeats the process until the file is exhausted.**

# WDI Solution

- **WDI 3.2 needs to reduce its consumption of main storage in processing large messages, while retaining its use of memory for small messages to obtain the performance desired.**
- **A change to the Message Broker allows output returned from the components to be passed back to the Utility for each transaction. The Utility then outputs data, write messages, updates the TS and then returns to the Broker. The storage in the lists can then be freed. The Broker indicates to the Utility that this return is not the final return or a "flush" of the Interchange. The Utility returns to the Broker and allows resumption of transaction parsing and processing. At the last transaction, a "flush" indicator is passed and the Utility then reads the next input unit of work. (P8010321)**
- **A second change is to the EDI Parser and the ROD Parser. This change reduces the amount of memory allocated for a single Interchange. As was mentioned before, the entire EDI Interchange is passed from the Logical Message Adapter to the Message Broker. The Message Broker, based on a PAGE(Y) option and a buffer size > xxMB, writes the Interchange to a work file, and then the Parser reads the work file and get buffers big enough for a transaction. This is similar to the technique used by the XML parser in the XML Split enhancement.**
- **A third change creates a "pageable AMM". It requires a type of "paging subsystem" similar to that with Pageable Translation in the Send / Receive Translator.**

# P8010477 PK06404 on z/OS

- **10K Claim input file**
- **Before PK06404 enhancement:**
  - System area used: 364K/10M
  - Virt storage used: 712K/354M
  - Step completion code: 0000
  - Total CPU time used: 00:04:20.91
- **With PK06404 enhancement:**
  - System area used: 360K/10M
  - Virt storage used: 728K/122M
  - Step completion code: 0000
  - Total CPU time used: 00:04:18.02
- **Note: Virt storage comparison shows a 230MB improvement.**

- **60K Claim input file:**
- **Before PK06404 enhancement:**
  - System area used: 364K/10M
  - Virt storage used: 8820K/1517M
  - Step completion code: 0012
  - Total CPU time used: 04:24:01.75
- **With PK06404 enhancement:**
  - System area used: 360K/10M
  - Virt storage used: 728K/515M
  - Step completion code: 0000
  - Total CPU time used: 00:29:25.59
  - SRB CPU time used: 00:00:03.96
- **Note: Did not complete successfully before PK06404 -- ran out of virt storage. Nonetheless, virt storage comparison shows at least 1GB improvement.**

# 8010477 with PAGETHRESHOLD

- **Pageable AMM feature with PAGETHRESHOLD(500) in the PERFORM command and the 60MB test case without the PAMM feature. The following are the CPU time when both test cases were executed on z/OS**
- **With PAMM feature**
  - TOTAL CPU TIME= 20.53  TOTAL ELAPSED TIME=113.73
- **Without PAMM feature**
  - TOTAL CPU TIME= 40.46  TOTAL ELAPSED TIME=117.99

# Performance Problems

- **Using STROBE on z/OS and Rational Quantify on Windows, the following performance problems were uncovered:**
  - Memory allocation and freeing memory consumed 40% of the CPU time
  - DB2 usage was 20% of the Windows CPU usage and 45% of the z/OS CPU usage
  - Search times for some DB2 lookups were longer than others
  - Reallocation of XML4C resources used an extensive amount of CPU
  - Repeated loading of DLLs used an extensive amount of CPU

# Performance Tuning

- **To correct the problems, the following changes were made:**
  - Add an index to the EDIDTDHDR table by ROOTELEM; this reduced the CPU time used by the SQL statements accessing the EDIDTDHDR table
  - Add a caching subsystem for WDI objects such as Control Strings, Rules, and Trading Partners. This subsystem allowed subsequent calls to DB2 to be eliminated as long as the same object was being requested. Up to 5 instances of an object can be cached in a PERFORM cycle
  - On z/OS, eliminate the DB Connection request in each node, this reduced the calls to the STEPLIB and DSNLOAD from 222,000 to 46. Each DB Connect accessed the Call Attachment Facility module (CAF). Up to 5 connects were being done for each employee
  - Add a Node memory pool, so that the AMM would not have to issue "new" and "free" instructions, but could reuse existing, acquired storage.

