Protect sensitive Hadoop data using InfoSphere BigInsights Big SQL and InfoSphere Guardium

An integrated solution for data security and compliance

Sundari Voruganti
InfoSphere Guardium Evangelist
IBM

Kathy Zeidenstein (krzeide@us.ibm.com)
InfoSphere Guardium Evangelist
IBM

Roopa Sivakumar (rosivaku@in.ibm.com)
InfoSphere BigInsights QA Architect
IBM

10 November 2014

Major advantages of using Big SQL, the SQL interface to Hadoop data within InfoSphere® BigInsights™, are its enterprise-ready capability for speed, functionality, and security. This tutorial provides a brief overview of the built-in security capabilities of Big SQL and then goes into greater depth to highlight the integration with InfoSphere Guardium®, which provides automated compliance reporting, real-time alerting, dynamic data masking, and much more.

Supported release
Support for Big SQL requires InfoSphere Guardium 9.1 (GPU 300). InfoSphere Guardium does not support any version of Big SQL prior to InfoSphere BigInsights 3.0.

Overview
InfoSphere BigInsights 3.0 comes with a highly enhanced version of its SQL processing engine, Big SQL, which offers rich SQL language support over Hadoop and a massively parallel processing architecture to provide better performance over traditional Hive with MapReduce. See Resources for more information about Big SQL.

We will briefly describe the InfoSphere BigInsights security architecture and the fine-grain access control capabilities in Big SQL. Then we describe how InfoSphere Guardium Activity Monitoring for Hadoop provides an end-to-end monitoring, reporting, and alerting solution to help integrate InfoSphere BigInsights with the rest of your security and compliance infrastructure. InfoSphere
Guardium reduces audit overhead, provides enterprise-level reporting, real-time alerting, and other features designed to reduce the operational overhead of collecting and analyzing data for audit and compliance purposes and to mitigate risk.

**Note:** InfoSphere Guardium can monitor and audit all access to InfoSphere BigInsights (and other Hadoop distributions), whether through HBase, Hive, or direct to HDFS. For information about the benefits and how to configure InfoSphere BigInsights Hadoop with Guardium, see Resources. This tutorial focuses on the latest capability of using InfoSphere Guardium with BigInsights Big SQL 3.0.

### An InfoSphere BigInsights security and auditing ecosystem

The InfoSphere BigInsights Knowledge Center has an excellent overview of the security ecosystem. We provide an abbreviated explanation here. The diagram below helps illustrate the discussion.

**Figure 1. An InfoSphere BigInsights security architecture**

*Authentication* is the first level of security on any system, and it involves validating the identity of the user trying to access the system. InfoSphere BigInsights is typically configured with a pluggable authentication module (PAM), which delegates authentication to the OS. Via these PAM modules, InfoSphere BigInsights supports LDAP, Kerberos, and whatever special authentication is configured via PAM.

**Recommendation:** Use the same authentication mechanism for Big SQL as you use for the InfoSphere BigInsights Console to simplify identity and access management.

Now, let's consider the next level of access, which is *authorization*. Following a successful authentication, a user should only be allowed access to objects and functions that the user is authorized to. To simplify access control management, the BigInsights Web console includes predefined roles to separate out tasks that system administrators can do, for example, from what a user can do.

With Big SQL, there are many granular authorities and privileges, which are also grouped under predefined authorizations. After the initial installation, the user "bigsql" has all the privileges and authorities over all Big SQL tables and objects. You can control access more finely by using the predefined authorizations. For example, one such predefined role is SECADM, who can grant and revoke all privileges and authorities. Another is DBADM, who can create, alter, and drop all non-security related objects. Administrators can grant the SECADM and DBADM roles to real users, and can even create custom roles to suit their own organizational structure.

Big SQL goes one step further for security access control for Hadoop data by providing fully implemented form of row and column-level fine-grained access control. Row-level access control...
and dynamic data masking of columns provides for stronger enforcement of security principles, such as least privilege and need-to-know basis. This is defined and managed by SECADM authority for separation of duties.

