



**HPE Shadowbase SOLV Product Suite –
Online Database Loading, Copying, Synchronizing, and
Unloading**

A Gravic, Inc. White Paper



Executive Summary

Before a database is useful, it first must be loaded with data. The HPE Shadowbase Online Loader (SOLV) and the HPE Shadowbase Extract, Transform, and Load (ETL) Toolkit are utility products designed for this.



SOLV provides online loading of a source database into a target database; the source and/or target databases can be open for reading and updating while the load (or copy) occurs. SOLV can perform the load while regular replication is occurring, consuming the source database changes and merging them with the data being loaded to keep the target fully synchronized with the source while the load takes place. This synchronization avoids the need to queue (and ultimately drain) the replication changes after the load completes.

SOLV works in both HPE NonStop Guardian and OSS environments. SOLV can load audited and non-audited HPE NonStop Enscribe source files and HPE NonStop SQL tables into any target [environment and database combination](#) supported by the HPE Shadowbase line of data replication products (e.g., HPE NonStop Enscribe or SQL targets, or Shadowbase Other Server targets including Oracle, SQL Server, IBM Db2®, SAP HANA, SAP Sybase, and MySQL). SOLV works in both HPE NonStop Guardian and Open System Services (OSS) environments.

The SOLV utility plays an important role in eliminating *planned* downtime via zero downtime migration (ZDM). In the case of a platform or hardware refresh, the utility loads the database onto the new system from the database on the existing system while the existing system's database remains online and application services remain available. Therefore, the utility enables the building and synchronizing of the new database without outage of the existing production database while the data and application conversion/migration effort occurs.

This utility also provides an option for replicating non-audited HPE NonStop Enscribe and SQL data: SOLV *Snap-Shot Loads* (also known as *refreshes*). Using the SOLV solution, the user can periodically load (or refresh) the entirety (or a portion) of the non-audited file or table into the target environment.

The HPE Shadowbase ETL Toolkit is used to extract database changes or initial load data into flat files for subsequent ETL loading into a data warehouse using comma-separated value (CSV), fixed-position, and tab-delimited formats. More specifically, for certain applications, the user may want to utilize an existing vendor's ETL utility to load either initial data or change data into a data warehouse. The Toolkit enables a Shadowbase software user to perform these tasks. For example, when coupled with the SOLV loader, the Toolkit can be used to extract select data from a source database, transform/cleanse it into the proper format, and save it into flat files that the ETL utility can understand and load into a target database environment not directly supported by Shadowbase replication. Using this Toolkit, the user can also periodically extract source database change data from the source database's audit trail (change log), and process it into a flat file format that can be incrementally loaded into the target environment using a micro-batch incremental update approach to keep the target environment synchronized with the source.

The Shadowbase Online Loader and the Shadowbase ETL Toolkit are utility products designed to remedy this situation by performing initial database loads including any subsequent target database resynchronization, providing online and offline loading (or unloading) of a source database into a target database. A variety of features and databases are supported.

Table of Contents

Executive Summary	2
Table of Contents	3
Table of Figures	3
HPE Shadowbase SOLV Product Suite – Online Database Loading, Copying, Synchronizing, and Unloading	4
HPE Shadowbase Online Loading and Verification (SOLV)	4
<i>SOLV Online Loading</i>	4
<i>Key Features of SOLV</i>	5
HPE Shadowbase Online Loading and Verification (SOLV) for Offline Loading	5
HPE Shadowbase SOLV for HPE NonStop Open System Services (OSS)	5
HPE Shadowbase ETL Toolkit	6
<i>Key Features of the ETL Toolkit</i>	6
HPE Shadowbase ETL Toolkit and Online Loading and Verification	7
<i>Load-Time Consistency for Snapshot Loads</i>	7
<i>Data Consistency</i>	7
<i>ETL and SOLV Utilize the Shadowbase Data Replication Engine in Certain Situations</i>	7
<i>Typical SOLV Unloading Scenarios</i>	8
<i>HPE Shadowbase SOLV Manager (SOLVMGR)</i>	9
Summary	10
International Partner Information	11
Gravic, Inc. Contact Information	11

