IBM
DB2 Analytics Accelerator for z/OS, v2.1

Providing extreme performance for complex business analysis

Willie Favero

IBM Silicon Valley Lab
Data Warehousing on System z Swat Team

Thursday, March 15, 2012
1:30 PM-2:30 PM
Session Number: 10504

my TWITTER

Please Note:

IBM's statements regarding its plans, directions, and intent are subject to change or withdrawal without notice at IBM's sole discretion.

Information regarding potential future products is intended to outline our general product direction and it should not be relied on in making a purchasing decision.

The information mentioned regarding potential future products is not a commitment, promise, or legal obligation to deliver any material, code or functionality. Information about potential future products may not be incorporated into any contract. The development, release, and timing of any future features or functionality described for our products remains at our sole discretion.

Performance is based on measurements and projections using standard IBM benchmarks in a controlled environment. The actual throughput or performance that any user will experience will vary depending upon many factors, including considerations such as the amount of multiprogramming in the user's job stream, the I/O configuration, the storage configuration, and the workload processed. Therefore, no assurance can be given that an individual user will achieve results similar to those stated here.

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.

Extreme Performance for Complex Business Analysis

IBM DB2 Analytics Accelerator powered by Netezza Technology

Business Analytics on System z

**IBM DB2 Analytics Accelerator for z/OS Announcement 211-454
October 12, 2011**

SHARE in Atlanta
2012

Generate More Revenue

Reduce Risk

Predict Future Outcomes with Greater Confidence

Lower Costs

Knowing what happened is no longer adequate.

Leaders say they need to know

what is happening now,

what is likely to happen next and

what actions they should take.

Modernize Your Decision Systems

Reengineering your information infrastructure

IBM Smart Analytics System 9700

Production Data

Low Latency Data

Historical Data

DB2 Analytics Accelerator

Transactional Information

Operational Analytics

Enhancing the operational systems by bringing business insights pervasively across the organization embedded in the operational applications

Operational Applications
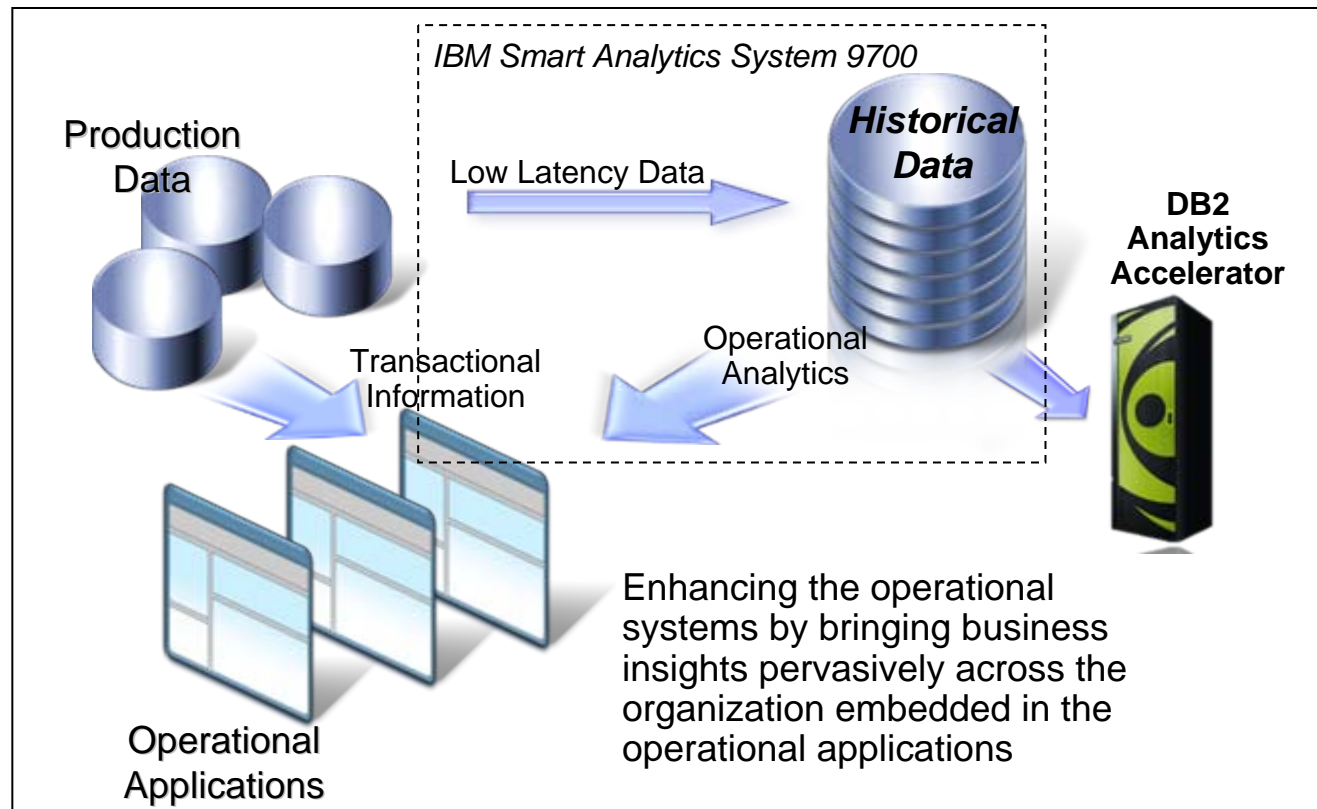
- Facilitate and manage transaction-oriented applications with analytics

- 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

SHARE in Atlanta
2012

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%

Actual customer results, October 2011

SHARE in Atlanta 2012

Performance & Savings

Query	Total Rows Reviewed	Total Rows Returned		DB2 Only			DB2 with IDAA			Times Faster
				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