**Recommendation:** Visit the InfoSphere BigInsights Knowledge Center for recommendations and best practices on the interaction between InfoSphere BigInsights and Big SQL authorities.

**Real-time monitoring and auditing** is required for detecting real-time security events and logging relevant activity for later forensic analysis. Big SQL includes native auditing capabilities that log certain events based on policies. However, given that Big SQL is only one interface to the base Hadoop data you want to make sure you do not have holes in your auditing solution. There is only one way of doing this efficiently and effectively, which is using InfoSphere Guardium. InfoSphere Guardium complements its monitoring and audit capabilities with real-time alerting as well. The integration with InfoSphere Guardium is the subject of this article.

**Encryption** of data at rest and data in transit are critical components of security architecture. For data in transit, Big SQL supports SSL for encrypting data movement between the database client and the database engine.

For encryption of data at rest on the file system, InfoSphere Guardium Data Encryption provides the capability to encrypt data while allowing transparent access to administrators and applications (see Resources).

The InfoSphere BigInsights platform and surrounding ecosystem has a solid security infrastructure that provides the tools you need to implement defense in depth.

**Benefits of using InfoSphere Guardium with BigInsights Big SQL**

Most organizations also have regulatory requirements that require demonstrable proof of adherence to particular security and auditing practices, such as monitoring privileged user activity or auditing all changes to financial data. For these purposes, InfoSphere Guardium complements the native security capabilities quite well. Guardium can reduce the overhead for organizations that must comply with regulatory compliance mandates, such as the Payment Card Industry (PCI) or Sarbanes-Oxley (SOX) requirements. But beyond just compliance, InfoSphere Guardium provides the analytical capabilities to help organizations respond in near real time to potential data breaches.

InfoSphere Guardium is deployed in large and small organizations around the world to address the issues of data protection and compliance for a variety of databases, including data warehouses, NoSQL systems, and Hadoop distributions. The solution is flexible, powerful, and effective.

The InfoSphere Guardium architecture is designed to minimize overhead and support separation of duties. The architecture for InfoSphere Guardium is designed for minimal performance impact on production data processing. Let's briefly review how it works using the simple diagram below.
**Figure 2. InfoSphere Guardium architecture**

InfoSphere Guardium consists of two key components:

- A lightweight server-side agent known as an **S-TAP** (software TAP), that sits in the OS
- A hardened hardware or software or virtual appliance known as a **collector**, which includes a repository for audit data as well as a web-based console that can be configured for different roles

When the database server (such as Big SQL) receives a data request from a client, the S-TAP copies the network packet and sends it to the collector, where that packet is compared against the security policy that's in place.

As a matter of fact, the real intelligence of the InfoSphere Guardium system is on the collector. That is where the message is broken down into its component parts and logged into an internal repository on the collector, and any necessary actions are taken, such as generating a real-time alert, logging the activity, or blocking a particular local user. By offloading the bulk of the activity (analysis, parsing and logging) to a hardened collector, the performance impact on the database server is minimized (not more than 2-4 percent in most cases).

The types of activities that are logged include session-level information (database user, client IP, server IP, login time, etc.) and the actual SQL that the user issues. Information returned from the database server can also be logged, including SQL errors and login failures. In certain cases, the actual result set can be examined for particular patterns, such as credit card data. Depending on the rules set in the policy, different actions can be taken such as logging a policy violation, real-time alerting, and, in some circumstances, blocking or quarantining users.

The InfoSphere Guardium role-based web console provides centralized management of alerts, report definitions, compliance workflow processes, and settings (such as archiving schedules) without the involvement of database administrators, thus providing the separation of duties required by auditors and streamlining compliance activities.

**Detailed auditing and flexible reporting**

InfoSphere Guardium logs detailed and granular information, as you can see in the sample report shown below. It includes details of the client and server IPs, the source program, the database user, and the SQL command text. (Sensitive data in command text can be masked.)
An alternative to creating reports for ad hoc investigation is to use the quick search capability in Guardium. Quick search is a combination of text search and faceted browsing capability to help you narrow in on an area of interest.

