An Introduction to IBM’s pureQuery

Bryan F. Smith  bfsmith@us.ibm.com
IBM

Session Code: Xxx
Thursday, March 3, 2011: 6:00 PM-7:00 PM
ACC, Room 211A
Disclaimer

© Copyright IBM Corporation [current year]. All rights reserved.
U.S. Government Users Restricted Rights - Use, duplication or disclosure restricted by GSA ADP Schedule Contract with IBM Corp.

THE INFORMATION CONTAINED IN THIS PRESENTATION IS PROVIDED FOR INFORMATIONAL PURPOSES ONLY. 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. IN ADDITION, THIS INFORMATION IS BASED ON IBM’S CURRENT PRODUCT PLANS AND STRATEGY, WHICH ARE SUBJECT TO CHANGE BY IBM WITHOUT NOTICE. IBM SHALL NOT BE RESPONSIBLE FOR ANY DAMAGES ARISING OUT OF THE USE OF, OR OTHERWISE RELATED TO, THIS PRESENTATION OR ANY OTHER DOCUMENTATION. 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 ANY AGREEMENT OR LICENSE GOVERNING THE USE OF IBM PRODUCTS AND/OR SOFTWARE.

IBM, the IBM logo, ibm.com, DB2, Optim, Tivoli, Rocket Software, and Data Studio are trademarks or registered trademarks of International Business Machines Corporation in the United States, other countries, or both. If these and other IBM trademarked terms are marked on their first occurrence in this information with a trademark symbol (® or ™), these symbols indicate U.S. registered or common law trademarks owned by IBM at the time this information was published. Such trademarks may also be registered or common law trademarks in other countries. A current list of IBM trademarks is available on the Web at “Copyright and trademark information” at www.ibm.com/legal/copytrade.shtml

Other company, product, or service names may be trademarks or service marks of others.
Abstract

• There's a surge of interest from people who are becoming aware of the cost savings and management benefits of pureQuery. The client optimization capability of pureQuery enables DBAs to enable static execution for any existing Java™, .NET, or CLI application without changing any application code. This presentation will show you how to client optimize an existing application and provide hints and tips for managing the process and making it more efficient. DB2 for z/OS and DB2 for LUW environments are covered.
Agenda

- Static SQL binding for Java, .NET, and CLI
- Correlate SQL to java source code
- SQL execution statistics
- SQL replacement and lockdown
- Literal consolidation and replacement
- Privacy visibility
pureQuery

- How Java, .NET and CLI applications can be managed like static SQL applications
  - What problems do these application platforms present to the DBA?
    - ORM frameworks obscure the SQL generated
    - Dynamic SQL
  - Capturing SQL for review/tune/revisement
  - Converting to static SQL and the benefits of doing so
Frequently Cited Concerns

I have more and more Java workload coming into my data server driving up costs, but the budget is not keeping pace.

I don’t even want to allow framework-generated SQL on my database. If I can’t see it, I don’t know how it will impact me.

Java, .NET, and CLI performance problems are a real pain to resolve because I can’t even tell what application issued the SQL.
Contemporary Application Stack Challenges

- Simplify development, but …
  - Challenge problem resolution
  - Impact performance
  - Obscure impact analysis
  - Impede capacity planning

![Application Server and Database Server Diagram]
Introducing pureQuery

A high-performance, data access platform to simplify developing, managing, securing, and optimizing data access for new and existing applications.

