Reduce Latency and Cost – Query IMS Data directly with DB2 Analytics Accelerator

Deepak Kohli
IMS Product Management
deepakk@us.ibm.com

Acknowledgements and Disclaimers

Availability. References in this presentation to IBM products, programs, or services do not imply that they will be available in all countries in which IBM operates.

The workshops, sessions and materials have been prepared by IBM or the session speakers and reflect their own views. They are provided for informational purposes only, and are neither intended to, nor shall have the effect of being, legal or other guidance or advice to any participant. While efforts were made to verify the completeness and accuracy of the information contained in this presentation, it is provided AS-IS without warranty of any kind, express or implied. IBM shall not be responsible for any damages arising out of the use of, or otherwise related to, this presentation or any other materials. Nothing contained in this presentation is intended to, nor shall have the effect of, creating any warranties or representations from IBM or its suppliers or licensors, or altering the terms and conditions of the applicable license agreement governing the use of IBM software.

All customer examples described are presented as illustrations of how those customers have used IBM products and the results they may have achieved. Actual environmental costs and performance characteristics may vary by customer. Nothing contained in these materials is intended to, nor shall have the effect of, stating or implying that any activities undertaken by you will result in any specific sales, revenue growth or other results.

Agenda

- Background / History
- Overview of the DB2 Analytics Accelerator for z/OS
- Existing client interest & use cases
- The solution for IMS data in DB2 Analytics Accelerator
- Implementation Steps
- IMS Lab POT
  - InfoSphere DataStage
  - DB2 Analytics Accelerator Loader for z/OS Tool
- Futures Investigation
- Resources
- Hands on Lab

Background / History

Ah, The Good Ol' Days

- ETL IMS data into data warehouses
- Queries submitted to the data warehouse
- Queries ran forever
- DBAs spent hours analyzing & fine tuning the SQL queries

Why we did what we did

- Back then IMS had no query capability
  - Query capability via JDBC started in IMS V7 & further solidified in IMS V10 with IMS Open database feature
- Netezza – high performance data warehouse appliances
  - Founded in 2000 and in 2010 IBM announced its acquisition

Overview of
the DB2 Analytics Accelerator for z/OS

The IBM DB2 Analytics Accelerator (1)

- What is it?
  - An integration of IBM PureData technology (formerly Netezza) with z Systems technology that delivers dramatically faster business analytics
- What does it do?
  - Accelerates complex queries, up to 2000x faster
  - Improves access to and lowers the cost of storing, managing and processing historical data
  - Minimizes latency
  - Reduces z Systems capacity requirements
  - Improves security and reduces risk

IBM DB2 Analytics Accelerator (2)

Enabling data-driven insight

- Transparent to applications and reporting tools that access DB2 data
- Inherits DB2 for z/OS data security attributes
- Simplified database design – no need for indexes
- Eliminates need for tuning of analytic queries
- Fast deployment and time-to-value
- Join data between IMS, DB2 & other sources in order of magnitude faster

Why it's transparent

A DB2 Analytics Accelerator is an extension of a DB2 for z/OS system

BI applications connect to DB2 server as always – no change

DBAs interact with Analytics Accelerator via Data Studio, stored procedures – only access to Accelerator is via DB2

DB2 for z/OS

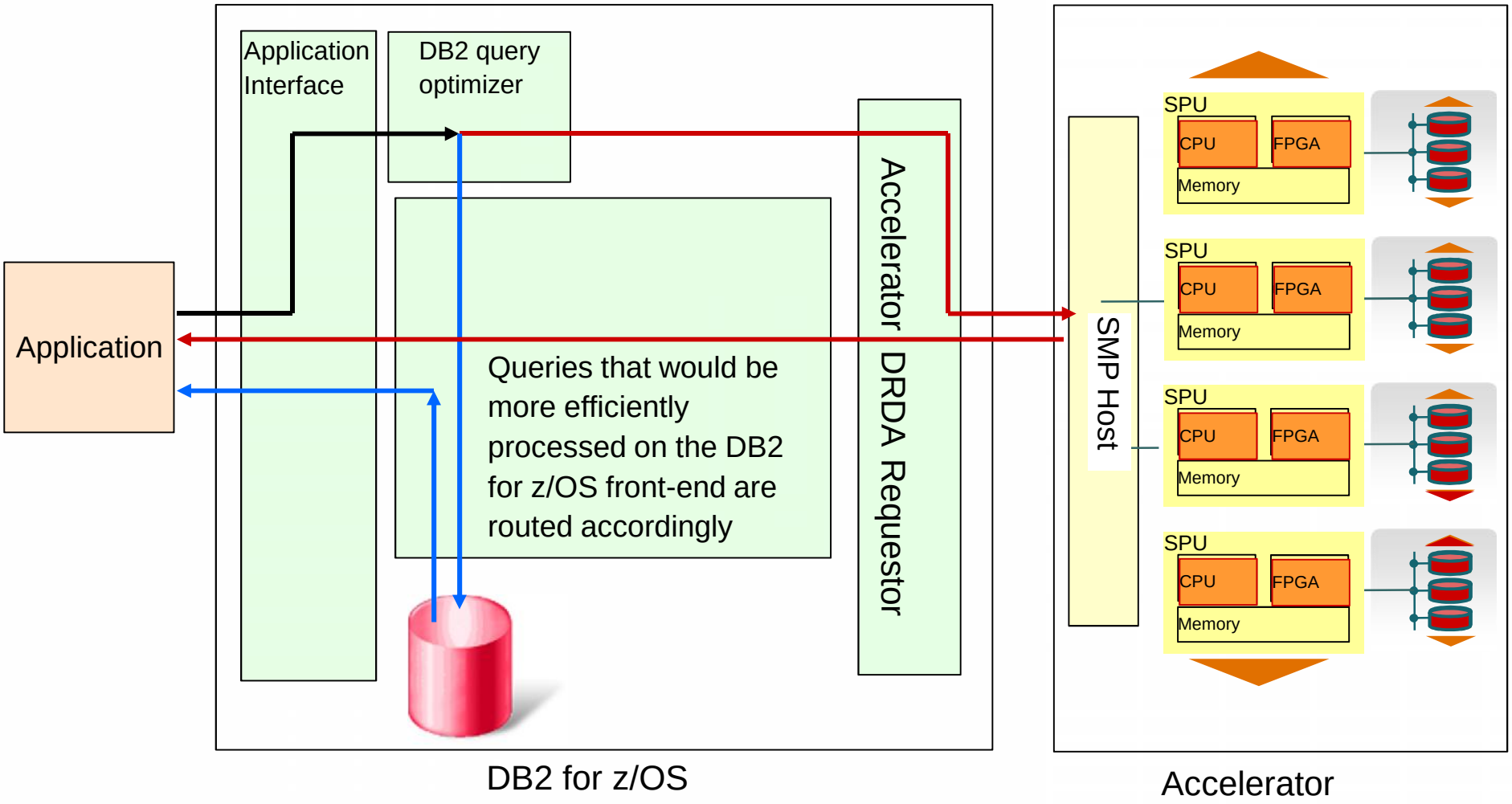
Deep integration with DB2 for z/OS:

- Automatic query routing
- EXPLAIN output
- Messaging
- Monitoring

IBM DB2 Analytics Accelerator

Superior performance for complex, data-intensive queries

Query execution process flow



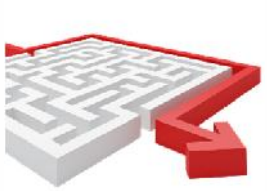
Existing Client Interest

Existing Client Interest …

- Major Insurance Co with both IMS and DB2
  - Recently did a very successful POC:
    - Loaded IMS & DB2 data into DB2 Analytics Accelerator
    - Ran queries against both IMS & DB2 data
    - Moving to production
- Financial Institution in the middle east
  - For years: ETL IMS data to Exadata
  - Successful POC putting that data into DB2 Analytics Accelerator
- Financial Institution in Japan & another in Europe
  - Expressed interest

DB2 Analytics Accelerator use cases with IMS data

Make better decisions faster



Large volume reporting of combined IMS and DB2 assets

Better understand your customers

Leverage full breadth of transactional data for analytics

Trust your data

Ensure consistency of data relationships between IMS and DB2

IBM Software Group
Technical White Paper

August 2014

IBM

Accelerate business insights with IMS transactional data

Introduction

IBM® System z® servers are home to 80 percent of the world's data,[1] a large part of which is Information Management System (IMS) transactional data. IMS applications and data continue to form the core of the world's critical online transaction processing (OLTP) workloads. IMS data on System z servers is widely viewed as the single version of the truth, from which data stores and warehouses are populated.

The System z platform offers high availability, scalability, security and performance for OLTP workloads. The data that resides on the System z platform and is targeted for analytics on the System z platform has those same qualities of service. Bringing analytics directly to the data, rather than moving data to the analytics, makes sense. It preserves these qualities of service to ensure timely, accurate, and secure insights, and reduces costs of moving data off the platform.

The insights that analytics provide must be delivered efficiently across an organization, with high quality and proper governance. Moving data off of a System z platform and across platforms is not only costly, it literally ages data out of its productive use and impedes the delivery of business insights. Where transactional data originates plays an increasingly important role in analytics. Some percentage of that 80 percent of data is not being analyzed at all due to the performance and cost challenges of moving data.