Actual customer results, October 2011
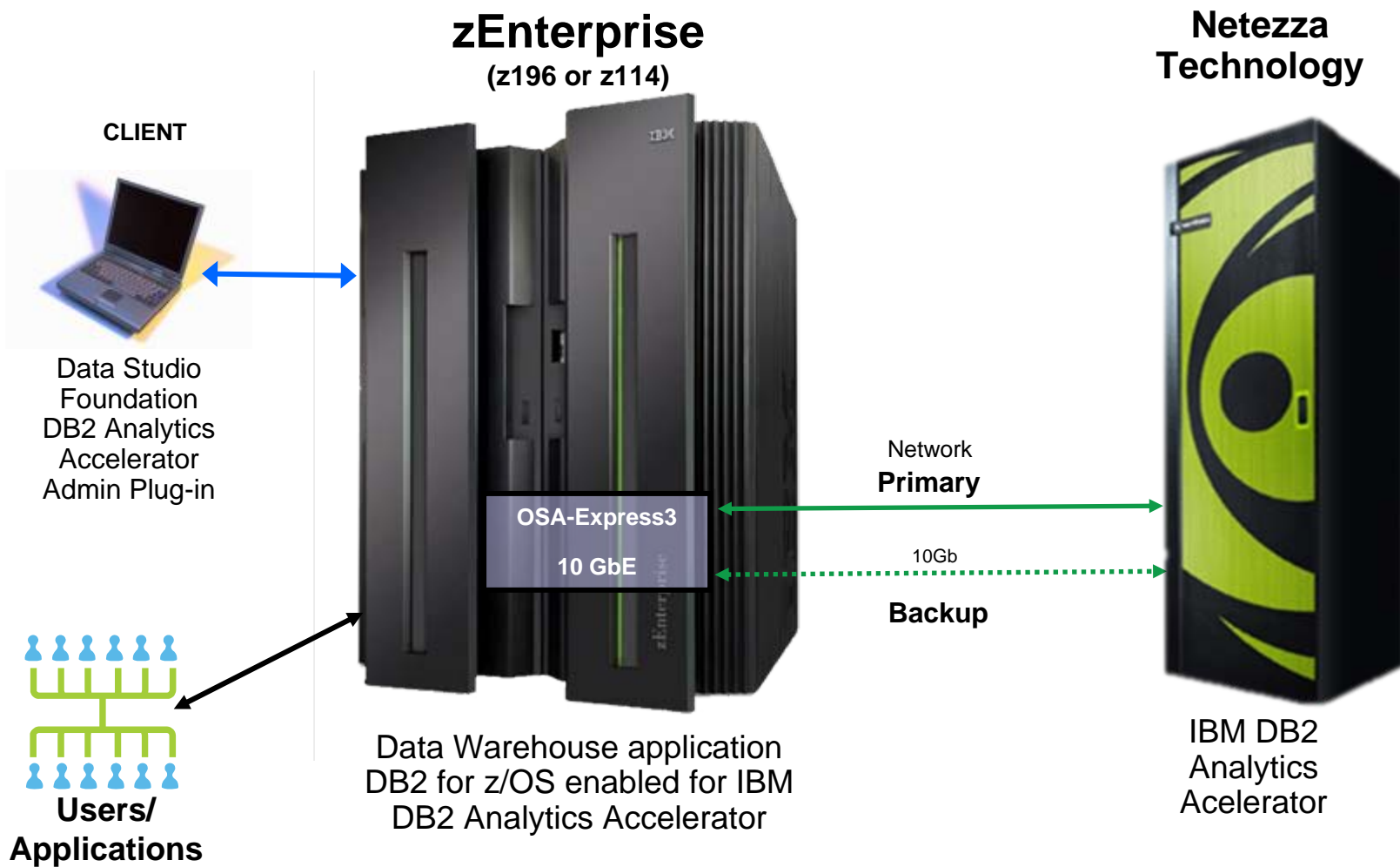
Queries run faster

- **Save CPU resources**
- **People time**
- **Business opportunities**

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

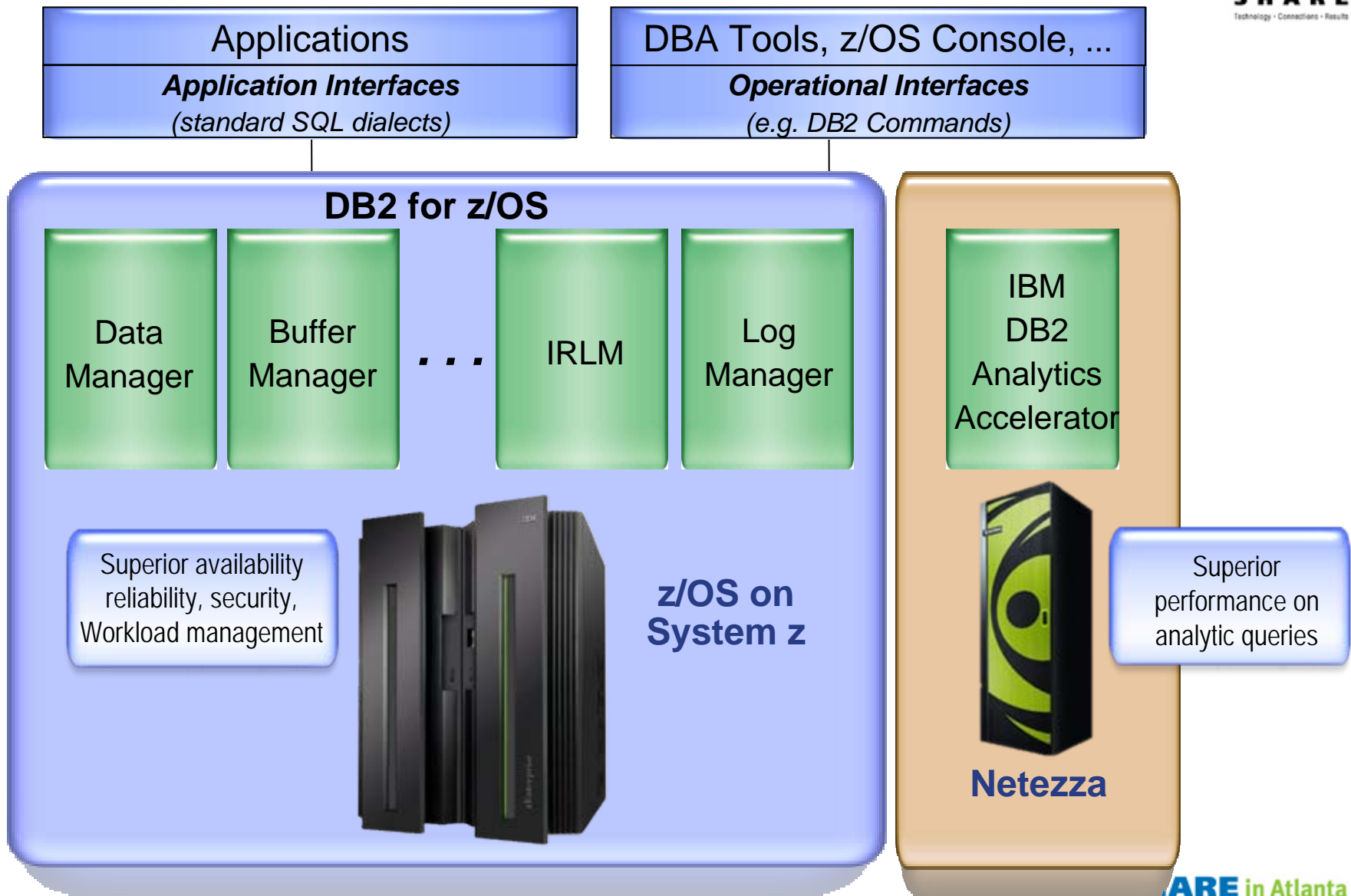DB2 Analytics Accelerator: "we expect ROI in less than 4 months"

