Building a Data Bridge Catalog: protecting valuable data from the bad guys

By Bill Wunder

One of the important initiatives within the Runbook of knowledge and information shared among the skilled workers in a data center is to document data protection measures. Almost always *de facto* business rules will exist about who can and cannot view some or all of the data stored in organization’s databases. Recording these rules in the Runbook is the first step in documenting data protection measures.

At some point after the rules are recorded, it is vital to examine all known database access in the context of these business rules. A highly effective way to describe the multiple ways and means that data is accessed is by building a Data Bridge Catalog.

A data bridge is any complete logical path through which data can be moved in or out of the database. A data bridge will have 2 endpoints – a data store and a consumer - and a transport span.

There is no security context association for a data bridge. Security is applied within the data bridge context. There may be any number of security contexts within a data bridge. In fact examining the data bridge critically to uncover unintentional bridges is one of the most interesting aspects of data bridge cataloging.

Just as road bridges have many designs, Data Bridges have many designs. So much so that critical data bridges can go unidentified or unattended and as a result out of band or malicious acess to the organizations valuable data is possible completely unbeknownst to the DBA. In the same way that a road bridge might be exploited by the power company, the utility company, the bus and even the sewer department to get their product to the other side, good application designers and developers will look to existing infrastructure to get data from the database to the data consumer. Unfortunately, bad guys can use the same approach to move important data where it ought never go.

USB ports are a simple yet classic example of an endpoint that may not be considered when designing a database security architecture. Query results and even screen shots of confidential reports are easily moved without detection using these ubiquitous port. Its politically risky to say, but any one with administrator rights is another often overlooked and unmonitored enabler of illicit data bridges.

This article describes a collaborative thought project to identify and catalog the data bridges.

It is not realistic to expect to build a complete data bridg catalog in a week or a month. Expect this to be an ongoing and iterative initiative.

Data security models can be defined in the context the data bridge. That is, until someone has access to the data, it remains safe: a bridge to nowhere. From that point, those given access to the data must be included in a *principle-of-least-privledge* based security scheme. All others are excluded from access through the data bridge if at all possible.

Just the though that one of the SPIDs connected to a SQL Instance may be the competition or the enemy gives a good DBA a seeminly instinctive urging to get a backup. How else to keep that data safe? And then, given the opportunity to adequately contemplate the security, the DBA may realize that creating the backup creates a new and even more opaque data bridge. The file or tape image itself becomes a second data bridge. Moreover, this data bridge is outside the scope of the application layer security schemes and possibly even organizational permissioning schemas. From the time the backup is created, every security question asked about the database must also be asked about the backup file. And if a copy of the backup file is created or the the backup is restored to another SQL Instance, the once doubled security requirement is now tripled. From this is is apparent how quickly the risk of unintented exposure of data to outside world can snowball out of control.

A data bridge is a conceptual architectural element not a process label. Many processes may use a data bridge. Conversely, a data bridge can link several data bridges into an over arching conduit fordata. By the way, such a linked collection of data bridges is collectively referred to as a data bridge.

The storyboard surface in *Figure 1* depicts a conceptual representation of on organizational database environment. The surface is useful to describe data bridge connectivity scenarios. The storyboard serves as a visual queue to get the conversation started between the data team, system architects, security officers, the network administrators and the business stakeholders concerning unplanned/unintended data bridges. It also serves as a documentation of that conversation by adding notes and markings to explain the know endpoints as well as potential risks for the data bridge. (Some example post-conversation storyboards are shown at the end of the article.)

Remote/Mobile/Branch Office

Database Network

DMZ

Firewall

Other Networks “in the building”

Administrator’s Console



Figure : a data bridge storyboard template

This article explores a methodology for cataloging of data bridges. Along the way, the document will also explain why a data bridge catalog is vital to the data team and to the organization in the quest to secure the data assets. It is left to the reader to define the form and fit of the Data Bridge Catalog as it might be used in the reader’s scenario. To wit; some catalogers may find it more useful to introduce the storyboard as an artifact created by all in the conversation while others may find it more useful to make personal notes during or after a conversation involving the data team. The storyboard need not be a part of the dialog as long as adequate information to recreate the data bridge is cataloged. The storyboard should not inhibit the concise description of the data bridge.