Table of Figures

Figure 1 – SOLV Online Loading Integrated with HPE Shadowbase Change Data Replication	5
Figure 2 – HPE Shadowbase Online Loader (SOLV) Offline Loading	5
Figure 3 – HPE Shadowbase SOLVUTIL for OSS Regular Files.....	6
Figure 4 – SOLV Unloading with HPE Shadowbase ETL Output	8
Figure 5 – SOLV Unloading Integrated with HPE Shadowbase Replication and ETL Output	9
Figure 6 – HPE Shadowbase SOLVMGR Injecting Flat File Data into the Replication Engine	10

HPE Shadowbase SOLV Product Suite – Online Database Loading, Copying, Synchronizing, and Unloading

Online Database Loading Architectures

HPE NonStop to HPE NonStop

This is common during a Zero Downtime Migration, and for an initial installation and configuration of a DR system.

HPE NonStop to Other Servers

This type of load is common for initial loading, refresh loading, and for other miscellaneous types of heterogeneous loads..

Other Servers to HPE NonStop

This feature is currently unavailable.

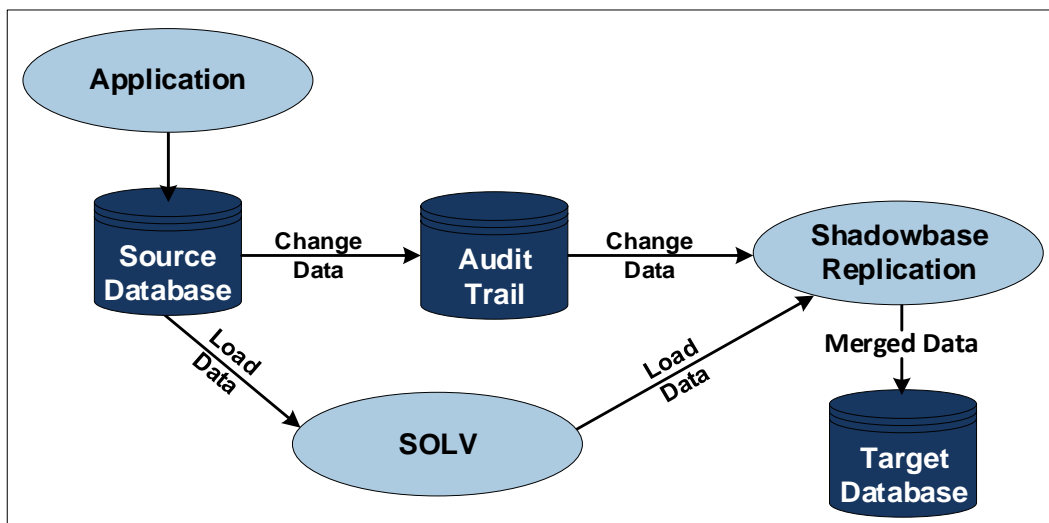
HPE Shadowbase Online Loading and Verification (SOLV)

The SOLV utility provides online loading of a source database into a target database, meaning that the source and/or target databases can be open for reading and updating while the load (or copy) occurs. The SOLV utility enables applications to access either a source or target online database while a load occurs.

SOLV Online Loading

SOLV loading utilizes special [patented](#) features that allow it to properly merge the data being loaded with the data being replicated. However, note that SOLV loading can also occur without HPE Shadowbase replication being active/in-use¹. In this case, SOLV acts like a stand-alone data transformation utility, reading and converting the source data format into the target data format, and applying that information into the target environment.

The SOLV utility can load audited and non-audited HPE NonStop Enscribe source files and HPE NonStop SQL tables into any target [environment and database combination](#) supported by the Shadowbase line of data replication products (e.g., HPE NonStop Enscribe or SQL targets, or other server targets such as Oracle, SQL Server, IBM Db2®, SAP HANA, SAP Sybase, and MySQL). The utility can perform the load while regular Shadowbase replication is also taking place, consuming the source database changes and merging them with the data being loaded to keep the target fully synchronized with the source while the load proceeds (Figure 1).



