IBM DB2 Analytics Accelerator
Trends and Directions

Namik Hrle
IBM Fellow

Peter Bendel
IBM STSM
IBM’s statements regarding its plans, directions, and intent are subject to change or withdrawal without notice and 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.
Agenda

- Level-set: What is IDAA and the recent enhancements
- Big picture: Long-term strategy
- Work in progress: Top themes
Agenda

- Level-set: What is IDAA and the recent enhancements
  - Big picture: Long-term strategy
  - Work in progress: Top themes
IBM DB2 Analytics Accelerator

The hybrid computing platform on System z

- Supports transaction processing and analytics workloads concurrently, efficiently and cost-effectively
- Delivers industry leading performance for mixed workloads
- The unique heterogeneous scale-out platform
- Superior availability, reliability and security

DB2 Accelerator personalities

- Turbo charged access path with hardware assisted early filtering
- Full-width index
- Specialty engine
- Archive
- Tablespace
- ETL/ELT and in-database analytics acceleration
- Integration hub: fast federated joins across heterogeneous sources
- Hybrid cloud
V5.1 Highlights

- **Accelerator-only tables** can benefit statistics and analytics tools that use temporary data for reports. They enable acceleration of data transformations implemented via SQL statements. Storing interim results in accelerator-only tables enables subsequent queries or data transformations to process all relevant data on the accelerator with high speed.

- **In-database analytics** capabilities enable acceleration of predictive analytics applications. This enables SPSS/Netezza Analytics data mining and in-database modeling to be processed within the Accelerator.

- **System Temporal and Bi-temporal tables support**
  - Active and history table loaded to Accelerator
  - Timestamp(12) data type now supported
    - Truncation to timestamp(6), new options for ZPARM QUERY_ACCEL_OPTIONS
  - Queries using temporal SQL expressions can be routed to Accelerator

- **Transparent archive tables support (since V4PTF3)**
  - Archive-enabled and archive table table loaded to Accelerator
  - Archive table older partitions can be archived in IDAA

- **Encryption of data-in-motion**
  - Network traffic includes:
    - DRDA traffic (e.g. Queries, Table Loads)
    - Configuration Console traffic
    - Incremental Update traffic
  - Requires configuration and enablement on z Systems and on the Accelerator

- **Incremental update operation enhancements**
  - Ability to disable query acceleration automatically for suspended tables
  - Continuous replication and re-load of replicated tables without taking a table lock on the source

- **Improved installation and maintenance process**
  - Extended compatibility between stored procedure version level and accelerator version level
  - Enhanced product packaging

- **More supported SQL functions**
  - Date/Time arithmetic functions
  - Inline scalar UDFs
V6.1 Highlights

- **IDAA on Cloud**
  - Overwhelming general market trends and IBM focus
  - Demonstrating z analytics vitality
  - Fast deployment, quick time to value
  - Basis for hybrid cloud offering
  - Basis for change of accelerator engine (dashDB)
  - Gaining necessary new skills

- **Mix and match**
  - Simultaneous deployment of on-premises and cloud IDAA servers

- **R support in IDAA on prem**
  - Support for data scientists most popular tool
  - Enabling in-database analytics

---

© 2017 IBM Corporation
IBM DB2 Analytics Accelerator Loader for z/OS v2.1

- **Extending analytic capabilities by bringing non-DB2 data to the accelerator and z Systems**
  - Significant cost and time reduction by eliminating manual ETL processes of non-DB2 data
  - Insight into more data types such as IMS, VSAM, sequential files, Oracle, Adabas, etc.
  - Accelerate operational analytics by supporting SMF, RMF, Syslog, ... data

- **Ensure High Availability for critical analytics applications**
  - Loading multiple accelerators in parallel: a single invocation and a single unload from DB2
  - Applies to non-DB2 sources as well

- **Add data to existing table to avoid reloading entire table**

- **z Systems synergy: zIIP usage**
Agenda

- **Level-set:** What is IDAA and the recent enhancements
- **Big picture:** Long-term strategy
- **Work in progress:** Top themes
1. Bring analytics to z data
   - DB2 for z/OS as the control point
     - Deeply integrated and transparent to typical personas that use database
     - Integration hub for z and non-z data
   - Industry leading:
     - Complex query performance
     - In-database transformation acceleration
     - Archiving

2. Hybrid platform architecture
   - Workload-optimized ‘system of systems’
   - Best of breed for each component
   - Hybrid analytics/transactional processing (HTAP)
Strategy in 2016 and beyond

1. Bring analytics to z data
   - DB2 for z/OS as the control point
     - Deeply integrated and transparent to typical personas that use database
     - Integration hub for z and non-z data
   - Industry leading:
     - Complex query performance
     - In-database transformation acceleration
     - Archiving

