

DB2 Data Warehousing at *KBC*

Dirk Beauson
dirk.beauson@kbc.be

Agenda

- **Our DataWarehouse environment**
- **Adding data into our DWH**
- **Exploiting the data in our DWH**
- **Maintaining our DWH**
- **The challenges and the future of our DWH**
- **Some UNIX/ORACLE DWH stuff**
- **Conclusions**



Our DataWarehouse environment

- **Most of our DWH info is on zOS/DB2**
- **But also some specific DWH data is stored in UNIX/ORACLE**

- **First focus on zOS/DB2**
- **Then on UNIX/ORACLE**

Our DataWarehouse environment : zOS/DB2

→ Hardware :

- z196 2817/719

→ Platform

- CPU model 2817
- # online CPU : 9
- Max : 13
- # zIIPs online : 3
- Central Storage : 52224 MB
- Storage infrastructure : 8.400 Gb

→ Filesystem/Database

- zOS DB2

Our DataWarehouse environment : zOS/DB2

- **How important is the DWH environment for our company ?**
 - **Lots of reports are generated from our Datawarehouse Environment**
 - **Business Data analysis on DWH**
 - **By default no (or CPU Limited Access) query access on the Base Tables**
 - **Decision taking data**
 - **So, very important**
 - **Example : Client Profile (Rendement per Cliënt)**

Adding data into our DWH : zOS/DB2

→ Sources

- No fancy stuff, only
 - DB2 base tables / Files

→ Design :

- Also no fancy stuff, only
 - DB2 extracts

→ Enriched Data :

- Some data is enriched,
- Some Datamarts are designed,
- but most of the data is a 1 to 1 replication of our DB2-base tables (most create cost effective)