¹More specifically, SOLV loading can be used with any replication engine, or without a replication engine, since SOLV can coexist with other applications and products that access and update the database. If a solution uses an existing replication engine, then care must be taken during the load to either pause the replication engine, or otherwise ensure that the load path stays ahead of the replication path during the load.

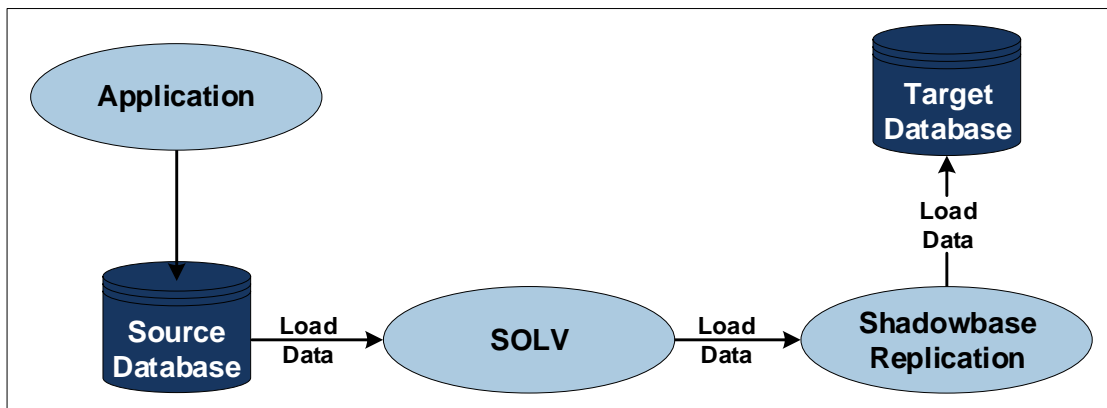
Figure 1 – SOLV Online Loading Integrated with HPE Shadowbase Change Data Replication**Key Features of SOLV**

- Initiate and refresh target data-loading
- Bulk load from a source's point-in-time to a target
- Extract, transform, filter, and cleanse the data being loaded
- Integrate with the Shadowbase replication engine to keep the target synchronized during and after the load sequence
- Automatically keep the target synchronized with the Shadowbase replication engine during and after the load function
- Consume/merge replicated events with the loading activity, so replication operations continue during the load
- Dynamically adjust loading activity and efficiently utilize available resources
- Keep the source and target online for read-write activity while the load or refresh occurs
- Completely and efficiently load online source data into a target environment with one operation
- Utilize SOLV's offline loading feature, if desired

HPE Shadowbase Online Loading and Verification (SOLV) for Offline Loading

SOLV also has a function for offline loading. Figure 2 depicts a typical SOLV *offline* loading architecture. In this context, offline means the source database is not being actively changed by the application during the load; if it is, then the change data are either being queued or ignored.

In Figure 2, an application posts its changes to a source database. SOLV reads a copy of the source data at a point-in-time and optionally performs data transformations on it; then the Shadowbase replication engine prepares and applies the data into the target database. In this example, a point-in-time copy of the source database (load data) is taken and applied into a target database environment by SOLV and then the Shadowbase replication engine.

**Figure 2 – HPE Shadowbase Online Loader (SOLV) Offline Loading
(Change Data Replication is not Active)**

This form of loading is *offline* because any application changes made to the source database by the application after the load starts (and after the data has been loaded) will not subsequently be replicated into the target database. This form of loading is most suitable when the source database is not actively being updated, or when a point-in-time snapshot of the source database is required. SOLV allows specifications for the entire source file/table (or even key ranges) for the load. Source data filtering (e.g., by partition, or data content) is also supported.

HPE Shadowbase SOLV for HPE NonStop Open System Services (OSS)

SOLV for OSS supports loading/copying OSS *regular* files and validating the data consistency, and can create a backup copy of the data and synchronize it with the source. It leverages the SOLVUTIL process (a low-level Shadowbase replication process that runs in the Guardian space) to copy a particular file (or set of files) from the source to the target.

