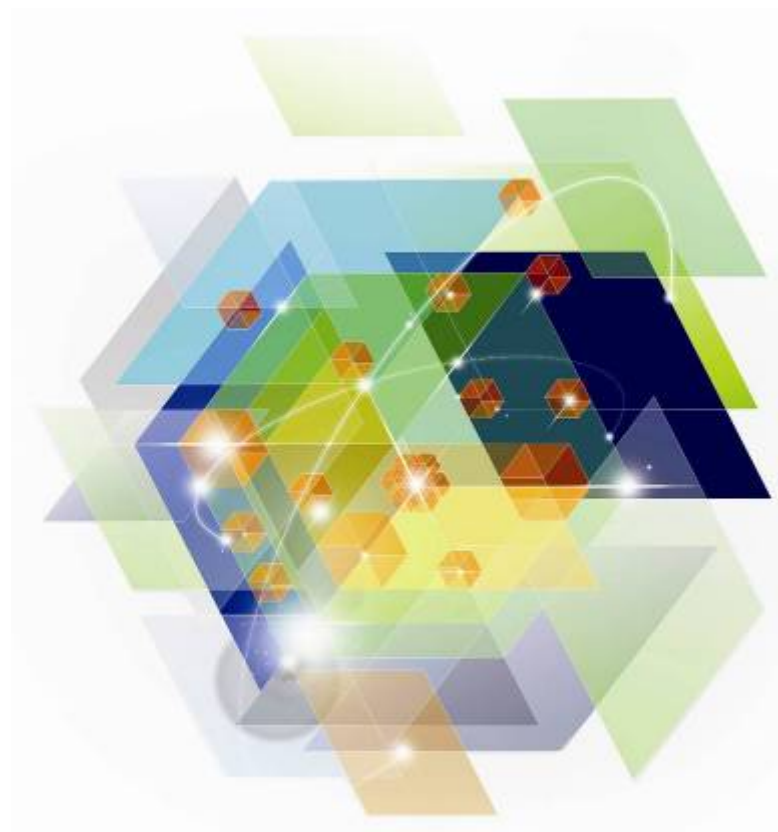


Extreme Performance for Complex Business Analysis

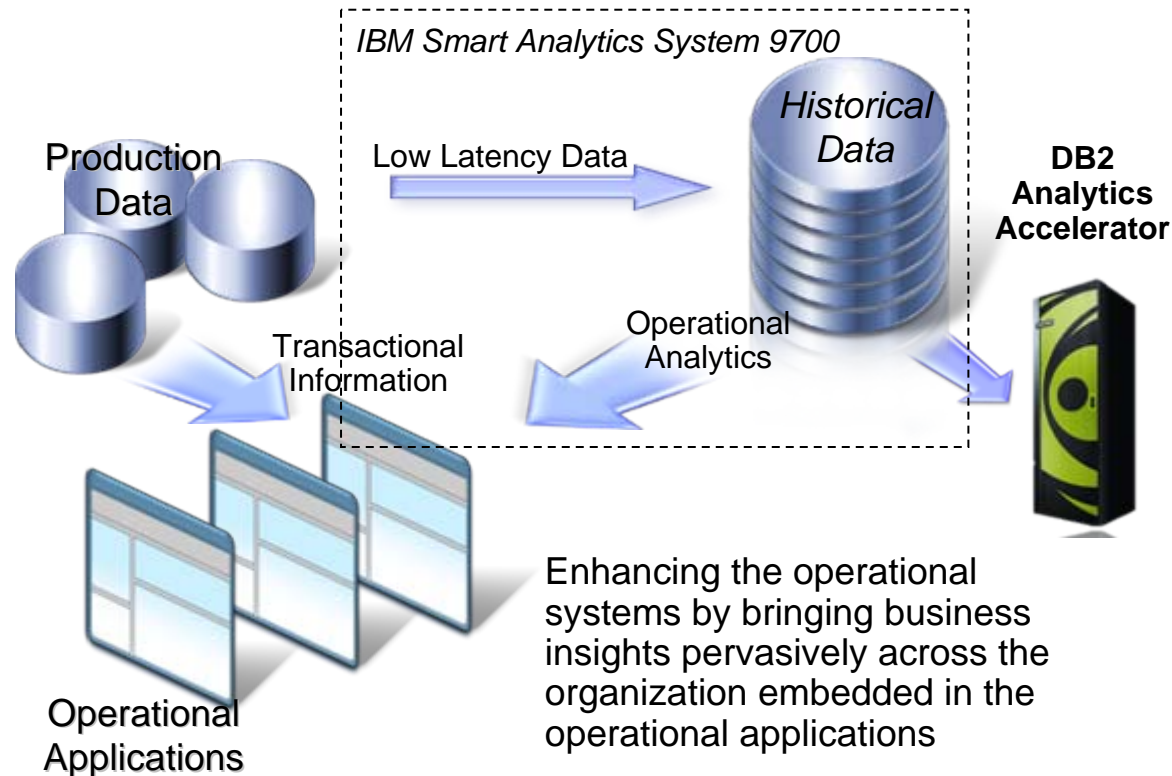
An in-depth look at the

*IBM DB2 Analytics Accelerator
powered by Netezza Technology*



Modernize Your Decision Systems

Reengineering your information infrastructure



- Facilitate and manage transaction-oriented applications with analytics
- Access to customer purchase histories, customer behaviors and real time sales trends
- Sift through massive amounts of data and make the information relevant and actionable almost immediately

Time is money

“The more data we have, the longer our analysis takes!”

- **Waiting for fact based information to drive key decisions**
- **Waiting for key reports to complete**
- **Paying analysts to perform query analysis**
- **Adding indices and MQTs in an attempt to speed up queries**
- **Consuming MIPS for long running queries**



DB2 Analytics Accelerator

Accelerating decisions to the speed of business

Blending System z and Netezza technologies to deliver unparalleled, mixed workload performance for complex analytic business needs.



Get more insight from your data

- Fast, predictable response times for “right-time” analysis
- Accelerate analytic query response times
- Improve price/performance for analytic workloads
- Minimize the need to create data marts for performance
- Highly secure environment for sensitive data analysis
- Transparent to the application

Fast Time to Value

- **IBM DB2 Analytics Accelerator (Netezza 1000-12)**
 - ➔ Production ready - 1 person, 2 days
- **Table Acceleration Setup ... 2 Hours**
 - DB2 “Add Accelerator”
 - Choose a Table for “Acceleration”
 - Load the Table (DB2 copy to Netezza)
 - Knowledge Transfer
 - Query Comparisons
- **Initial Load Performance ...**
 - ➔ 400 GB “Loaded” in 29 Min
 - 570 million rows (Loads of 800GB to 1.3TB/Hr)
- **Actual Query Acceleration ... 1908x faster**
 - ➔ 2 Hours 39 Minutes to 5 Seconds
- **CPU Utilization Reduction**
 - ➔ 35% to ~0%



Performance & Savings

Query	Total Rows		DB2 Only		DB2 with IDAA		Times Faster
	Reviewed	Returned	Hours	Sec(s)	Hours	Sec(s)	
Query 1	2,813,571	853,320	2:39	9,540	0.0	5	1,908
Query 2	2,813,571	585,780	2:16	8,220	0.0	5	1,644
Query 3	8,260,214	274	1:16	4,560	0.0	6	760
Query 4	2,813,571	601,197	1:08	4,080	0.0	5	816
Query 5	3,422,765	508	0:57	4,080	0.0	70	58
Query 6	4,290,648	165	0:53	3,180	0.0	6	530
Query 7	361,521	58,236	0:51	3,120	0.0	4	780
Query 8	3,425.29	724	0:44	2,640	0.0	2	1,320
Query 9	4,130,107	137	0:42	2,520	0.1	193	13

Queries run faster

- Save CPU resources
- People time
- Business opportunities

Actual customer results, October 2011

DB2 Analytics Accelerator: “we had this up and running in days with queries that ran over 1000 times faster”

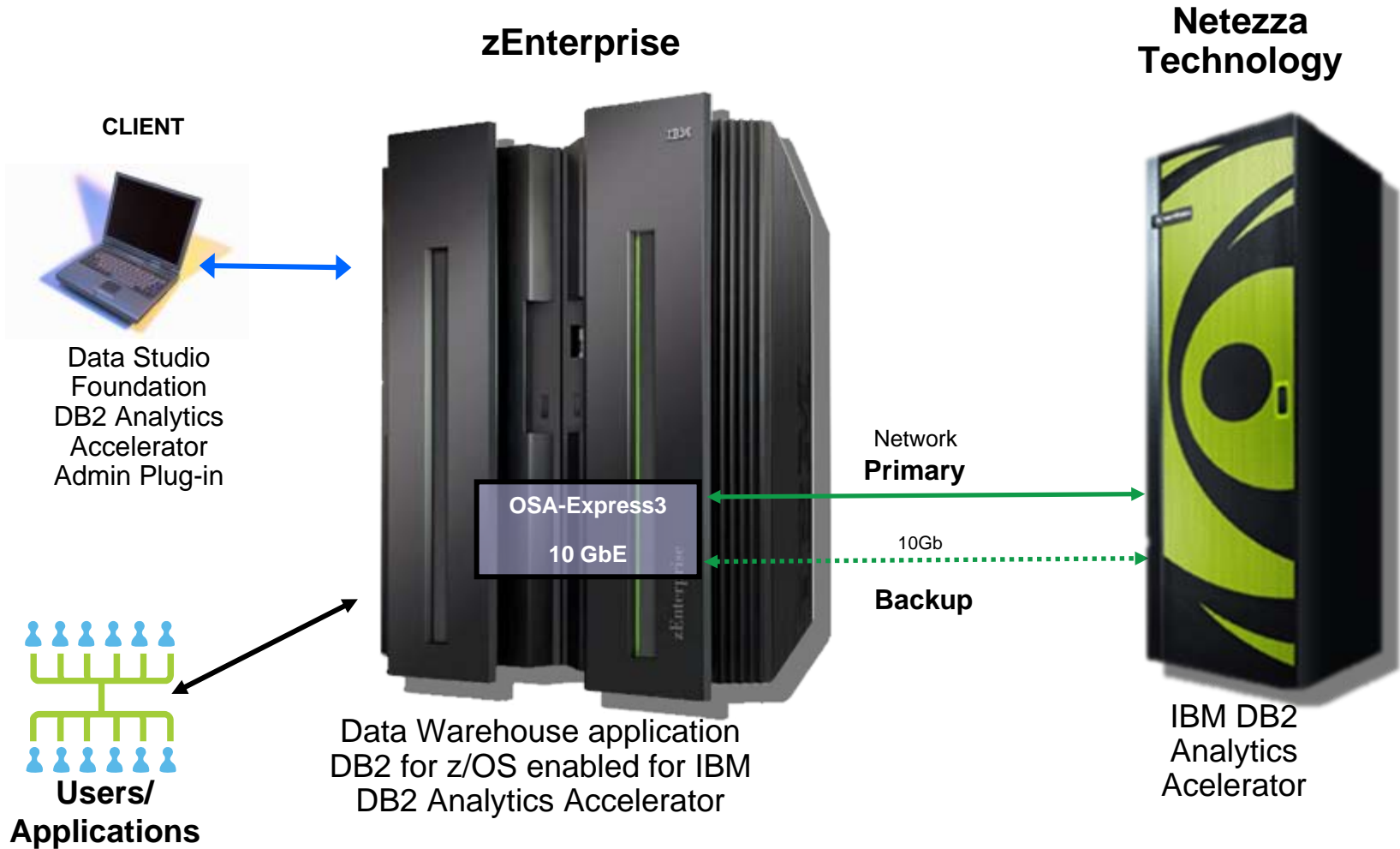
DB2 Analytics Accelerator: “we expect ROI in less than 4 months”