For example, in Figure 4:

1. We simply entered the search term `bigsql` in the InfoSphere Guardium search bar and were automatically brought into the quick search tool, where you can see on the left the returned data in groups of `Where`, `Who`, `What`, `When` (not shown in screenshot), and `Exceptions`. 
2. We click on **Object** to retrieve all objects (tables, views, files) audited in the timeframe we specified (last day).

3. Then we clicked on the object of interest, which narrows the search results to just that object.

**Outlier detection**

So much activity data is generated that it can be difficult to know where to focus one's attention. Built-in machine learning capabilities let you train the system over time to typical workload patterns and to provide a graphical dashboard of outliers. You can drill down and investigate the details behind the outliers to determine if indeed there is a real problem. The screenshot below shows the case where a user started using different kinds of statements (DML of insert, update, and delete) which was outside the normal pattern of basic selects that were used in the past.

**Figure 5. Outliers can identify activity that is outside of normal behavior as observed over time**

![Outlier Detection Screenshot]

**Real-time and threshold-based alerting**

You can audit and alert on database exception conditions, such as detecting when a user has more than five failed logins in three minutes, which can be an indicator of a brute force attack.

**Figure 6. Excessive failed logins policy violation**

![Real-time Alert Screenshot]

Real-time policy alerts can be specified for a wide variety of conditions. Organizations have a responsibility to do whatever they can to avoid embarrassing or damaging data breaches. Demonstrating compliance gets you part of the way there, but when breaches do occur, being able to detect and react quickly — within minutes or hours rather than days or weeks — can mean the...
The difference between a hugely damaging loss and a minor inconvenience. Real-time alerting and alerts on threshold breaches help you detect suspicious behavior within seconds or minutes.

Here is an example of an alert-type policy violation caused by a privileged user accessing sensitive data. After you define customer data as a sensitive object, all monitored accesses to that object will be recorded, whether through Big SQL or if someone comes through the back door using HDFS commands. Those alerts can be sent using email or to a security information event monitoring (SIEM), such as IBM QRadar® or HP ArcSight. The infrastructure is flexible so events can also be sent to other tools, such as Splunk.

You can create threshold alerts that track activity over time. Threshold alerts work on the same queries on the audited data you use for reports. For example, InfoSphere Guardium tracks how many records are affected by a particular Big SQL statement. That information can be included in a report, as shown below.

**Figure 7. Report including the number of records affected**

<table>
<thead>
<tr>
<th>Timestamp</th>
<th>Client IP</th>
<th>Server IP</th>
<th>Records Affected</th>
<th>DB User Name</th>
</tr>
</thead>
</table>

Using threshold alerting, you can set up an alert based on the same query used to create the report above triggered whenever the number of records affected for each statement is over a predefined limit over a certain period of time. This can help you detect malicious users who "trickle"-download sensitive data over multiple statements or sessions to avoid detection.

The following alert excerpt was generated because our threshold alert specified that an alert be raised whenever a user downloads (selects) more than 200 records in 2 minutes.

**Figure 8. Excerpt from alert when a user downloads more records than allowed**

<table>
<thead>
<tr>
<th>Timestamp</th>
<th>Client IP</th>
<th>Server IP</th>
<th>Records Affected</th>
<th>DB User Name</th>
</tr>
</thead>
</table>

**Audit and compliance workflow**

Demonstrating compliance can be time-consuming and burdensome, as these often require some level of regular review and signoff. InfoSphere Guardium not only lets you create the reports you need to satisfy audit requirements but it also has a robust workflow capability that integrates into your business processes and saves all signoffs and reviews as part of the audit trail.

