Business-critical insight for proven competitive advantage

Gary Crupi, IBM Distinguished Engineer, Big Data & Analytics

September 2014

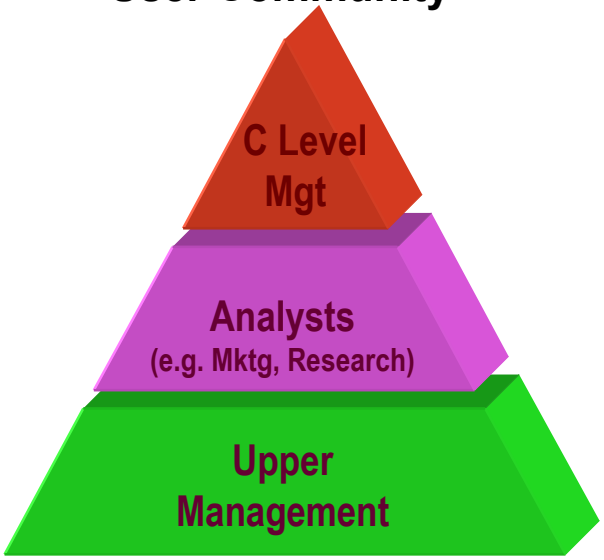
Goals for Today

- **Understand…**
  - Business Critical Analytics
  - The role System z and the DB2 Analytics Accelerator play in delivering business outcomes
  - Customer Stories
- **Act…**
  - Deepen your knowledge of the DB2 Analytics Accelerator
  - Socialize the value of the Accelerator to address business / application requirements
  - Engage with IBM to determine the best path for adopting the technology

Business Critical Analytics

"Traditional" Business Analytics

User Community



Goals:

- Increase Revenue
- Reduce Costs
- Reduce Risk
- Track Trends

Workload Characteristics

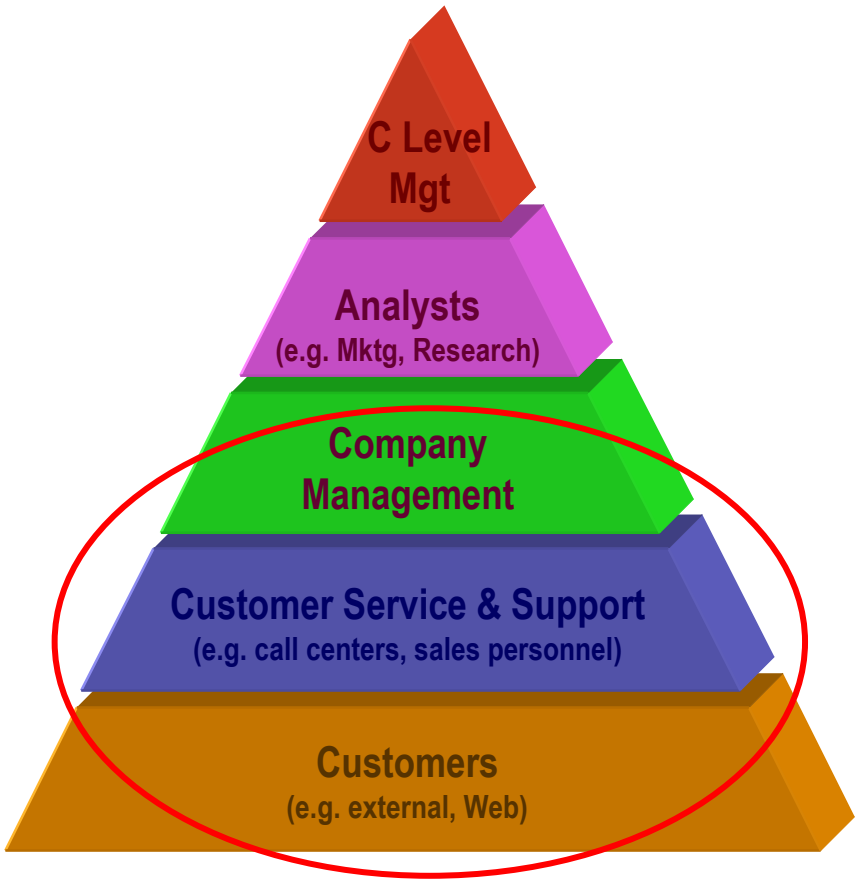
Number of Users	*Transaction Volume*	*Transaction Latency*	*Availability*	*Transaction Type*
Few	Small	Less Important	Less Important	Complex

Business Benefits

Objective	Activity
React to Changes	Track Market & Consumer Trends
Cost Management	Supply Chain Reporting
Consumer Protection	Post Transaction Fraud Detection
Financial Management	End of Period Reporting
Business Management	Batch Summary & Detail Reports

"Today's" Business Analytics Requirements – "Business Critical Analytics"

User Community



Workload Characteristics

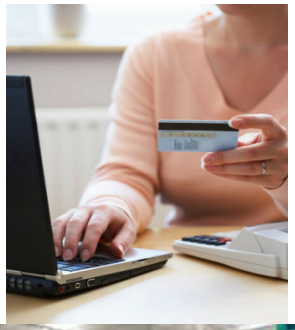
Number of Users	*Transaction Volume*	*Transaction Latency*	*Availability*	*Transaction Type*
Few	Small	Less Important	Less Important	Complex
Many	Very Large	Critical	Critical	Simple to Complex

Advent of Operational & Business Critical Analytics

Highest Qualities of Service Required

What is Business Critical Analytics?

- Any analytic application critical to optimally running a business
- If this application fails for any length of time you can lose business
- Infrastructure Matters
  - When “good enough” is not enough
  - When “mainframe like” is just not a mainframe



Preventing Credit Card Fraud

Reducing Customer Churn

Cross-selling, up-selling customers

The Hybrid Approach
Traditional Approach to Reporting & Analytic Systems

Delivering business critical analytics

Operational Applications

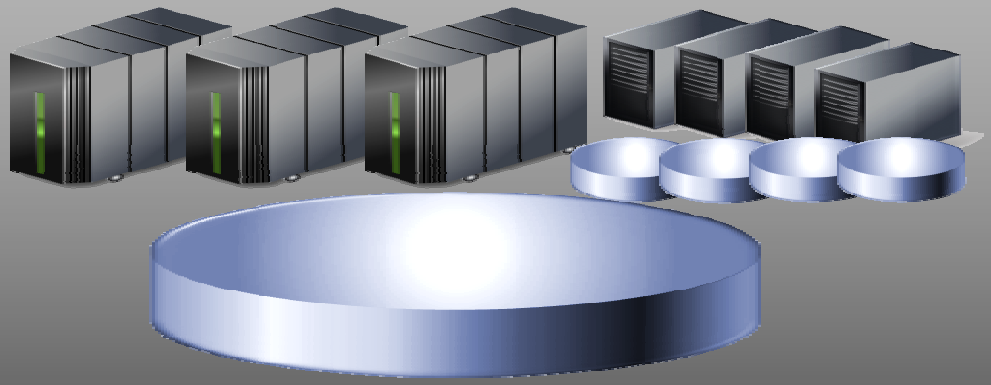
Reporting & Analytic

True Mixed Workloads

Transaction Proc

Shared Every

Transactional Processing, Traditional Analytics & Business Critical Analytics



Hybrid DB

Reduced Latency. Greater Security.
Improved Data Governance. Reduced Complexity.