**pureQuery Components:**

- Optim Development Studio
  - Integrated development environment with Java and SQL support
  - Improve problem isolation and impact analysis

- Simple and intuitive API
  - Enables SQL access to databases or in-memory Java objects
  - Facilitates best practices
  - Enables heterogeneous batching

- Optim pureQuery Runtime
  - Flexible static SQL deployment for DB2
Three steps

1. Capture the SQL
   - Use pureQuery API, generate from WebSphere JPA, or capture while executing
   - Use with custom-developed, framework-based, or packaged applications
   - Translation of literals to host variables (new in 2.2)

2. Bind SQL to DB2 (Client Optimization)
   - Use tooling in Data Studio Developer, WAS console or command line

3. Choose execution mode
   - Dynamic or static
   - Choose at deployment time instead of development time

Static SQL value
- Make response time predictable
  - Lock in the SQL access path pre-execution
- Limit user access and reduce SQL injection
  - Grant execute privileges on the query packages rather than access privileges on the table
- Accelerate problem resolution
  - Trace SQL execution to a specific package and the originating source
- Improve impact analysis and capacity planning
  - Visualize application SQL and correlation metadata
- Increase system capacity
  - Drive down DB cycles

"The ability to use static SQL with pureQuery is huge. Recently, I worked with a client who could reduce CPU usage by 7 percent thanks to this one feature."

— David Beulke, Pragmatic Solutions Inc.
Dynamic vs. Static Execution

**Dynamic SQL Full Prepare**
- Check auth for package / plan
- Parse SQL Statement
- Check Table / View Auth
- Calculate access path
- Store access path in a temporary package
- Execute SQL statement

**Dynamic SQL Short Prepare**
- Check auth for package / plan
- Copy skeleton from cache to local DB2 thread storage
- Execute SQL statements

**Static SQL**
- Check auth for package / plan
- Extract access path from SYSIBM.PACKAGES and STATEMENTS
- Execute SQL statements
Improving Throughput with pureQuery, a z/OS Example

- In-house testing shows **over 40% reduction** in CPU costs over dynamic JDBC using pureQuery and DB2 for z/OS (type-2)
  - Read [IBM Optim pureQuery Runtime for z/OS Performance](#)
  - IRWW – an OLTP workload, cache hit ratio between 70 and 85%

Preset application path never varies at runtime

Reduced CPU per transaction means more growth on existing resources
Improving Throughput with pureQuery

- In-house testing shows almost a doubled throughput over the existing JDBC application.
- The new application that uses pureQuery method style API more than doubled the database transaction throughput.

90% cache hit ratio

95% cache hit ratio
Java Database Access and pureQuery
Many on-ramps for new and existing apps

DB2 (z/OS, i & LUW), Informix, and Oracle now
More coming
Deploying with pureQuery Runtime

**Application tier:**
- z/OS, Linux, UNIX, Windows

**Database tier:**
- DB2 for z/OS
- DB2 for i
- DB2 for Linux, UNIX, and Windows
- Informix Dynamic Server
- Oracle
Unique Package Names Improves PD

- Most dynamic Java or .NET applications use packages SYSLNx00 or SYSshxxyy making it hard to identify specific programs.
Static SQL for security – administering table privileges

- Dynamic SQL
  - Table privileges granted directly to users(groups)/secondary authids
    - Security exposure and administrative burden

```
GRANT SELECT ON TABLE PAYROLL TO DEPT_D47
```

```
<table>
<thead>
<tr>
<th>NAME</th>
<th>POSITION</th>
<th>SALARY</th>
<th>...</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
```

- Static SQL (pureQuery)
  - Users get no table privileges

```
GRANT EXECUTE ON PACKAGE POSITION_REPORT TO DEPT_D47
```

```
<table>
<thead>
<tr>
<th>NAME</th>
<th>POSITION</th>
<th>SALARY</th>
<th>...</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
```

```
PACKAGE SELECT NAME, POSITION FROM PAYROLL...
```

```
<table>
<thead>
<tr>
<th>NAME</th>
<th>POSITION</th>
<th>SALARY</th>
<th>...</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
```
More Visibility and Control of Application SQL

- Capture SQL
- Share, review, and optimize SQL
- Revise and validate equivalency
- Bind for static execution or run dynamically
- Restrict SQL to eliminate SQL injection
Capture Application SQL: At Development or Later

Capture → Review → Optimize → Revise → Restrict

Three methods
1. Use pureQuery API
2. Use JPA and generate the pureQuery file
3. Set captureMode=ON and execute the program