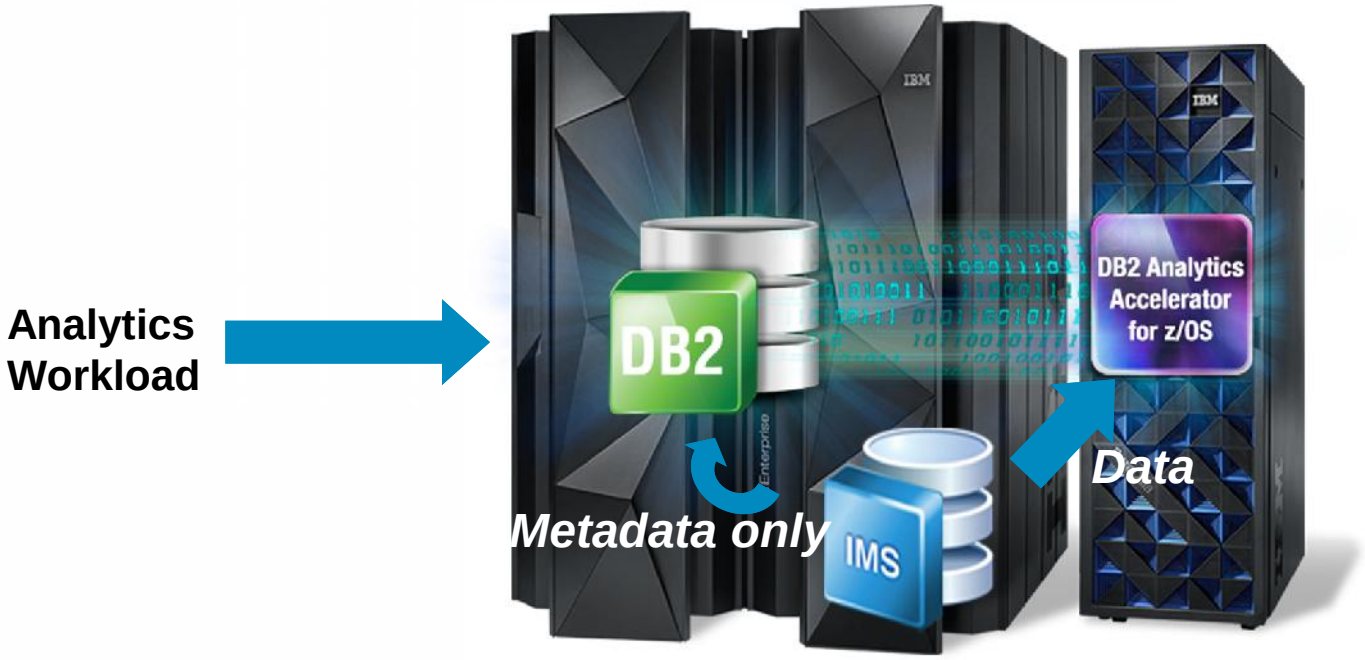
IBM

Get Started Today!

- Technical Whitepaper and "how-to" guide available here

14

The Solution: Routing IMS Queries through DB2



Implementation Steps (the simplistic view):

1. Extract and Transform IMS data
   - Transform because of special data types e.g. Packed decimal Date types in IMS

2. Define the IMS Tables (segments) to both DB2 and DB2 Analytics Accelerator

3. Load the IMS data into DB2 Analytics Accelerator only (no data needs to be loaded into DB2)

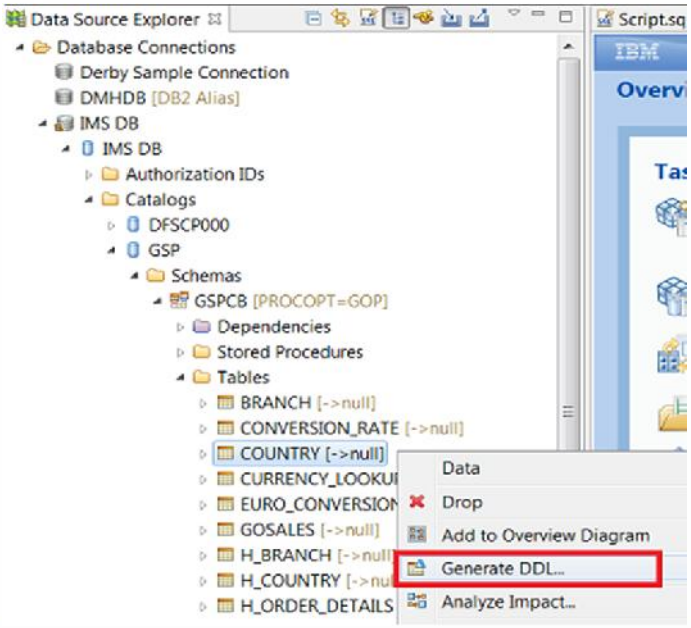
4. Enable query for acceleration

Implementation Steps (a little more detail):

1. Extract and Transform IMS data
   - Options:
     - User Written Program
     - ETL Tool: IBM InfoSphere DataStage Product

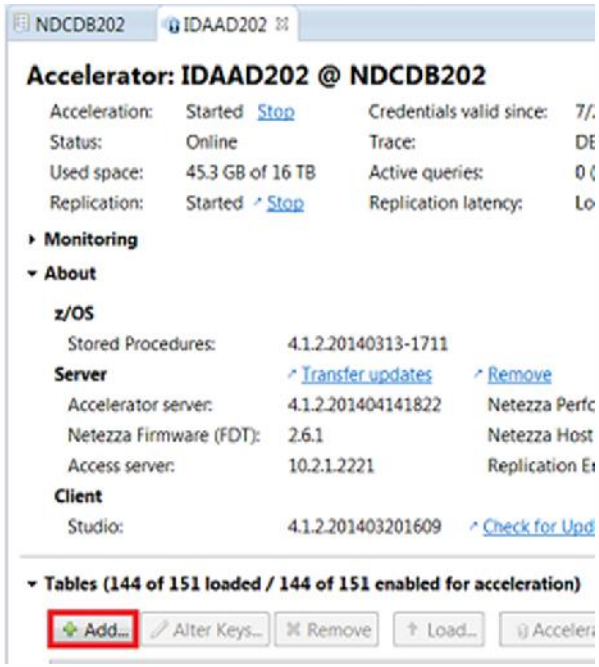
Implementation Steps (a little more detail):

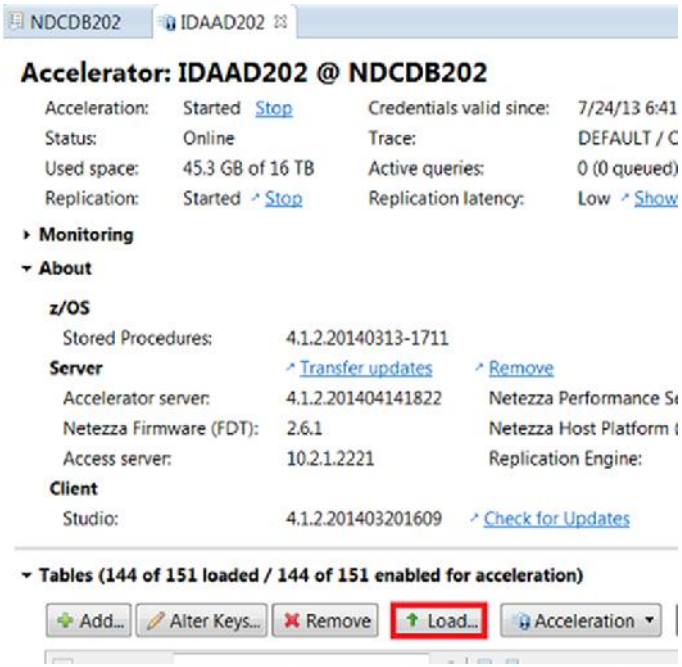
1. Extract and Transform IMS data
   - Options:
     - User Written Program
     - ETL Tool: IBM InfoSphere DataStage Product
2. Define the IMS Tables (Segments) to DB2 & DB2 Analytics Accelerator
   - Defining to DB2 done via execution of DDL (thru SPUFI for example)
   - The DDL can be generated by the IMS Explorer for development