High volume business transactions and batch reporting running concurrently with complex queries

High volume transactions, c reporting, ar processing concurr

iness dictive

g DB

omplex batch g

The Analyst Community Has Taken Notice!

Gartner®

- "By eliminating analytic latency and data synchronization issues, hybrid transaction/analytical processing will enable IT leaders to simplify their information management infrastructure"

- "This architecture will drive the most innovation in real-time analytics over the next 10 years via greater situation awareness and improved business agility"

Gartner Research Note G00259033: Gartner 01-2014 Hybrid Transaction Analytical Processing Will Foster Opportunities

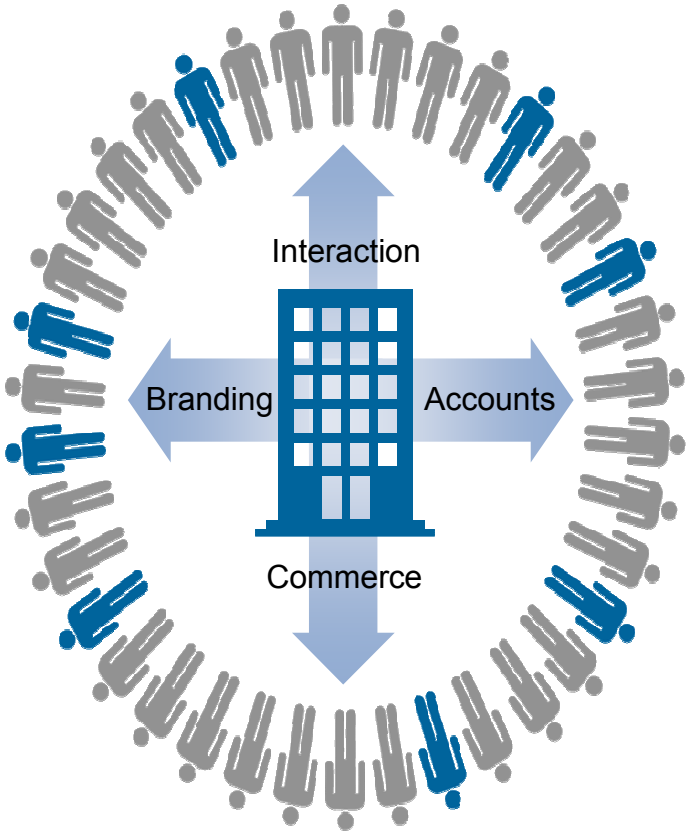
Hybrid Transaction and Analytics Processing (HTAP)

Gartner Research Note, 28 January 2014

Real Time Analytics that minimizes or eliminates analytics latency or synchronization issues by eliminating the divide between operational and analytical systems.

The business relationship has changed forever

Then: *"I have an offer – let me find a customer I can sell to"*



Now: *"I have a customer – what do they need most?"*

***Customer experience** is the competitive advantage for top-line growth*

Leaders must leverage data to outperform

Drive **top-line growth** via transformational new services

80%

of marketers send **same content** to all subscribers

6%

of businesses **"extremely satisfied"** with ability to use customer data for decisions

+7.6%

annual increase in **customer lifetime value** for firms that use engagement analytics

Dramatically **"maximize yield"** on investments in standard electronic business practices

$226B

estimated **fraud loss** to healthcare

16%

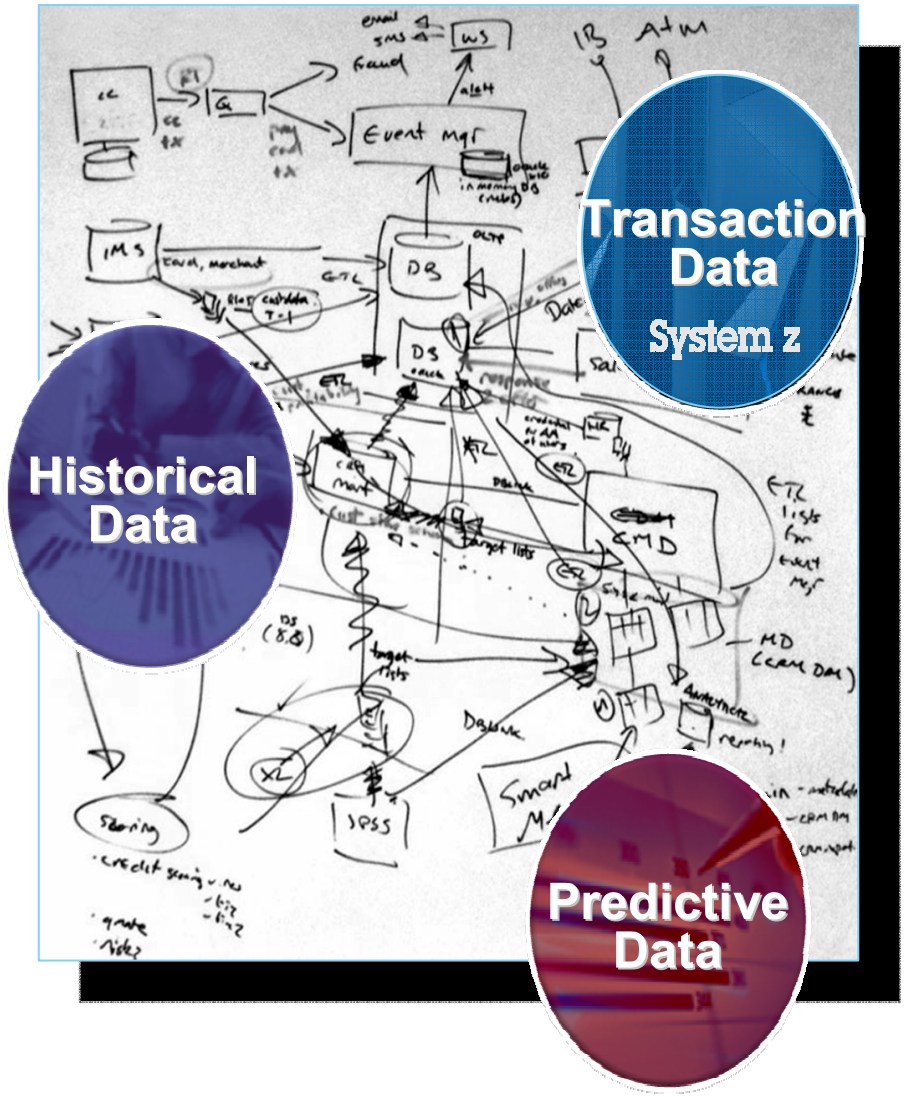
government **tax revenues lost** to non-compliance

$100M

typical banking regulation **non-compliance fine**

IT must be exploited as a business strategy

Challenges with traditional analytics processing



Significant complexity

Data is move from operational databases to separated data warehouses/data marts to support analytics

Analytics latency

Transactional data is not readily or easily available for analytics when created

Lack of synchronization

Data is not easily aggregated and users are not assured they have access to "fresh" data

