The Magic Box of Healthcare Analytics

Scott Davis, Senior Systems Engineer/DBA

Blue Cross and Blue Shield of Alabama

Who we are…

- Only plan to have won 19 Brand Excellence Awards from the Blue Cross and Blue Shield Association
- Over 2.1 million customers within Alabama with an additional 900,000 customers outside of the state
- Industry leader in administrative costs. In 2014, over 92 percent of all revenue went toward treating patients.

Agenda

- Problem returning large amounts of data
- The IBM DB2 Analytics Accelerator Solution
- The “Magic Box” in the real world of Healthcare Analytics

What's the issue?

Data Performance

- Business areas run ad-hoc queries against databases, but results are often slow to return or don't return at all
- Time should be spent on analyzing the data rather than gathering the data

Why are query speeds slow?

- Large amounts of data being analyzed
- Joining multiple tables/views to accomplish required analysis
- Non-technical analysts using software to create queries/reports

WHERE
PRD.ENDDT < PRD.SNPSHT_DT GROUP BY PRD.SNPSHT_DT , PRD.UNQ_MEM_ID , PRD.CNTRCT_PFX , PRD.CNTRCT_ID , PRD.MEM_SEQ_NR) G1 JOIN SUBSME.CNTRCT_MEM_EFF_CNDNSD_SNPSHT G2 ON G2.SNPSHT_DT = G1.SNPSHT_DT AND G2.UNQ_MEM_ID = G1.UNQ_MEM_ID AND G2.CNTRCT_PFX = G1.CNTRCT_PFX AND G2.CNTRCT_ID = G1.CNTRCT_ID AND G2.MEM_SEQ_NR = G1.MEM_SEQ_NR AND COALESCE(G2.CNTUS_COVG_PRD_NR, 0) = G1.COVGPRD GROUP BY G1.SNPSHT_DT , G2.UNQ_MEM_ID , G2.CNTRCT_PFX , G2.CNTRCT_ID , G2.MEM_SEQ_NR , G1.COVGPRD) G3 JOIN SUBSME.CNTRCT_MEM_EFF_CNDNSD_SNPSHT G4 ON G4.SNPSHT_DT = G3.SNPSHT_DT AND G4.UNQ_MEM_ID = G3.UNQ_MEM_ID AND G4.CNTRCT_PFX = G3.CNTRCT_PFX AND G4.CNTRCT_ID = G3.CNTRCT_ID AND G4.MEM_SEQ_NR = G3.MEM_SEQ_NR AND G4.CNTUS_COVG_PRD_NR = G3.COVGPRD AND G4.EFF_TO_DT = G3.G2_ENDDT JOIN SUBSME.UNQ_MEM_ID_SUMM_SNPSHT AE1 ON AE1.SNPSHT_DT = '04/27/2012' AND AE1.UNQ_MEM_ID = G4.UNQ_MEM_ID AND AE1.CNTRCT_PFX = G4.CNTRCT_PFX AND AE1.CNTRCT_ID = G4.CNTRCT_ID AND AE1.MEM_SEQ_NR = G4.MEM_SEQ_NR LEFT JOIN REF.DIV_CAN_CD DC ON G4.DIV_CAN_CD = DC.DIV_CAN_CD AND G4.DIV_CAN_CD IS NOT NULL LEFT JOIN REF.CAN_RSN_CD CC ON G4.CNTRCT_CAN_RSN_CD = CC.CAN_RSN_CD AND G4.CNTRCT_CAN_RSN_CD IS NOT NULL LEFT JOIN REF.MEM_RMVL_CD MC ON G4.MEM_RMVL_CD = MC.MEM_RMVL_CD AND G4.MEM_RMVL_CD IS NOT NULL) G LEFT JOIN REF.ACTURL_ENRL_SUBC AS1 ON AS1.ACTURL_ENRL_SUBC = G.OLDAE LEFT JOIN REF.ACTURL_ENRL_CATGY AC1 ON AC1.ACTURL_ENRL_CATGY = AS1.ACTURL_ENRL_CATGY JOIN SUBS.SNPSHT_DT_TBL_XREF SX ON SX.TBL_NM = 'COVD_LVS_SNPSHT' AND SX.SNPSHT_DT_1 >= G.SNPSHT_DT AND SX.SNPSHT_DT_1 < DATE(G.SNPSHT_DT + 7 MONTHS) LEFT JOIN SUBSME.CNTRCT_MEM_EFF_CNDNSD_SNPSHT OTH ON OTH.SNPSHT_DT = SX.SNPSHT_DT_1 AND OTH.UNQ_MEM_ID = G.UNQ_MEM_ID AND (OTH.CNTRCT_PFX = G.CNTRCT_PFX OR OTH.CNTRCT_ID = G.CNTRCT_ID) AND (OTH.EFF_TO_DT IS NULL OR (OTH.EFF_TO_DT > OTH.EFF_DT AND OTH.EFF_TO_DT > G.OLDEND)) AND OTH.CNTRCT_PFX = 'WRI' LEFT JOIN SUBSME.UNQ_MEM_ID_SUMM_SNPSHT AE2 ON AE2.SNPSHT_DT = OTH.SNPSHT_DT AND AE2.UNQ_MEM_ID = OTH.UNQ_MEM_ID AND AE2.CNTRCT_PFX = OTH.CNTRCT_PFX AND AE2.CNTRCT_NR = OTH.CNTRCT_NR LEFT JOIN SUBSME.CNTRCT_COVG_CNDNSD_SNPSHT CCH ON CCH.SNPSHT_DT = OTH.SNPSHT_DT AND CCH.CNTRCT_PFX = OTH.CNTRCT_PFX AND CCH.CNTRCT_NR = OTH.CNTRCT_NR AND (CCH.EFF_TO_DT IS NULL OR (CCH.EFF_TO_DT > CCH.EFF_DT AND CCH.EFF_TO_DT > G.OLDEND)) AND AE2.CNTRCT_PFX IS NULL AND CCH.TYP_BUS IN ('1', '2', '4') LEFT JOIN SUBSME.CNTRCT_COVG_CNDNSD_SNPSHT CCD ON CCD.SNPSHT_DT = OTH.SNPSHT_DT AND CCD.CNTRCT_PFX = OTH.CNTRCT_PFX AND CCD.CNTRCT_NR = OTH.CNTRCT_NR AND (CCD.EFF_TO_DT IS NULL OR (CCD.EFF_TO_DT > CCD.EFF_DT AND CCD.EFF_TO_DT > G.OLDEND)) AND AE2.CNTRCT_PFX IS NULL AND CCD.TYP_BUS = '7' LEFT JOIN SUBSME.CNTRCT_COVG_CNDNSD_SNPSHT CCR ON CCR.SNPSHT_DT = OTH.SNPSHT_DT AND CCR.CNTRCT_PFX = OTH.CNTRCT_PFX AND CCR.CNTRCT_NR = OTH.CNTRCT_NR AND (CCR.EFF_TO_DT IS NULL OR (CCR.EFF_TO_DT > CCR.EFF_DT AND CCR.EFF_TO_DT > G.OLDEND)) AND AE2.CNTRCT_PFX IS NULL AND CCR.MAJ_LN_BUS = '6' LEFT JOIN SUBSME.CNTRCT_COVG_CNDNSD_SNPSHT CCX ON CCX.SNPSHT_DT = OTH.SNPSHT_DT AND CCX.CNTRCT_PFX = OTH.CNTRCT_PFX AND CCX.CNTRCT_NR = OTH.CNTRCT_NR AND (CCX.EFF_TO_DT IS NULL OR (CCX.EFF_TO_DT > CCX.EFF_DT AND CCX.EFF_TO_DT > G.OLDEND)) AND AE2.CNTRCT_PFX IS NULL AND CCX.TYP_BUS NOT IN ('1', '2', '4', '7') AND CCX.MAJ_LN_BUS = '6'