Advance to 32 minute mark for DB2 Analytics Accelerator section of keynote

Accelerating decisions to the speed of business

IBM DB2 Analytics Accelerator V2 Product Components



Deep DB2 Integration within zEnterprise


Applications
Application Interfaces
 (standard SQL dialects)

DBA Tools, z/OS Console, ...
Operational Interfaces
 (e.g. DB2 Commands)

DB2 for z/OS


Data Manager Buffer Manager ... IRLM Log Manager

Superior availability, reliability, security, Workload management



z/OS on System z

IBM DB2 Analytics Accelerator



Netezza

Superior performance on analytic queries

DB2 Analytics Accelerator V2

Powered by Netezza 1000 Appliance



Disk Enclosures

SMP Hosts

Snippet Blades™
(S-Blades, SPUs)

Slice of User Data

Swap and Mirror partitions

High speed data streaming

High compression rate

EXP3000 JBOD Enclosures

*12 x 3.5" 1TB, 7200RPM, SAS (3Gb/s)
max 116MB/s (200-500MB/s compressed data)*

e.g. TF12:

8 enclosures → 96 HDDs

32TB uncompressed user data (→ 128TB)

IDAA Server

*SQL Compiler, Query Plan, Optimize
Administration*

2 front/end hosts, IBM 3650M3

clustered active-passive

2 Nehalem-EP Quad-core 2.4GHz per host

Processor &

streaming DB logic

High-performance database

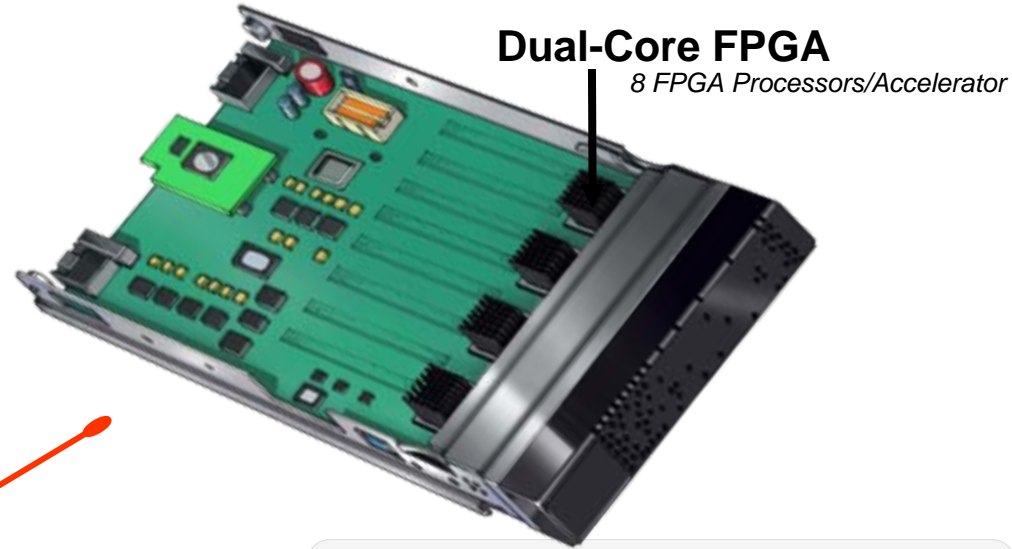
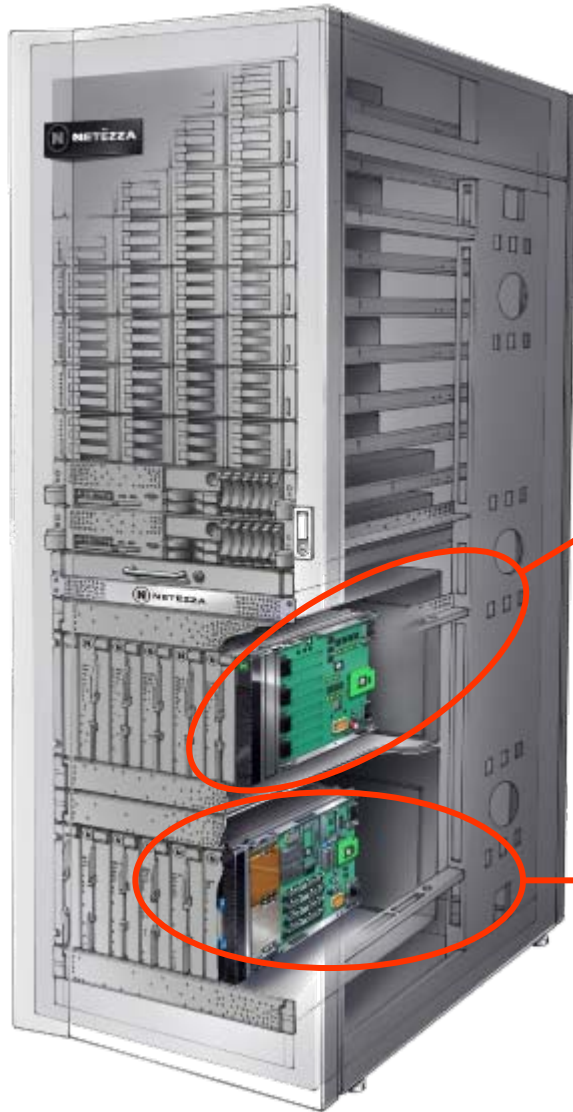
engine streaming joins,

aggregations, sorts, etc.

e.g. TF12: 12 back/end SPUs

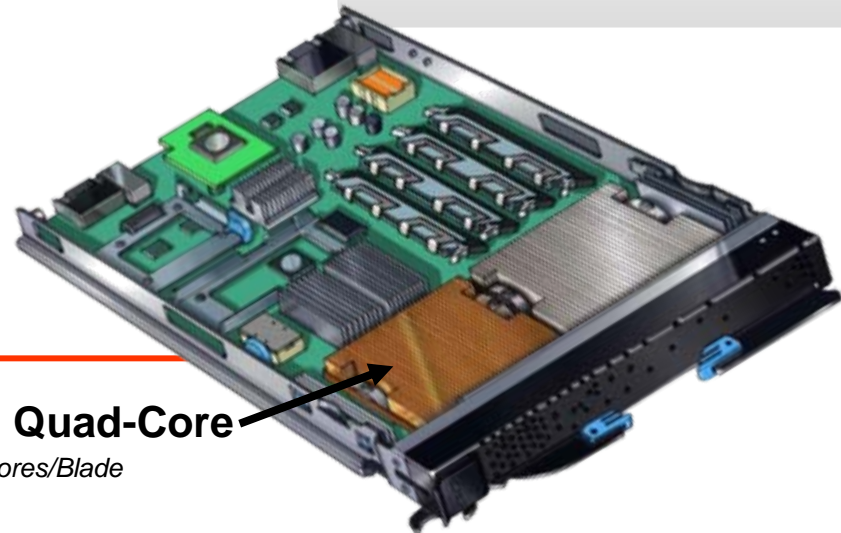
(more details on following charts)