Data duplication

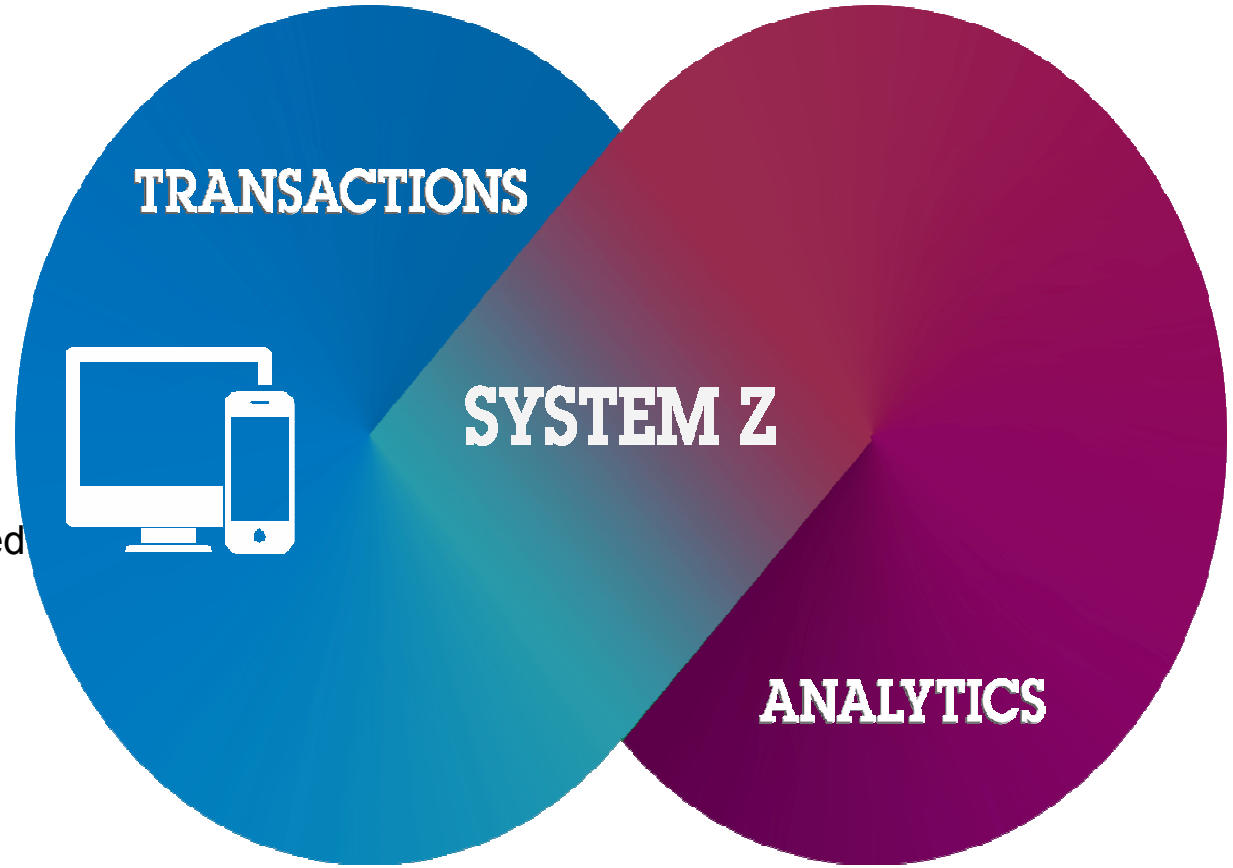
Multiple copies of the same data is proliferated throughout
the organization

Excessive costs

An IT infrastructure that was not designed nor can support real-time analytics

Transactions & analytics processed together

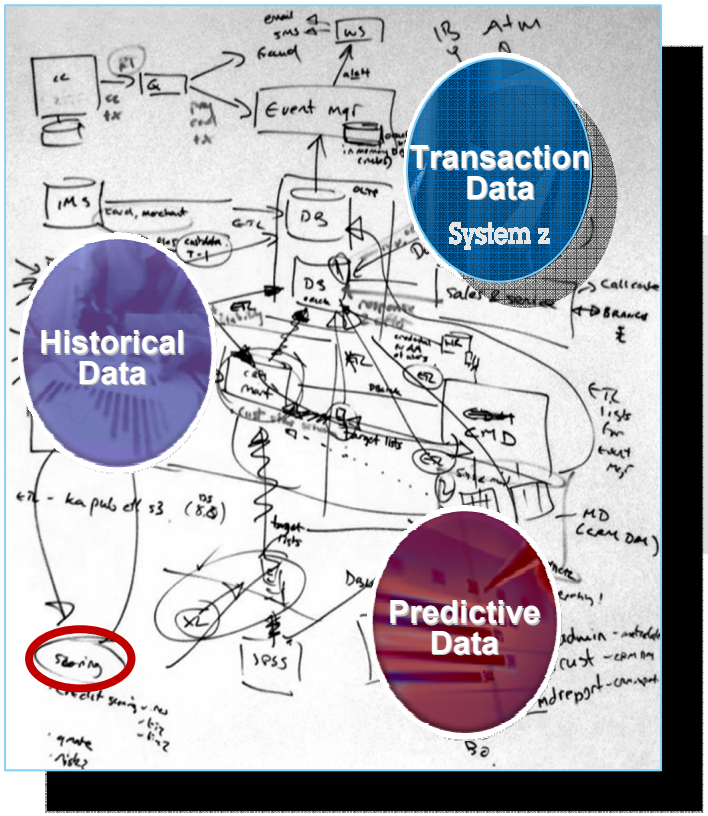
- Purchase made
- Resources consumed
- Bill paid
- Claim submitted
- Information updated
- Call center contacted