# Performance Tuning

- Allow reuse of Parser instances; instead of creating a new instance of a parser for each time a parser is used, use a previously initialized, loaded parser
- Avoid repeatedly loading DLL's. This was done by not destroying the Message Flow until the message broker is terminated.
- Suppress repetitive messages from the printfile (RU0003, UT0008, FF0007); sppression is based on the attribute VERBOSE_NO; several million print lines were suppressed as 5-7 lines are written for each split "transaction"
- Avoid repeated allocation of output buffer by changing XML_Writer to keep output buffer intact between calls.
- Avoid repeatedly loading the XML namespace DB2 table if the dictionary name value has not changed between split "transactions".

# Tuning Results

- **The results of the executions after all the changes**
- **AIX, 43p machine, 1.26 GHz**
  - 13 MB took 1 minute 20 seconds
  - 300 MB took 46 minutes
  - 2 GB took 7 hours
- **Windows, 1.6 GHz Pentium processor**
  - 13 MB took 1 minute
  - 300 MB took 22 minutes
  - 2 GB took 2 hours

# 8010510 Results

- **The memory high water mark went down dramatically**
  - With change:
    - System area used: 352K 10M
    - Virt storage used: 672K 106M
  - Before change:
    - System area used: 352K 10M
    - Virt storage used: 672K 892M
- **There is a slight difference in elapsed and cpu time, but this could be due to system load.**
  - With change:
    - Time elapsed: 00:27:52.57 Total CPU time used: 00:10:46.84
  - Before change:
    - Time elapsed: 00:22:06.51 Total CPU time used: 00:13:29.94

# 8010534 Results

## CPU Timings

| | 25 Envelopes (appx 33 K) | "Small" test case (appx 5 Meg) | "Big" test case (appx 28 Meg) |
|---|---|---|---|
| Receive map | 0.05 min | 0.70 min | 1.40 min |
| Baseline DT map | 0.65 min | 10.29 min | 51.79 min |
| DT map with EDIUPECM changes | 0.27 min | 4.56 min | 19.12 min |
| DT map with all changes | 0.24 min | 3.92 min | 16.33 min |

# 8010321 w/o PAMM

- **For 25+ MB input file :**
  - real 5m42.59s
  - user 5m34.30s
  - sys 0m4.89s
- **For 75+ MB input file :**
  - real 18m45.12s
  - user 18m29.38s
  - sys 0m14.69s
- **For 100+ MB input file:**
  - real 9h29m51.52s
  - user 9h28m41.99s
  - sys 0m44.25s

WebSphere Data Interchange

# A Case Study about Throughput

# Case Study

- **Following are test results from the IBM Lab, interjecting variables of translation types, volumes, mapping complexities, and file sizes. The team also varied the number of concurrent threads associated with each message type to demonstrate and document the impact based on this customers environment allocating different threads to different processes.**
- **This information is NOT a benchmark. It was not designed to determine maximum performance or throughput. Instead, it was an exercise to try to simulate real-world activities, and assist a customer in their tuning efforts. As a result, typical benchmark numbers would have exceeded the results of these tests. The test data was limited to three trading partners, which restricts the horizontal scaling effects.**
- **All tests were run on AIX 5.2 ML2 4 way box, with 4 gig of RAM, DB2 7.1.10a WDI 3.2.1.16.**

# Case Study (cont.)

- **There were 3 Message/Mapping types that were injected into the process simultaneously**
  - XMLEDI:
    - Used 4 message types, to generate 20 different messages by changing Trading Partner ID.
    - XML to EDI transformation
    - No delayed enveloping - transactions enveloped as soon as they are transformed not enveloped into multiple transactions per envelope.
    - Average 2K XML input message
  - XMLADF:
    - Used one message to generate the 20 different messages by changing Trading Partner ID.
    - XML to Flat File transformation
    - Average 2K XML message
  - XMLEDIDELAY:
    - Used one message to generate the 20 different messages by changing Trading Partner ID.
    - XML to EDI transformation
    - Batched files
    - Delayed enveloping - all similar transactions to one TP23o onegroup, (allgroupns toe) Tj 26.25 0 TD -0.3282 Tc 033412 Tw one TP23o or