IBM Optim Development Studio
Visualize Application and SQL Metadata

- Review the captured SQL
- View metrics about execution frequency and duration
- Share captured SQL with DBA
SQL Outline

Speed up PD – even when using frameworks

- Capture application-SQL-data object correlation (with or without the source code)
- Trace SQL statements to using code for faster problem isolation
- Enhance impact analysis identifying application code impacted due to database changes
- Answer “Where used” questions like “Where is this table used within the application?”
- Use with modern Java frameworks e.g. Hibernate, Spring, iBatis, OpenJPA
Optimize SQL

Capture → Review → Optimize → Revise → Restrict

- Launch Visual Explain
- Copy SQL to Optim Query [Workload] Tuner

View SQL detail

Launch Query Tuner for frequently run queries
Optimization Service Center (the fate of) (DB2 for z/OS only)

- Current OSC remains supported through DB2 9
- OSC functionality will be split among the following offerings:
  - Data Studio (no charge)
  - Query Tuner (single query tuning capabilities already overlap with Q[W]T today)
  - Base function in DB2 for z/OS customers
    - SQL environment capture
    - Profile monitoring
## Contents of Eclipse-based Query Tuning offerings

<table>
<thead>
<tr>
<th>Feature</th>
<th>Data Studio</th>
<th>Optim Query Tuner for z/OS</th>
<th>Optim Query Workload Tuner for z/OS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Queries from all sources</td>
<td><img src="true" alt="Yes" /></td>
<td><img src="true" alt="Yes" /></td>
<td><img src="true" alt="Yes" /></td>
</tr>
<tr>
<td>Reports</td>
<td><img src="true" alt="Yes" /></td>
<td><img src="true" alt="Yes" /></td>
<td><img src="true" alt="Yes" /></td>
</tr>
<tr>
<td>Query Formatter</td>
<td><img src="true" alt="Yes" /></td>
<td><img src="true" alt="Yes" /></td>
<td><img src="true" alt="Yes" /></td>
</tr>
<tr>
<td>Access Plan Graph</td>
<td><img src="true" alt="Yes" /></td>
<td><img src="true" alt="Yes" /></td>
<td><img src="true" alt="Yes" /></td>
</tr>
<tr>
<td>Query Statistics Advisor</td>
<td><img src="true" alt="Yes" /></td>
<td><img src="true" alt="Yes" /></td>
<td><img src="true" alt="Yes" /></td>
</tr>
<tr>
<td>Query Annotation</td>
<td></td>
<td><img src="true" alt="Yes" /></td>
<td><img src="true" alt="Yes" /></td>
</tr>
<tr>
<td>Visual Plan Hint</td>
<td></td>
<td><img src="true" alt="Yes" /></td>
<td><img src="true" alt="Yes" /></td>
</tr>
<tr>
<td>Query Index Advisor</td>
<td></td>
<td><img src="true" alt="Yes" /></td>
<td><img src="true" alt="Yes" /></td>
</tr>
<tr>
<td>Query Advisor</td>
<td></td>
<td><img src="true" alt="Yes" /></td>
<td><img src="true" alt="Yes" /></td>
</tr>
<tr>
<td>Access Path Advisor</td>
<td><img src="true" alt="Yes" /></td>
<td><img src="true" alt="Yes" /></td>
<td><img src="true" alt="Yes" /></td>
</tr>
<tr>
<td>Workload Statistics Advisor</td>
<td></td>
<td><img src="true" alt="Yes" /></td>
<td><img src="true" alt="Yes" /></td>
</tr>
<tr>
<td>Workload Index Advisor</td>
<td></td>
<td><img src="true" alt="Yes" /></td>
<td><img src="true" alt="Yes" /></td>
</tr>
<tr>
<td>Workload Query Advisor</td>
<td></td>
<td><img src="true" alt="Yes" /></td>
<td><img src="true" alt="Yes" /></td>
</tr>
</tbody>
</table>
Revise SQL Without Modifying the Application