Accelerating decisions to the speed of business

SHARE in Atlanta 2012

DB2 Analytics Accelerator V2 Components

zEnterprise
(z196 or z114)

Netezza
Technology

CLIENT

Data Studio
Foundation
DB2 Analytics
Accelerator
Admin Plug-in

OSA-Express3

10 GbE

Network
Primary

10Gb

Backup

Users/
Applications

Data Warehouse application
DB2 for z/OS enabled for IBM
DB2 Analytics Accelerator

IBM DB2
Analytics
Acelerator

SHARE in Atlanta
2012

Deep DB2 Integration within zEnterprise

SHARE
Technology · Connections · Results

Applications	DBA Tools, z/OS Console, ...
Application Interfaces *(standard SQL dialects)*	*Operational Interfaces* *(e.g. DB2 Commands)*

DB2 for z/OS

Data Manager	Buffer Manager	. . .	IRLM	Log Manager

IBM DB2 Analytics Accelerator

Superior availability reliability, security, Workload management

z/OS on System z

Superior performance on analytic queries

Netezza

SHARE in Atlanta
2012

DB2 Analytics Accelerator V2

Powered by Netezza 1000 Appliance
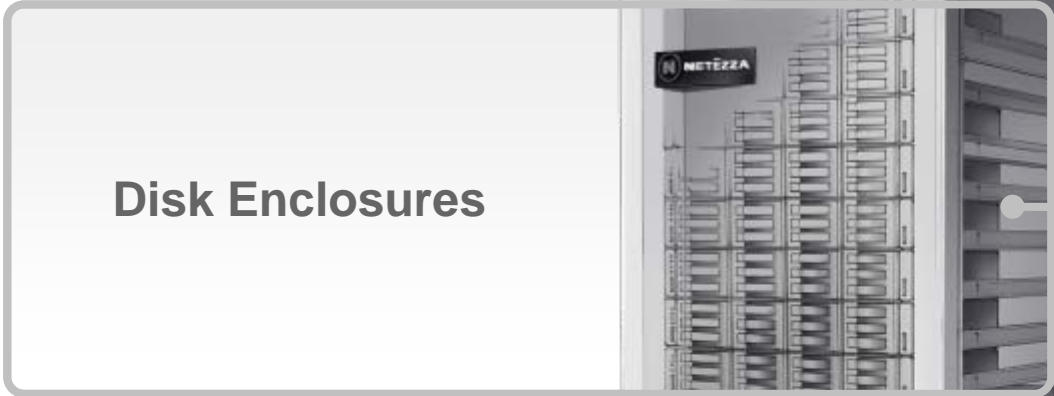
Disk Enclosures

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)*

SMP Hosts

DB2 Analytics Accelerator 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
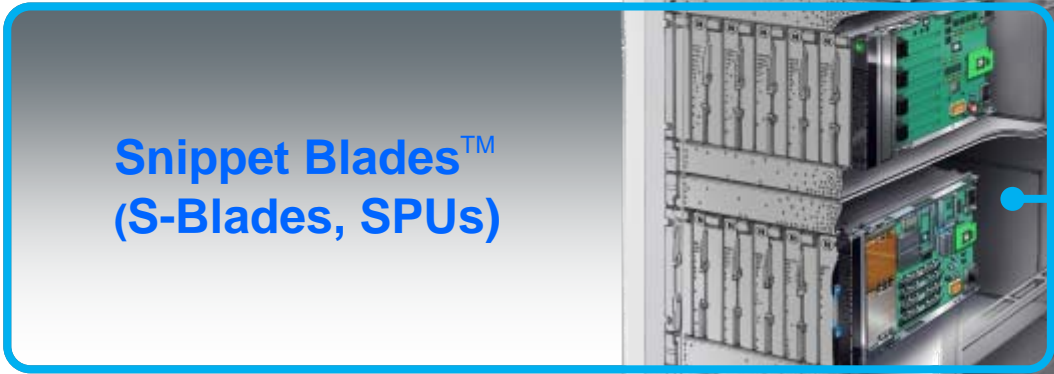
Snippet Blades^{TM}
(S-Blades, SPUs)

Processor &
streaming DB logic
High-performance database
engine streaming joins,
aggregations, sorts, etc.
e.g. TF12: 12 back/end SPUs
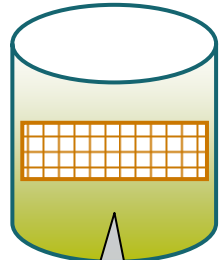 *(more details on following charts)*