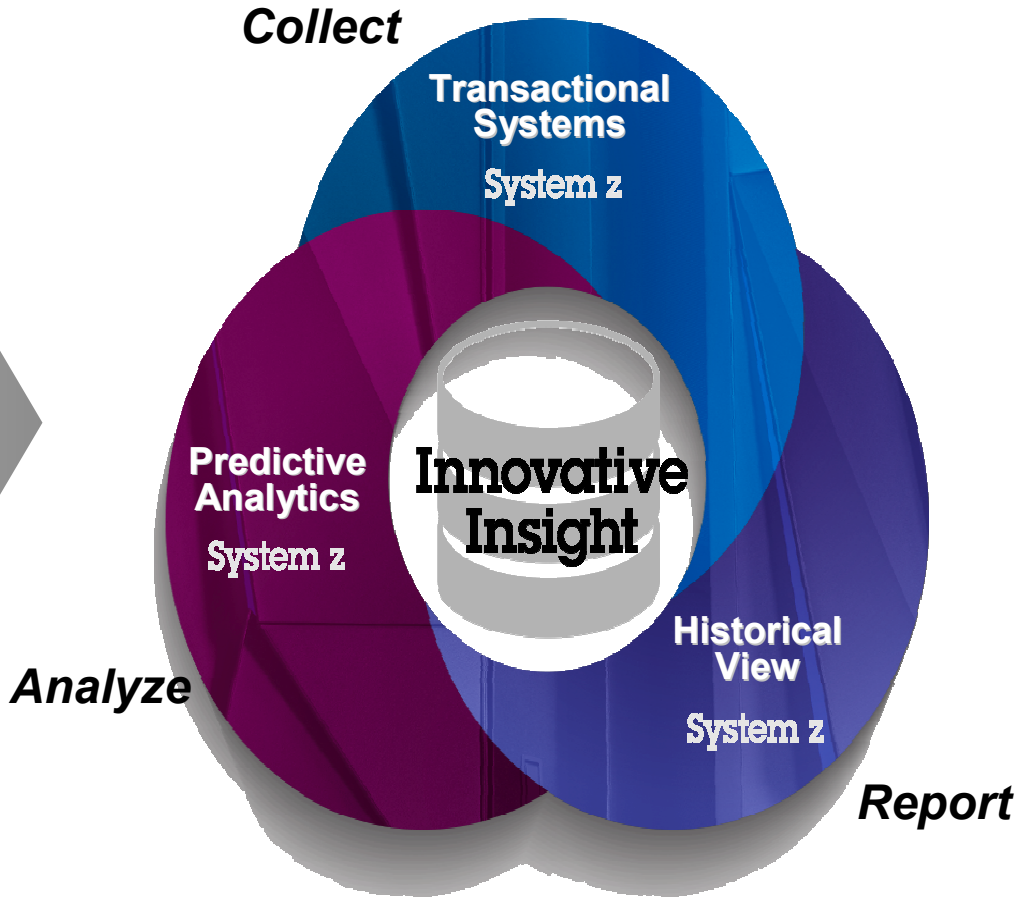


- What happened?
- How many, how often, where?
- What actions are needed?
- What will happen if?
- What will produce the best outcome?

Analytics as part of the flow of business; insights on every transaction

Enabling innovative insight

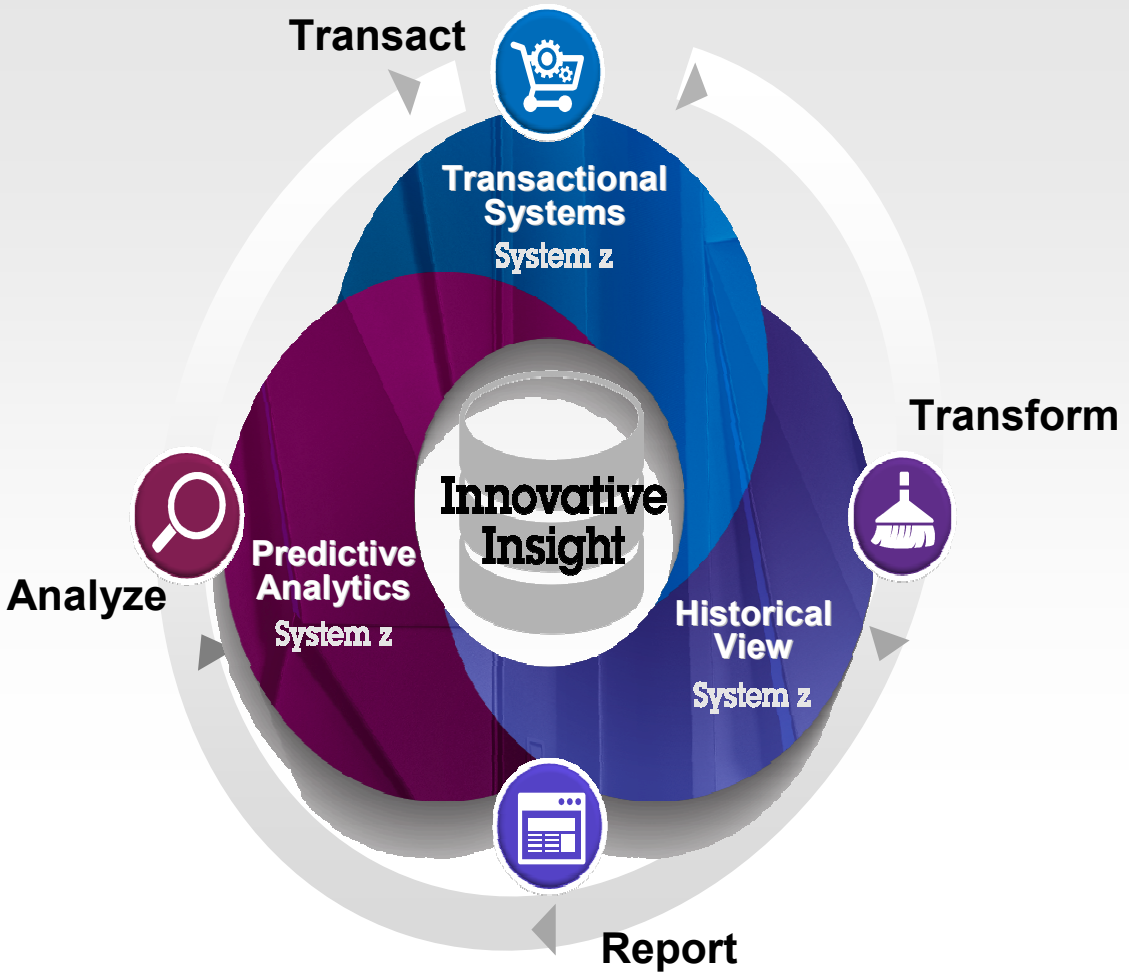




The System z strategy

Integrating operations & analytics in one streamlined, end-to-end data lifecycle

- Better business response
- Reduced data movement, reduced complexity, reduced configuration resources
- More accurate, more secure, more available



5 key takeaways

- **Many organizations are trying to deliver instantaneous**, **on-demand customer service** with IT systems designed to provide after-the-fact intelligence
- **Achieving insight with every transaction demands** a holistic implementation of an integrated data lifecycle with business-critical systems
- **System z has the vision, strategy and technology** to fuse transactions and analytics by eliminating the latency and complexity pitfalls that develop with a distributed approach
- **System z "operational analytics" builds advanced decision management support** on this integrated data platform injecting intelligence into operations without sacrificing performance
- **Truly transformational business opportunities** require truly transformational infrastructure - and that infrastructure is System z

The Critical Innovation: DB2 Analytics Accelerator

IBM DB2 Analytics Accelerator

Do things you could never do before!

What is it?

The IBM DB2 Analytics Accelerator is a workload optimized, appliance add-on to DB2 for z/OS, that enables the integration of analytic insights into operational processes to drive business critical analytics and exceptional business value

What does it do?