S-Blade™ Components



Dual-Core FPGA
8 FPGA Processors/Accelerator

Netezza DB Accelerator

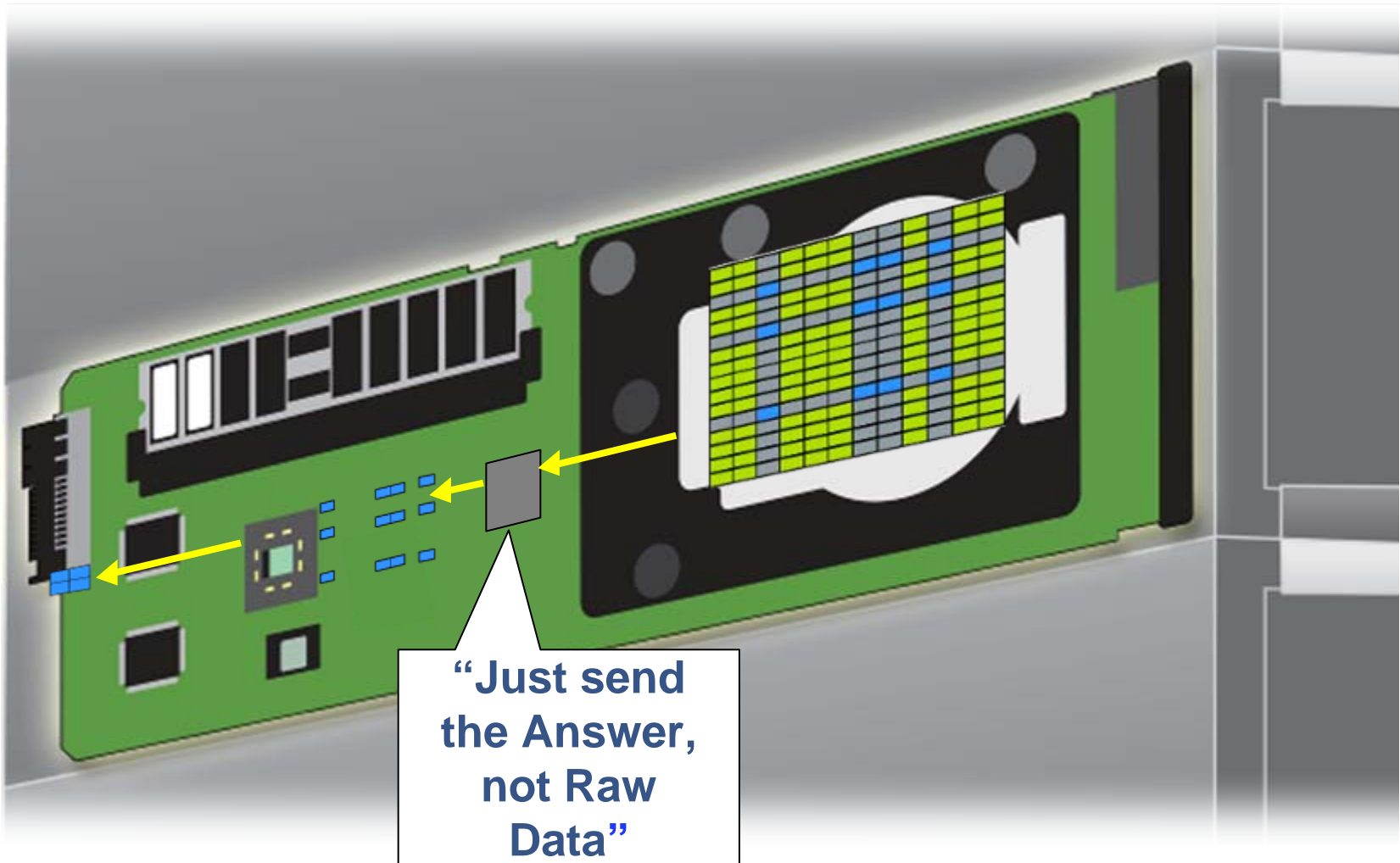


Intel Quad-Core
8 Cores/Blade

IBM BladeCenter Server

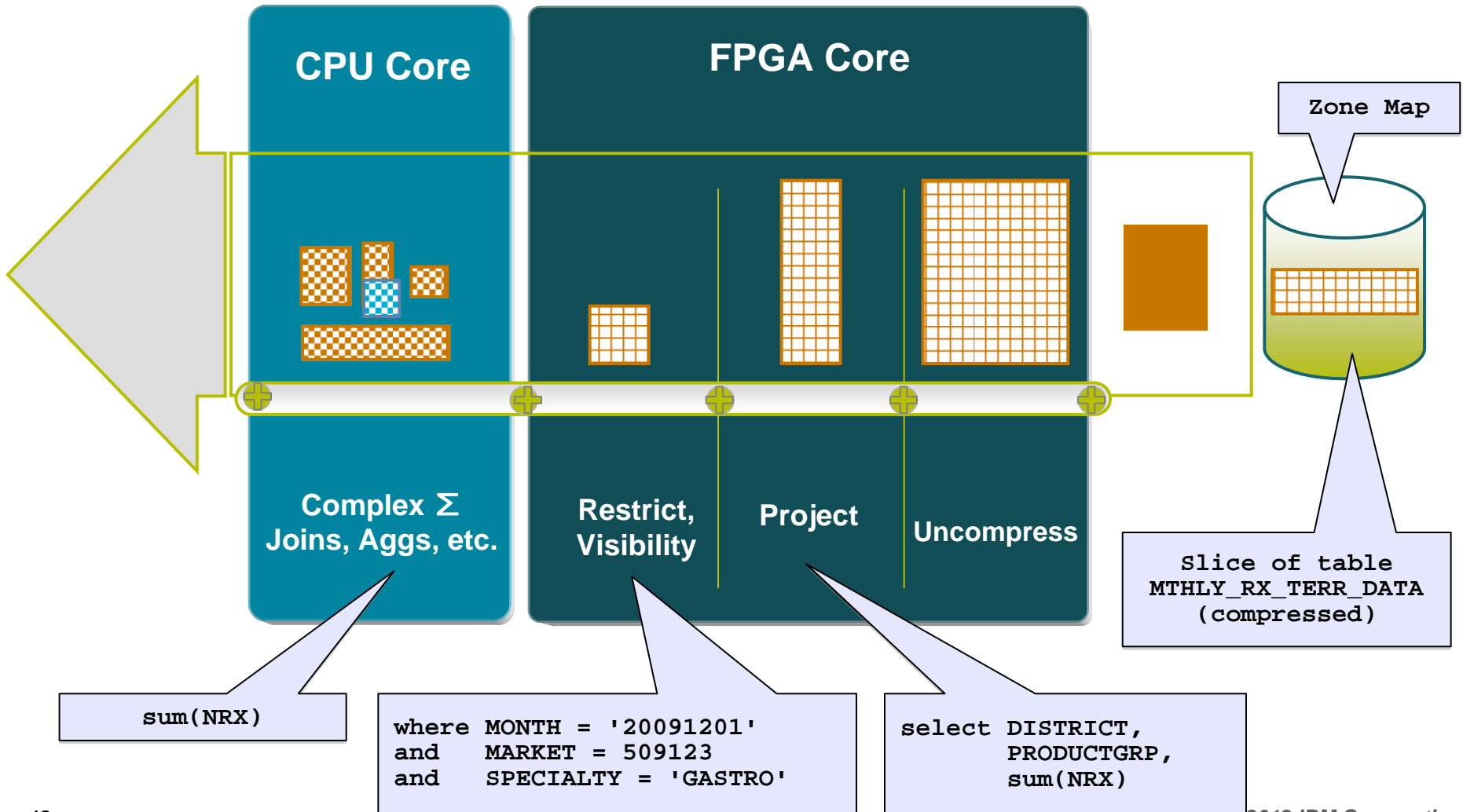
Eliminating the I/O Bottleneck

Move the SQL to the hardware... to where the data lives



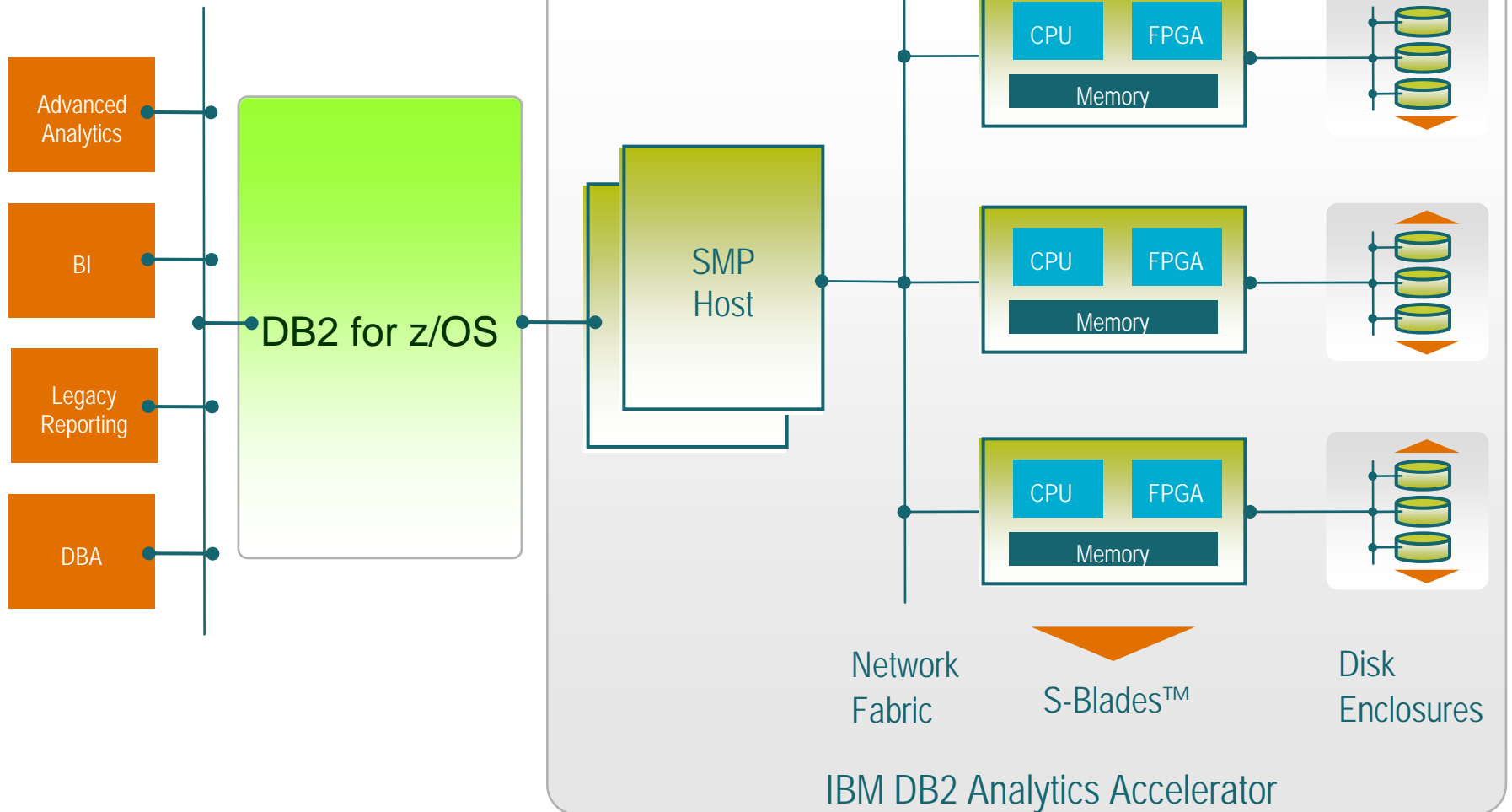
The Key to the Speed

```
select DISTRICT,PRODUCTGRP, sum(NRX)
from MTHLY_RX_TERR_DATA
where MONTH = '20091201'
and MARKET = 509123
and SPECIALTY = 'GASTRO'
```