BCBSAL Environment

2 EC12 CPCs (1 model 705, 1 model 706)

z/OS 1.13

DB2 V10 NFM

3 main DB2 z/OS subsystems – Processing, Warehouse and Development

Enterprise Data Warehouse utilizes 3 way data sharing with one member hosting a majority of the workload

Information Management

Information Management program established to manage information/data cohesively and comprehensively

'Data Performance' identified as a major area for improvement

Information Management Strategy developed with the assistance of IBM

DB2 Analytics Accelerator Solution (2013)

IBM DB2 Analytics Accelerator

- Netezza (IBM PureData for Analytics) based solution

Netezza Data Warehouse appliance connected to IBM mainframe through 10 gig dedicated fiber cable

DB2 Analytics Accelerator software on mainframe integrates with DB2 to determine which queries to accelerate

DB2 Analytics Accelerator Solution (continued)

Netezza appliance built on Massive Parallel Processing (MPP) architecture

- Several times faster than native DB2 processing

IBM DB2 Analytics Accelerator for z/OS V2.1

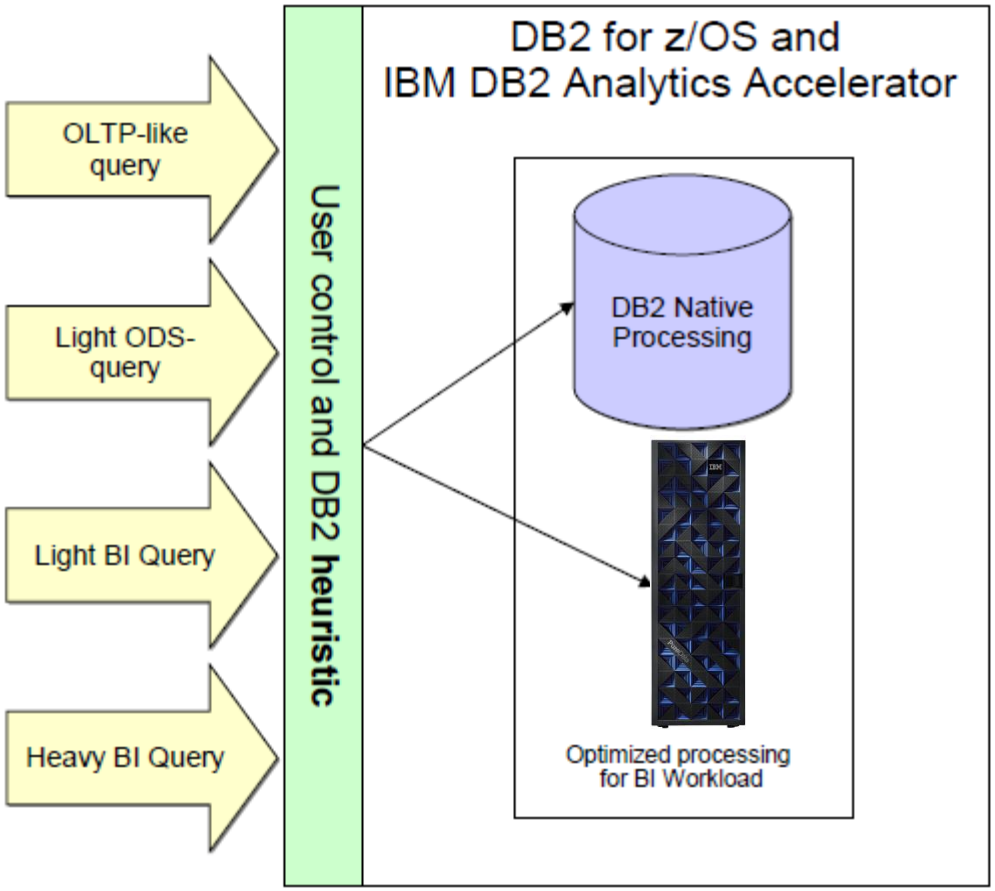
What it is / fit – validate agenda assumption

What is it?

The IBM DB2 Analytics Accelerator is a workload optimized, appliance add-on to a DB2 z/OS environment that services long-running, complex queries.