The data bridge catalog does not document only the known data paths within the organization, but also serves to uncover potential unintended or unexpected data bridging that can be secured as appropriate. Typically, cataloging will at first be mostly a documentation effort. Over time, with continued scrutiny, review of the catalog may identy exposures and to proactively improve data security. And of course if the business actually pipes data across know bridges for malicious use, expect the evolution of the Data Bridge Catalog to accelerate. At least temporarily.

Even a somewhat out of date Data Bridge Catalog will provide data team experts with an eye toward data security an effective and suprisingly complete index for a methodical analysis provided each entry is adequately described. The more the team works with concept of data bridges, the more likely any risks and exposures will be uncovered.

A completed data bridge story board will include a minimum of four fundamental pieces of information:

* The data bridge - applications, scripts, tools or services used to connect Data Consumers to the store for that data. A complete data bridge defintion would include the technical detail or organization contact information necessary to install, configure, update, enable/disable, and uninstall the data bridge.
* The data consumer – the business interest or target of the data supplied by a data bridge. The consumer could be a user, organizational groups, database role, customer, client, domain, partner, system, etc. that a data bridge purposefully connects to the database. More than one Business Interest might have a legitimate use for the same data bridge. All data consumers, including those that can transparently piggyback on a data bridge purposefully built for another consumer, should be explicitly identified and cataloged. Think of the data consumer as the scope of intended trust for the data.
* The data store - a concise functional collection of data. All narrowing attributes - those that aid in uniquely identifying a data store – should be included to correctly and accurately describe a data store (e.g. database NorthWinds, schema Sales, PROD1, etc.). Think of the data store as all data presented to the data bridge. It is not relevant to the data store if the consumer is actually using accessible data at any given moment or not.
* Exposures - Identified ways that the data can be used for other than the intended purpose or by any but the intended data consumers. It is not necessary to enter exposures when a data store is cataloged. It is even possible that none exist I presume, though it is far more likely that exposures are difficult to detect than that there are none to be found

Figure 2 depicts the data bridge story board template as it might be loaded to OneNote as template or projected to a screen for collaborative use.

Remote/Mobile/Branch Office

Database Network

DMZ

Firewall

Other Networks “in the building”

Administrator’s Console



Use the storyboard to design a data bridge or to document an existing data bridge.

Add notes, labels, sketches, questions, and other information anywhere on the storyboard to help describe the data bridge. Modify the storyboard in any way that more accurately describes the data store. Try to list the reasons and requirements for the data bridge in terms of:

* data bridge
* data consumers
* data stores

Include enough of the implementation details to be able to recreate the data bridge’s infrastructure in the event that might become necessary. Details like port numbers or ACL assignments need not be included; however the fact that a data bridge requires a port assignment or ACLs is vital to the data bridge documentation.

Figure 2: data bridge storyboard annotated template

Add details

Add details

Add details

One value that a data Bridge Catalog adds is to keep the pathways to unintended consumers visible along with the intentional endpoints. This provides support for a methodical security monitoring practice and data protection that begins at the moment each data bridgeis added to the catalog.

Some of the more common and important data access patterns to describe in the catalog are:

the primary business supporting application(s)

internal reporting mechanisms

backup & restore files and tapes

data imports and conjunct source files/databases

data exports and conjunct target files/databases

monitoring

*ad hoc*. access

There are many out of the box SQL Server data bridge building blocks. Here are a few of the more common components that can enable new and amend existing data bridges.

SQL Server Management Studio

SQLCMD

Notification Services

Integration Services

Reporting Services

OLE automation

SQL CLR methods

BCP

Bulk Load

Backup/Restore

Detatch/Attach

Command line, console and desktop scripts and applications - built with technologies that enable the access and manipulation of SQL Server data or the execution of T-SQL Scripts will be the consumer on most data bridges identified. The use of data access technologies indicates a potential SQL Server Data Bridge application.

ASP.NET

ADO.NET

DBLIB

Java

http.sys

ODBC

OLE Db

SQL Client

VDI

Web Service

Windows Service

Cataloging the data bridges are currently in intentional use creates visibility to organizational usage patterns that are useful to establish consistent design, scalability and security best practices. It also defines a white list that monitors can use to identify unauthorized data bridge use. This is the task to take on before attempting to find those pesky unintentional bridges. Otherwise the task becomes overwhelming from the onset. The Data Bridge Catalog is supposed to help get things done correctly, not bury already busy people in chaos. Creating a catalog of what is supposed to exist is a project that can be done as a reasonable project that can be published in less two weeks in most scenarios. The second project can be to flesh out the highest potential unintentional pathways identifiable from the work of the first iteration.