For example, we’ve set up an audit workflow process that sends a daily failed login report (a task) to the administrator.
Figure 9. Report with failed logins information can be added as a scheduled task to an audit process

This can be sent via email as a PDF or CSV, and it also appears in their to-do list in the InfoSphere Guardium UI, as shown in Figure 10. If the administrator is required to sign off that they have reviewed it, that signoff is stored along with any comments along with the audit task data in case you are required to show that information for an audit.

Figure 10. Reports can be automatically added to a reviewer's to-do list

Configuring InfoSphere Guardium to collect Big SQL traffic (S-TAPs and inspection engines)

This section does not cover how to configure S-TAPs for monitoring non-Big SQL activity. For that, see the developerWorks article listed in the Resources. As shown below, for maximum security and protection, install S-TAPs on all nodes, the master and the compute nodes, to ensure that direct access on the compute nodes is also captured.

Figure 11. Red dots indicate where S-TAPs should be installed and configured to monitor Big SQL activity
On the Big SQL Server: Optional configuration step when encryption in transit, Kerberos, or GPFS is used

A special communication exit is used on the Big SQL nodes to pass information after decryption, such as user names encrypted using Kerberos, to the S-TAP for proper auditing of that information. Loading of the communication exit library is relatively simple but must be done on all Big SQL nodes as follows:

1. As root, enter the following commands:

   Listing 1. Sample code listing at maximum width

   ```
   # Make the directory in Big SQL for the communications exit
   mkdir /home/bigsql/sqllib/security64/plugin/commexit
   
   # Copy the Guardium library to this new directory
   cp /usr/local/guardium/lib64/libguard_db2_exit_64.so /home/bigsql/sqllib/security64/plugin/commexit
   
   # Make bigsql and biadmin owners of the exit
   chown bigsql:biadmin /home/bigsql/sqllib/security64/plugin/commexit
   
   # Make bigsql and biadmin owners of the library
   chown bigsql:biadmin /home/bigsql/sqllib/security64/plugin/commexit/libguard_db2_exit_64.so
   
   # run the Guardium guardctl utility to add bigsql to the Guardium authorization group
   /usr/local/guardium/guard_stap/guardctl db-instance=bigsql db-type=db2 db-user=bigsql authorize-user
   ```

2. As the bigsql user, enter the following commands to update the configuration file for Big SQL:

   Listing 2. Sample code listing at maximum width

   ```
   db2 UPDATE DBM CFG USING COMM_EXIT_LIST libguard_db2_exit_64
   db2stop
   db2start
   ```

On InfoSphere Guardium: Configure the inspection engines

Inspection engines are how you specify to the InfoSphere Guardium analysis engine (sometimes known as the *sniffer*) what type of message traffic to expect and what ports that traffic is coming in on. These inspection engines can be defined using the GUI or by using an API.

The Big SQL protocol is extended only slightly from that used by DB2®, so you can use the same inspection engine protocol for Big SQL as you do for DB2. Similarly, if you use the DB2 communications exit described in the previous section, you would use the DB2 Exit protocol.

Here is the inspection engine definition if you are not using Kerberos, encryption in transit, or GPFS.
Protect sensitive Hadoop data using InfoSphere BigInsights Big SQL and InfoSphere Guardium

**Figure 12. Configuring the Big SQL inspection engine from the InfoSphere Guardium web console (no DB2 Exit used)**

Here is how you configure the DB2 Exit for secure clusters (when Kerberos, encryption in transit, or GPFS is used). Specify DB2 Exit as the protocol.

**Figure 13. Configuring the Big SQL inspection engine from the InfoSphere Guardium web console when DB2 Exit is used**

**Using the InfoSphere Guardium API**

The InfoSphere Guardium API is issued from the command line (or script). This first example is a way to configure the inspection engine for unencrypted traffic (when the DB2 communications exit is *not* used as the protocol).