SHARE in Atlanta
2012

The Appliance Connected to a System z

Disk Enclosures

S-Blades™

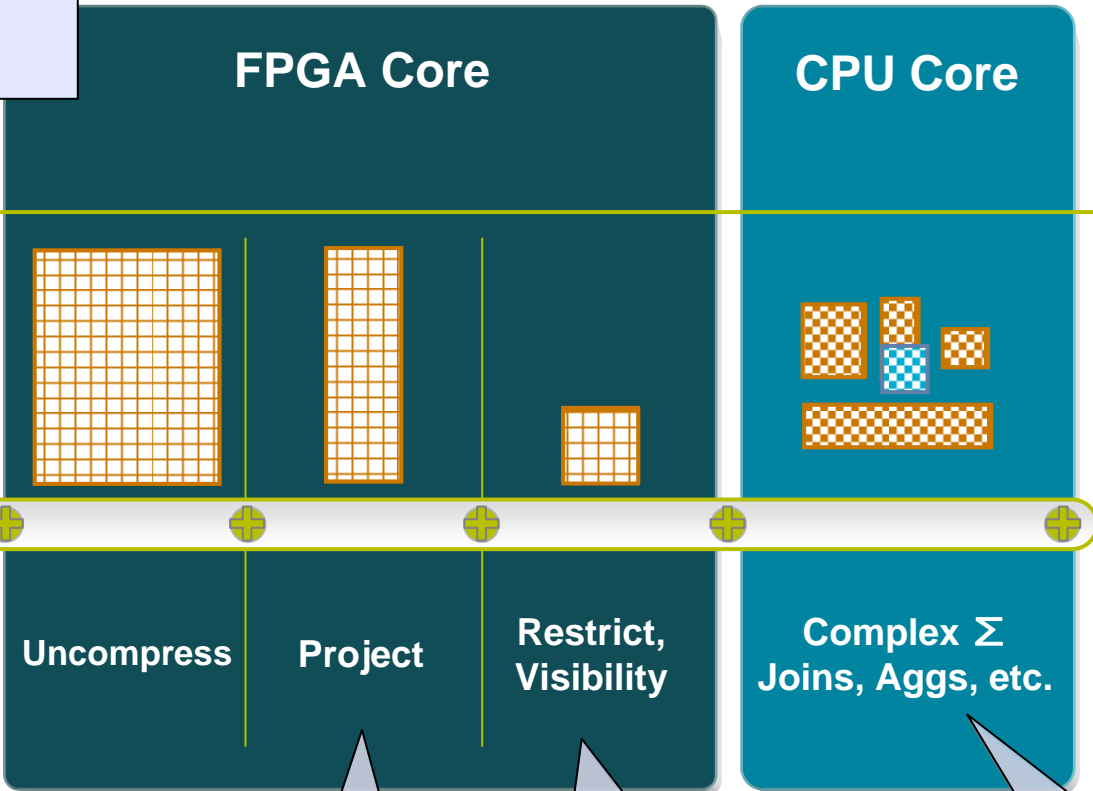**Network Fabric**
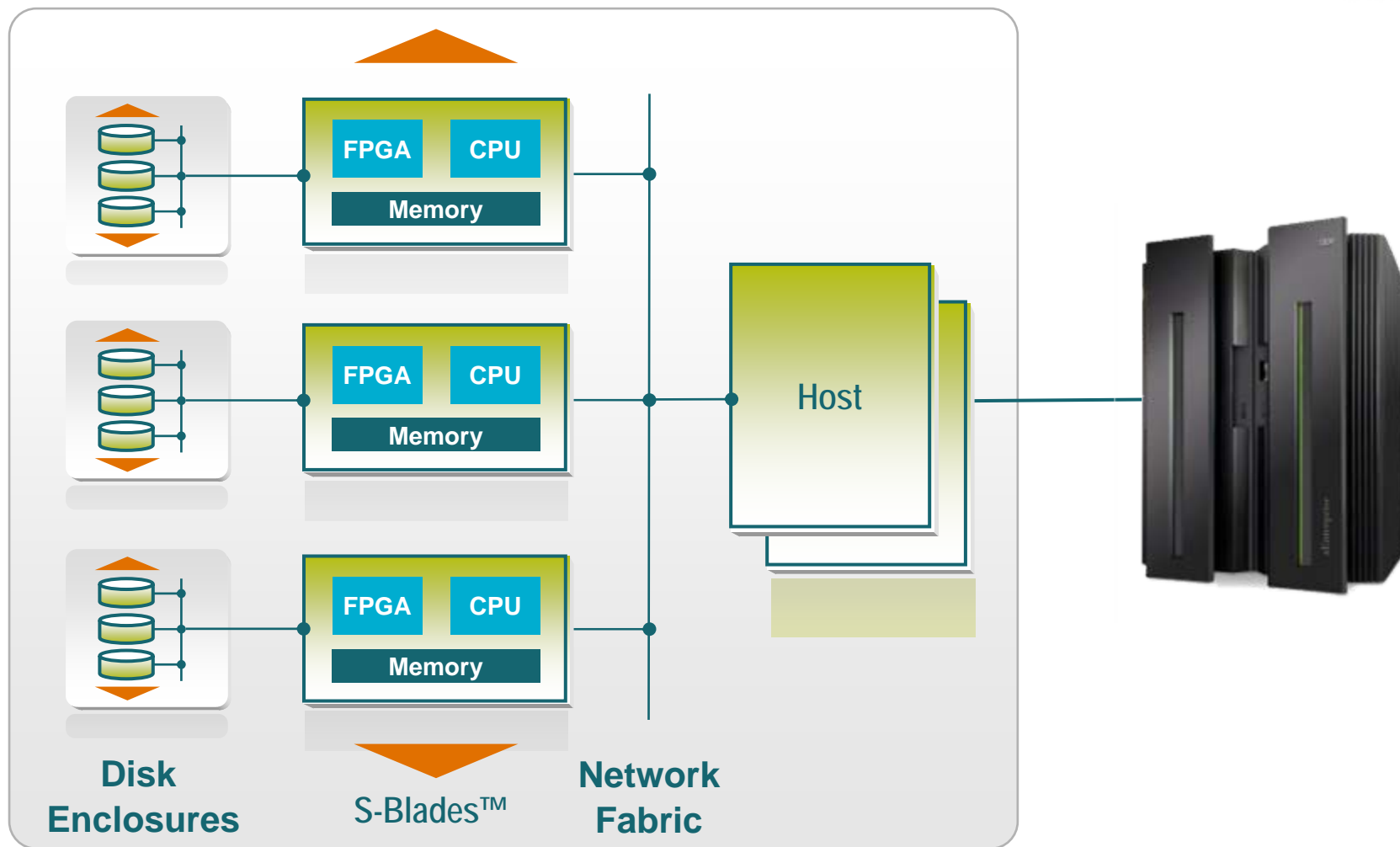
Netezza Appliance