- Accelerates complex queries, up to 2000x faster
- Lowers the cost of storing, managing and processing historical data
- Minimizes latency
- Reduces zEnterprise capacity requirements
- Improves security and governance
- Reduces operational costs and risk
- Complements existing investments

DB2 Analytics Accelerator - not "just an appliance"

An extension of the DB2 for z/OS system

Analytic applications connect to DB2 server as always – no change

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

DB2 for z/OS

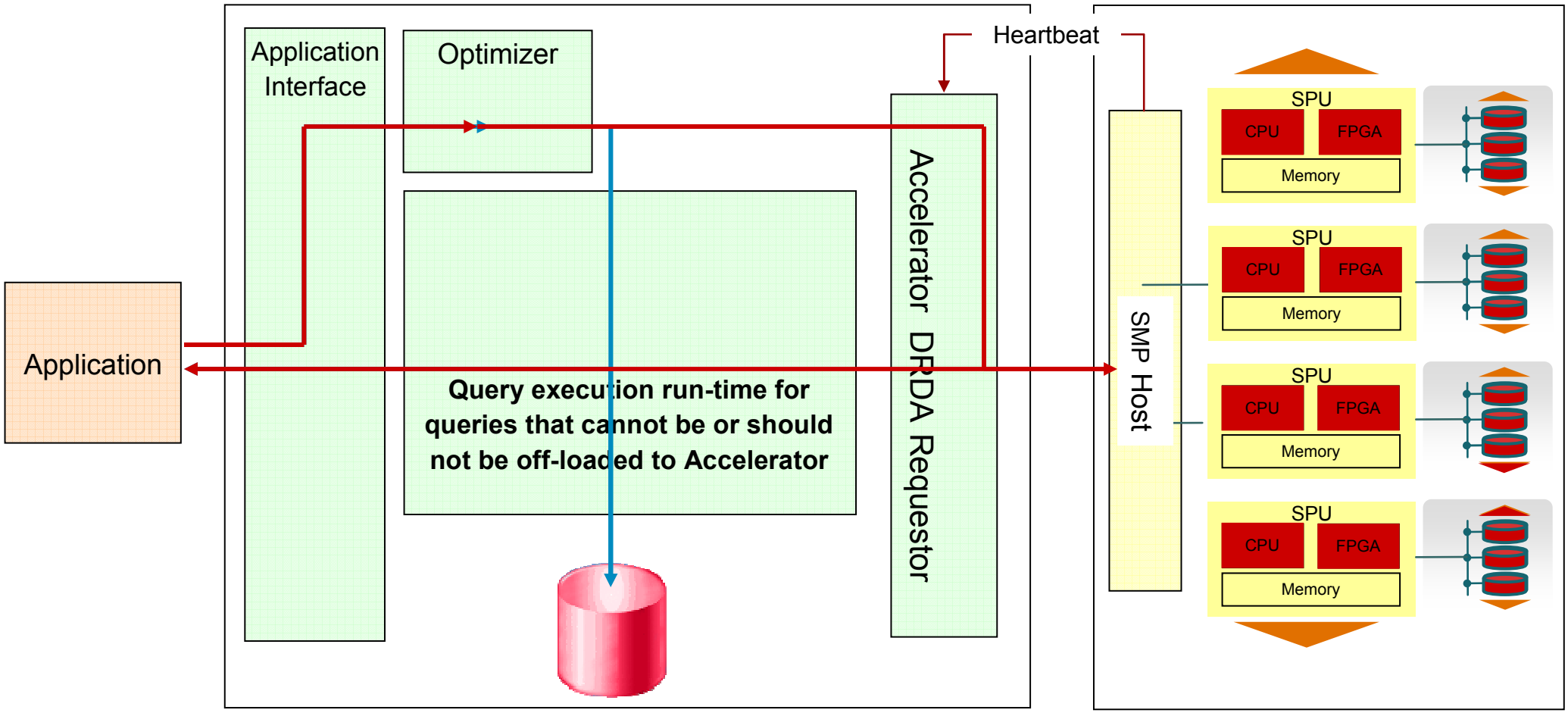
Deep integration with DB2 for z/OS:

- Automatic query routing
- EXPLAIN output
- Messaging
- Monitoring

IBM DB2 Analytics Accelerator

Superior performance for complex, data-intensive queries

Query Execution Process Flow



DB2 for z/OS

DB2 Analytics Accelerator

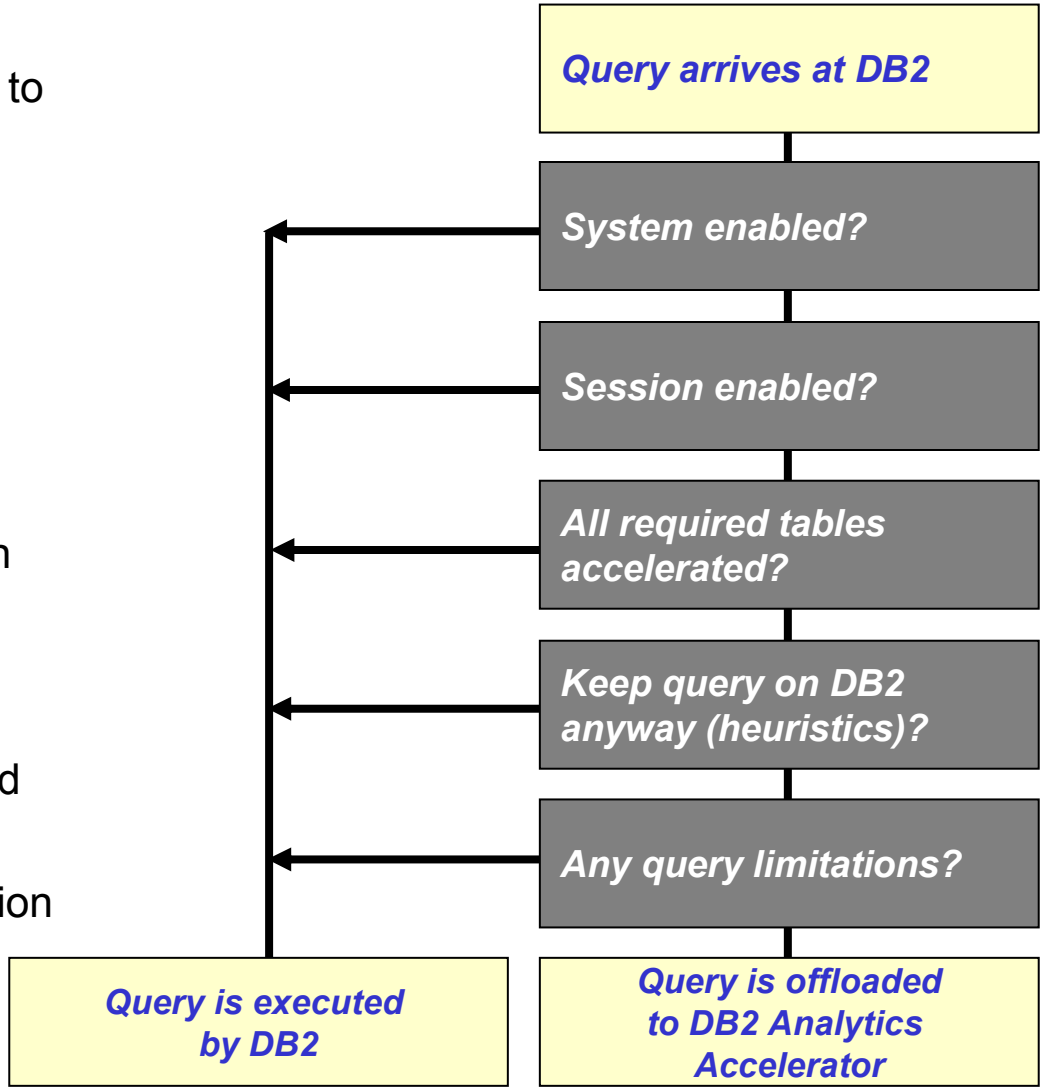
Queries executed without Accelerator

Queries executed with Accelerator

Heartbeat (availability and performance indicators)