→ All our DWH tables are designed using MEGA

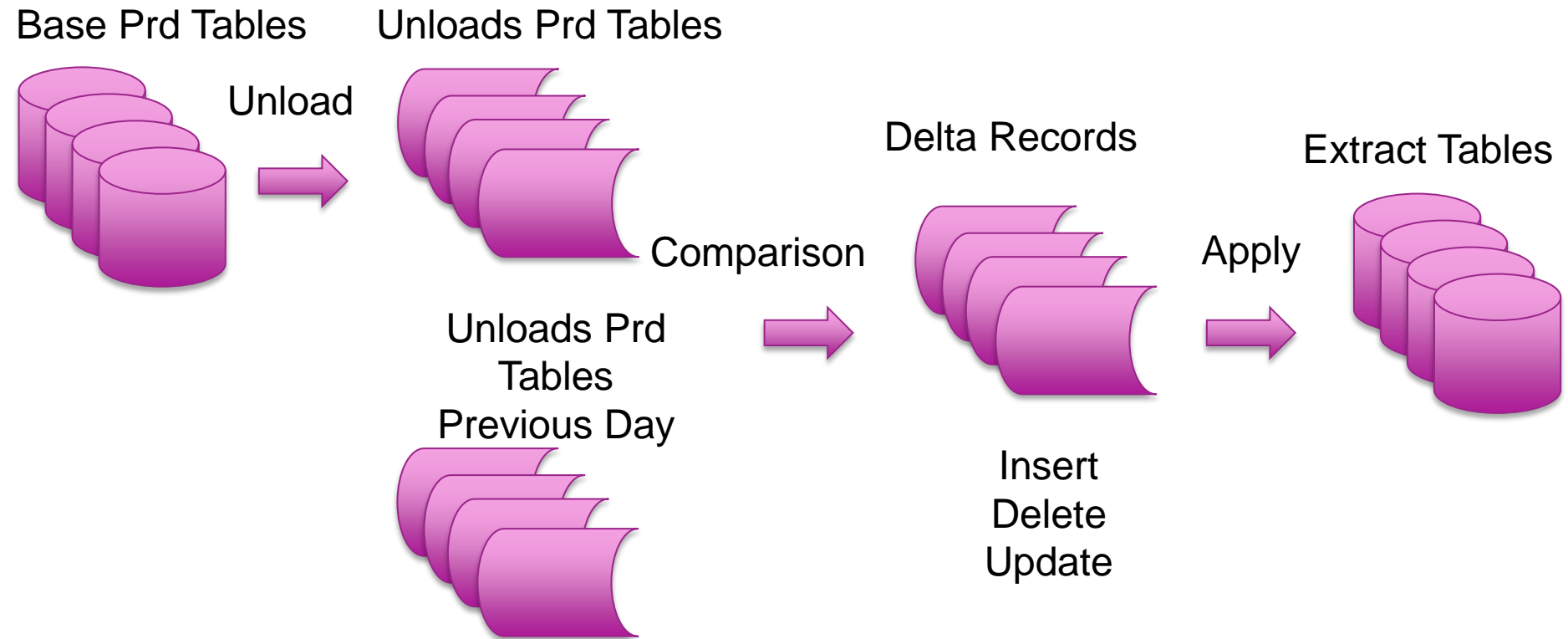
→ No RI

Adding data into our DWH : zOS/DB2

- **Different sorts of DB2-extracts :**
 - **Daily DB2-extracts (+- 600)**
 - **Weekly DB2-extracts (+-50)**
 - **Monthly DB2-extracts (+- 2000)**
 - **History**
 - **IL-DB2 extracts (+- 100)**

Adding data into our DWH : Daily DB2-extracts

→ Feeding :

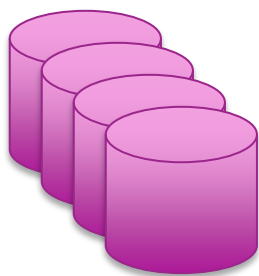


- **6 times a year we have a planned maintenance window. During that window all DAILY DB2-extracts are refreshed with the data of the DB2-Base tables via unload-(conversion)-load**

Adding data into our DWH : Monthly DB2-extracts

→ Feeding :

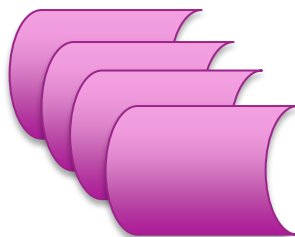
Base Prd Tables



Unload



Unloads Prd Tables



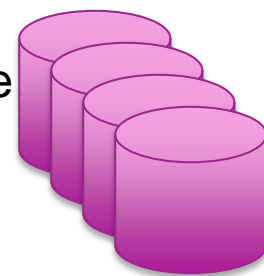
Conversion



Load Replace



Extract Tables



– Exception :

- **Some of the monthly DB2-extracts are used as weekly DB2-extract, so loaded every week instead of once a month.**

Adding data into our DWH : History of Monthly DB2-extracts

→ 2 types of History

– Semester history : (10 years)

- June
- December

– Monthly history : (2 years)

• extra runs in :

- January
- February
- March
- April
- May
- July
- August
- September
- October
- November