The Key to the Speed

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

FPGA Core

CPU Core

Uncompress | **Project** | **Restrict, Visibility** | **Complex Σ Joins, Aggs, etc.**
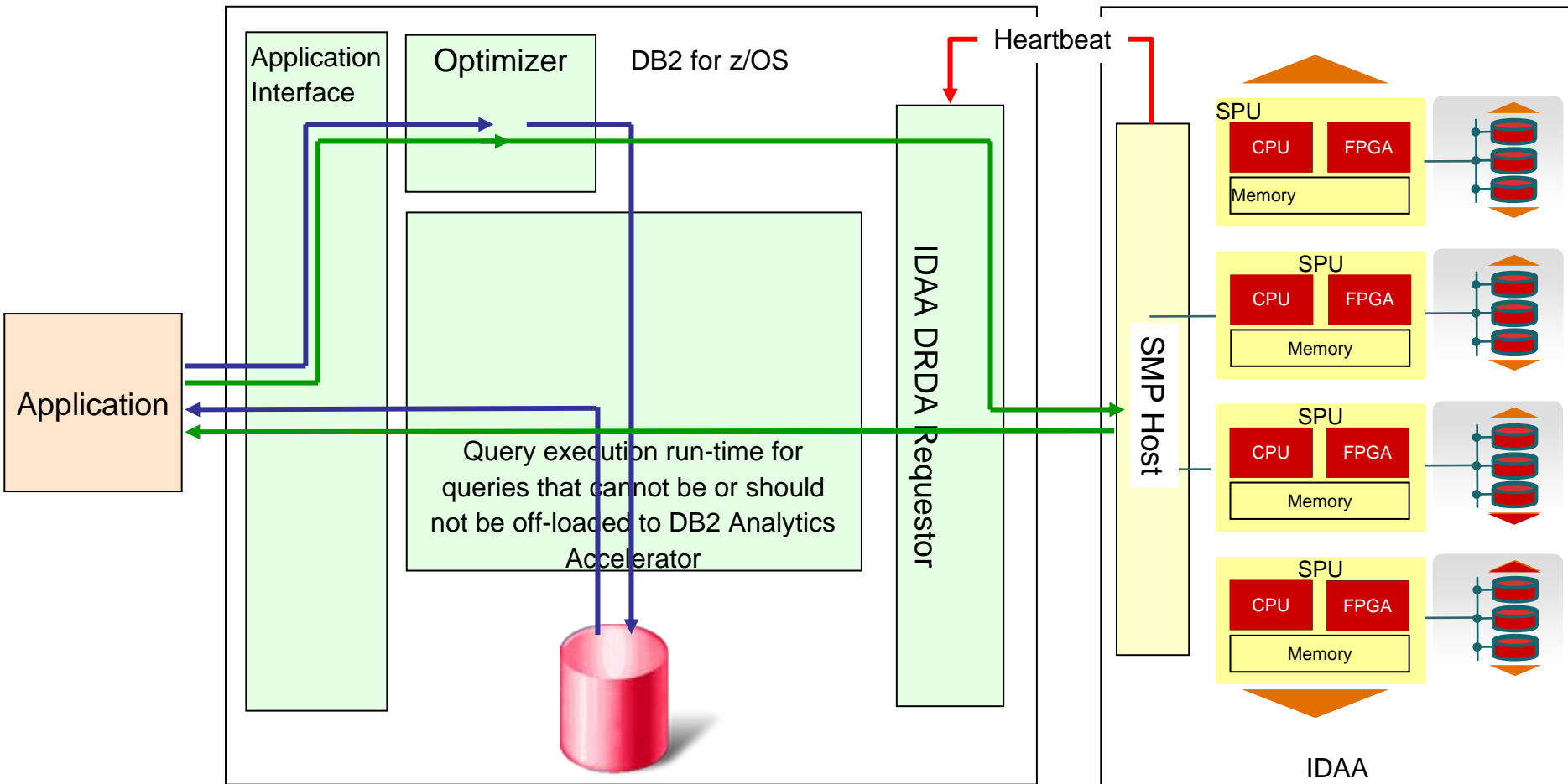
```
Slice of table
MTHLY_RX_TERR_DATA
   (compressed)
```

```
select DISTRICT,
       PRODUCTGRP,
       sum(NRX)
```

```
where MONTH = '20091201'
and   MARKET = 509123
and   SPECIALTY = 'GASTRO'
```

```
sum(NRX)
```

SHARE in Atlanta
2012

Query Execution Process Flow

S H A R E
Technology · Connections · Results

Heartbeat

Application Interface

Optimizer

DB2 for z/OS

IDAA DRDA Requestor

SMP Host

Application

Query execution run-time for queries that cannot be or should not be off-loaded to DB2 Analytics Accelerator

SPU
CPU FPGA
Memory

SPU
CPU FPGA
Memory

SPU
CPU FPGA
Memory

SPU
CPU FPGA
Memory

IDAA

Queries executed without DB2 Analytics Accelerator

Queries executed with DB2 Analytics Accelerator

Heartbeat (DB2 Analytics Accelerator availability and performance indicators)

