Addressing Security, Performance, and Usability Concerns in SAP® Data Extraction via Winshuttle QUERY

Summary
An easy and secure way to extract live SAP® ERP data, allowing business users to do ad hoc data analytics and answer specific business questions rapidly would have a significant impact on the enterprise. Despite this need, currently the tools that allow such ad-hoc data extracts: (1) do not completely preserve SAP security, (2) may adversely affect the SAP transactional system performance, and (3) may not be easy to use for business users. This white paper describes Winshuttle QUERY and how it addresses performance, security, and usability concerns while giving users a way to extract live SAP data for ad hoc analysis.

A white paper by:
Winshuttle LLC.
+1 (800) 711-9798
www.winshuttle.com
Introduction

Even with the variety of business intelligence (BI) systems installed and operational at most companies, one of the top challenges reported by SAP end-users is the lack of easy access to real-time SAP data in the transactional system. The ability of end-users to easily query and extract transactional SAP data is critical for many purposes beyond BI for corporate performance measurement, including:

- Ad hoc operational reporting
- Firefighting driven by business changes
- Prototyping steady-state BI reports

The three main concerns against users' querying live SAP data are: security, performance, and usability.

- Security concerns are based on the fact querying live SAP data does not always respect a user's organizational level authorization, and the user may be able to view more data than they have access to.
- Performance concerns are based on the fact that a user may create a query that puts too much stress on the transactional system which in turn may affect the routine transactional performance of the SAP system.
- Usability concerns arise related to users' lack of knowledge of the SAP database schema and the metadata. Users are typically more familiar with transactional screens and fields than with underlying database schema and may not be able to create valid queries.

These concerns are valid and create obstacles when allowing access to critical data necessary to accurately measure business performance. This paper focuses on Winshuttle QUERY and how it addresses some of these performance, security, and usability concerns. The intended audience for this white paper are business and technical decision makers looking to provide users with an easy way to access SAP data.

What is Winshuttle QUERY?

Winshuttle QUERY is an SAP data extraction and analysis tool that allows users to shuttle live data from the SAP transactional system into Microsoft Excel, Access and other databases for immediate analysis and action. Winshuttle RUNNER, a run-only version of QUERY, is a stand-alone application or an Excel add-in that executes previously created query files from within Excel, enabling business users to get the data they need quickly and easily.

Built into the Winshuttle QUERY product, the Winshuttle Function Module (WFM) is an ABAP component that is installed on your SAP server and is designed to allow QUERY or RUNNER users to securely extract information while minimizing the performance impact on the SAP system. Tested and certified by SAP, the WFM governs Winshuttle QUERY interaction with the SAP system.
Winshuttle QUERY Security Features

• **Preserve Native SAP Table-Level Security**
Winshuttle QUERY goes through the application layer instead of going directly into the database layer via ODBC or other direct database connections. The main advantage of going via the application layer is that the security is completely transparent because users’ existing SAP roles, profiles, and authorization objects are respected.

The standard SAP authorization object used to control table display authorization is called S_TABU_DIS and as a minimum security, QUERY respects this authorization. While this same authorization is used for SAP tools such as Table Browser (SE16), Quick Viewer (SQVI), etc., a QUERY user is not required to have access to those tools.

• **Enhance Security with Organization Level Security**
Once the basic table display authorization has been checked, Winshuttle QUERY further checks a user’s organization assignments, e.g., company code, plant, sales organization, etc., and automatically restricts the user to only records relevant to their organization. The SAP query tools – SE16, SQVI, or SQ01 do NOT respect these organizational assignments and therefore many companies do not allow users to use these tools.

To enhance the standard SAP security for query tools, in addition to S_TABU_DIS, Winshuttle QUERY checks for standard SAP authorization objects, such as M_MATE_WRK, F_BKPF_BUK, M_BEST_EKO, etc. for Plant, Company Code, and Purchasing Org, etc.-level access. Since these organizational elements exist in specific tables, only access to those tables needs to be protected.

As an example, to protect plant-level access to material master data, the table MARC needs to be secured via the authorization object M_MATE_WRK. So, if a user for whom the M_MATE_WRK object is set to plant “DR01”, wants to query the MARC table, QUERY will only retrieve the records relevant to plant DR01 and nothing else.