The Data Bridge Catalog is an aid in troubleshooting connectivity related application errors and describing the data access topology to those less experienced or unfamiliar with the organizations SQL Server's. This is important in terms of knowledge sharing and business continuity assurance. As the workforce changes, removing the need to repeatedly re-discover the extant organizational data bridges will unlock the door to a high performance IT team.

A comprehensive Data Bridge Catalog enables knowledge and information sharing among the IT personnel responsible for the organizations SQL Servers. By extension, the Data Bridge Catalog also helps assure business continuity at the data layer.

A small but significant amount of time is required from those most knowledgable with a system to initialize the Data Bridge Catalog. Less time is needed to maintain the catalog, although the accuracy of the data bridge catalog can easily be lost over time without an adequately complete and continuous effort to catalog. The business benefits of the ablity to quickly and accurately explain the Data Bridge configuration aspects, even in the absence of any of the individuals that helped to establish the catalog must be clear to anyone that has had to support an unfamilair database environment. It is rarely possible to get a complete answer to the question “who uses this data”. Without the Data Bridge Catalog, the most realistic approach may be to work with the well know data users as necessary and then to identify the remaining marginal usages as they become problems. With the catalog the data security breaches, barriers and problems can most easily be avoided. Understanding the way data might unexpectedly propogate through the organization and in some cases beyond the organization is possible by building and maintaining a Data Bridge Catalog.

As examples, consider the data bridge story depicted in *Figure 3* and *Figure 4*.

Figure 3: data bridge storyboard example: Production/QA/Test DBA Group Access

Remote/Mobile/Branch Office

Database Network

DMZ

Firewall

Other Networks “in the building”

Administrator’s Console

PROD-DBA



SSL Terminator

TS Gateway

TS Web Access

DBA Desktop

DBA Notebook

TS Remote App

RDP

RDP

Any RDC 6.1 Client

RDP in

HTTPS

Description: Provide a single-sign on administrative access to all SQL Servers for the DBA group from the office, home or while traveling. DBAs use this data bridge for all direct administrative activities within the Production, Q or Test environments: code promotions, maintenance, monitoring and troubleshooting. The data bridge should not restrict the DBAs ability or authority for performing all essential activities. The Applications required are defined in the DBA groups SQLClue Runbook Topic “Application Requirements for the Shared Environment Console”.

Implementation notes: Console is Windows Server 2008 R2 with RDS enabled configured in the TS VPC cluster. The Hardware Support team uses a standard to image with the required Terminals Services configuration and policies applied to build the console. The existing TS WebAccess web server and SSL terminator will be used. TS RemoteApp configured for DBA applications. TS Gateway configured signed rdp file with resource sharing and for TS Web Access with no resource sharing. Machines not known to the Gateway must use TS Web Access to open the RDP session. DBA Notebooks are secured to Hardware support team standards…

Figure 4: data bridge storyboard example: Database Backups

All automated backups are configured and scheduled through the console Master Server console. SQL Instance is configured for multi-server administration and central management services

Remote/Mobile/Branch Office

Database Network

DMZ

Firewall

Other Networks “in the building”

Administrator’s Console



Only the assigned SQL Agent Backup Proxy account can write to the database share on the backup file server. Readers of the share include the proxy and DBA group

Description: All backups are written a central backup server provided by Network Operations.

The Backup File Server is swept by the tape library once every 24 hours and new backups are written to tape. Tapes are managed by the Data Center Operations Policy.

The Customer database has Transparent Data Encryption (TDE) enabled to prevent unintended or unauthorized visibility to this data in the backup.

Exposures –

* Ad hoc or one off Backups are not described by this data bridge although TDE would provide customer data protection in all scenarios. Users with BACKUP command rights are listed in Exhibit 1
* A fairly large group of internal users has read access to the backup file folder. (Listed in Exhibit 2)
* Data Center Operations group are local administrators on the backup file server.
* The network administrators can move and copy backu files with imopunity.

For more information about Data Bridges and creating a Runbook stop by my web site at http://www.bwunder.com