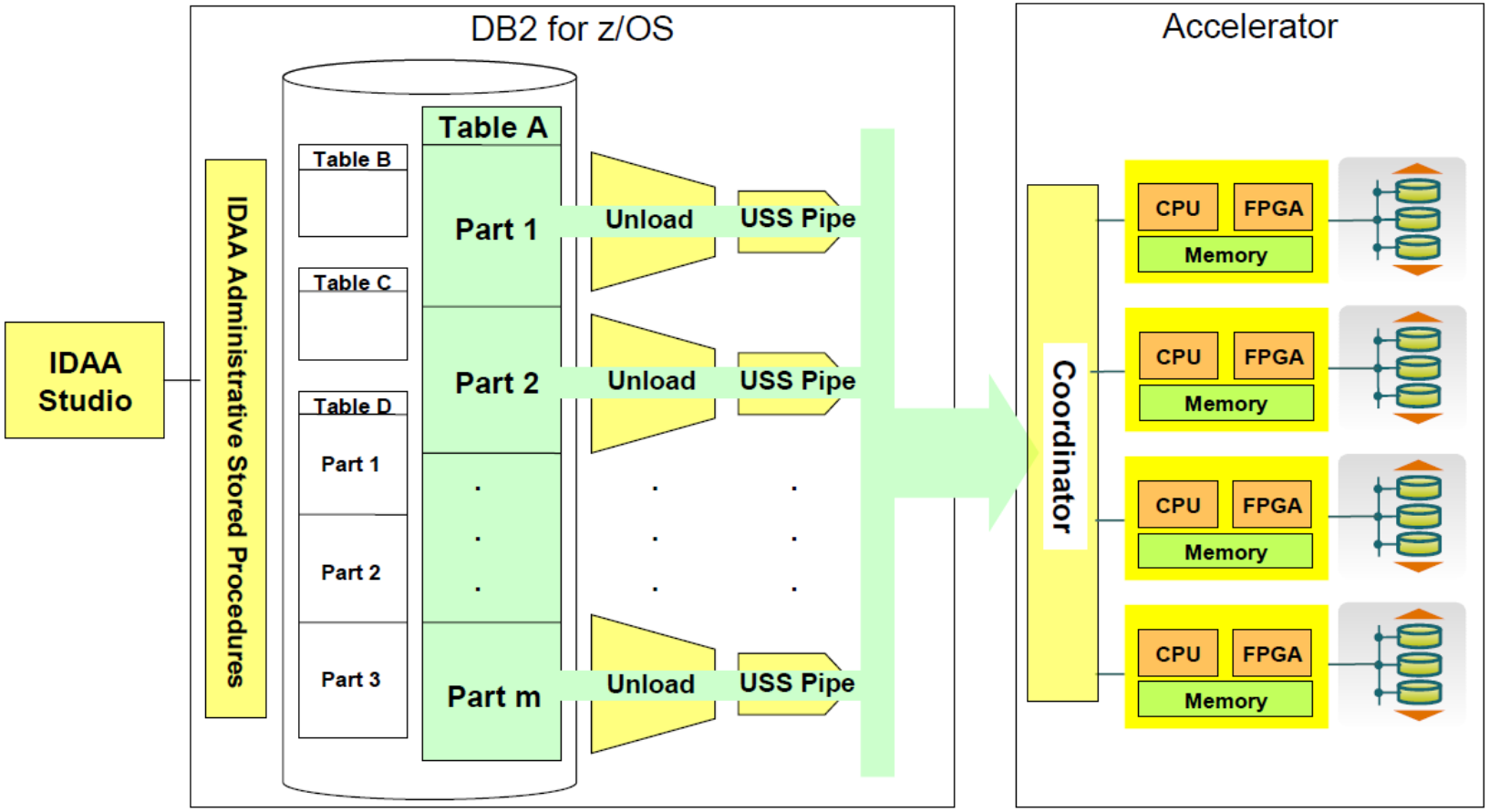
- **Breakthrough technology enabling new opportunities**
- **Extreme performance for complex analytics (aka Train of Thought Analysis)**
- **Integrated with DB2 for z/OS V9 and V10 as a dedicated appliance exclusive to the System z environment**
- **Transparent to DB2 applications and users**

Workload-Optimized Query Execution



- Single and unique system for mixed query workloads
- Dynamic decision for most efficient execution platform
- New special register QUERY ACCELERATION
  - NONE
  - ENABLE
  - ENABLE WITH FAILBACK
- New heuristic in DB2 optimizer
- **Combines the strengths of both System z and Netezza**
- **Merging operational and data warehouse into a single optimized environment**