Implementation Steps (a little more detail):

- Defining to DB2 Analytics Accelerator is a simple matter of executing ACCEL_ADD_TABLES stored procedure (can use Accelerator Studio GUI)





Implementation Steps (a little more detail):

- Step 3: Load the IMS data into DB2 Analytics Accelerator.
  - Use the DB2 Analytics Accelerator Loader for z/OS Tool (to load into Accelerator only)
  - Note the extracted data has to be in a format that the DB2 Analytics Accelerator Loader for z/OS tool expects

Implementation Options

- Option 1:
  - Extract and transform IMS data via a custom application and then
  - Load the data into DB2 Analytics Accelerator only using the DB2 Analytics Accelerator Loader for z/OS tool

Implementation Options

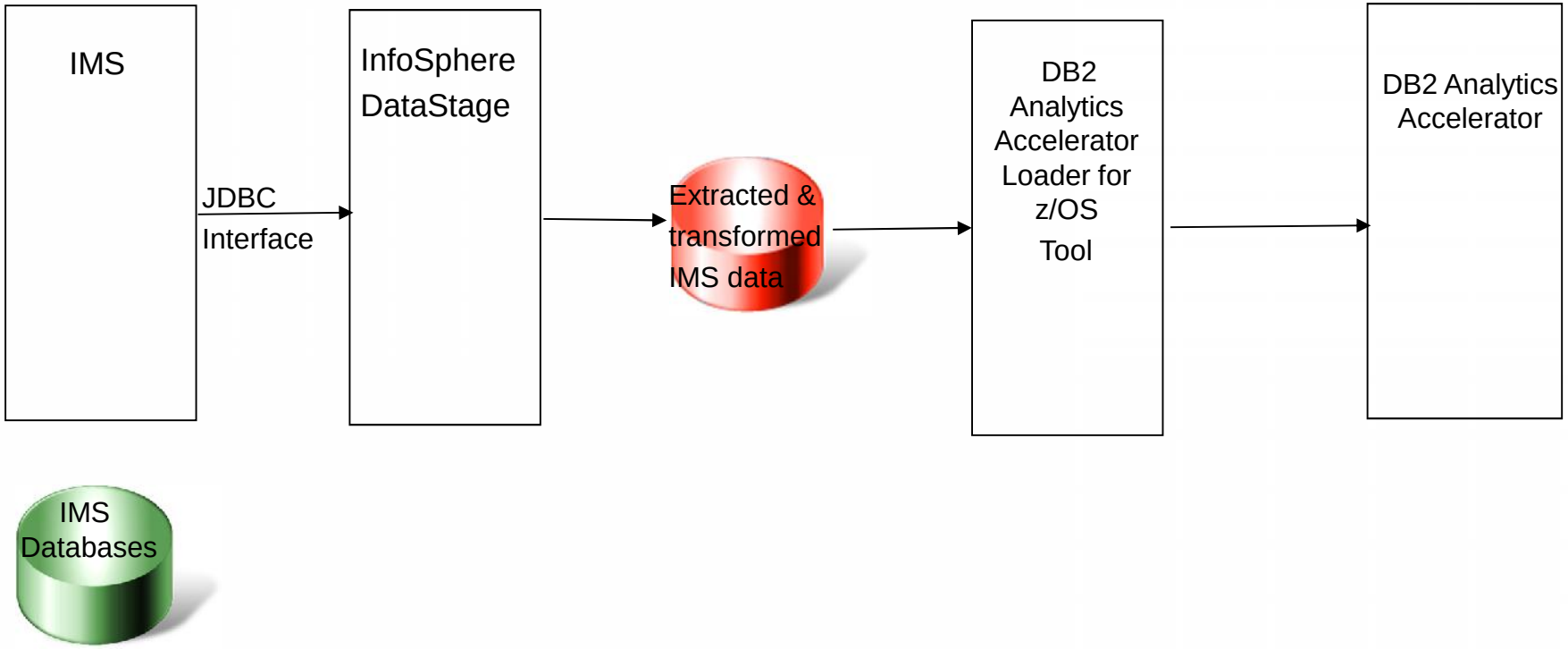
- Option 2:
  - Extract and transform IMS data using IBM InfoSphere DataStage or similar ETL tool
  - Load the data into DB2 Analytics Accelerator only using the DB2 Analytics Accelerator Loader for z/OS tool

Implementation Options

- Option 3:
  - Extract and transform IMS data using IBM InfoSphere DataStage tool
  - Load the extracted data into DB2 from DataStage using the DB2 Connector stage
  - Load the data from DB2 into DB2 Analytics Accelerator using the Accel_Load_Tables stored procedure