The core list of authorization objects and tables that are checked by QUERY include the following:

<table>
<thead>
<tr>
<th>Authorization Object</th>
<th>Description</th>
<th>Table</th>
</tr>
</thead>
<tbody>
<tr>
<td>F_BKPF_BUK</td>
<td>Accounting Document: Authorization for Company Codes</td>
<td>BKPF</td>
</tr>
<tr>
<td>F_BNKA_BUK</td>
<td>Banks: Authorization for Company Codes</td>
<td>BNKA</td>
</tr>
<tr>
<td>F_KNA1_BUK</td>
<td>Customer: Authorization for Company Codes</td>
<td>KNB1</td>
</tr>
<tr>
<td>F_KNKA_KKB</td>
<td>Credit Management: Authorization for Credit Control Area</td>
<td>KNA1</td>
</tr>
<tr>
<td>F_LFA1_BUK</td>
<td>Vendor: Authorization for Company Codes</td>
<td>LFB1</td>
</tr>
<tr>
<td>F_SKA1_BUK</td>
<td>G/L Account: Authorization for Company Codes</td>
<td>SKB1</td>
</tr>
<tr>
<td>M_BEST_EKO</td>
<td>Purchasing Organization in Purchase Order</td>
<td>EKKO/EKPO</td>
</tr>
<tr>
<td>M_BEST_WRK</td>
<td>Plant in Purchase Order</td>
<td>EKKO/EKPO</td>
</tr>
<tr>
<td>M_MATE_WRK</td>
<td>Material Master: Plants</td>
<td>MARC</td>
</tr>
<tr>
<td>P_ORGIN</td>
<td>HR: Master Data</td>
<td>Depend on InfoType</td>
</tr>
<tr>
<td>V_KNA1_VKO</td>
<td>Customer: Authorization for Sales Organizations</td>
<td>KNA1</td>
</tr>
<tr>
<td>V_VBAK_VKO</td>
<td>Sales Document: Authorization for Sales Areas</td>
<td>VBAK/VBAP</td>
</tr>
<tr>
<td>V_VBRK_VKO</td>
<td>Billing: Authorization for Sales Organizations</td>
<td>VBRK/VBRP</td>
</tr>
</tbody>
</table>

In addition to the above list of core objects, a list of over 300 additional objects is shipped with the product and can be added to further enhance the security model. Furthermore, this list is entirely customizable, and custom security objects and tables can also be added here.
Winshuttle QUERY Performance Features

- **Protect the types of joins**
  When allowing queries from multiple tables, one of the main performance concerns is related to a user’s ability to create meaningful table joins. Giving non-technical users the ability to create any kind of joins may lead to malformed queries that can severely degrade system performance. Winshuttle QUERY imposes several limits on the kinds of joins that are allowed by non-technical users and allows an administrator to remove these limits for technical users. The following join limits are imposed by QUERY by default for non-technical users:
    - Allow only inner or natural joins between tables
    - Allow only the joins that are specified in the SAP Dictionary and not allow removal or existing joins or forced joins between other fields
    - Do not allow joins of transparent tables with cluster tables, pool tables, or views

- **Limit the number of tables in a join**
  Another factor that can have an important effect on system performance is the number of tables included in a join. Again, Winshuttle QUERY imposes a maximum limit on the number of tables that can be joined by default and allows an administrator to remove these limits for technical users. The default maximum number of tables allowed for non-technical users of QUERY is 4 tables.

- **Limit the number of records to be extracted**
  The total number of records returned by a query can have an impact on the system performance, especially if the number of returned records is very large. A large number of returned records can cause both memory and network bottlenecks; therefore, QUERY places a limit on the maximum number of records that can be downloaded at 100,000 records. An administrator is able to change or eliminate this limit.

- **Data chunking**
  One of the pitfalls of running a query that returns a large amount of data is the available network bandwidth. A sequential process where results of a query are transferred only after a query has finished executing can lead to the creation of a network bottleneck depending on the amount of data being returned. Sending partial results to the client application as they become available is called Data Chunking and is an effective way to avoid network bottlenecks.

  Winshuttle QUERY supports Data Chunking and allows users to set a size of the chunk while retrieving data from SAP.

- **Timeout and forced query stops**
  Depending on the type of a query and the size of a table being queried, the execution of a single query may take a very long time. QUERY allows a long-running query to be terminated after a certain time has elapsed – the timeout value. There is a default timeout value which is tailored by an administrator. In addition, a long-running query can also be forcibly filled by an SAP BASIS administrator via the SAP transaction SM50 – Process Overview.

- **Adaptive query throttling**
  One major performance concern around giving users the ability to create queries against the live transactional system is that the query execution could take substantial computing resources that might compete against regular SAP transactions.