Accelerator Data Load



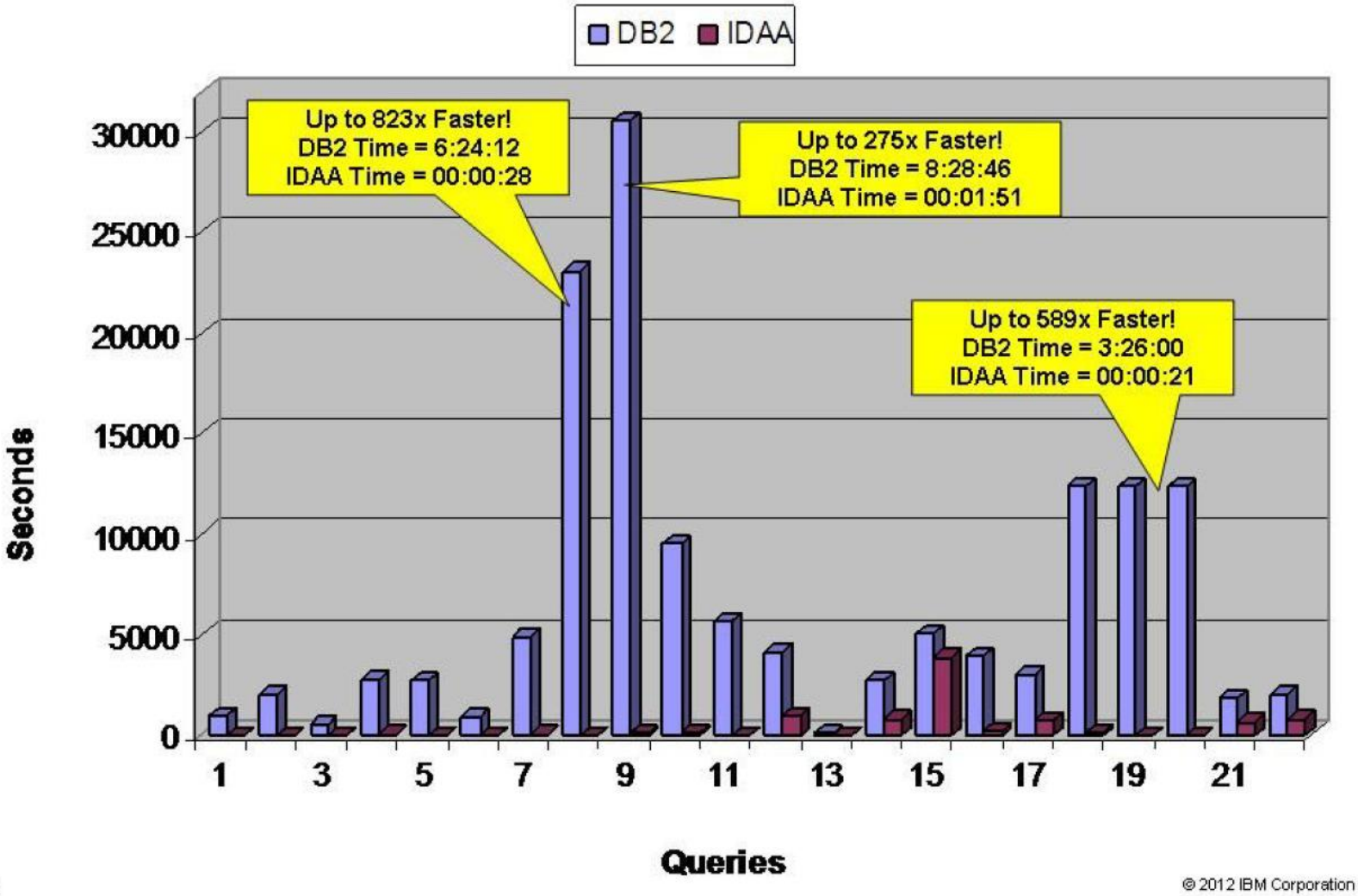
- Load speed up to 1 TB/hr are common
- Trickle-feed update under development