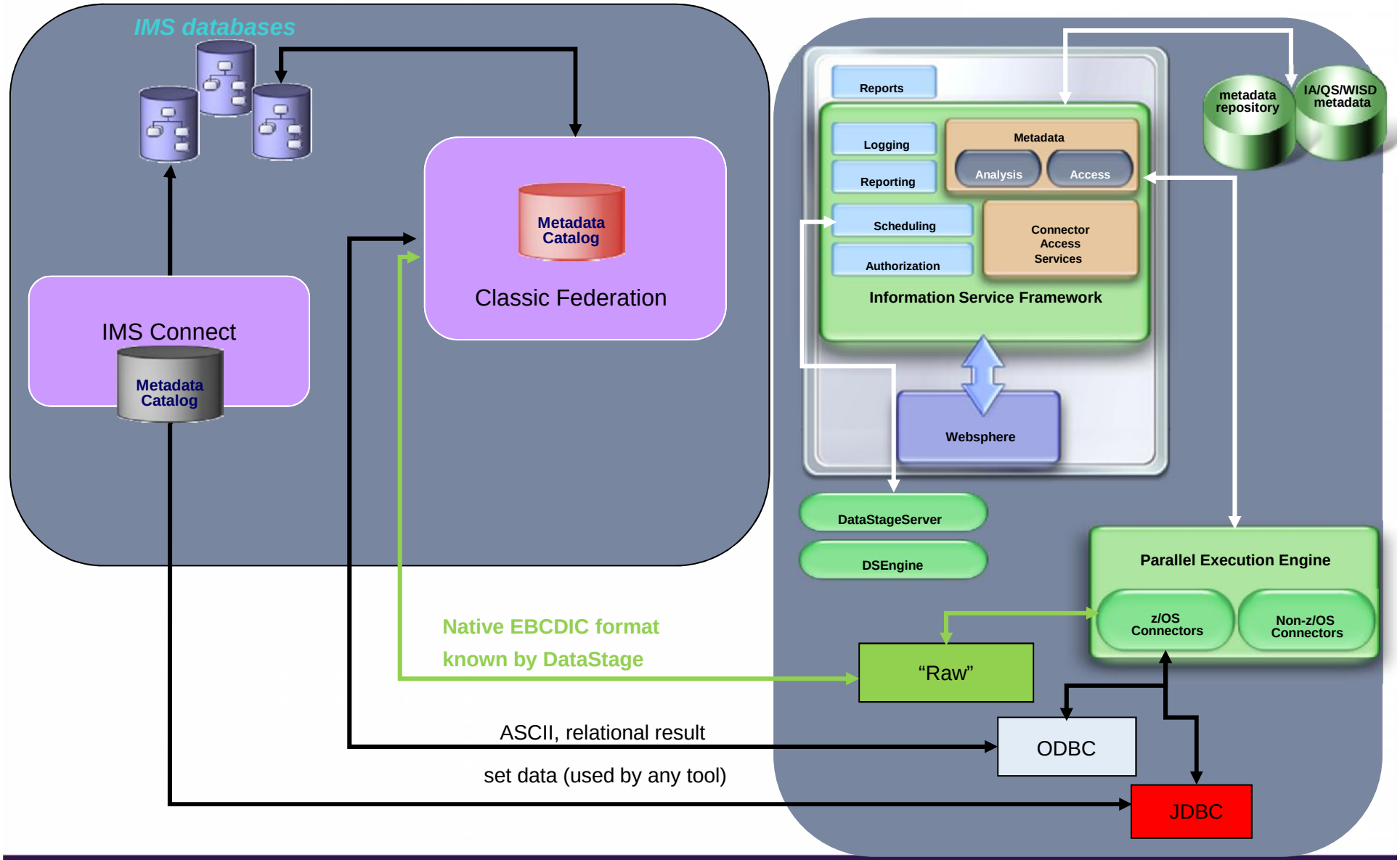
- Cons: Data Duplication
- Pros:
  - No need for DB2 Analytics Accelerator Loader for z/OS tool
  - For near real time analytics, could propagate changes to IMS data (using CDC) from IMS to DB2 to DB2 Analytics Accelerator.
    - For this case, don’t need Datastage and don’t need the DB2 Analytics Accelerator Loader for z/OS tool

IMS Lab POT


IMS
JDBC
Interface
InfoSphere
DataStage
Extracted &
transformed
IMS data
DB2
Analytics
Accelerator
Loader for
z/OS
Tool
DB2 Analytics
Accelerator
IMS
Databases

Extraction and transformation via DataStage

IMS Access Integration with DataStage – Classic Federation



Setting up DataStage

- Need to use DataStage 9.1.2 or later (to use the JDBC interface to IMS)
- DataStage has a client piece & a server piece
  - Client piece: DataStage Designer (on Windows)
  - Server piece: ran it on SUSE Linux Enterprise Server for System z but could be run on distributed
- Make the IMS JDBC drivers available to the DataStage server
  - http://pic.dhe.ibm.com/infocenter/iisinfsv/v9r1/topic/com.ibm.swg.im.iis.conn.jdbc.usage.doc/topics/jdbc_config_driver.html
- And that's it!!!

Using DataStage

- Use InfoSphere Metadata Asset Manager (IMAM) to import IMS metadata from the IMS Catalog – you need IMS Catalog installed!!
- Using the DataStage Designer client piece to design a job to do the Extract & transform
- Compile & run the job
- The job will produce a file of extracted & transformed data.