**Listing 3. GRDAPI to configure Big SQL inspection engine**

```
grdapi create_stap_inspection_engine client=0.0.0.0/0.0.0.0 protocol=DB2  dbInstallDir=/home/bigsql procName=/home/bigsql/sqllib/adm/db2sysc ktapDbPort=51000 portMax=51000 portMin=51000 stapHost=10.19.232.21
```

When communications exit *is* used:

**Listing 4. GRDAPI to configure Big SQL inspection engine (DB2 exit is used as protocol)**

```
grdapi create_stap_inspection_engine client=0.0.0.0/0.0.0.0 protocol=DB2_exit dbInstallDir=/home/bigsql procName=/home/bigsql/sqllib/adm/db2sysc ktapDbPort=51000 portMax=51000 portMin=51000 stapHost=10.19.232.21
```

At this point, you are ready to configure an InfoSphere Guardium security policy, run a Big SQL workload and ensure that you are seeing results in InfoSphere Guardium.
Define a security policy

For the security policy we use in this article, we cloned the existing Hadoop policy included with InfoSphere Guardium and modified it slightly to include a rule for detailed monitoring of privileged users.

Figure 14. Sample InfoSphere BigInsights security policy

The screenshot below shows you how we created rule 3 of the policy.

Figure 15. Editing rule 3, log full details of privileged users

If you edit the group, which you can do directly from the policy builder, you can add IDs of privileged users.
Note that this policy rule does not include any conditions to filter out specific types of traffic. The rule will be tested against any incoming activity from the server, whether it's HDFS or Big SQL or Hive.

You can validate your results using activity reports, such as that shown in Figure 3.

### Monitoring Cognos end users

Big SQL enables SQL developers and business users to access Hadoop data in a familiar way. Business intelligence tools like Cognos can seamlessly connect to Big SQL using ODBC and JDBC connections. Cognos uses a pooled connection to access the database, which means that access to the database is done through a pooled ID. This means that the DB username in InfoSphere Guardium reports is always that pooled ID. An example of this is shown in the activity report below.

#### Figure 17. Access to data is through pooled ID

In this scenario, we demonstrate a simple scenario of how to configure Cognos to send the Cognos logged-in application user to InfoSphere Guardium, along with the pooled ID.

### To capture end user from Cognos: Configure command blocks

Some applications, such as SAP, PeopleSoft and Cognos, use a pooled ID to access the database. A developerWorks tutorial titled "InfoSphere Guardium application user translation" describes the various methods to capture the real application end user with InfoSphere Guardium. When built-in support is not available, the Guardium AppEvents API call can be used.
As stated in the referenced article, the Guardium Application Event API calls are simple non-operational SQL statements executed before and after an application user takes charge of a connection in a connection pool. The no-op calls tell InfoSphere Guardium what application user is using the database session. The calls perform no change on the database, but can be detected by InfoSphere Guardium and can be used in reports and policies where it is known as Application User. With Big SQL, this no-op call can be executed on the SYSDUMMY1 table.

With Cognos, you can use its connection wizard to specify database commands that run when a database connection is opened or closed. The figure below shows the SQL in the open session command block you can use to obtain the session parameter of the logged-in Cognos BI user and how to release that when the session is closed.

**Figure 18. Use Cognos command blocks to obtain Cognos user information**

![Figure 18. Use Cognos command blocks to obtain Cognos user information](image)

The figure below shows you where the command blocks are entered in the Cognos administration console.

**Figure 19. Add command blocks in the Cognos administration console**

![Figure 19. Add command blocks in the Cognos administration console](image)

Following are the command blocks in text format.

**Listing 5. Cognos open session command block**

```xml
<commandBlock>
  <commands>
    <sqlCommand>
      <sql>
        SELECT #"GuardAppUser:" + $account.personalInfo.userName + ":" FROM sysibm.sysdummy1
      </sql>
    </sqlCommand>
  </commands>
</commandBlock>
```
Listing 6. Cognos close session command block

```
<commandBlock>
  <commands>
    <sqlCommand>
      <sql>
      SELECT #"GuardAppEvent:Released"#FROM SYSIBM.SYSDUMMY1
      </sql>
    </sqlCommand>
  </commands>
</commandBlock>
```

You need to log out and back in to have the changes take effect.