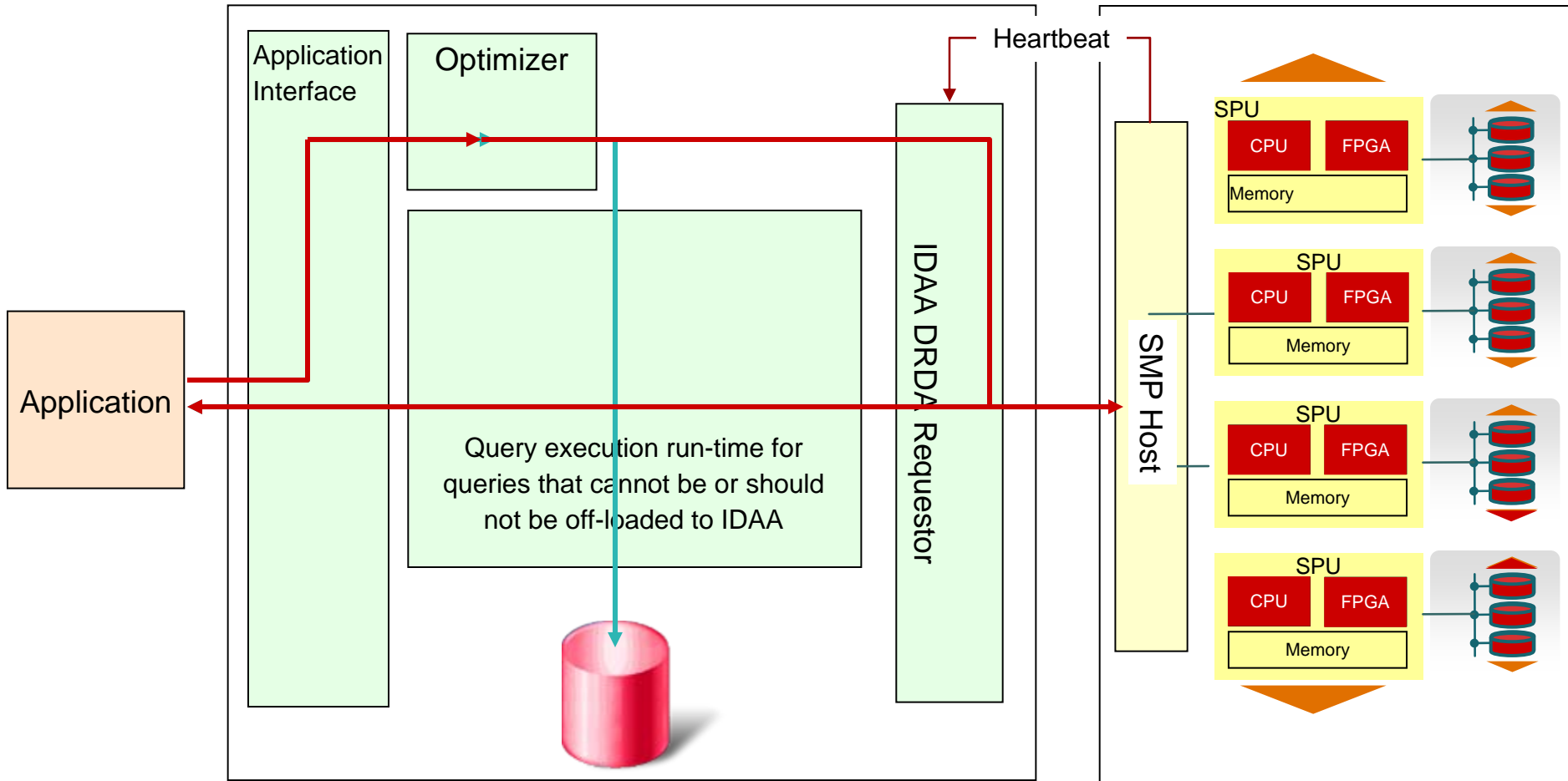


Bringing Netezza AMPP™ Architecture to DB2 for z/OS

AMPP = Asymmetric Massively Parallel Processing



Query Execution Process Flow



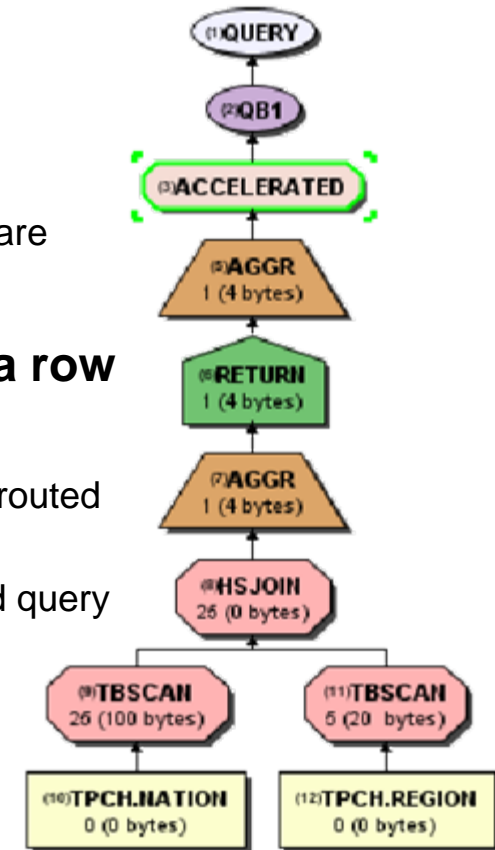
DB2 for z/OS

IDAA

- Queries executed without IDAA
- Queries executed with IDAA
- Heartbeat (IDAA availability and performance indicators)

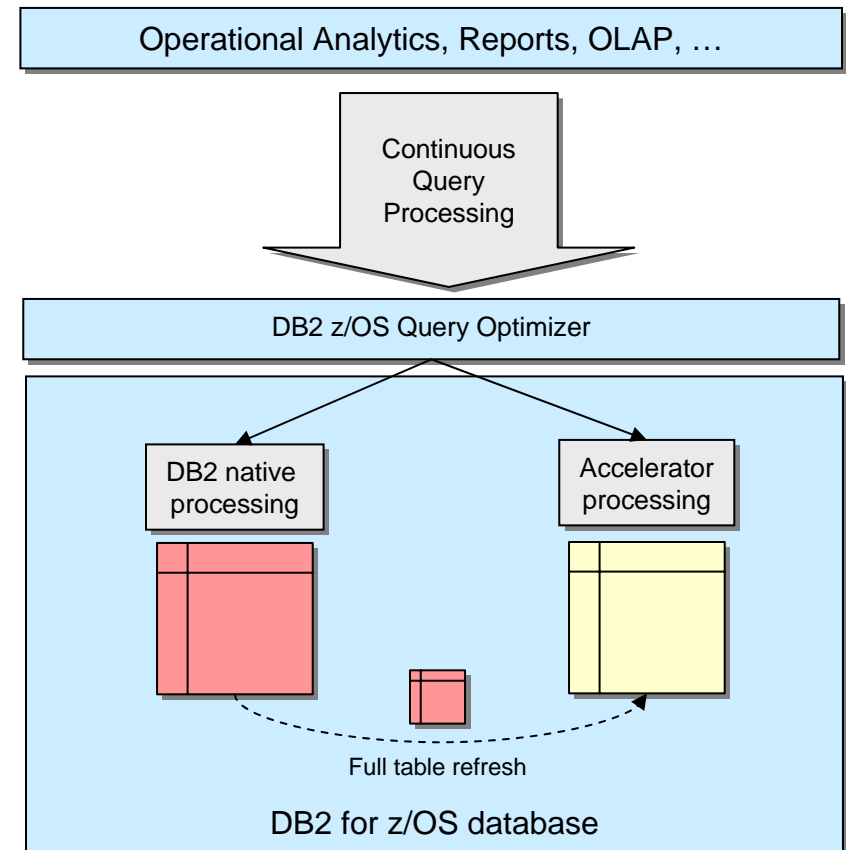
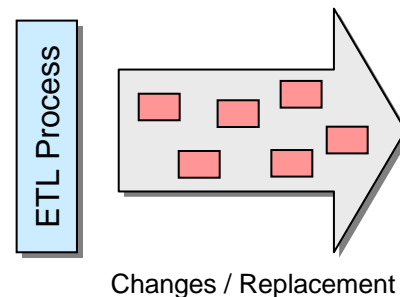
EXPLAIN

- **DB2 EXPLAIN function is enhanced to provide basic information about accelerator usage**
 - Whether query qualifies for acceleration and, if not, why
 - The access path details associated with the query execution by Netezza are provided independently of DB2 EXPLAIN by the IDAA Studio.
- **For each query (irrespective of the number of query blocks) a row is inserted in the following tables:**
 - in both PLAN_TABLE and DSN_QUERYINFO_TABLE, if the query is re-routed
 - PLAN_TABLE's ACESSTYPE column is set to a value of 'A'
 - DSN_QUERYINFO_TABLE's QI_DATA column shows the converted query text
 - in DSN_QUERYINFO_TABLE only, if the query is not re-routed
 - REASON_CODE and QI_DATA columns provide details
- **Note that the EXPLAIN tables can be populated with above described information even if there is no accelerator connected to DB2**
 - Specifying EXPLAINONLY on START ACCEL command does not establish any communications with an actual accelerator, but enables DB2 to consider its presence in the access path selection process



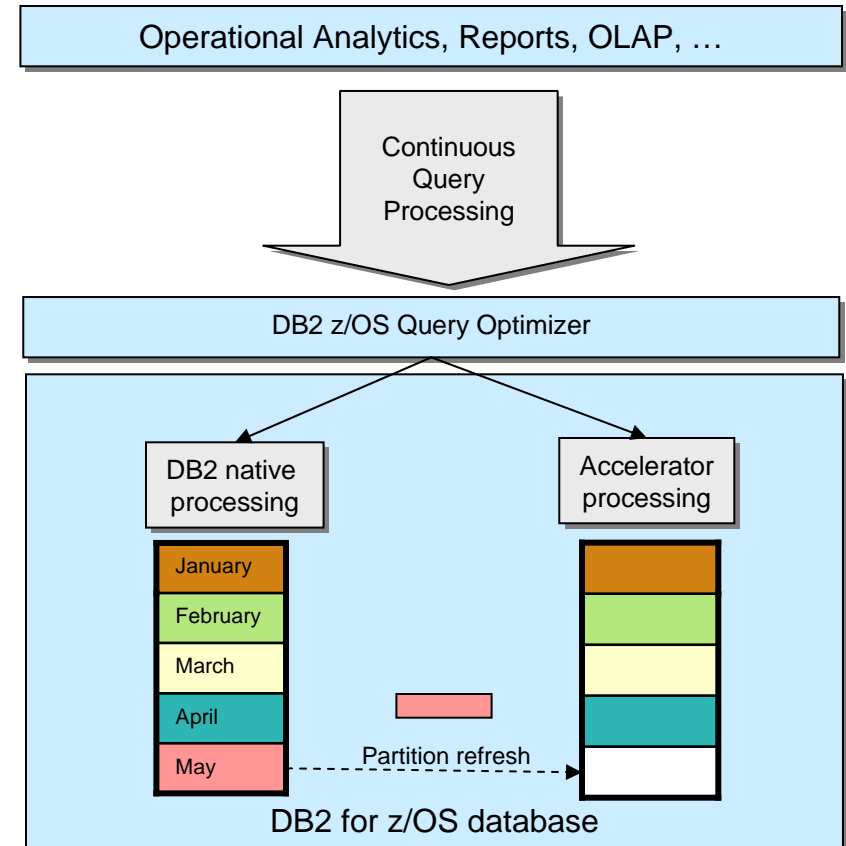
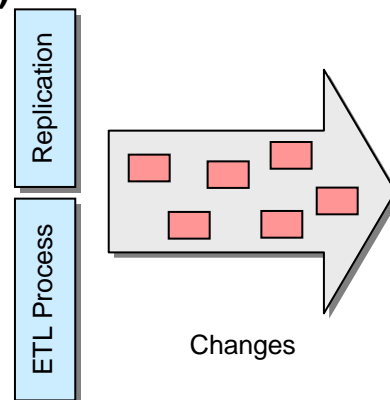
Option 1: Full Table Refresh