Tools Used by Analysts

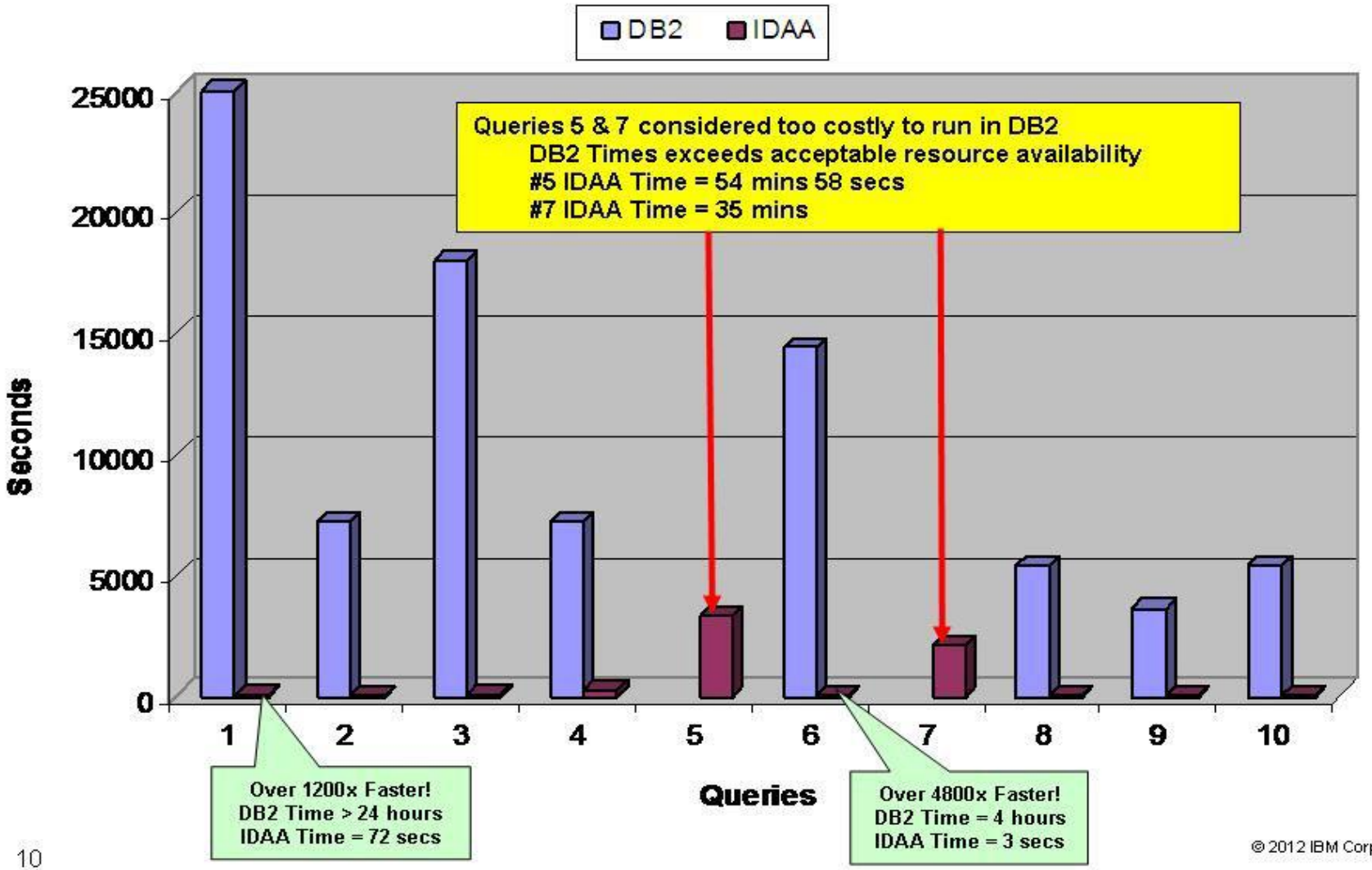
Reporting environment

- Business Objects Crystal Reports
- Business Objects Web Intelligence
- Business Objects Enterprise InfoView 3.1
- Tableau
- Toad Data Point
- AQT
- WinSQL
- SPUFI
- DB2 Connect 9.7 Fixpack 7

Comparison of Query Elapsed Time



Elapsed Time Comparison for Exception Queries



DB2 Analytics Accelerator Implementation

Purchased Twinfin 12 in first quarter of 2013.

- PoC was conducted on Twinfin 6

Acceleration enabled in warehouse subsystem on June 9, 2013. (Software version 2.1)