Capture → Review → Optimize → Revise → Restrict

Tooling validates equivalent SQL

Shows replaced

Edit or delete SQL
Eliminate SQL Injection

- Restrict SQL execution to only those statements captured
- Set capturedOnly=true in pdq.properties
- pureQuery Runtime looks for it in the classpath
Optimize for WebSphere and DB2 with pureQuery

Capture metadata from existing applications
• Capture from JPA without executing
• Derive performance, costs, security and manageability value

Reduce HW and SW costs
• Up to 42% lower CPU/Trans
• Move workload to zIIP and zAAP

Replace SQL without changing the source
• Editor validates equivalency

Speed up problem resolution
• Trace SQL back to line of code in the application

Prevent SQL Injection
• Lock down SQL for dynamic or static execution

Jump start application design
• Generate SQL and Code from Database Objects
• Setup basic DAO Pattern

Enhance development productivity
• Code generation, content assist
• Database aware, Java SQL Editor

Simplify impact analysis
• Categorize by Java, SQL, Database, Packages, track back to line of code

Focus tuning efforts
• Find and sort by query elapsed time from Java

Enhance performance
• Leverage best practices, automatically for JPA
• Use static execution, automatically for JPA
• Lock in access plans for consistent performance

Reduce security exposure
• Grant access to queries, not tables
Trial: IBM Optim Development Studio and Optim pureQuery Runtime

Download Optim Development Studio, which provides an integrated database development environment for Oracle, DB2, and Informix. Optim Development Studio improves development productivity up to 50% for developing and testing SQL and XQuery queries, stored procedures, Web services, and Java data access layers.

**Trial download**

The evaluation period for this trial is 30 days. All of the product’s features are enabled.

A trial version of Optim pureQuery Runtime V2.2 is available within the Optim Development Studio V2.2 trial package for use on the same Windows or Linux machine. To try the pureQuery Runtime on another platform, contact your IBM sales representative or IBM Business Partner.

System requirements

The estimated download time using Download Director over a 1.5Mbps connection is 1 hour and 24 minutes; 43 minutes using Installation Manager.

Choose one of these options to download the trial. A Web install downloads and installs the product from a Web based repository. A local Install downloads and installs the product to your local hard drive.

<table>
<thead>
<tr>
<th>Operating system</th>
<th>Version</th>
<th>Size</th>
<th>Download method</th>
</tr>
</thead>
<tbody>
<tr>
<td>Web install using IBM installation Manager (recommended)</td>
<td>V2.2</td>
<td>480MB</td>
<td></td>
</tr>
<tr>
<td>Windows, Linux</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Web install if you already have IBM Installation Manager installed</td>
<td>V2.2</td>
<td>372MB</td>
<td></td>
</tr>
<tr>
<td>Windows</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Download it today!
Query management enhancements in version 2.2.0.3 of pureQuery Runtime available June 10th, 2010

- Configuration Database to store captured SQL and pureQuery properties. Properties and captured SQL can be stored and managed centrally in a database.
- Control pureQuery behavior by updating records in the configuration database.
- Add SQL from SQL script files and Java source files to captured SQL data. Applications can realize benefits of static package without first running capture.
- Organize SQL into separate database packages during configure. SQL that shares use the same tables can be grouped together.
- Automatically remove unused SQL during merge or configure. Obsolete SQL no longer take up space in the database packages.
- Command line and GUI to manage data stored in the repository such as create, update, remove, grant and revoke privileges to the repository tables.
**DB2 Developer Workbench vs. Data Studio**

**before**

**IBM DB2 Developer Workbench V9.1**
- SQL Query Editor
- SQLJ Editor
- SQL Builder
- XQuery Builder
- SQL Routine Debugger
- Java Routine Debugger
- XML Editor
- XML Schema Editor
- Data Management
- Visual Explain
- Project Management

Data Studio is a full replacement of DB2 Developer Workbench plus much more

- DB2 for Linux, Unix, Windows v8.x, v9.x
- DB2 for z/OS v8, v9, v10
- DB2 for i5/OS v5r2, v5r3, v5r4
- Informix Dynamic Server (IDS) v9.x, v10.x, v11