Transformation with DataStage

- DataStage has many built-in functions:
  - DateFromJulianDay – returns a date from the given Julian date
  - DecimalToString – returns the string representation of the given decimal
- Transformer Routines
  - Custom developed routines
  - Written in C++

Defining IMS tables to DB2 & DB2 Analytics Accelerator

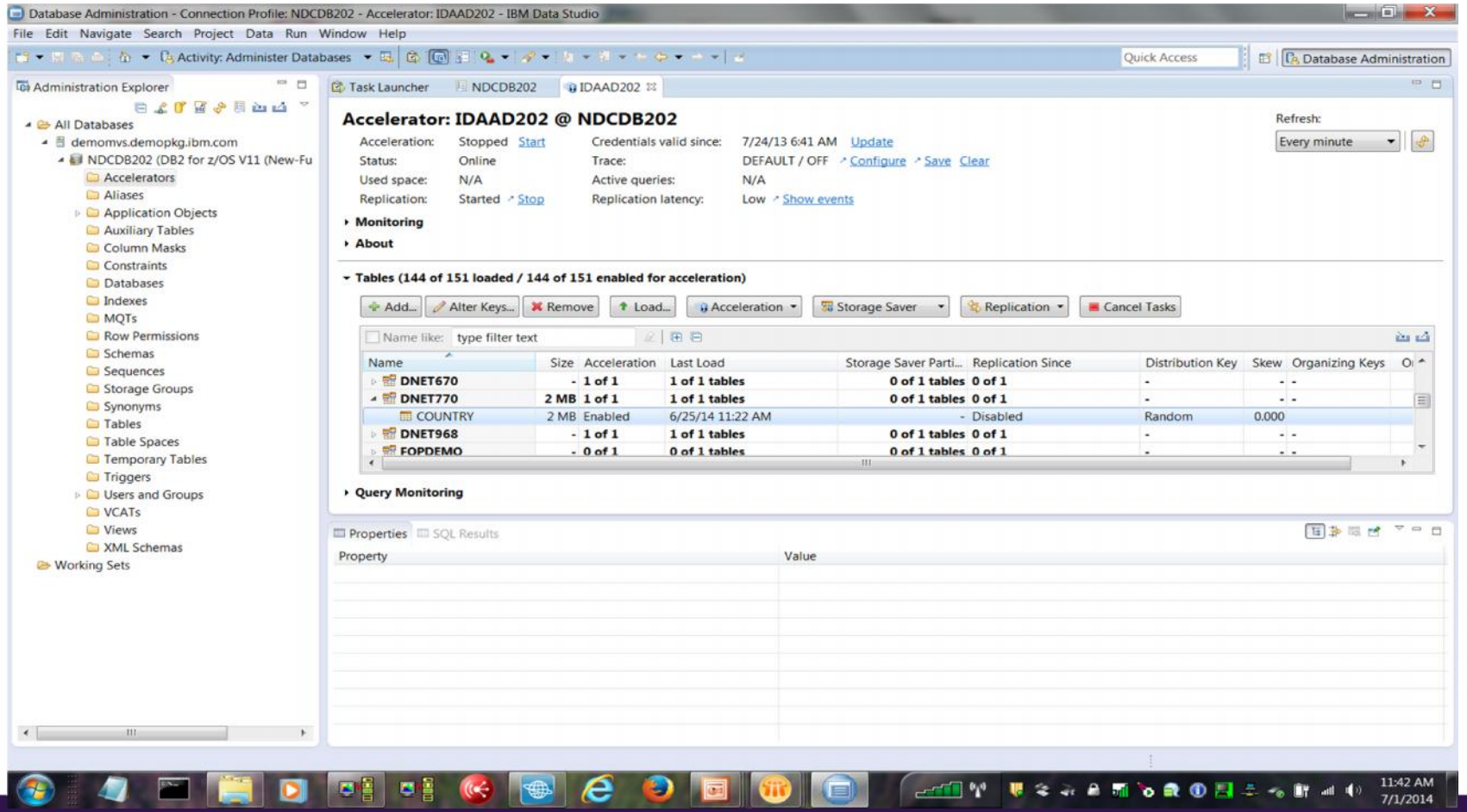
Defining IMS tables to DB2

- Executing DDL – via SPUFI for example

```
CREATE TABLE DNET770.COUNTRY (
GOSALES_ROOT                CHAR(12),
COUNTRY_CODE                INTEGER,
COUNTRY_EN                  CHAR(180),
SALES_REGION_CODE           INTEGER,
ISO_THREE_LETTER_CODE       CHAR(18),
ISO_TWO_LETTER_CODE         CHAR(12),
ISO_THREE_DIGIT_CODE        CHAR(18) );
```

Defining IMS tables to DB2 Analytics Accelerator

Matter of executing ACCEL_ADD_TABLES stored procedure (can use Accelerator Studio GUI)