- Changes in data warehouse tables typically driven by scheduled (nightly or more frequently) ETL process
- Data used for complex reporting based on consistent and validated content (e.g., weekly transaction reporting to the central bank)
- Multiple sources or complex transformations prevent from propagation of incremental changes
- Queries may continue during full table refresh for accelerator
- Full table refresh may be triggered through DB2 stored procedure (scheduled, integrated into ETL process or through GUI)



Option 2: Table Partition Refresh

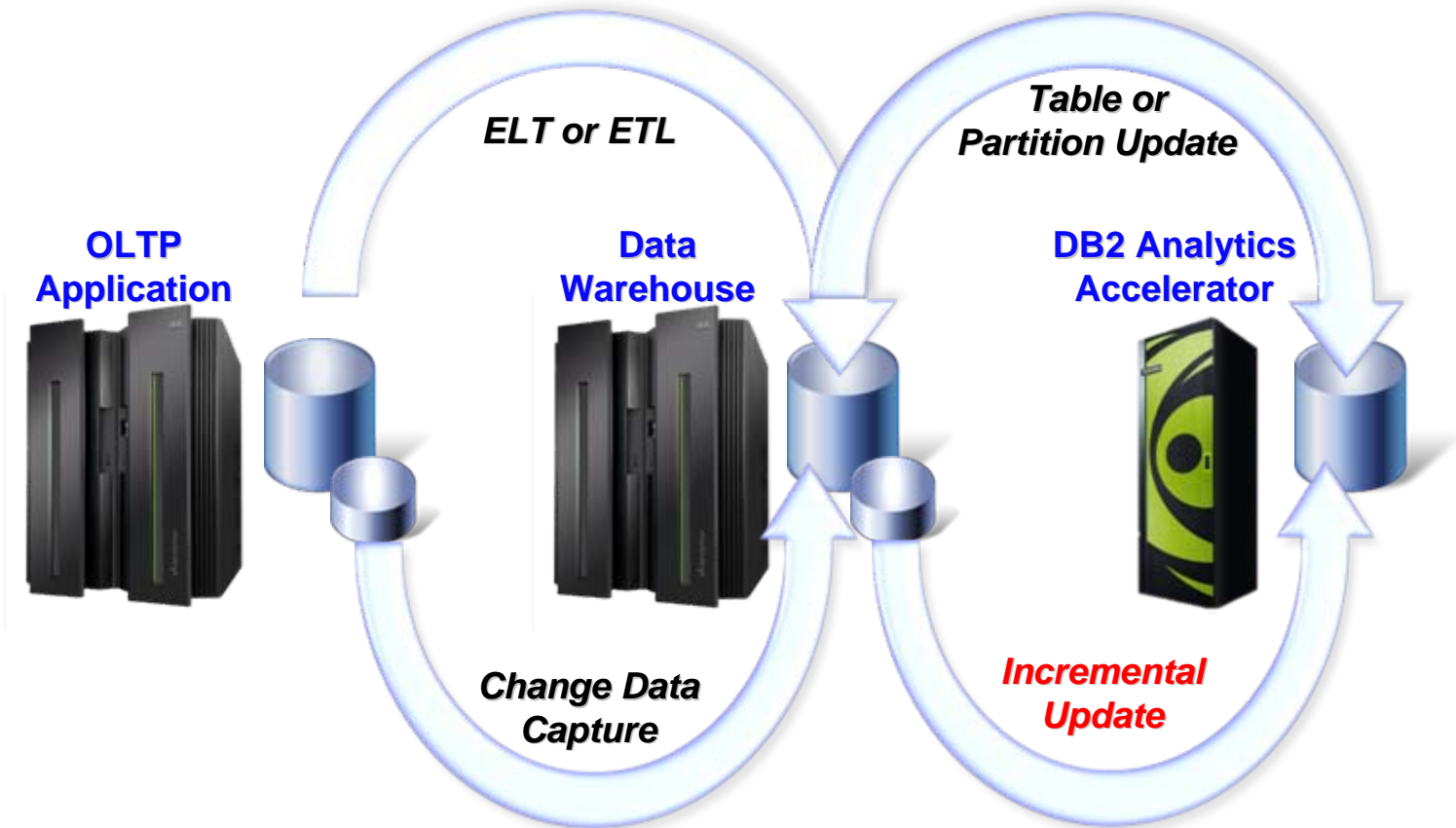
- Changes in data warehouse table typically driven by “delta” ETL process (considering only changes in source tables compared to previous runs) or by more frequent changes to most recent data
- Optimization of Option 1 when target data warehouse table is partitioned and most recent updates are only applied to the latest partition
- Maintains snapshot semantics for consistent reports
- Queries may continue during table partition refresh for accelerator
- Table partition refresh may be triggered through DB2 stored procedure (scheduled, integrated into ETL process or through GUI)



Incremental Update

- An alternative to a full table load or table partition load. Refreshes only the records of the table that have been recently modified in the data warehouse.
 - This capability keeps the data on the DB2 Analytics Accelerator in sync with the data on the mainframe DB2.
 - This is an initial release with the full release of incremental update being made available in the next major release of the offering.
 - Incremental Update is a capability most customers want, and therefore is being added to the offering for all customers, and *not* a separately orderable feature

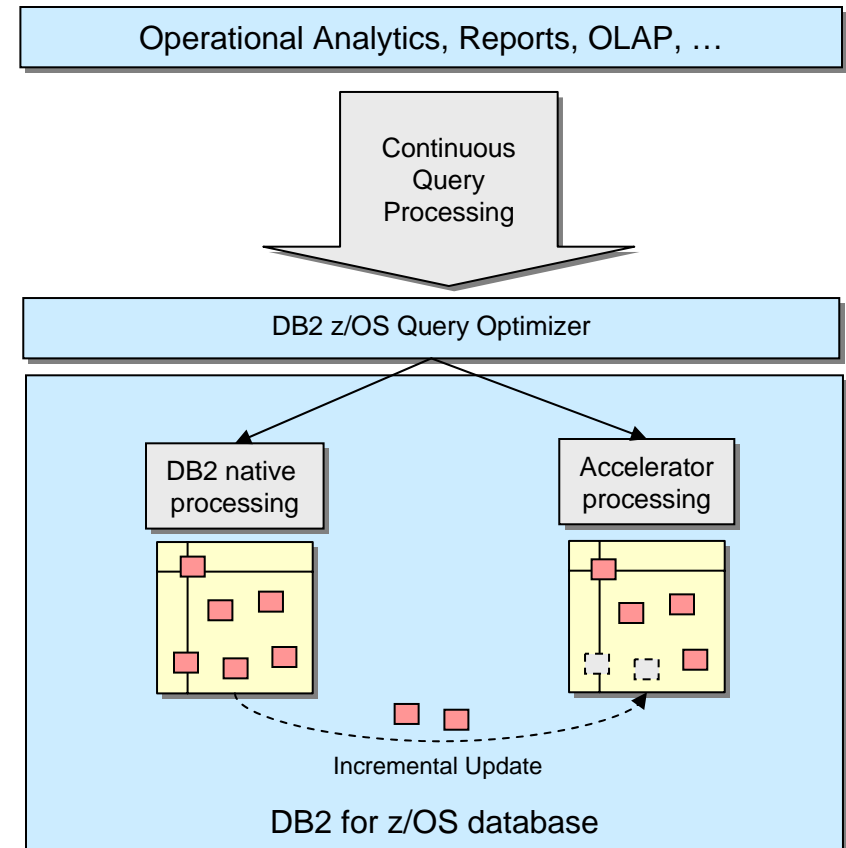
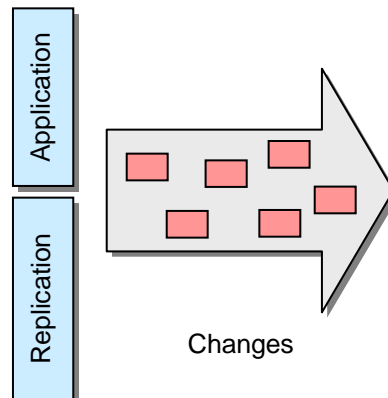
Introducing Incremental Update



Synchronizing data to lower data latency from days to minutes/seconds

Option 3: Incremental Update (Controlled Availability)

- **Changes in data warehouse tables typically driven by replication or manual updates**
 - Corrections after a bulk-ETL-load of a data warehouse table
 - Continuously changing data (e.g. trickle-feed updates from a transactional system to an ODS)
- **Reporting and analysis based on most recent data**
- **May be combined with Option 1 & 2 (first table refresh and then continue with incremental updates)**
- **Incremental update can be configured per database table**



Now expandable to 960 cores and 1.28 petabytes



	TF3	TF6	TF12	TF24	TF36	TF48	TF72	TF96	TF120
Cabinets	1/4	1/2	1	2	3	4	6	8	10
Processing Units	24	48	96	192	288	384	576	768	960
Capacity (TB)	8	16	32	64	96	128	192	256	320
Effective Capacity (TB)*	32	64	128	256	384	512	768	1024	1280

Accelerator Platforms

Predictable, Linear Scalability throughout entire family

Capacity = User Data space
 Effective Capacity = User Data Space with compression