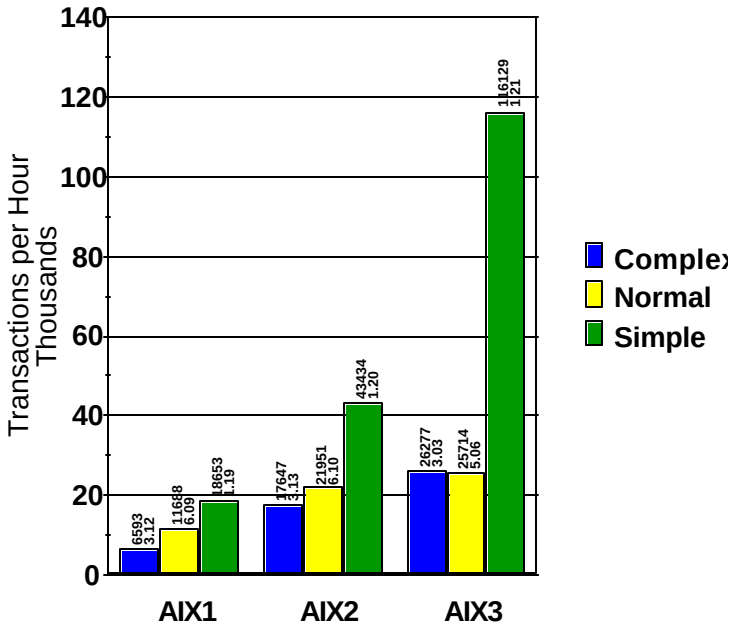
# Case Study

- **The wait time is at about 80% and CPU usage is very low.**
- **Memory is steady at 12MB per translator.**
- **CSD18 + interim**
- **Execution 1: Total of 5 threads, achieving highest throughput**
- **Execution 2: Total of 7 threads, minimizing unused threads**
- **Execution 3: Total of 7 threads, held one thread to zero transactions to process**

| Message Type | Number of Threads | 15 Minute Total | 1 Hour Total XMLEDI |
|---|---|---|---|
| XMLEDI | 4 | 7,743 | 30,972 |
| XMLADF | 2 | 3,387 | 13,548 |
| XMLEDIDELAY | 1 | 1,844 | 7,376 |
| GRAND TOTAL | | 12,951 | 51,808 |

# Machine Sizing

- **Using the anticipated peek volume and the transaction complexity, match it to a column using the scale on the left. For example, a system which anticipates 20,000 moderately complex requests in an hour could run on AIX3. A slightly smaller platform, AIX2, may be able to keep up with the requests if the implementation is "simple", but there is a risk that requests will be delayed during translation.**



# Machines Used

- **AIX1**
  - 9076-260
  - 2 x 200 MHz Power3 processor
  - Gb Ethernet Adapters
  - IBM SSA 160 Serial RAID Adapter
  - 8 x SSA Disks
- **AIX2**
  - IBM eServer pSeries 690
  - (model 7040-681)
  - 4 x 1.1 GHz Power IV Processors
  - 4 GB RAM
  - Gb Ethernet Adapters
  - IBM SSA 160 Serial RAID Adapters
  - Fiber Channel Adapters
  - 4 x SSA Disks
- **AIX3**
  - IBM eServer pSeries 690
  - (model 7040-681)
  - 8 x 1.1 GHz Power IV Processors
  - 4 GB RAM
  - Gb Ethernet Adapters
  - IBM SSA 160 SeriaL RAID Adapters
  - Fiber Channel Adapters
  - 4 x SSA Disks

WebSphere Data Interchange

# A Comparison with Send / Receive Maps

# Send / Receive Performance

- **The test runs on a z/OS machine.**
- **The input is a file of 235 Transactions. There are 2 maps being used.**
- **The maps are complex.**
- **Functional Acknowledgments are being generated.**
- **Command**
  - PERFORM DEENVELOPE AND TRANSLATE WHERE FILEID(INPUT) DUPENV(Y) RAWDATA(Y) RECOVERY(E) PURGINT(-1) FUNACKFILE(FACK)