This method of replication is called *file chasing* – or more accurately *end of file (EOF) chasing* – because SOLVUTIL will replicate any data appended into the source file to the target file, applying it to the end of the target file. When SOLVUTIL reaches EOF on the source, it sleeps for a short period, and then wakes up and checks if there is more data to replicate. This process continues the entire time the source file is appended by the source application. SOLVUTIL contains a function to automatically track when the application switches from one source file to another (e.g., the source log file rolls to a new log file), and starts replicating into a new target file. To validate the data, the SOLVUTIL process can be given two filenames and it will compare the contents of both files (Figure 3).

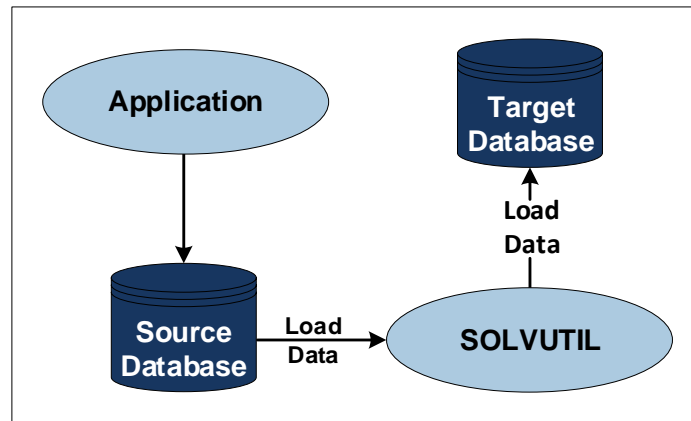


Figure 3 – HPE Shadowbase SOLVUTIL for OSS Regular Files

The SOLVUTIL process recognizes two situations:

- *Mismatched Data* – It provides the user with the address of the first mismatch, and options to perform a loading session to correct the target's mismatched data.
- *Row Not Found* – This situation occurs when a row of data is missing in the target table. It provides the user options to perform a loading session to insert values from the source to resolve the situation.

HPE Shadowbase ETL Toolkit

The HPE Shadowbase ETL Toolkit is particularly useful for feeding high-volume quantities of loading data into a data warehouse, and when feeding row-based data into columnar databases. The Toolkit contains a sample programmatic user exit set of functions that uses and extends the SOLV loading capabilities. This approach enables reading and injecting events from flat files into the Shadowbase replication engine for processing. It also produces flat files of database data or database change events that can then be subsequently processed by an ETL tool. In this mode, Shadowbase SOLV and the Toolkit are used to extract database changes or initial load data into flat files for subsequent ETL loading into a data warehouse using comma-separated value (CSV), fixed-position, and tab-delimited formats.



For example, with certain applications, the user may want to use an existing vendor's ETL loading utility to load either initial or change data into a data warehouse. The Toolkit allows a Shadowbase user to perform these tasks. For example, when coupled with SOLV, the Toolkit can be used to extract select data from a source database, transform/cleanse it into the proper format, and save it into flat files that the ETL loading utility can understand and load into a target database environment not directly supported by Shadowbase replication. Using this approach, the user can also periodically extract source database change data from the source database's audit trail (change log), and process it into a flat file format that can be periodically loaded into the target environment using a micro-batch incremental update approach to keep the target environment synchronized with the source.

Key Features of the ETL Toolkit

- Provide a Shadowbase user exit example set of 'C' language functions that can be adapted for specific scenario/needs

- Trickle-feed replicated-events loading and UN-loading utility
- Snapshot point-in-time bulk loading and UN-loading utility
- Extract, transform, filter, and cleanse data before loading and unloading
- Read or generate CSV and fixed-position flat files for data loading and unloading
- Keep the source or target database online for read-write activity while loading and unloading
- Leverage the SOLV features for ETL processing

HPE Shadowbase ETL Toolkit and Online Loading and Verification

The ETL utility uses and extends the SOLV loading capabilities to allow for reading and injecting events from flat files into the Shadowbase engine for processing, as well as producing flat files of database data or database change events that can then be subsequently processed by an ETL tool.