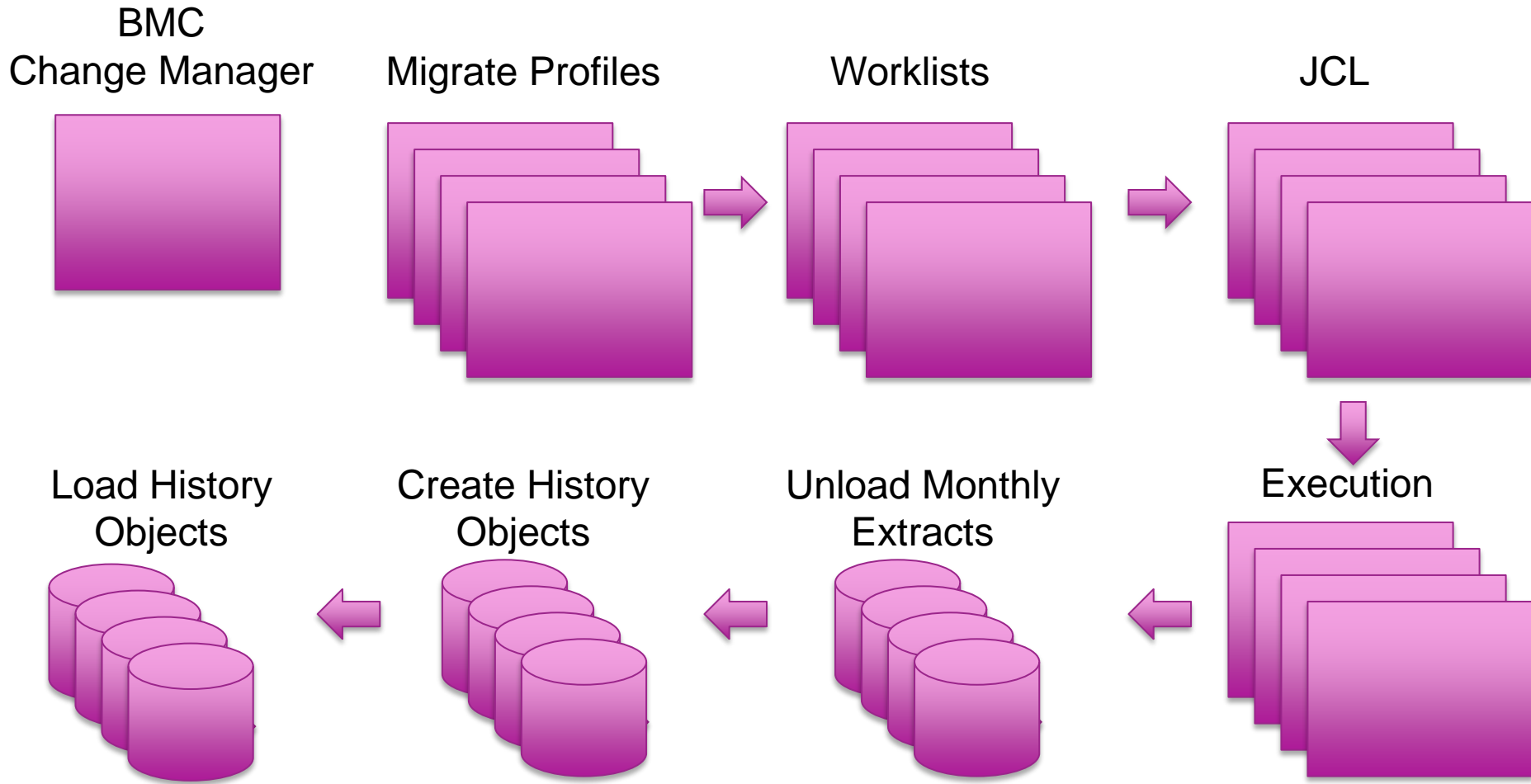
→ VSAM migrated to cassette if not used for 5 days

Adding data into our DWH : History of Monthly DB2-extracts

→ Feeding

– Via Change manager BMC

• Using Migrate Profiles

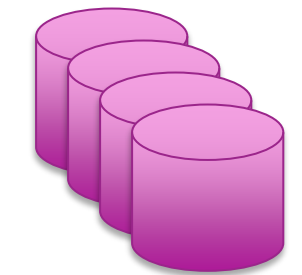


Adding data into our DWH : IL DB2-extracts

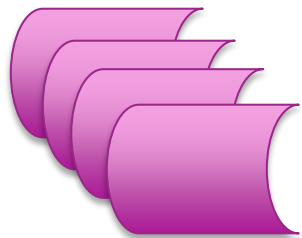
→ Feeding :

- Technically = Daily DB2-extract tables
- But using other procedures

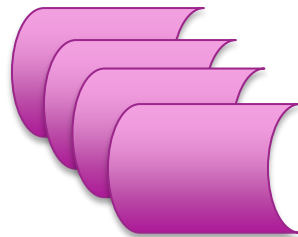
Base Prd Tables



Files



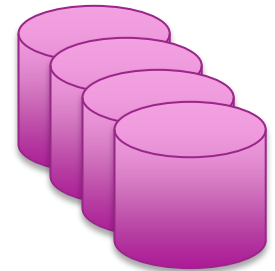
Load Files IL Tables



Load Replace



IL Tables



Exploiting the data in our DWH : zOS/DB2

→ Reporting tools

- QMF (Business, IT)
- SAS (Business)
- BO (Specialized team)
- Microsoft Suite
 - Excell
 - Access
- ...