# Send / Receive Performance

- **Send / Receive Map on Data Interchange Version 3.1**
- **Job report on next page**

```
*-------------------------------------------------------------------------------------------------------------------------*
*       SYSTEM NAME - SCND                 S T E P   C O M P L E T I O N   R E P O R T            SYSTEM LEVEL - MVS/SP 7.0.4  *
*-------------------------------------------------------------------------------------------------------------------------*
*     JOB NAME : RPOPEA62     REGION        <16MB        >16MB                          CIO: 00:00:00.08                     *
*     STEP NAME: RUNDI        AVAILABLE:    10,216K   1,597,440K   PRV MOUNT: 0        WLM: BATCH                           *
*     STEP NO  : 3            REQUESTED:    10,216K           0K   SCR MOUNT: 0        SCN: BATCH3                          *
*     PGM NAME : IKJEFT01     USED USER:     1,628K      45,888K                       GRN:                                 *
*     COND CODE: 0002         USED SYS :       516K      10,232K                       RCN:                                 *
*-------------------------------------------------------------------------------------------------------------------------*
*   DDNAME    UCB   VOLUME    I/O TIME    MAX BLKSZ   EXCP COUNT    DDNAME    UCB   VOLUME    I/O TIME    MAX BLKSZ   EXCP COUNT  *
*                                                                                                                         *
*   STEPLIB   762E  SN0210  00:00:00.00     19,069            2    STEPLIB   721E  S60312  00:00:00.03     19,069           36  *
*   STEPLIB   7519  S60301  00:00:00.00     19,069           60    STEPLIB   741E  S60371  00:00:00.00     19,069          248  *
*   STEPLIB   7619  S60347  00:00:00.16     19,069          274    STEPLIB   7000  S60100  00:00:00.00     19,069           31  *
*   SYSTSIN   7040  S60438  00:00:00.00      6,160            2    FFSWORK   7B24  S60503  00:00:00.68     32,760        2,156  *
*   TTABLE01  7621  S60451  00:00:00.00     27,920            2    FACK      721A  S60296  00:00:00.12     23,440          156  *
*   INPUT     7519  S60301  00:00:00.46     27,000          265    INPUT     7719  S60306  00:00:00.46     27,000          267  *
*   IN811     7621  S60451  00:00:03.82     27,000        2,191                                                               *
*-------------------------------------------------------------------------------------------------------------------------*
                                        IEF373I STEP/RUNDI   /START 2005145.0946
IEF374I STEP/RUNDI   /STOP  2005145.0957 CPU    2MIN 20.98SEC SRB    0MIN 00.15SEC VIRT  1628K SYS   516K EXT   45888K SYS   10232K
*-------------------------------------------------------------------------------------------------------------------------*
*       SYSTEM NAME - SCND                 J O B   C O M P L E T I O N   R E P O R T            SYSTEM LEVEL - MVS/SP 7.0.4  *
*-------------------------------------------------------------------------------------------------------------------------*
*            HIGHEST CONDITION   JOB START   JOB START    JOB END      JOB END     JOB ELAPSED TIME        PRIVATE     SCRATCH   *
*  JOBNAME          CODE            DATE        TIME         DATE        TIME       (HHHH:MM:SS.TH)         MOUNTS      MOUNTS   *
*                                                                                                                         *
*  RPOPEA62         0002         05/25/2005  09:46:15.91  05/25/2005  09:57:25.19      00:11:09.28             0           0    *
*-------------------------------------------------------------------------------------------------------------------------*
```