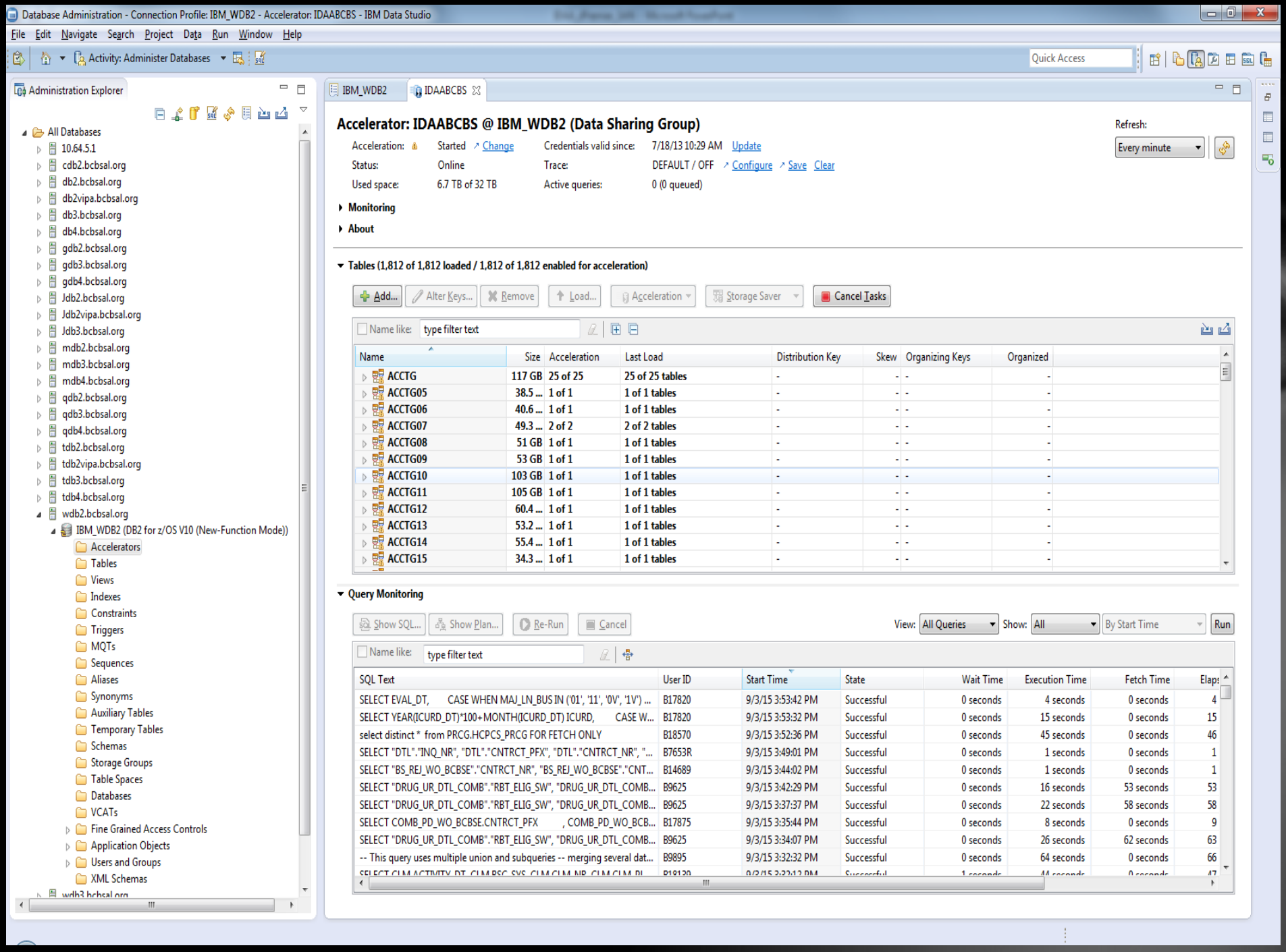
Upgrade to software version 3.1 on November 23, 2013

Upgade to software version 4.1 on April 11, 2015

Technology areas involved in implementation

Who is responsible for the "Magic"???

- Behind the scenes work performed by Technology Support Team
  - Systems Programmers
  - Database Administrators
  - Information Management Team (Warehouse)
  - Networking Team
  - Data Center Team