→ Query parallelism is activated, only in our Datawarehouse environments

Exploiting the data in our DWH : zOS/DB2

→ Problems

- We see more and more development @ Business on DWH
 - So more and more Dynamic SQL
 - Very long running queries
 - Very complex
 - » Readability ???
 - » Functionality ???
 - Bad Accesspaths
 - Lots of joins
 - Other Accesspaths thx to DB2 10

Exploiting the data in our DWH

- **Too long running queries**
 - **No result**
 - **Used a lot of CPU for nothing**
 - **Pure Waste**
- **Design (copy of base tables)**
- **Indexes (mostly only key, and some basic IX copied from Base tables)**

Maintaining our DWH : zOS/DB2

→ Backup

- Only some tables of the IL DB2-tables are backup-ed
 - Some tables have to be recoverable in a very short period
- All other tables are NOT backup-ed
 - Load files
 - Cumul Delta files

Maintaining our DWH : zOS/DB2

→ Reorg

- Std none of the DWH tables are reorganized
- Most of them are loaded in the correct sequence
- Some exceptions

→ Runstats

- Because most of the tables are loaded with stats, no extra runstats are sheduled
- Some exceptions

Maintaining our DWH : zOS/DB2

→ Performance monitoring

- No active performance monitoring for our DWH systems
- More ad hoc, reactive way of servicing our clients

→ How do you Delete data / Archive data out of the DWH

- Monthly/Weekly/IL Extracts
 - Only data needed on the extract is loaded
 - No need of Delete Process
- Daily Extracts
 - Maintenance Data via Delta Processing
- Archive data -> History monthly data

The challenges and the future of our DWH : zOS/DB2

- **More and more development @ Business**
 - Lots of onpredictable Dynamic SQL to cope with
 - More complex queries
- **Need of storing more and more data, also much longer lifecycle of data**
 - So more Storage needed
 - More complex, longer Maintenance Windows
- **Follow up the IBM Accelerators !!!**
 - In the near future we will look of some of our workload is accelerator eligible
 - If so, we will look @ the IBM Accelerators

Our DataWarehouse environment : UNIX/ORACLE

→ Hardware :

- HPUX 11.31

→ Platform

- CPU : Application Dependent
- Central Storage : Application Dependent
- Memory : Application Dependent

→ Filesystem/Database

- Oracle 11g on a virtualised environment

Some UNIX/ORACLE DWH Applications

→ **MISKBC**

→ **IL Risk**

→ **IL Markets**

→ **...**

Adding data into our DWH : UNIX/ORACLE

→ **MISKBC : Project Management**

– **Physical Standby Database**

- Active Standby using Data Guard
- Provides a physically identical copy of the primary database, with on disk database structures that are identical to the primary database on a block-for-block basis. The database schema, including indexes, are the same. A physical standby database is kept synchronized with the primary database, through Redo Apply, which recovers the redo data received from the primary database and applies the redo to the physical standby database.
- A physical standby database can be used for business purposes other than disaster recovery on a limited basis.

– **Why on ORACLE ?**

- LOLA application by default NOT MF DB2

Adding data into our DWH : UNIX/ORACLE

→ **IL Risk :**

- KBC Risk Information Layer: targets like FERMAT, ALM, CRIMS, CPM are depending on it.
- Database setup is very specific
- Dynamic scheme, generated queries, hard to tune

- **Why on ORACLE ?**
 - **Historically build on UNIX/ORACLE**

Adding data into our DWH : UNIX/ORACLE

→ **IL Markets :**

- KBC Markets Informational Layer:
 - Sources: operational systems like IBS, Murex, ...
 - Targets: IL Risk, InfoCenter, Intellimatch, ...

- **Why on ORACLE ?**
 - Historically build on UNIX/ORACLE

Questions?

 **THANKS!**