Automatic Routing Criteria

- DB2 Optimizer decides if query should be sent to accelerator
  - Special Register
  - QUERY ACCELERATION
    - NONE
    - ENABLE
    - ENABLE WITH FAILBACK
    - ELIGIBLE
    - ALL
  - Set via zPARM, Set Statement, Connection Properties / Driver, SQL Pre-Pend, DB2 PROFILEs, BIND Option
- Whole query, not parts of query are accelerated
- Only read queries are considered for acceleration
- Both static and dynamic queries can be accelerated*



* Routing for static queries determined at bind time

Customer "A" Example:

270 of the Mixed Workload Queries

Executes in DB2 returning results in seconds or sub-seconds

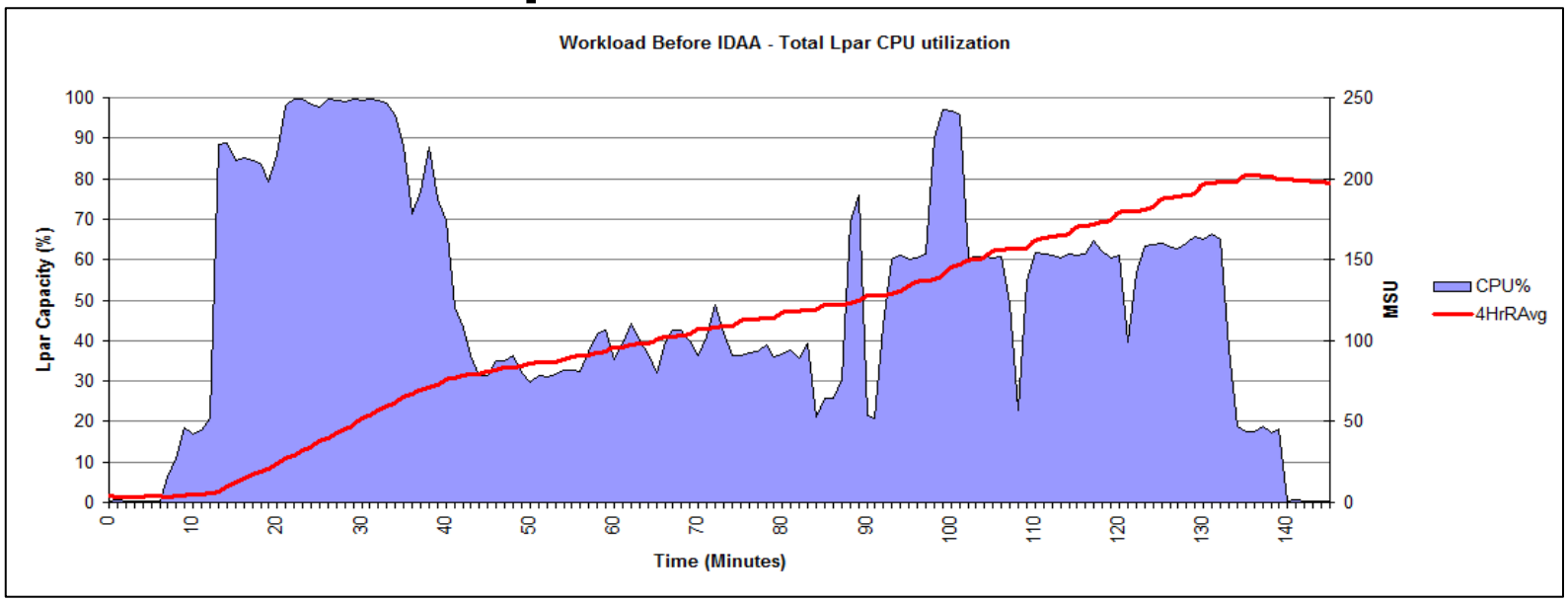
30 of the Mixed Workload Queries took minutes to hours

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

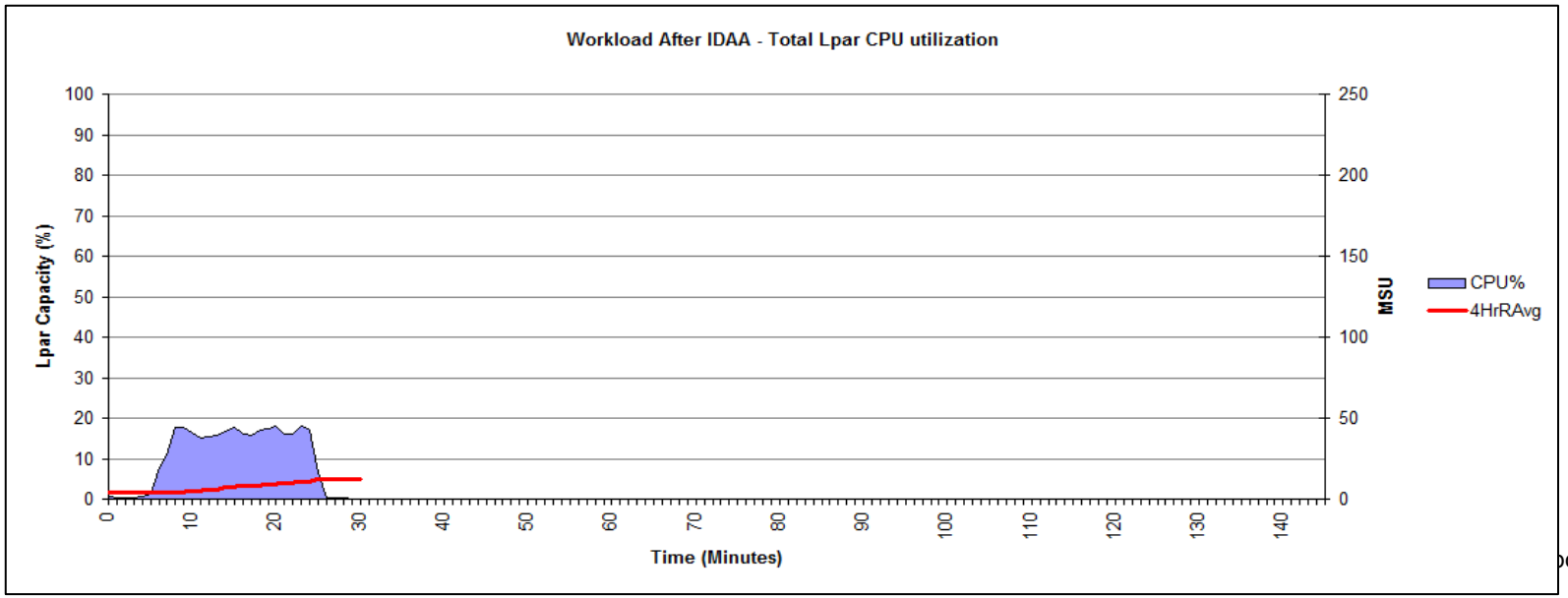
Successfully accelerated the problem queries without affecting the rest

LPAR CPU utilization comparison

Without accelerator



With accelerator



Fast Evolution of IBM DB2 Analytics Accelerator

- Version 1
  - IBM Smart Analytics Optimizer
  - In-memory, column-store, multi-core and SIMD algorithms
  - Discontinued and replaced by IBM DB2 Analytics Accelerator