# Send / Receive Performance

- **Send / Receive Map on WebSphere Data Interchange Version 3.2.1**
- **Job report on next page**

```
*-------------------------------------------------------------------------------------------------------------------------*
*       SYSTEM NAME - SCND                  S T E P   C O M P L E T I O N   R E P O R T            SYSTEM LEVEL - MVS/SP 7.0.4  *
*-------------------------------------------------------------------------------------------------------------------------*
*     JOB NAME : RPOPEA51     REGION        <16MB        >16MB                          CIO: 00:00:00.06                     *
*     STEP NAME: RUNDI        AVAILABLE:    10,216K   1,597,440K   PRV MOUNT: 0        WLM: BATCH                           *
*     STEP NO  : 3            REQUESTED:    10,216K          0K   SCR MOUNT: 0        SCN: BATCH3                          *
*     PGM NAME : IKJEFT01     USED USER:     1,580K      46,032K                       GRN:                                 *
*     COND CODE: 0002         USED SYS :       484K       9,836K                       RCN:                                 *
*-------------------------------------------------------------------------------------------------------------------------*
*   DDNAME    UCB   VOLUME    I/O TIME    MAX BLKSZ   EXCP COUNT    DDNAME    UCB   VOLUME    I/O TIME    MAX BLKSZ   EXCP COUNT  *
*                                                                                                                         *
*   STEPLIB   762E  SN0210  00:00:00.00     19,069            2    STEPLIB   721E  S60312  00:00:00.05     19,069           45  *
*   STEPLIB   7414  S60299  00:00:00.35     19,069          541    EDITSIN   7040  S60438  00:00:00.00      6,160            2  *
*   SYSTSIN   7040  S60438  00:00:00.00      6,160            2    FFSWORK   7FFF  VIO     00:00:00.00     32,760          888  *
*   FACK      7A43  S60586  00:00:00.13     23,440          159    INPUT     7519  S60301  00:00:00.46     27,000          265  *
*   INPUT     7719  S60306  00:00:00.46     27,000          267    IN811     7621  S60451  00:00:03.42     27,000        1,955  *
*-------------------------------------------------------------------------------------------------------------------------*
                                        IEF373I STEP/RUNDI   /START 2005144.1930
IEF374I STEP/RUNDI   /STOP  2005144.1937 CPU    1MIN 26.19SEC SRB    0MIN 00.11SEC VIRT  1580K SYS   484K EXT   46032K SYS    9836K
*-------------------------------------------------------------------------------------------------------------------------*
*       SYSTEM NAME - SCND                  J O B   C O M P L E T I O N   R E P O R T            SYSTEM LEVEL - MVS/SP 7.0.4  *
*-------------------------------------------------------------------------------------------------------------------------*
*            HIGHEST CONDITION   JOB START   JOB START    JOB END      JOB END     JOB ELAPSED TIME        PRIVATE     SCRATCH   *
*  JOBNAME         CODE            DATE        TIME         DATE         TIME       (HHHH:MM:SS.TH)         MOUNTS      MOUNTS    *
*                                                                                                                         *
*  RPOPEA51        0002         05/24/2005  19:30:01.61  05/24/2005  19:37:14.82     00:07:13.21             0           0     *
*-------------------------------------------------------------------------------------------------------------------------*
```

# Send / Receive Performance

| | CPU Time | Elapsed Time |
|---|---|---|
| **Send / Receive Map V3.1** | **2 Min 21 Secs** | **11 Min 09 secs** |
| **Send / Receive Map V3.2.1** | **1 Min 26.19 Sec** | **7 Min 13 secs** |

- **Conclusion: Send / Receive maps run at least as well in V3.2 as in V3.1, and could be significantly better.**

# Questions

# The difference between Deluxe and Basic lies in the intended audience

- **Blue Onyx Deluxe, Blue Pearl Deluxe: Generally for "customer-facing" presentations**
  - Blue Pearl Deluxe is useful for one-on-one laptop presentations and for easy printing. Textures on the opening screen carry through the blue bands on text slides.
  - Blue Onyx Deluxe relies heavily on black for maximum contrast, particularly in projection.
- **Blue Onyx Basic, Blue Pearl Basic: Intended for basic internal presentations. May also be used for customers.**
  - Blue Onyx Basic uses black throughout for maximum contrast, particularly in projection.
  - Blue Pearl Basic works well for one-on-one laptop presentations and makes printing easy.
- **Additional usage tips in Notes page below**