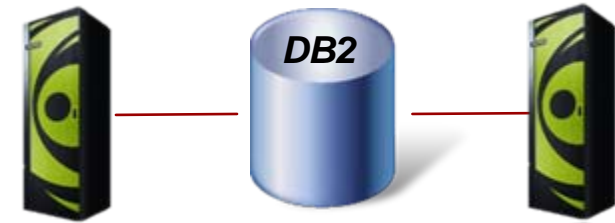
*: 4X compression assumed

Connectivity Options

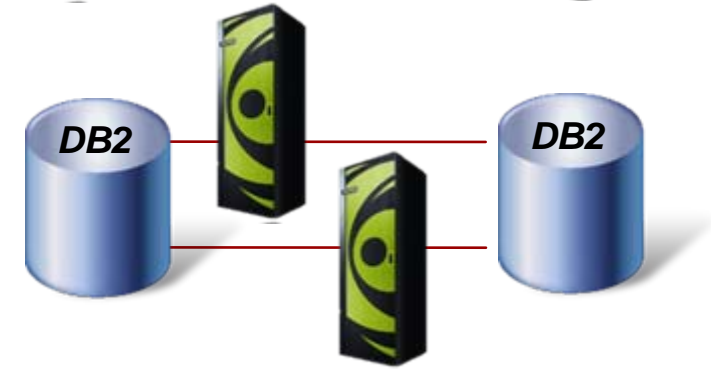
Multiple DB2 systems can connect to a single IDAA



A single DB2 system can connect to multiple IDAAs



Multiple DB2 systems can connect to multiple IDAAs

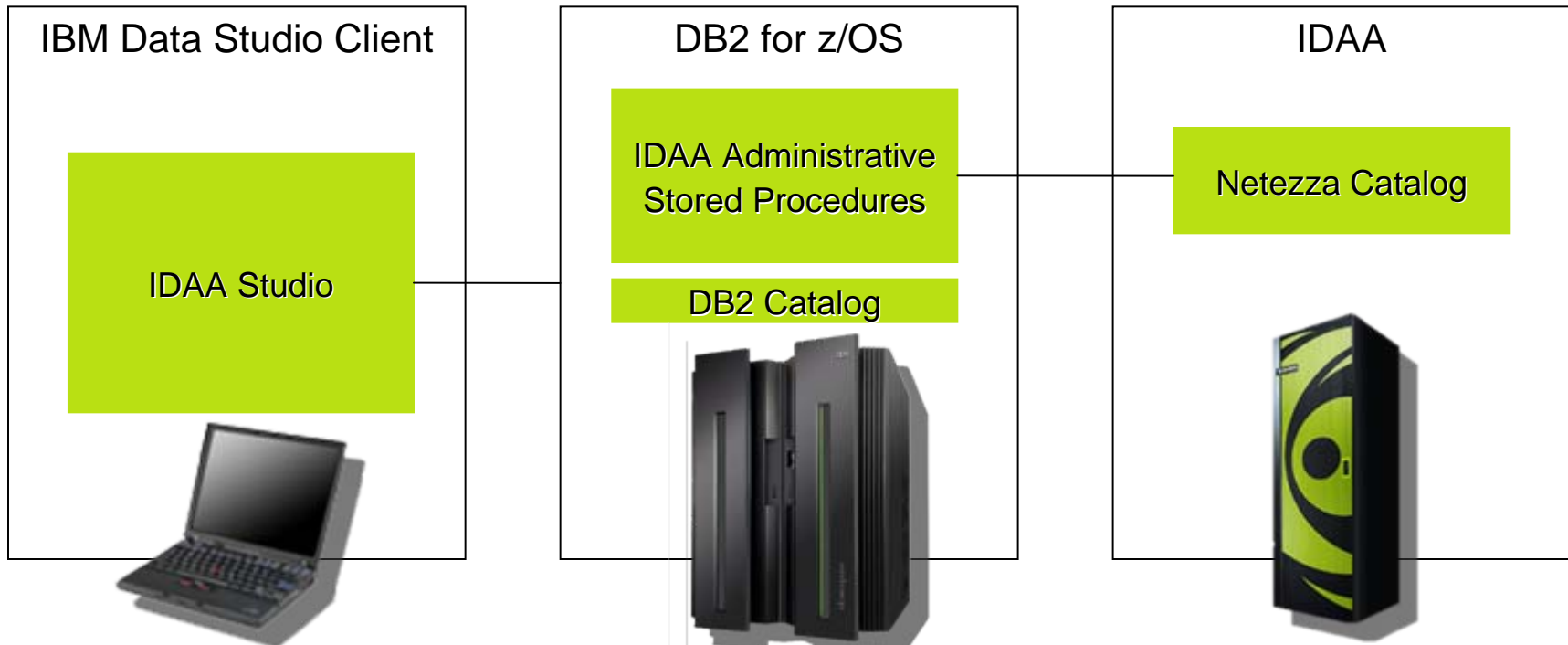


Better utilization of IDAA resources
Scalability
High availability

Full flexibility for DB2 systems:

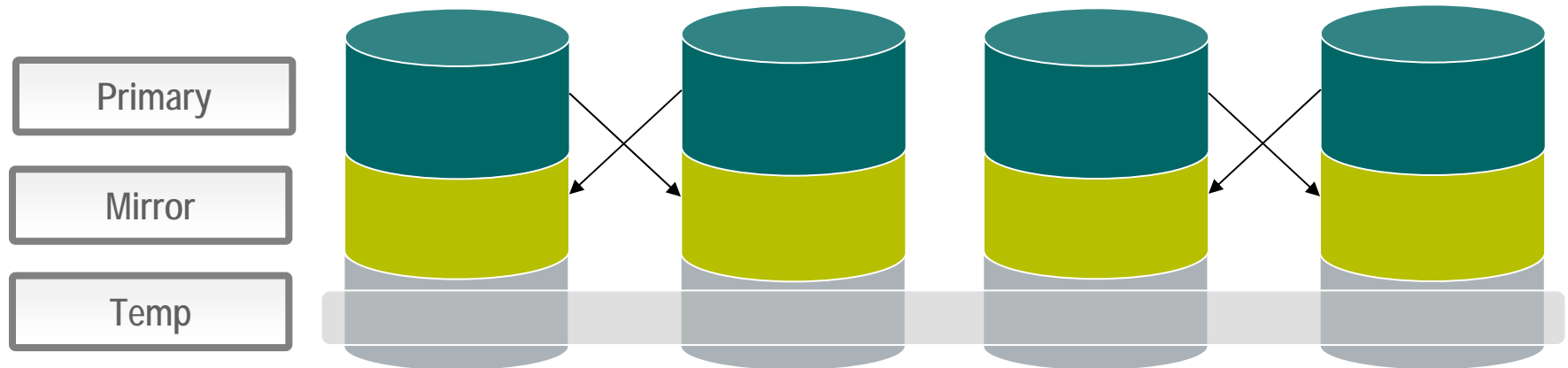
- residing in the same LPAR
- residing in different LPARs
- residing in different CECs
- being independent (non-data sharing)
- belonging to the same data sharing group
- belonging to different data sharing groups

Analytics Accelerator Table Definition and Deployment



- The tables need to be defined and deployed to IDAA before data is loaded and queries sent to it for processing.
 - Definition: identifying tables for which queries need to be accelerated
 - Deployment: making tables known to DB2, i.e. storing table meta data in the DB2 and Netezza catalog.
- IBM DB2 Analytics Accelerator Studio guides you through the process of defining and deploying tables, as well as invoking other administrative tasks.
- IBM DB2 Analytics Accelerator Stored Procedures implement and execute various administrative operations such as table deployment, load and update, and serve as the primary administrative interface to IDAA from the outside world including IDAA Studio.

Shielding Against Disk Failures



- All user data and temp space mirrored
- Disk failures transparent to queries and transactions
- Failed drives automatically regenerated
- Bad sectors automatically rewritten or relocated

Why Both?

Marrying the best of both worlds

IBM Netezza



Focused Appliance

IBM System z



Mixed Workload System

Capitalizing on the strengths of both platforms while driving to the most cost effective, centralized solution - destroying the myth that transaction and decision systems had to be on separate platforms

Very focused workload

Very diverse workload

Tailored to your needs

A Hybrid Solution

IBM Netezza

IBM System z with IBM DB2 Analytics Accelerator

Focused Appliance

Mixed Workload System

- Appliance with a streamlined database and HW acceleration for performance critical functionality
- Price/performance leader
- Speed and ease of deployment and administration
- Optimized performance for deep analytics, multifaceted, reporting and complex queries

- Mixed workload system z with operational transaction systems, data warehouse, operational data store, and consolidated data marts.
- Unmatched availability, security and recoverability
- Natural extension to System z to enable pervasive analytics across the organization.
- Speed and ease of deployment and administration

Simplicity

The right mix of simplicity and flexibility

Flexibility

Accelerator use case:

Traditional BI



Benefits

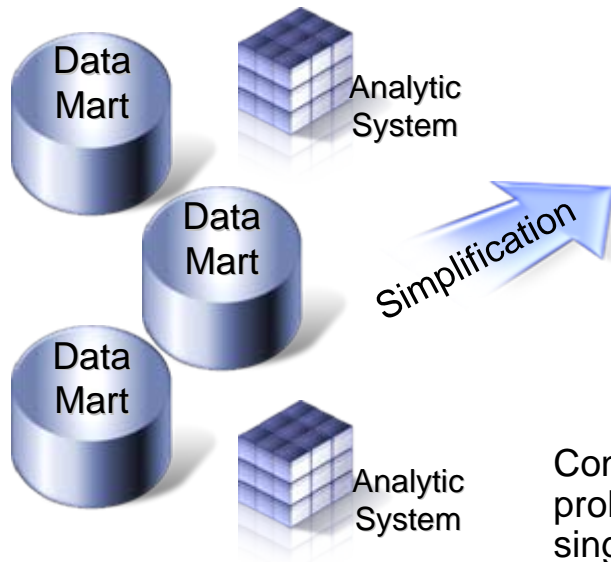
- Reduce the complexity of multi platform, multi product environments
- Realize significant increases in availability, security, recoverability, and virtualization
- Work with the latest data instead of out of date copies
- Accelerate long running DB2 for z/OS queries from minutes to seconds for greater business value
- Avoid costs and efforts to tune individual queries
- The forgotten query: consider queries previously set aside due to performance challenges?