- Version 2
  - New name: IBM DB2 Analytics Accelerator
  - Incorporates Netezza query engine
  - Preserves key V1 value propositions and adds many more

- Version 3
  - Better performance, more capacity
  - Incremental update
  - High Performance Storage Server

- Version 4
  - Much broader acceleration opportunities
  - More enterprise features

Version 4.1 at a glance

More Query Acceleration

- Static SQL
- DB2 11 (2)
- Multi-row fetch from local applications
- EBCDIC and Unicode in the same DB2 system and Accelerator
- NOT IN and ALL predicates (3)
- FOR BIT DATA support (3)
- 24:00:00 time value (3)
- MEDIAN support (3)
- SELECT INTO for static SQL support (4)

Enhanced Capabilities

- Improved scalability of Incremental Update
- Better performance of Incremental Update
- Improved performance for large result sets (2)
- Better access control for HPSS archived partitions
- HPSS archiving to multiple Accelerators
- Extending WLM support to local applications
- Rich system scope monitoring
- Reporting prospective CPU cost and elapsed time savings
- Separation of duties for Accelerator system administration operations
- Support for N2002 hardware (3)
- Incremental Update continues replicating even for tables in AREO state (3)
- Loading from flat file or image copy or as of any point in time (1)
- Loading in parallel to DB2 and Accelerator or into Accelerator only (1)
- Improved load performance (4)
- EXPLAIN of dynamic queries against a specific Accelerator (4)

Improved Transparency

- Automatic workload balancing with multiple Accelerators
- New RTS 'last-changed-at' timestamp (2)
- Automated NZKit installation
- Built-in Restore for HPSS
- Protection for image copies created by HPSS archiving process
- Profile controlled special registers (2)
- Improved continuous operations for Incremental Update
- Refreshing DB2 Analytics Accelerator table without table lock even if incremental update active (3)
- Static SQL and workload balancing enablement migration tool (3)

Enabling new use cases

(1) – delivered by a separate tool
(2) – DB2 11 only
(3) – DB2 Analytics Accelerator Version 4.1 PTF2
(4) – DB2 Analytics Accelerator Version 4.1 PTF3

Customer Observations and Experiences

What customers are saying…

- IBM says queries can run up to 2000x faster with the Accelerator, but we had one query run 4800x faster – from 4 hours to 3 seconds
- Our users call DB2 Analytics Accelerator the Magic Box
- Without acceleration, queries would take from several minutes to never returning – with DB2 Analytics Accelerator, queries return in less than 1 minute (usually 15 seconds)
- Whatever you paid for this, it was well worth it!
- It is unbelievable that there are still DB2 for z/OS shops out there without IBM DB2 Analytics Accelerator

Trends…

in the current customer set…

65%

- Are from the Finance and Insurance industries

39%

- Chose to purchase a full rack Accelerator or larger

25%

- Purchase more than one Accelerator

DB2 Analytics Accelerator – Four Usage Scenarios

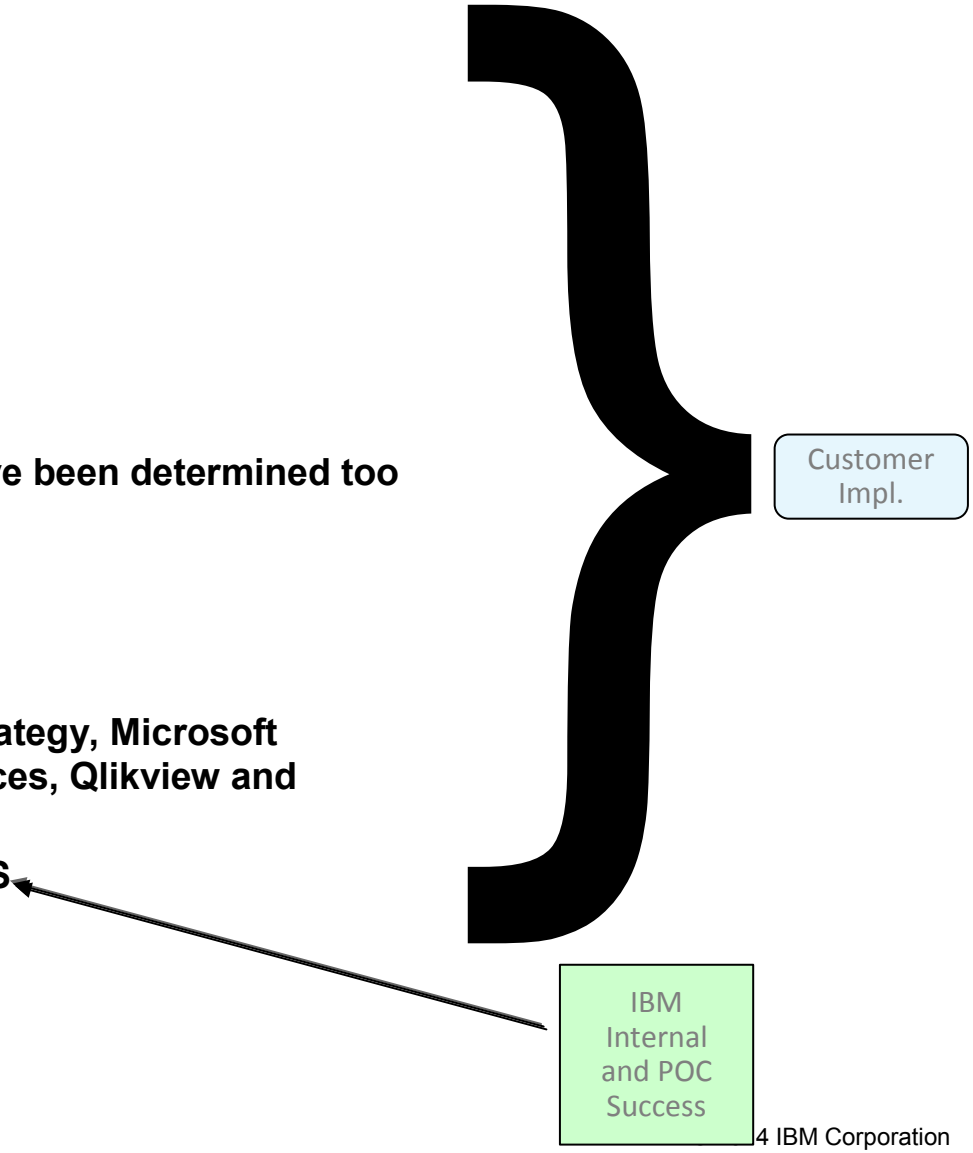
Understand your workload and data:

On average, **70%** of the data that feeds data warehousing and business analytics solutions **originates on the System z platform** (financial information, customer lists, personal records, manufacturing…)