2. Hybrid platform architecture
   - Workload-optimized ‘system of systems’
   - Best of breed for each component
   - Hybrid analytics/transactional processing (HTAP)

3. Support hybrid cloud
   - Uniform experience, simultaneous use and easy transition between on-premises, private cloud and public cloud implementations

4. Make z data simple
   - Make z data easily consumed by data scientists, data engineers, business analysts and application developers by their tools of choice
   - Watson Data Platform integration
Agenda

- **Level-set:** What is IDAA and the recent enhancements
- **Big picture:** Long-term strategy
- **Work in progress:** Top themes
  - Hybrid Transactional/Analytical Processing
  - New generation accelerator engine
  - Federated access support
Best of breed HTAP

- Hybrid Transactional/Analytical Processing
  - Real-time processing on real-time data

- Internal code name *Unified Store*
  - DB2/IDAA team has pioneered this technology
  - Based on patented, research prototype of *just-in-time* currency replication protocol

- Unique in industry, heterogeneous scale-out solution for enterprise-grade HTAP
  - Superior to the competition
Basic idea

Reading most recent committed data during asynchronous replication
Introducing a new zparm / bind option / special register
- CURRENT QUERY ACCELERATION WAITFORDATA = n.m
- 0.0 - 3600.0 seconds, 0.0 = no wait, default: 0.0
  0.0:
  - immediately execute in accelerator without waiting
  > 0.0:
  - wait for most recently committed changes before query to be applied by asynchronous replication and then run the query
  - or: fail the query if max wait time specified exceeded
  - or: execute query in DB2 if “WITH FAILBACK” is specified in CURRENT QUERY ACCELERATION special register
CURRENT QUERY ACCELERATION = ENABLE WITH FAILBACK
and query is routable to IDAA

1. Execute in IDAA
2. Wait for data
3. Check if most recent data is available in less than n seconds
   - Yes: Proceed with query acceleration
   - No: Execute in DB2
Implementation roadmap

- **Stage 1:**
  - Phase 1: Delay protocol
  - Phase 2: Conditionally skip delay protocol if only unchanged tables referenced in query
  - Phase 3: Partial apply

- **Stage 2:**
  - Phase 1: DB2 to IDAA optimized replication protocol
  - Phase 2: dashDB apply component optimizations (goal: few seconds staging)
  - Phase 3: Read-own-writes semantics support
Agenda

- **Level-set:** What is IDAA and the recent enhancements

- **Big picture:** Long-term strategy

- **Work in progress:** Top themes
  - Hybrid Transactional/Analytical Processing
  - New generation accelerator engine
  - Federated access support
Powered by dashDB: The new acceleration engine for IDAA

- IBM’s premier analytics engine
- Latest technology innovations
- Improved SQL compatibility and performance
- Potential for faster ingest for incremental updates
- Higher degree of concurrent users and queries
- Compatible to existing DB2 Analytics Accelerator installations (co-existence)
- Rich in-database analytics capabilities
- Future Spark integration
Scale Up CPU, Memory, IO according to your requirements and infrastructure availability
Hardware appliance

IDAA based on next generation PDA

Delivered as part of IDAA Solution
IBM Secure Service Container

Docker container

IDAA server

dashDB engine

Additional future functionality

Systems Manager

Workload Monitoring

Authentication

Linux + management

PR/SM LPAR

CPU

Memory

Storage (SAN, NAS)

Filesystem

Customer’s Storage Management

Existing Components

Delivered as part of IDAA Solution
Public cloud offering

- Docker container
  - IDAA server
  - dashDB engine
  - Additional future functionality
  - Systems Manager
  - Workload Monitoring
  - Authentication

- Linux
- SL Virtual or physical server
  - CPU
  - Memory

- Storage (local, SAN, NAS)
  - (Clustered) Filesystem

Delivered as part of IDAA Solution
Software Defined Environment

Hosting IDAA on private cloud or on-premises data center

- **Docker container**
  - IDAA server
  - dashDB engine
  - Additional future functionality
  - Systems Manager
  - Workload Monitoring
  - Authentication

- **Linux**
- **Virtual or physical server**
  - CPU
  - Memory
- **Storage (local, SAN, NAS)**
  - (Clustered) Filesystem

**Existing Components**
**Delivered as part of IDAA Solution**
Uniform experience, simultaneous use, and easy transition between on-premises, private cloud and public cloud implementations

Common accelerator engine across all the platforms: dashDB
Agenda

- Level-set: What is IDAA and the recent enhancements
- Big picture: Long-term strategy

- Work in progress: Top themes
  - Hybrid Transactional/Analytical Processing
  - New generation accelerator engine
  - Federated access support
Federated access support

... what works today
Federated access support

… what works today
Federated access support

… work in progress
Thank You