Accelerator use case:

Data Mart consolidation

Utilize virtualization to optimize the use of resources while reducing costs and gaining new agility

A single platform to manage and administer



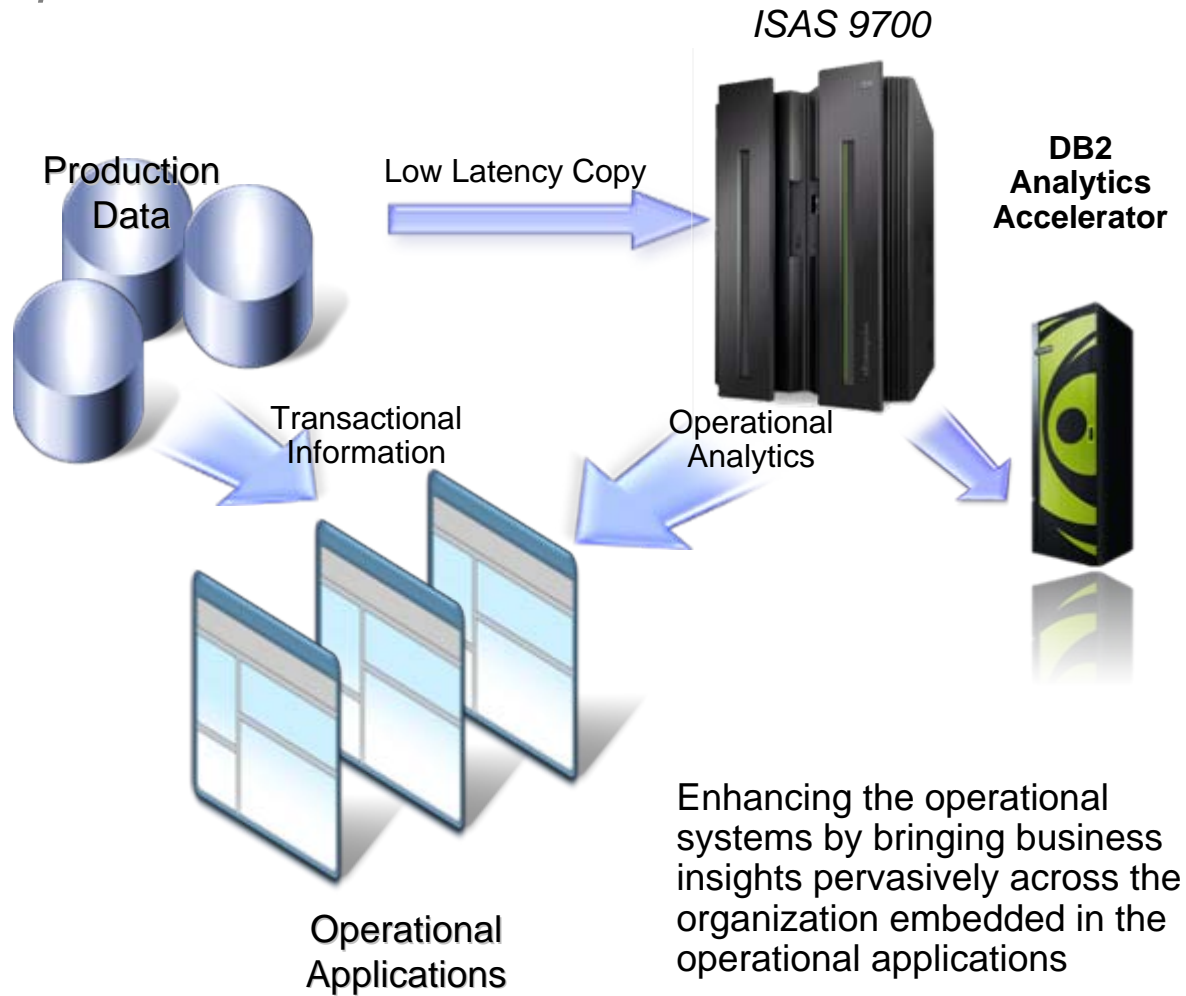
Consolidate the ever growing proliferation of data marts onto a single, easily managed platform

Benefits

- Consolidate isolated islands of data on one secure Hub
- Simplify management of costly/complex Data Marts while retaining the isolation benefits of individual platforms
- Provide consistency of sources of information
- Increase time to value to deploy new Data Marts
- Enable application queries which would you prefer to run with more real-time data on System z

Accelerator use case:

Operational BI



Benefits

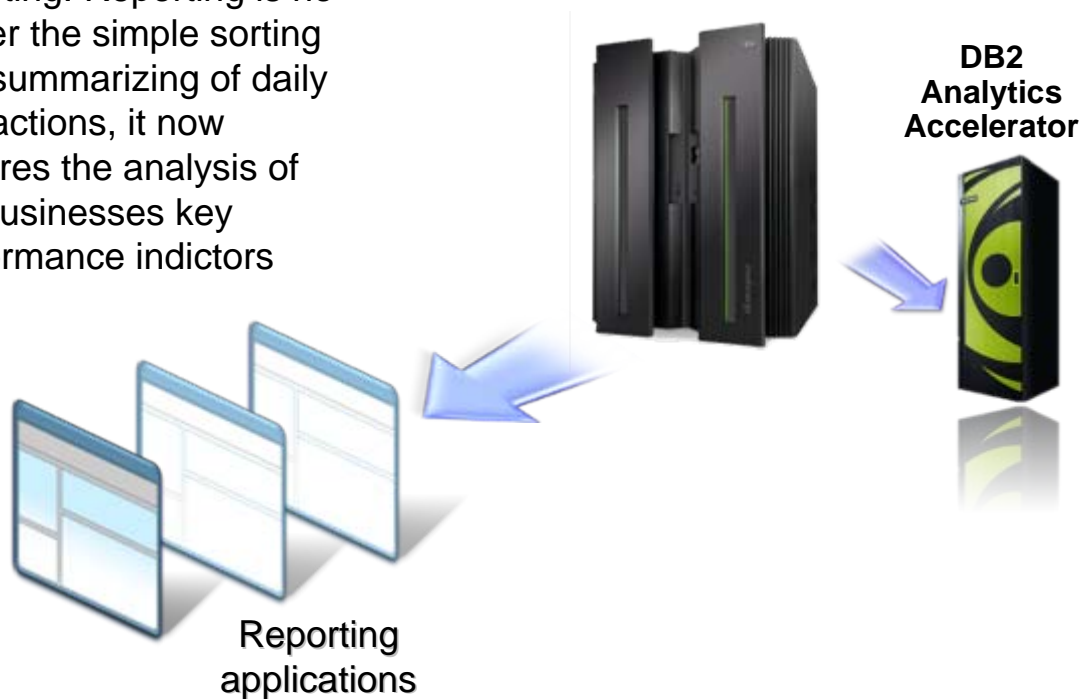
- Provide analytic information at the point of decision enabling fact-based decisions
- Deliver new insights that help business users in an operational application
- Pervasively enable decision makers and other end users across the organization
- Enable customers with historical information to increase loyalty and sales
- Accelerate long running DB2 for z/OS queries from minutes to seconds for greater business value with IDAA

Accelerator use case:

Operational Reporting

Dramatically improve complex operational reporting. Reporting is no longer the simple sorting and summarizing of daily interactions, it now requires the analysis of the businesses key performance indicators

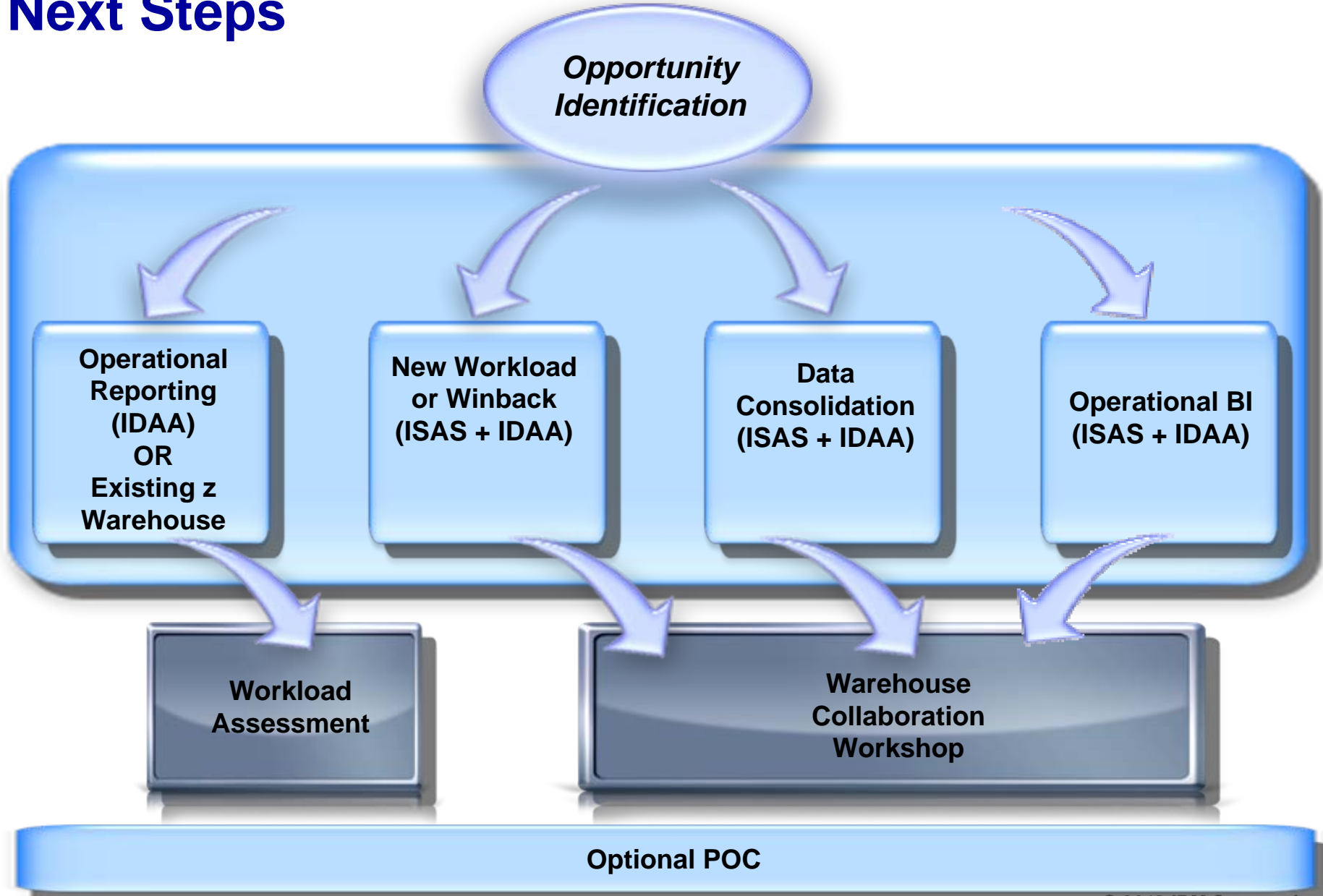
Keep new reporting capabilities onto a single, easily managed platform



Benefits

- Time and agility gained through more timely delivery of complex information to the business
- Consolidate reporting where the majority of data being analyzed resides (z/OS)
- Business benefits of analytics for queries previously set aside
- Fast time to value with transparent integration into existing applications
- Easy-to-install appliance add-on to DB2 for database query acceleration