Database Administration - Connection Profile: IBM_WDB2 - Accelerator: IDAABCBS - IBM Data Studio
File Edit Navigate Search Project Data Run Window Help
Activity: Administer Databases
Quick Access
Administration Explorer
All Databases
10.64.5.1
cdb2.bcbsal.org
db2.bcbsal.org
db2vipa.bcbsal.org
db3.bcbsal.org
db4.bcbsal.org
gdb2.bcbsal.org
gdb3.bcbsal.org
gdb4.bcbsal.org
Jdb2.bcbsal.org
Jdb2vipa.bcbsal.org
Jdb3.bcbsal.org
mdb2.bcbsal.org
mdb3.bcbsal.org
mdb4.bcbsal.org
qdb2.bcbsal.org
qdb3.bcbsal.org
qdb4.bcbsal.org
tdb2.bcbsal.org
tdb2vipa.bcbsal.org
tdb3.bcbsal.org
tdb4.bcbsal.org
wdb2.bcbsal.org
IBM_WDB2 (DB2 for z/OS V10 (New-Function Mode))
Accelerators
Tables
Views
Indexes
Constraints
Triggers
MQTs
Sequences
Aliases
Synonyms
Auxiliary Tables
Temporary Tables
Schemas
Storage Groups
Table Spaces
Databases
VCATs
Fine Grained Access Controls
Application Objects
Users and Groups
XML Schemas
IBM_WDB2
IDAABCBS
Accelerator: IDAABCBS @ IBM_WDB2 (Data Sharing Group)
Refresh:
Every minute
Acceleration: Started Change
Credentials valid since: 7/18/13 10:29 AM Update
Status: Online
Trace: DEFAULT / OFF Configure Save Clear
Used space: 6.7 TB of 32 TB
Active queries: 0 (0 queued)
Monitoring
About
Tables (1,812 of 1,812 loaded / 1,812 of 1,812 enabled for acceleration)
Add... Alter Keys... Remove Load... Acceleration Storage Saver Cancel Tasks
Name like: type filter text
Name Size Acceleration Last Load Distribution Key Skew Organizing Keys Organized
ACCTG 117 GB 25 of 25 25 of 25 tables
ACCTG05 38.5 ... 1 of 1 1 of 1 tables
ACCTG06 40.6 ... 1 of 1 1 of 1 tables
ACCTG07 49.3 ... 2 of 2 2 of 2 tables
ACCTG08 51 GB 1 of 1 1 of 1 tables
ACCTG09 53 GB 1 of 1 1 of 1 tables
ACCTG10 103 GB 1 of 1 1 of 1 tables
ACCTG11 105 GB 1 of 1 1 of 1 tables
ACCTG12 60.4 ... 1 of 1 1 of 1 tables
ACCTG13 53.2 ... 1 of 1 1 of 1 tables
ACCTG14 55.4 ... 1 of 1 1 of 1 tables
ACCTG15 34.3 ... 1 of 1 1 of 1 tables
Query Monitoring
Show SQL... Show Plan... Re-Run Cancel
View: All Queries Show: All By Start Time Run
Name like: type filter text
SQL Text User ID Start Time State Wait Time Execution Time Fetch Time Elaps
SELECT EVAL_DT, CASE WHEN MAJ_LN_BUS IN ('01', '11', '0V', '1V') ... B17820 9/3/15 3:53:42 PM Successful 0 seconds 4 seconds 0 seconds 4
SELECT YEAR(ICURD_DT)*100+MONTH(ICURD_DT) ICURD, CASE W... B17820 9/3/15 3:53:32 PM Successful 0 seconds 15 seconds 0 seconds 15
select distinct * from PRCG.HCPCS_PRCG FOR FETCH ONLY B18570 9/3/15 3:52:36 PM Successful 0 seconds 45 seconds 0 seconds 46
SELECT "DTL"."INQ_NR", "DTL"."CNTRCT_PFX", "DTL"."CNTRCT_NR", "... B7653R 9/3/15 3:49:01 PM Successful 0 seconds 1 seconds 0 seconds 1
SELECT "BS_REJ_WO_BCBSE"."CNTRCT_NR", "BS_REJ_WO_BCBSE"."CNT... B14689 9/3/15 3:44:02 PM Successful 0 seconds 1 seconds 0 seconds 1
SELECT "DRUG_UR_DTL_COMB"."RBT_ELIG_SW", "DRUG_UR_DTL_COMB... B9625 9/3/15 3:42:29 PM Successful 0 seconds 16 seconds 53 seconds 53
SELECT "DRUG_UR_DTL_COMB"."RBT_ELIG_SW", "DRUG_UR_DTL_COMB... B9625 9/3/15 3:37:37 PM Successful 0 seconds 22 seconds 58 seconds 58
SELECT COMB_PD_WO_BCBSE.CNTRCT_PFX , COMB_PD_WO_BCB... B17875 9/3/15 3:35:44 PM Successful 0 seconds 8 seconds 0 seconds 9
SELECT "DRUG_UR_DTL_COMB"."RBT_ELIG_SW", "DRUG_UR_DTL_COMB... B9625 9/3/15 3:34:07 PM Successful 0 seconds 26 seconds 62 seconds 63
-- This query uses multiple union and subqueries -- merging several dat... B9895 9/3/15 3:32:32 PM Successful 0 seconds 64 seconds 0 seconds 66

Healthcare Analytics Example

Report on each customer group detailing age band and gender of insured (contract holder and dependents), relationship to contract holder and zip code.

Report returns 72,000,000 rows.

Before IDAA: Information returned with multiple queries and combined for final report. (One Week)

Healthcare Analytics Example (continued)

After IDAA: Single query returns 72,000,000. (20 Minutes)

Pros and Cons of IDAA

+++++++

No special security required

Transparent to customer

All dynamic sql can be accelerated

“Train of Thought” Analysis

Low maintenance

“Black Box” difficult to troubleshoot problem queries

Few tools available

In Closing…

Business customers now consider the “Magic Box” an essential tool for retrieving data for analysis!

Thank You!

sdavis@bcbsal.org