	Where transaction source data is being analyzed today	Use Case	Benefits
1	If the data is analyzed on the mainframe	Rapid Acceleration of Business Critical Queries	*Performance improvements* and *cost reduction* while retaining System z security and reliability
2	If the data is offloaded to a distributed data warehouse or data mart	Reduce IT Sprawl for analytics	*Simplify* and *consolidate* complex infrastructures, *low latency*, *reliability*, *security* and *TCO*
3	If the data is not being analyzed yet	Derive business insight from z/OS transaction systems	One integrated, *hybrid platform*, optimized to run *mixed workload*. *Simplicity* and *time to value*
4	If the analysis is based on a lot of historical data	Improve access to historical data and lower storage costs	*Performance improvements* and *cost reduction*

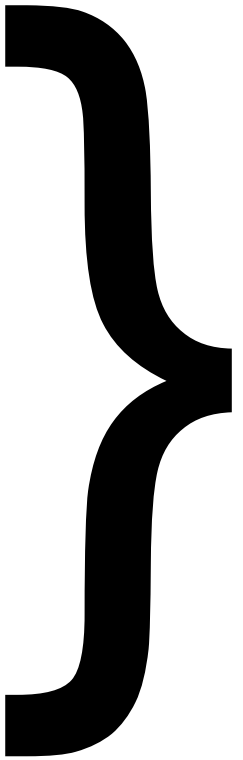
1 Rapid Acceleration of Business Critical Queries

- Accelerate data warehouse queries
- Accelerate scheduled reports
- Lower the cost of batch billing cycles
- Lower the cost of distributed DW and Mart extracts
- Lower the cost of SAS
- Meet ETL Service Level Agreements
- Accelerate operational reporting
- Improve customer ad-hoc experience
- Queries being killed by the Resource Limit Facility or have been determined too expensive to run
- Call Center Analytics
- Retail warehouse Inventory Analysis
- Cognos, Business Objects and QMF reports
- Customer is using other tools such as WinSQL, MicroStrategy, Microsoft Reporting Services, Microsoft SQL Server Analytic Services, Qlikview and Jaspersoft
- SAP and PeopleSoft customers with data on DB2 for z/OS



2 Reduce IT Sprawl for Analytics

- Bringing an ODS back to System z Accelerate scheduled reports
- Eliminate data marts and query the data warehouse directly
- Campaign Management (Unica)

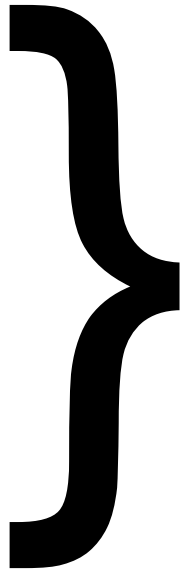


Customer Impl.

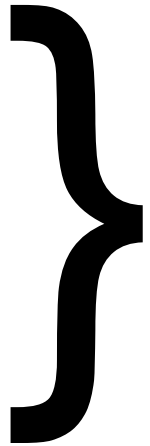
Less Bullets, But Huge Impact and Customer Commitment!

3 Derive business insight from z/OS transaction systems

- Analytics with the transactional detail data
- Virtual Data Warehouse (reporting via operational systems)
- Web Log Analytics
- Capacity Management Analytics
- Debit Card ODS Analysis
- Retail warehouse Inventory Analysis
- Store / Provider / Customer xyz “fuzzy” search



Customer Impl.



Sold / POC Success

4 Improve access to historical data and lower storage costs

- **High Performance Storage Saver**
  - Customer Investigation Noticeably Increasing
  - DB2 Performance Data / BigInsights coming into play too

Customer Impl.

zEnterprise Analytics System 9700 / 9710

A cost-competitive, integrated combination of hardware, software and services to deliver business reporting and business critical analytics

zEnterprise Analytics System 9700

zEnterprise EC12

zEnterprise Analytics System 9710

zEnterprise BC12

Preselected	Pretested	Solution Priced
To deliver a comprehensive, end-to-end, flexible solution	To meet business critical analytic demands	For deployment as an add-on to an existing zEnterprise system or as a new system

A Flexible Deployment Model

Forward Thinking Customers Leverage the 9700

9700 selected because…

- Workload Balancing across SMP and MPP
- Existing Security model
- Use of optimizer allows data to be used for both analytical and operational reporting purposes, if needed
- Leverage existing skill sets and tools
- Allows for quick ad-hoc research with existing operational tables, when needed

Forward Thinking Customers Leverage the 9700

9700 selected because…

Client Data Reporting… Dashboards, Detail and aggregated reports, Ad-hoc queries

- Improve report and ad-hoc experience where usability is tied to performance
- Improve Extract, Transform, Load (ETL) performance
  ETL jobs consistently missed SLA
- Reduce CPU consumption on z/OS
- And do it all without changing the data model or deploying on lower cost, but lower QoS platform…
- “What are you thinking? This is customer facing!”

Deepen Your Knowledge

Education and Essential Links

- Proof of Technology Events
- Primary Product Page
  - http://www.ibm.com/software/products/en/db2analacceforzos
- Redbook
  - http://www.redbooks.ibm.com/redpieces/abstracts/sg248213.html?Open&ce=ism0062&ct=swg&cmp=ibmsocial&cm=h&cr=im&ccy=us
- Customer Testimonials
  - https://engage.vevent.com/index.jsp?eid=556&seid=68284&code=brand
- zBig Data & Analytics Community
  - http://searchbusinessanalytics.techtarget.com/bigdataanalytics

Socialize

Socialization Steps

- Work with IBM to drive Business / Application Use Cases

- Companion documents available
  - One is a short summary (slides geared to application folks) that discusses the Accelerator from their point of view.
  - Second is a questionnaire regarding potential use cases that can be filled out and sent back.
    - The only thing a DBA would really have to do is add some data source examples in the questionnaire in case the recipients don't know if their application is touching DB2 z/OS data.
    - In some cases, these have been tailored by industry.

Example Slide: A wide variety of use cases

- Claims reporting
- Risk analysis for underwriting
- Fraud analysis
- Catastrophe analysis
- Provider searching
- Customer churn reduction
- Increased cross-sell
- Web interaction analysis
- General month end processing/ad-hoc querying capabilities
- Historical data analysis (data that is typically archived) for understanding trends

Example Slide: Where to look for workloads

- Extending the use of operational data for business analysis, embedding operational analytics in other applications, or daily business intelligence reporting (Cognos, QMF, Business Objects)
- Long-running DB2 for z/OS queries (e.g. > 5 seconds) with at least one of the following: WHERE, GROUP BY, ORDER BY, aggregate functions
  - Queries run from a business intelligence environment that provide important business information.
  - Queries that are scheduled in batch processes overnight to not affect corporate users during the day.
- Forgotten queries that are no longer run because of performance issues.
- Analytical and ad hoc queries that may not currently run in DB2 for z/OS but could provide significant value to the end user or organization.
- Consolidating data and analytic workloads to a single secure data environment, thereby reducing organizational costs from maintenance, ETL, administration, licensing.
- Move less frequently used tables or table partitions to the Accelerator and remove the data from DB2 for z/OS.
  - Reduce the cost of storing, managing, and processing historical data while achieving improved response times for queries against the historical data.
- Analyzing non-DB2 data (e.g. IMS, VSAM, flat files) on the accelerator.
  - Users can load non-DB2 data into the accelerator in cases where the user would like to analyze this data.

Demo Time, If Time Permits

Gary Crupi, IBM Distinguished Engineer, Big Data & Analytics
gary.crupi@us.ibm.com