  Winshuttle QUERY addresses this issue in a unique way by allocating only available computing resources to execute queries in the live SAP system. Using a patent-pending Adaptive Query Throttling™ technology, QUERY dynamically optimizes the execution of the queries based on the current load on the SAP server at any given time. The available server resources are calculated from the total number of open dialog processes. When over 50 percent of the dialog processes on the SAP system are available, the query is executed at full speed. However, when the available server load starts falling below 50 percent, the execution of the query is slowed down.
The graph below displays the query performance time against the system load with or without the adaptive query throttling algorithm applied. As this clearly shows, when the throttling algorithm is applied, the query execution slows down as the system load is increased. And without the algorithm, the query executes in constant time competing for resources with other transactions on the system.

- **Scheduled query extraction**
  Even with the Adaptive Query Throttling feature, a long-running query should ideally be scheduled to run at off-hours. Winshuttle QUERY allows users to schedule queries so that these queries can be run at off-hours or at pre-set times.

**Winshuttle QUERY Usability Features**

- **Query creator role versus query runner role**
  Winshuttle QUERY is provided in two versions for different user roles. The full QUERY product is intended for a query creator, a power-user, while CENTRAL is intended for end-users. The end-users are only able to execute queries that have been defined by a power-user who understands the SAP metadata and the database schema, and has the ability to create queries. The end-users might have the ability to select specific fields that they require and to create some limited search criteria, but they cannot change the tables or the joins between tables.

  This concept is very similar to SAP Infosets where the creator of an Infoset has full access to tables and table joins while a user of the an Infoset can only select the specific fields and add some search conditions.

- **User-friendly interface with 3-steps to create a query**
  Winshuttle QUERY provides a simple three-step interface to create and execute a query where the three steps are defined as follows:
  
  **Step 1:** Select - The user selects the tables and the fields that they want to extract from SAP. Also in this step, the user creates the selection criteria to allow selection of specific records.
  
  **Step 2:** Map - The selected fields are mapped to Excel or Access fields, or other output formats.
  
  **Step 3:** Run - Once the selection and mapping are defined, a query is ready to run and the Run tab allows users to pick run-time parameters for a query, e.g., the output filename, log column, etc.
• **Organized data dictionary view**
  One of the main usability issues with SAP data extraction is giving the users an easy way to find the appropriate tables and fields that are relevant to their reports. Winshuttle QUERY provides a local view (see figure below) of the SAP data dictionary organized by SAP modules and by transaction codes. This makes it easy to locate tables relevant to a user's query. While only the most commonly used tables are provided in this data dictionary view, users are able to add tables that are not present, including any custom tables to this view.

![Image of data dictionary view](image_url)

• **Organized data dictionary view**
  One of the main usability issues with SAP data extraction is giving the users an easy way to find the appropriate tables and fields that are relevant to their reports. QUERY provides a local view (see figure below) of the SAP data dictionary organized by SAP modules and by transaction codes. This makes it easy to locate tables relevant to a user's query. While only the most commonly used tables are provided in this data dictionary view, users are able to add tables that are not present, including any custom tables to this view.

• **Table search and field search functionality**
  The ability to easily search the data dictionary for tables and fields is a key component of finding tables and views. QUERY provides an easy-to-use way to search for tables or fields within the local data dictionary or within the full SAP data dictionary – as shown in the figure above.

**Future Performance, Security, and Usability Features of Winshuttle QUERY**

While Winshuttle QUERY already provides many performance, security, and usability features, there are other features that will become available in future versions of QUERY. Many of these features will be enabled via integration of QUERY with Winshuttle CENTRAL product. CENTRAL provides centralized visibility and governance for enterprise deployments of Winshuttle solutions. This not only enables companies to share files and best practices across the organization, but also addresses the organization’s requirements for specific controls and audit trails to meet SOX and corporate governance policies.
Some specific query-related governance and control activities that will be enabled by CENTRAL include:

- A central location which manages all of the parameters that are relevant to Winshuttle QUERY, such as query timeouts, maximum number of tables that can be joined, maximum number of records, etc.
- A central location to collaborate on query creation and share pre-built and approved queries.
- A workflow process to make sure that all of queries that are allowed to be run in a production environment are reviewed and approved by an expert.
- A mechanism to control allowable query execution hours, days, and times so that queries are only run during non-peak hours.

**Conclusion**

Winshuttle QUERY provides business users with an easy way to access SAP ECC data that does not compromise SAP