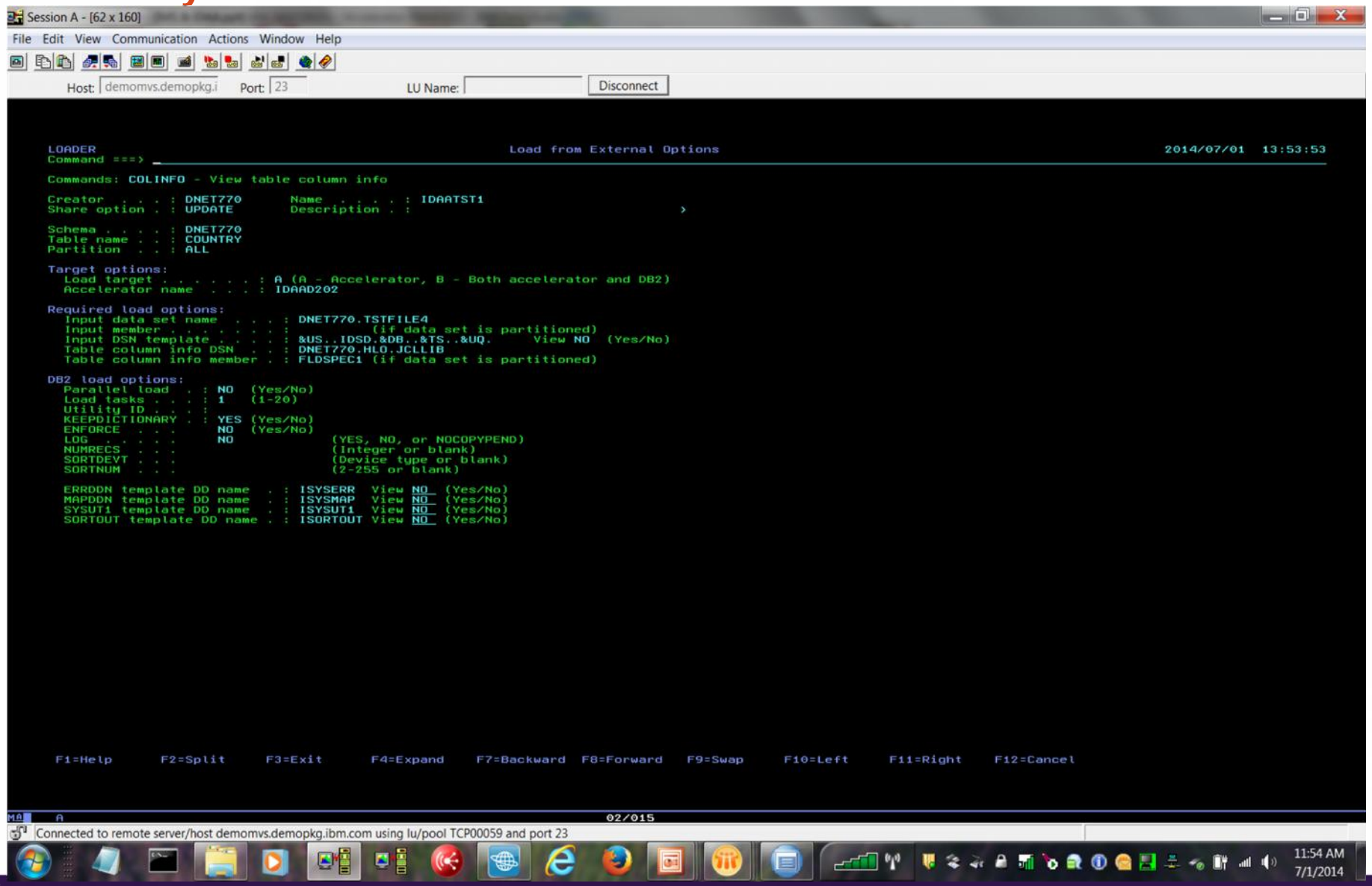


Loading IMS data into DB2 Analytics Accelerator using the DB2 Analytics Accelerator Loader for z/OS Tool

Loading IMS data into DB2 Analytics Accelerator

1. So we extracted & transformed the data using DataStage
2. We defined the IMS tables to DB2 & DB2 Analytics Accelerator
3. FTP the data over to z/OS
4. Run the DB2 Analytics Accelerator Loader for z/OS tool JCL to Load IMS data into DB2 Analytics Accelerator only
   - DB2 Analytics Accelerator Loader for z/OS tool ISPF Panels can generate the JCL
   - But user still needs to code the field specifications:
     - What fields are in which columns of the input dataset

DB2 Analytics Accelerator Loader for z/OS tool ISPF Panel



DB2 Analytics Accelerator Loader for z/OS tool example JCL:

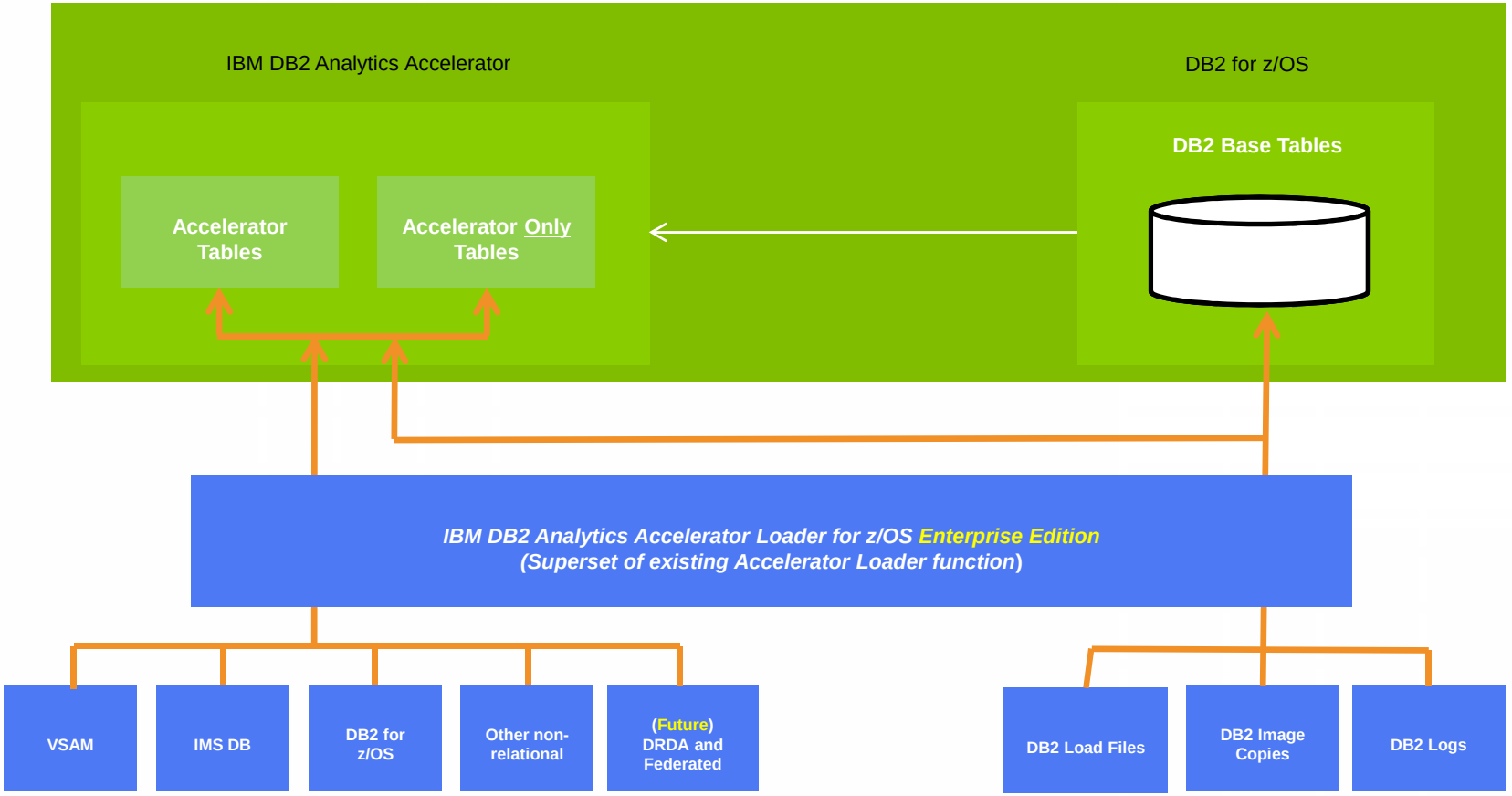
```
//HLOD0100 EXEC PGM=DSNUTILB,
//        REGION=0000M,
//        PARM=('DSNB')
//STEPLIB  DD DISP=SHR,DSN=DB2AAL.V1R1.SHLOLOAD
//        DD DISP=SHR,DSN=DB2.V11.DSNB.SDSNEXIT
//        DD DISP=SHR,DSN=DB2.V11.SDSNLOAD
//HLODUMMY DD DUMMY
//SYSPRINT DD SYSOUT=*
//UTPRINT  DD SYSOUT=*
//SYSIN    DD *
      TEMPLATE ISYSREC
        DSN 'DNET770.TSTFILE4'                          ← input dataset name
        DISP(SHR,KEEP,KEEP)
…
…
LOAD DATA
                IDAA_ONLY ON IDAAD202                   ← Accelerator
                INDDN ISYSREC
        …
                INTO TABLE
                        "DNET770"."COUNTRY"             ← Table
                                (
                                        GOSALES_ROOT POSITION ( 00001:00012) CHAR(00012),
                                        COUNTRY_CODE POSITION ( 00013:00016) INTEGER,
                                        COUNTRY_EN POSITION ( 00017:00196) CHAR(180),
                                        SALES_REGION_CODE POSITION ( 00197:00200) INTEGER,
                                        ISO_THREE_LETTER_CODE POSITION ( 00201:00218) CHAR(18),
                                        ISO_TWO_LETTER_CODE POSITION ( 00219:00230) CHAR(12),
                                        ISO_THREE_DIGIT_CODE POSITION ( 00231: 00248) CHAR(18)
                                )
//
```

Minimum Software versions required

- IMS
  - Custom application to do the ETL: no minimum requirement
  - DataStage to do the ETL
    - Using the JDBC Interface
      - DataStage 9.1.2 & IMS V12 or higher with IMS Catalog implemented
    - Using the ODBC interface or Raw interface: no minimum IMS requirement
      - But also need Classic Federation
- DB2 V10 or higher
- DB2 Analytics Accelerator Version 3 or higher

Futures Investigation

IBM DB2 Analytics Accelerator Loader for z/OS Enterprise Edition



Use Cases

- Load DB2 data and Non-DB2 data directly to IDAA
- Easy consolidation of enterprise data on secure z platform
- Exploit Accelerator to join DB2 non-DB2 data

- New analytic workload on DB2 for z/OS
- Trending, Fraud detection, Capacity Planning, etc.
- Analytics of SMF data, DB2/IMS performance data, etc.

IMS Data Load Process

- Map IMS data sources via copy books
- Create virtual tables that represent IMS data
- Build and test select statement that represents data desired in accelerator
- Test select to verify it returns desired data
- DB2 table created that matches select result set
- JCL built to perform direct load from IMS into accelerator
- Batch job will extract IMS data and load to accelerator
  - Data extracted and loaded directly – not landed in files
  - Can be placed in scheduler for refresh operation

Resources

Resources for IMS clients

- White paper
  - http://ibm.biz/accelerate_insights_ims_transactional_data
- Technical Implementation document (“Cook Book”)
  - https://ibm.biz/ims_idaa_technical_implementation
- IMS & IDAA demo on Youtube
  - https://ibm.biz/demo_ims_idaa

IMS Technical Resources …

- Kyle Charlet/Santa Teresa/IBM@IBMUS
  - IBM Distinguished Engineer
  - 1-408-463-4145

- Deepak Kohli/Silicon Valley/IBM@IBMUS,
  - 1-310-393-5902

Thanks for your time!