SHARE in Atlanta 2012

Accelerator Content Maintenance
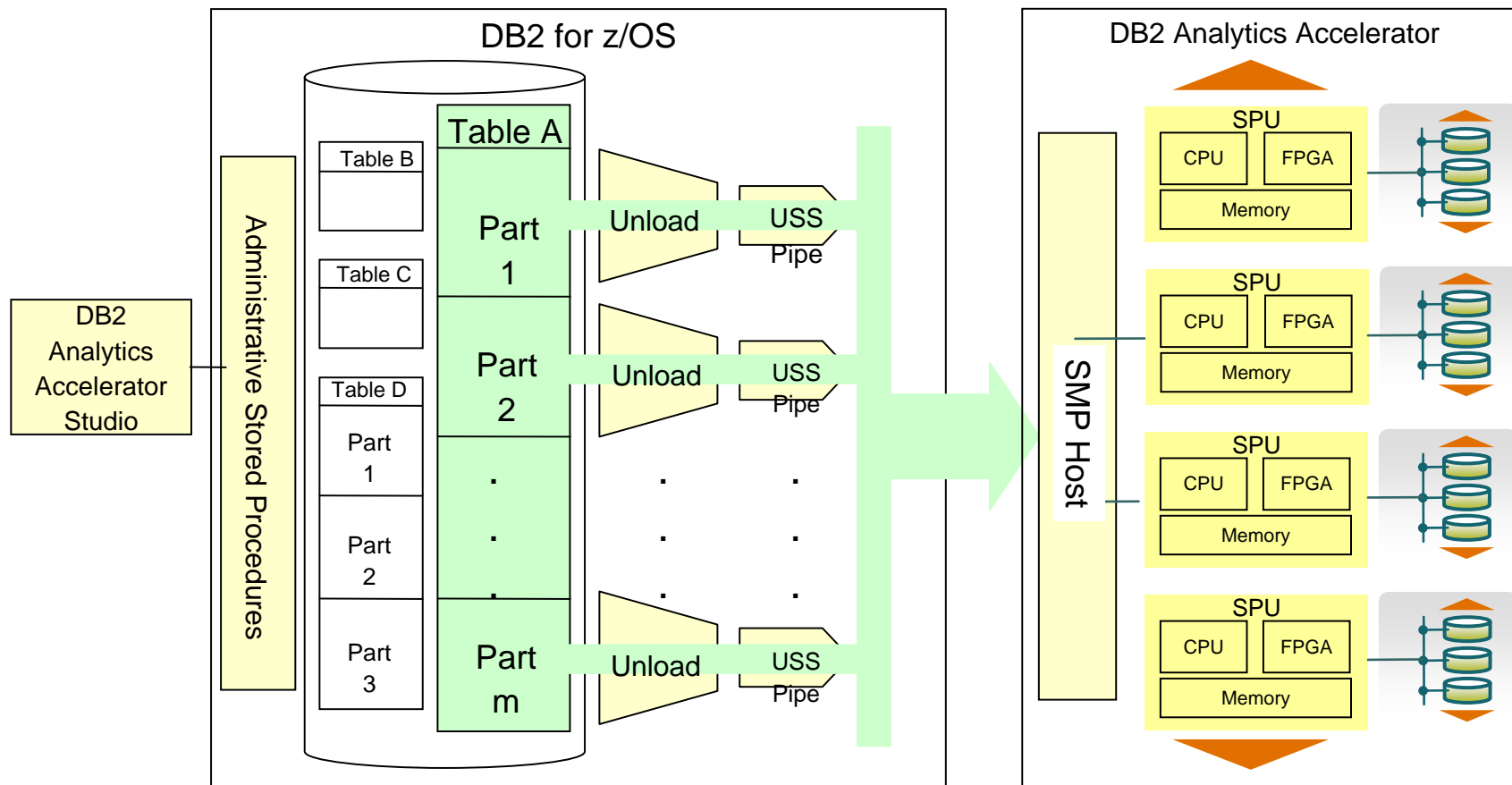
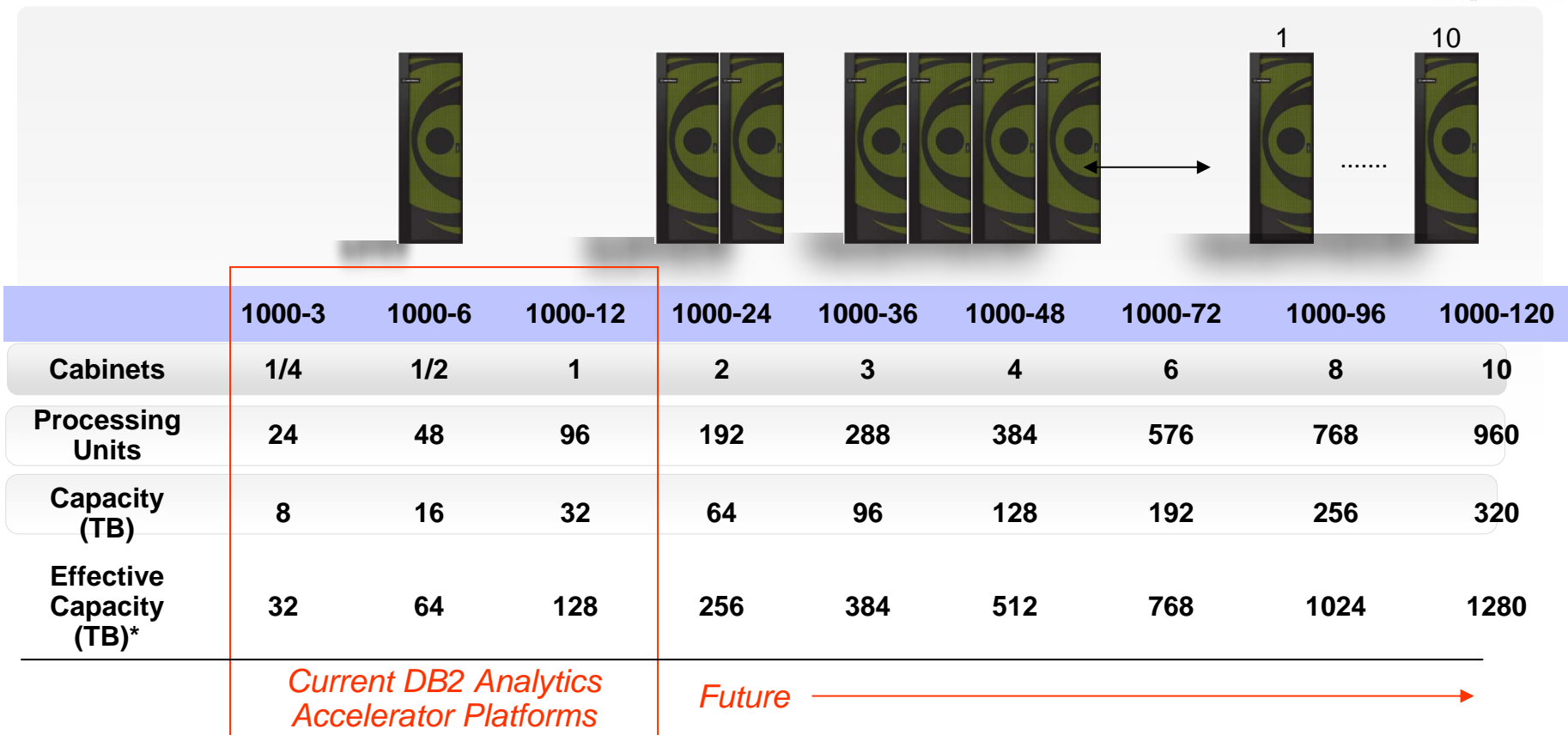- Partitions belonging to the same table can be loaded in parallel
  - User-defined degree of parallelism
- Updates are done on a per-table or per-partition level