**Security policy: Alert on unauthorized access to sensitive data**

Now that InfoSphere Guardium is able to collect the application end user, you can create security policies that can act based on intelligent understanding of that user. An organization would likely have users that have to access sensitive data as part of their jobs and those users can be added to a particular group. Everyone who is not part of that group, including privileged users, are not allowed to access the data. The figure below is an example of a security policy that will fire an alert if an unauthorized Cognos application user (including privileged users) accesses sensitive data.

**Figure 20. Send alert when unauthorized users access sensitive data**

The screenshots below from the InfoSphere Guardium group builder show the two groups used in this policy.
Figure 21. Sensitive objects group

Figure 22. Authorized Cognos application users

The policy now looks like this.

Figure 23. Policy with alert rule added

See the results

Now assume that a person with the user ID of administrator logs into Cognos and views sensitive information.

Figure 24. Privileged user viewing contents of sensitive data

That activity will be logged as shown below. The next figure shows the alert as recorded in the InfoSphere Guardium UI, but it's most likely that the alert will also be sent to a SIEM system, SYSLOG, or an email address.
Figure 25. Privileged user access in an activity report now shows logged in user

<table>
<thead>
<tr>
<th>Application User</th>
<th>DB User Name</th>
<th>Source Program</th>
<th>Full SQL</th>
<th>SQL Verb</th>
<th>Object Name</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cognos user</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>administrator</td>
<td>BIOSQL</td>
<td></td>
<td></td>
<td>SELECT</td>
<td>Object Name</td>
</tr>
<tr>
<td>Pooled user</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The alert shows in the Incident Management tab on the user interface. An excerpt is shown below.

Figure 26. Alert as displayed in the InfoSphere Guardium policy violations report

Monitoring Big SQL web console users

Similar to Cognos, the Big SQL web console also uses a single ID to connect to the Big SQL server. The Big SQL web console is using trusted context queries. This means that the console authenticates the end user and connects with a trusted context to Big SQL with the user "catalog." Big SQL will then behave as if the end user was connected. InfoSphere Guardium will pick up the trusted user in the Application User field. In the figure below, the user BIADMIN is logged into the web console and issues a query. In the resulting InfoSphere Guardium activity report, CATALOG is the database user, and BIADMIN is the application user.
Conclusion

The security capabilities in InfoSphere BigInsights, Big SQL, and InfoSphere Guardium integrate to help organizations with a full spectrum of data security challenges for Hadoop, including authentication, authorization, encryption, auditing and real-time alerting. The table below summarizes the capabilities available with the integrated solution.

Table 1. InfoSphere Guardium and BigInsights security features

<table>
<thead>
<tr>
<th>Feature</th>
<th>InfoSphere BigInsights</th>
<th>InfoSphere Guardium</th>
</tr>
</thead>
<tbody>
<tr>
<td>Authentication</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>Authorization</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>Row-level access control</td>
<td>X</td>
<td>X (not supported when communications exit is used)</td>
</tr>
<tr>
<td>Block access to privileged user (or from unknown connection)</td>
<td></td>
<td>X (not supported when communications exit is used)</td>
</tr>
<tr>
<td>Column-level dynamic masking</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>Redaction of returned data such as credit card data with no change to SQL</td>
<td></td>
<td>X (not supported when communications exit is used)</td>
</tr>
<tr>
<td>Native audit facility for SQL access</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>Integrated monitoring and audit platform across the Hadoop stack and other platforms</td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>Built-in reports for SOX, PCI, BASEL II, and data privacy (HIPAA)</td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>Flexible, powerful reporting facility and text search</td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>Audit compliance workflow</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>--------------------------</td>
<td>---</td>
<td></td>
</tr>
<tr>
<td>Encryption in transit</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>Encryption at rest</td>
<td>Guardium data encryption (HDFS only)</td>
<td></td>
</tr>
</tbody>
</table>