Next Steps



Project Definition Workshop

Overview

- IDAA Architecture & Concepts Presentation
- Review current customer BI architecture and current challenges
- Assess target workload for IDAA:
 - New project specific: IDAA architecture and use case for customer's new project. Possible PoC definition and next steps
 - Consolidation specific: IDAA architecture and use case for customer. Identify distributed workload for a possible POC and next steps

Objective

- Identify applications / files that will be accelerated with the IDAA technology

Client participants:

- Client DWH and BI
- Project sponsor
- Business users if new project

IBM Participants:

- IBM CITA (lead)
- IBM z/IM Tech. Sales
- IDAA CoE

Duration:

- Half day or full day
- Delivered with the support of the Data warehousing on System z team

Architecture Workshop

Overview

- IDAA Architecture & Concepts Presentation
- Review current customer BI architecture and current challenges
- Assess target workload for IDAA:
 - New project specific: IDAA architecture and use case for customer's new project. Possible PoC definition and next steps
 - Consolidation specific: IDAA architecture and use case for customer. Identify distributed workload for a possible POC and next steps

Objective

- Define the IM architecture for the customer build around IDAA technology

Client participants:

- Client DWH and BI
- Project sponsor
- Business users if new project

IBM Participants:

- IBM CITA (lead)
- IBM z/IM Tech. Sales
- IDAA CoE

Duration:

- Half day or full day
- Delivered with the support of the Data warehousing on System z team

Workload Assessment

Customer

- Collects information from dynamic statement cache, supported by step-by-step instruction and REXX script (small effort for customer)
- Upload compressed file (up to some MB) to IBM FTP server

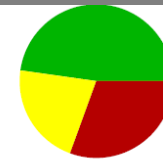
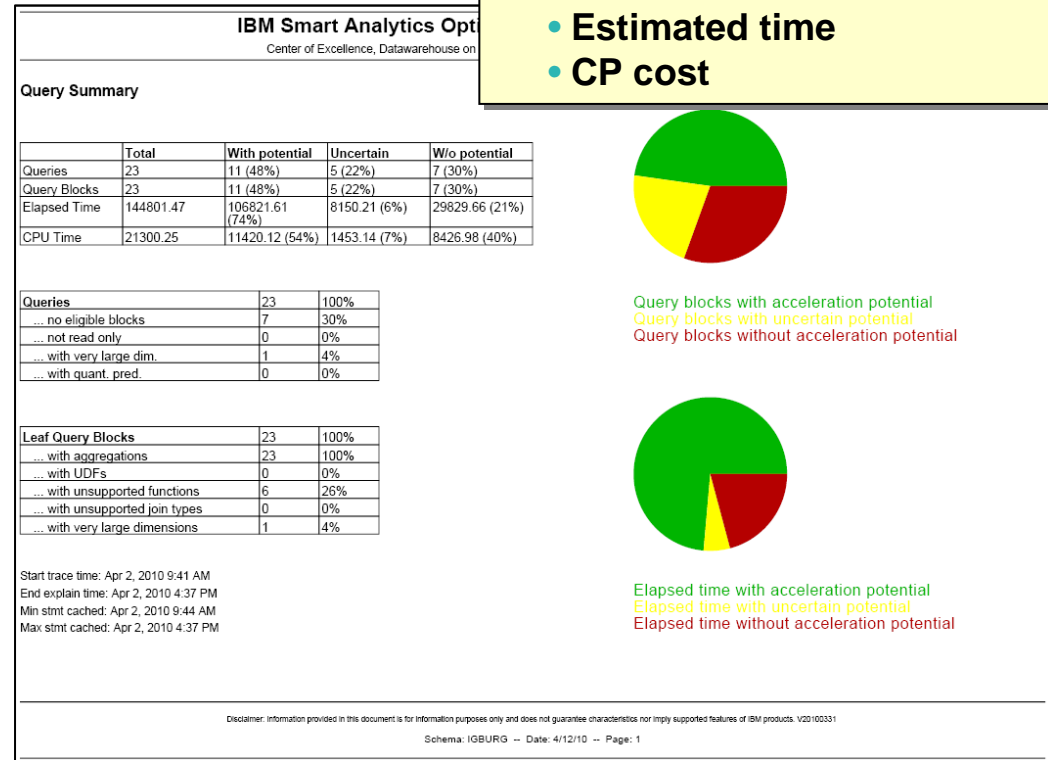
IBM

- Import data into local database
- Quick analysis based on known DB2 Analytics Accelerator capabilities

Key contact: Data Warehouse System z/Germany/IBM

Report for a first assessment:

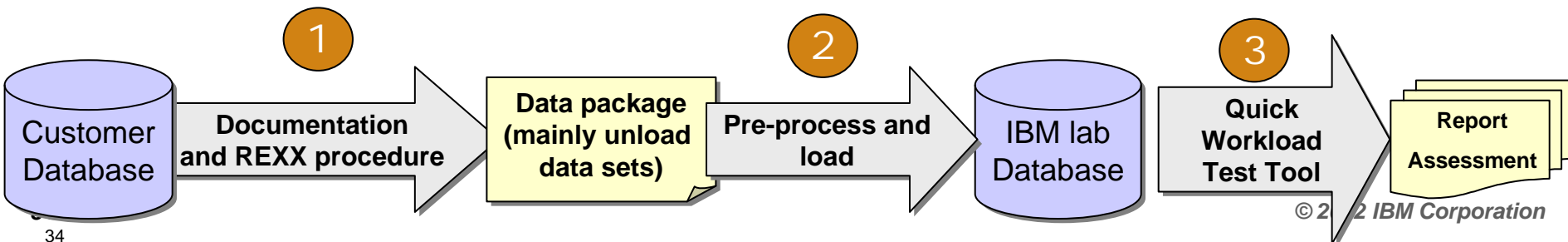
- Acceleration potential for
 - Queries
 - Estimated time
 - CP cost



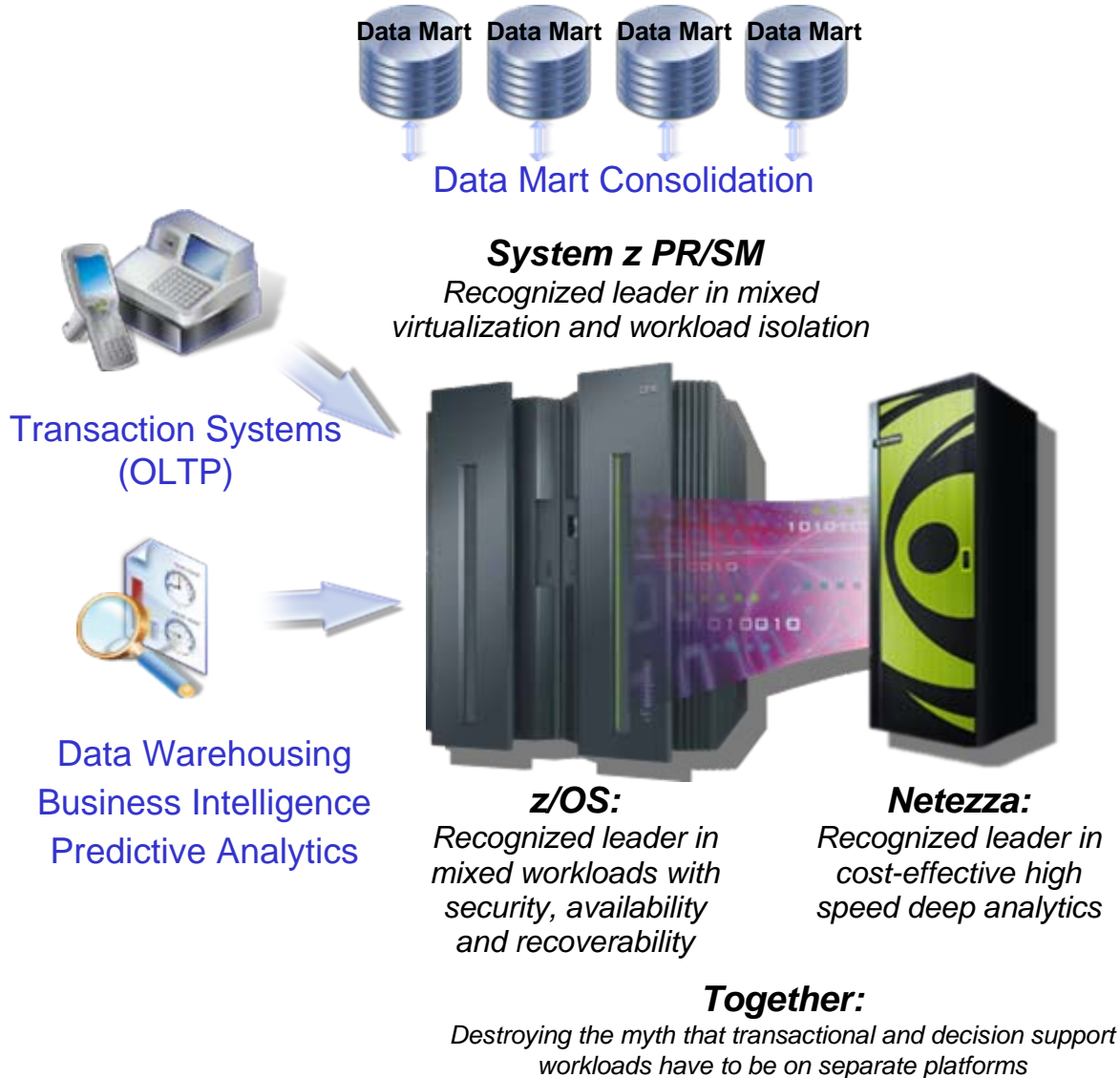
Query blocks with acceleration potential
Query blocks with uncertain potential
Query blocks without acceleration potential



Elapsed time with acceleration potential
Elapsed time with uncertain potential
Elapsed time without acceleration potential



The Ultimate Consolidation Platform



Bringing it all together

- *Better Business Response*
- *Reduced Costs*
- *More Available*
- *More Secure*
- *Reduced Data Movement*
- *Better Governance*
- *Reduced Data Latency*
- *Reduced Complexity*
- *Reduced Resources*

Learn More...

[Visit the Data Warehousing &
Business Analytics Webpage](http://www.ibm.com/software/data/businessintelligence/systemz/)

<http://www.ibm.com/software/data/businessintelligence/systemz/>

The text 'Thank You' is rendered in a large, 3D, metallic grey font. The letters are thick and have a slight shadow on the surface below them. The text is centered horizontally on the slide. Below the text, there is a faint, semi-transparent reflection of the words 'Thank You' on a light grey surface.