Netezza 1000 Appliance Scalability

1   10

	1000-3	1000-6	1000-12	1000-24	1000-36	1000-48	1000-72	1000-96	1000-120
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

Current DB2 Analytics Accelerator Platforms

Future ————————————————————————→

Predictable, Linear Scalability throughout entire family

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

*: 4X compression assumed

SHARE in Atlanta
2012

Connectivity Options

Multiple DB2 systems can connect to a single
DB2 Analytics Accelerator

A single DB2 system can connect to multiple
DB2 Analytics Accelerators

Multiple DB2 systems can connect to multiple
DB2 Analytics Accelerators

Better utilization of IBM DB2 Analytics
Accelerator resources

Scalability

High availability
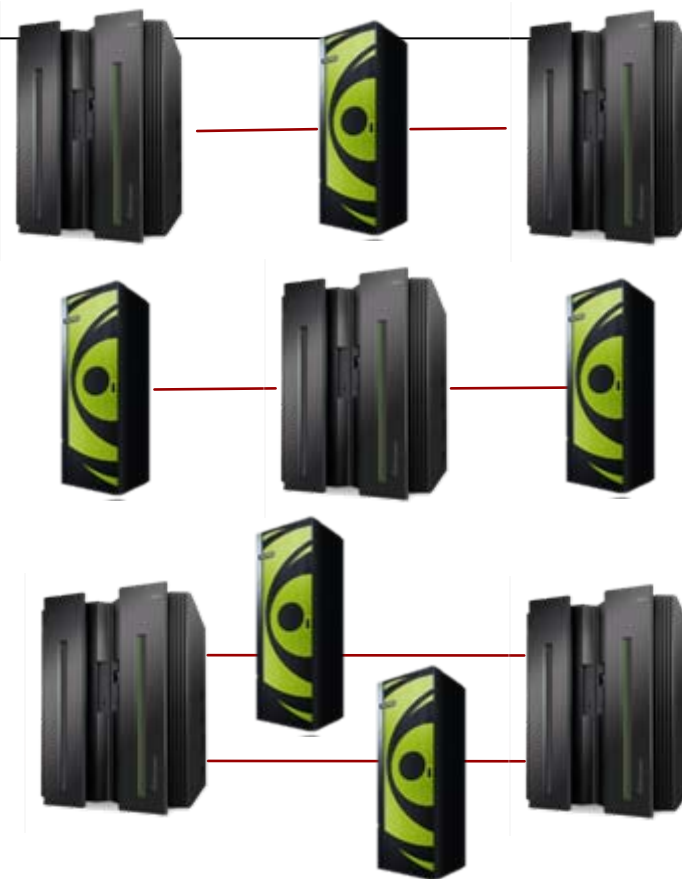
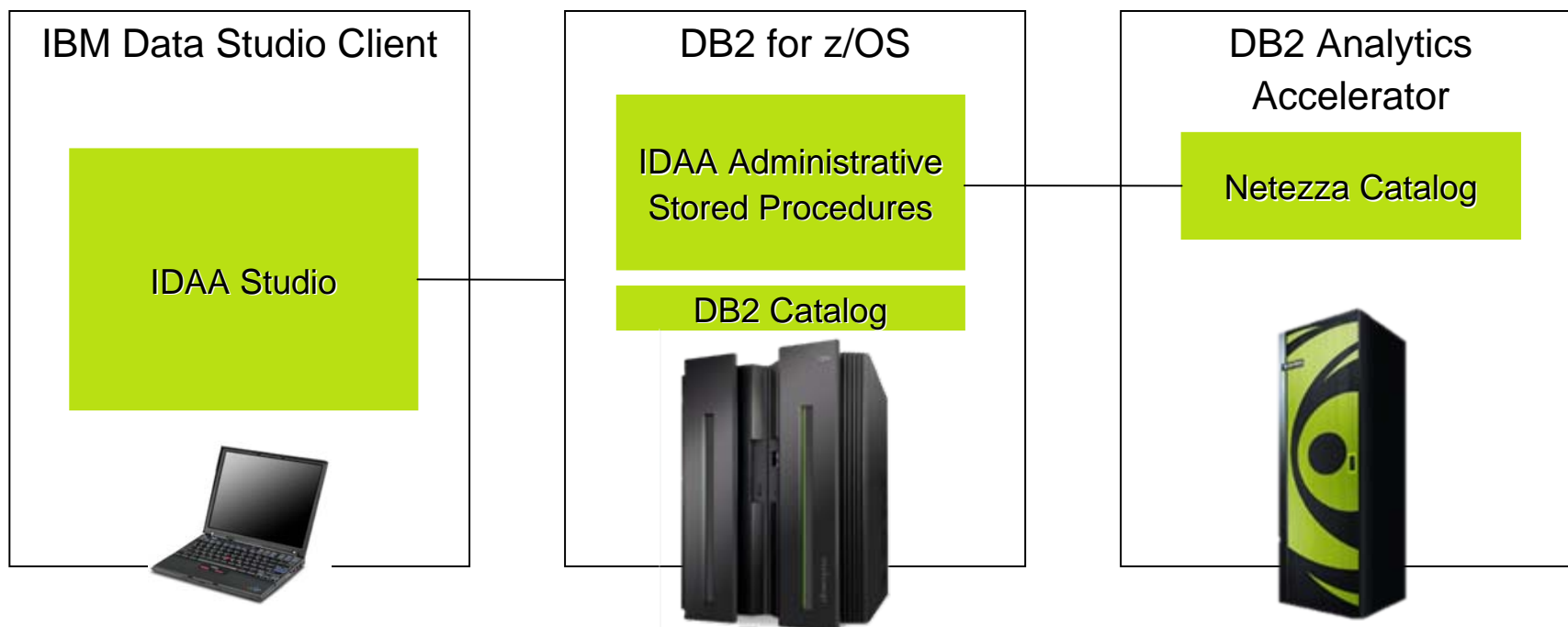Full flexibility for DB2 systems:

- 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

Accelerator Table Definition and Deployment

IBM Data Studio Client	DB2 for z/OS	DB2 Analytics Accelerator
IDAA Studio	IDAA Administrative Stored Procedures DB2 Catalog	Netezza Catalog

- 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.

SHARE in Atlanta
2012

Why Both?

Marrying the best of both worlds

**IBM
System z**

**IBM
Netezza**

Mixed Workload System

Focused Appliance

*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 diverse workload

Very focused workload

Tailored to your needs

A Hybrid Solution

IBM System z with IBM DB2 Analytics Accelerator	IBM Netezza
Mixed Workload System	**Focused Appliance**
• 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	• 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

Flexibility   *The right mix of simplicity and flexibility*   **Simplicity**

SHARE in Atlanta
2012

What is the value?

- Quickly delivers analytics to operational applications
- High speed analytics where the data is generated
- Enables train-of-thought analysis with high speed complex queries
- Substantially reduces operational costs by removing the need for complex query tuning
- Creates a highly secure environment for highly sensitive analysis (EAL5)
- Speeds batch reporting cycle to meet stricter SLAs
- Enables decision makers to perform business analysis they never dared in the past
- Enables query acceleration across multiple applications and systems
- Capitalizes on DB2 skills and certification removing the need to learn or convert to another SQL environment

"Back of the Envelope" ROI

Consider the MIPs of your z196 / z114

Consider software MLC reduction: z/OS, CICS, DB2…

Consider hardware value of MIPs redeployed

Examples of 6 month ROI:

- Avoiding   400 MIPs is roughly the cost of an IDAA Netezza 1000-3
- Avoiding   800 MIPs is roughly the cost of an IDAA Netezza 1000-6
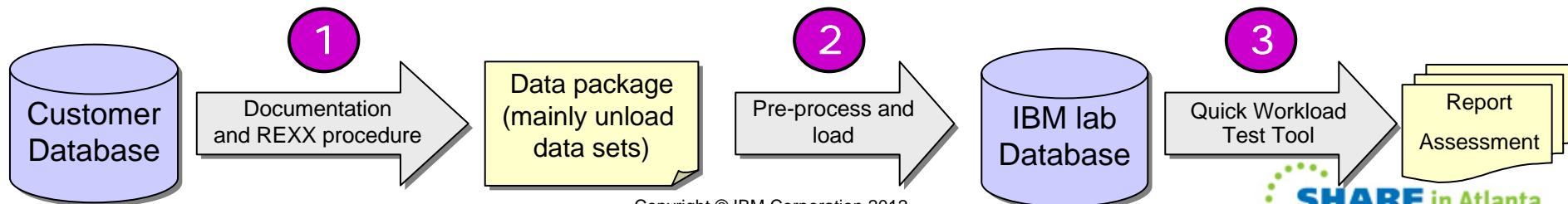- Avoiding 1600 MIPs is roughly the cost of an IDAA Netezza 1000-12

Next step: Quick Workload Test

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

- **IBM / Center of Excellence**
    - Importing data into local database
    - Quick analysis based on known DB2 Analytics Accelerator capabilities

Customer Database → ① Documentation and REXX procedure → **Data package (mainly unload data sets)** → ② Pre-process and load → **IBM lab Database** → ③ Quick Workload Test Tool → **Report Assessment**

SHARE in Atlanta 2012

The Ultimate Consolidation Platform

Data Mart **Data Mart** **Data Mart** **Data Mart**

Data Mart Consolidation

System z PR/SM
Recognized leader in mixed virtualization and workload isolation

Transaction Systems
(OLTP)

Data Warehousing
Business Intelligence
Predictive Analytics

z/OS:
Recognized leader in mixed workloads with security, availability and recoverability

Netezza:
Recognized leader in cost-effective high speed deep analytics

Together:
Destroying the myth that transactional and decision support workloads have to be on separate platforms

Bringing it all together

- ***Better Business Response***

- ***Reduced Costs***

- ***More Available***

- ***More Secure***

- ***Reduced Data Movement***

- ***Reduced Data Latency***

- ***Reduced Complexity***

- ***Reduced Resources***

SHARE in Atlanta
2012

Learn More...

Visit the Data Warehousing &

Business Analytics Webpage

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

Willie Favero

DB2 SME
Data Warehousing for System z Swat Team
IBM Silicon Valley Laboratory

My DB2 Blog
www.it.toolbox.com/blogs/db2zos/

http:www.WillieFavero.com

my TWITTER   in *my* LINKEDIN   My slideshare profile   f *my* FACEBOOK

SHARE in Atlanta
2012