**Acknowledgements**

The authors would like to thank Joe DiPietro, Dan Goodes, and Rosa Miroshnikov for technical assistance, and to Walid Rjaibi and Bert Van Der Linden for their review.
Resources

Learn

- Read the Verizon Data Breach Report for an analysis of data breach activity and causes.
- Read the Big SQL 3.0 White Paper for an overview of its low-latency parallel execution infrastructure.
- Read this blog for an example of fine-grained access control in Big SQL.
- The InfoSphere BigInsights 3.0 Knowledge Center includes an excellent overview of security features.
- "InfoSphere Guardium application user translation" (developerWorks, February 2012) explains the options for obtaining end users for applications such as SAP, PeopleSoft, Oracle EBS and so on.
- "Big data security and auditing with InfoSphere Guardium" (developerWorks, July 2013) provides an overview of configuration and policies for monitoring Hadoop activity. Note that this article is still at Hadoop 1.x level, but most of the material is still accurate for Hadoop 2.x.
- The Cognos Knowledge Center provides information on configuring command blocks.
- Listen to the InfoSphere Guardium tech talk on monitoring Hadoop for a deep dive on configuring the system to audit Hadoop data activity.
- Read more about Guard Application Event API in the InfoSphere Guardium Knowledge Center.
- Read more about Guard Application Event API in the InfoSphere Guardium Knowledge Center.
- Read the Redbooks publication titled "Deployment Guide for InfoSphere Guardium" for general information on using and deploying InfoSphere Guardium.
- Read the IBM Redbooks publication titled "Information Governance Principles and Practices for a Big Data Landscape" for a great overview of data governance considerations for big data, including Hadoop, and a big picture of capabilities available in the IBM portfolio.
- Check out Big SQL hands on labs for security and performance.
- Watch developerWorks demos ranging from product installation and setup demos for beginners, to advanced functionality for experienced developers.
- Visit the developerWorks Information Management zone to find more resources for DB2 developers and administrators.
- Stay current with developerWorks technical events and webcasts focused on a variety of IBM products and IT industry topics.
- Follow developerWorks on Twitter.
- Watch developerWorks on-demand demos ranging from product installation and setup demos for beginners, to advanced functionality for experienced developers.

Get products and technologies

- Evaluate IBM products in the way that suits you best: Download a product trial, try a product online, or use a product in a cloud environment.

Discuss
• Get involved in the **InfoSphere Guardium community on developerWorks**. Connect with other users while you explore the technical resources available on the wiki, blog, and forum.
• Get involved in the **developerWorks community**. Connect with other developerWorks users while exploring the developer-driven blogs, forums, groups, and wikis.
About the authors

Sundari Voruganti

Sundari Voruganti is a member of the InfoSphere Guardium QA team at IBM Silicon Valley Lab. Sundari has been with IBM over a decade and has a diverse background in both engineering as well as customer enablement roles. As a passionate technologist, she loves the challenge of learning and working with new technologies as well as helping customers understand and implement IM solutions. Sundari has a double Masters in Computer Science from Bangalore University and University of Alberta.

Kathy Zeidenstein

Kathy Zeidenstein has worked at IBM for a bazillion years. Currently, she is working as a technology evangelist for InfoSphere Guardium data activity monitoring, based out of the Silicon Valley Lab. Previously, she was an information development manager for InfoSphere Optim data lifecycle tools. She has had roles in technical enablement, product management, and product marketing within the Information Management and ECM organizations at IBM.

Roopa Sivakumar

Roopa Sivakumar is working as a software QA architect with the IBM Software Group in India. She has more than 14 years of IT industry experience working in different geographies in the areas of databases, data warehousing, business intelligence and big data. Her areas of expertise include DB2, InfoSphere Warehouse, InfoSphere BigInsights, ETL, and software quality assurance. She holds a master's degree in computer science and is an IBM Certified Advanced DB2 DBA.

© Copyright IBM Corporation 2014
Trademarks
(www.ibm.com/developerworks/ibm/trademarks/)