**now**

**IBM Data Studio V2.2**
- Integrated Query Editor – SQL + XQuery
- SQLJ Editor
- SQL Builder
- XQuery Builder
- SQL Routine Debugger
- Java Routine Debugger
- XML Editor
- XML Schema Editor
- Data Management
- Visual Explain
- Project Management
- ER Diagramming
- Data Distribution Viewer
- Object Management
- Browse & Update Statistics
- Security Access Control
- Connection Management integration with Kerberos and LDAP
- Data Web Services
- IDS Server Support

No-charge

Data Studio is no longer used as a brand
Data Studio: Data Modeling / App & DB Dev
A Consistent and Productive work environment

Create, Alter, Drop, Browse and Filter database objects

- Express yourself with optimal queries
  - Content assistance for database objects
  - Rapid interactive end-user feedback
  - Extensible templates
  - Multiple SQL statement testing
  - SQL assistance and XQuery assistance

ER Diagramming

Integrated Query Editor
Unleash SQL in your Java IDE

- SQL content assist

```java
// Select GOSALESCT.CUST by parameters
@Select(sql = "SELECT CUST_CODE, CUST_FRST_NAME, CUST_LAST_NAME, CUST_ADDR1," + " CUST_ADDR2, CUST_CITY, CUST_POST_ZONE, CUST_CTRY_CODE, CUST_PHN_NBR," + " CUST_INFO, CUST_EMAIL, CUST_GNDR_CODE, CUST_PROV_STATE" + " FROM GOSALESCT.CUST" + " WHERE ")
Cust getCust(int cust_code
```

Press CTRL-SPACE to invoke content assist or auto complete

- SQL validation

```java
// Select GOSALESCT.CUST by parameters
@Select(sql = "SELECT CUST_CODE, CUST_FRST_NAME, CUST_LAST_NAME, CUST_ADDR1," + " CUST_ADDR2, CUST_CITY, CUST_POST_ZONE, CUST_CTRY_CODE, CUST_PHN_NBR," + " CUST_INFO, CUST_EMAIL, CUST_GNDR_CODE, CUST_PROV_STATE" + " FROM GOSALESCT.CUST" + " WHERE CUST_CODE = ?")
Cust getCust(int cust_code
```

Table "CUST" does not contain column "CUST_COD".
Press 'F2' for focus.
Unleash SQL in your Java IDE

- Run SQL at design time without the need to compile the program
-Launch Visual Explain
Integrated editor and debugger for SQL PL and Java (JDBC or SQLJ)

Stored Procedure Support
SQL and Java Debugger
Data Management

Data Distribution Editor

- View data distribution skews with the Data Distribution Viewer

Edit Data

- Manage Roles
- Privileges By User
- Privileges By Object

Roles, Users, Privileges
IBM Data Studio

IBM Data Studio (stand-alone)

Downloads
To properly configure your download, please review the information below. Select the appropriate offering. When you are done, press the "Continue" button at the bottom.

<table>
<thead>
<tr>
<th>Offering</th>
<th>Platform</th>
<th>Format</th>
</tr>
</thead>
<tbody>
<tr>
<td>IBM Data Studio Version 2.2</td>
<td>Red Hat Linux, SUSE Linux Enterprise Desktop (SLED), SUSE Linux Enterprise Server (SLES)</td>
<td>download</td>
</tr>
<tr>
<td>Languages: English</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Languages: English</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

No-charge

Download it today!
More resources

• Community page: Integrated Data Management (Optim and Data Studio); articles, tutorials, blogs, forums …

• IBM Optim Development Studio and pureQuery Runtime forum

• pureQuery for DB2 for z/OS demo (Improving ROI for existing applications)

• Optim Performance Management Solution
Bryan F. Smith
IBM
bfsmith@us.ibm.com
Session Xxx
An Introduction to IBM’s pureQuery