By themselves, ETL and SOLV can each produce *snap-shot* loads of all or part of a source database into a target environment. These loads are also called *point-in-time loads* as they are initiated at a particular point-in-time. When used in this mode, any changes made to the source database while the load runs and after the load completes are typically not subsequently replicated to the target environment.

Load-Time Consistency for Snapshot Loads

During a SOLV and ETL load, the consistent data in the source database at the start time of the load is loaded into the target database. SOLV only reads and loads *consistent* source database data (i.e., unless configured to do so, it does not read through locked or otherwise inconsistent data, meaning data affected by transactions in process). Note, a typical load takes time, and the target will not contain the latest data that was updated while the load occurred.

When run in this mode, to ensure that the target matches the entire source after the load completes, the source database needs to stop updating when the snap-shot load is taken, or select the “lock file during the load” option. Otherwise, the change data that occurred during the load needs to be subsequently loaded into the target database to keep it synchronized with the source.

Data Consistency

ETL and SOLV offer several *data consistency* level options. These consistency options allow the user to tailor the loading parameters to the particular data consistency needs. These options can affect data concurrency while the load takes place. By default, the data being loaded is locally consistent (unlocked and consistent). However, SOLV also supports an inconsistent load level (that reads through locks) as well as a file/table level lock consistency mode, where it locks the entire source file or table before the load begins. This option ensures that the point-in-time load of the target is consistent with the *entire* source when the load completes.

ETL and SOLV Utilize the Shadowbase Data Replication Engine in Certain Situations

ETL and SOLV fully cooperate with the Shadowbase data replication engine to ensure that the data being replicated and the data being loaded are properly serialized when applied to the target database. ETL and SOLV leverage the Shadowbase core (the main data replication engine) for certain services. If the engine is being used for data replication of the source file or table being loaded to the target, then Shadowbase ETL and SOLV can integrate with it to load the target database while change data replication is occurring. Hence, while the load occurs, the Shadowbase core properly keeps that part of the target database (that has already been loaded) synchronized with any subsequent changes that the application makes to the source database. This approach is particularly useful for high-volume applications with a large database because change data replication is not suspended while the load takes place, (unlike other data loading products). Instead, the queue of database changes is consumed while the load takes place, and the target is always kept *current* with the source’s database changes during the load. The target does not grow stale after the load completes because the Shadowbase core continues to keep the target database synchronized with the source database.

SOLV can utilize the Shadowbase core’s powerful data transformation, data filtering, and data cleansing routines and scripts. Many data mappings are performed out-of-the-box. Other more complex data mapping functions must be developed using Shadowbase data transformation technology. Then, both the Shadowbase engine as well as ETL and SOLV can use these same functions during the load. The routines or mappings

only need to be written and tested once, since they are then used for both database change replication as well as the loading operation. They do not have to be re-implemented in a separate custom ETL tool, script, or program.

SOLV leverages the Shadowbase technology that allows the source and target to be homogeneous (such as like-to-like replication), or heterogeneous (such as loading a non-relational source file into a target SQL table or when loading a data warehouse). Once Shadowbase replication is configured, SOLV automatically uses this information to perform the load.

Typical SOLV Unloading Scenarios

Figure 4 depicts a typical HPE Shadowbase ETL scenario for producing flat files of source data. In this figure, an application is posting changes to a source database, which is being read, parsed, and formatted into a flat file, such as a typical fixed-position (FP) or comma-separated values (CSV) format for a subsequent program, application, or ETL tool to process. More specifically, SOLV is reading the source data (of interest) from the source database, and sending that data to the Shadowbase engine to transform, filter, and process it (via the embedded Shadowbase ETL Toolkit) into the desired flat file target format. The flat file is produced by the Shadowbase user exit functionality using the Toolkit, typically written as a *consumptive* user exit. A consumptive user exit receives the data changes, and parses them into the desired output format and writes them directly into the target flat file. Shadowbase data replication can also be active on the source data being processed, allowing for database changes to be added into the output file. This application has proven popular, for instance, for telco environments to capture their call detail record (CDR) data for loading into a downstream data warehouse system.

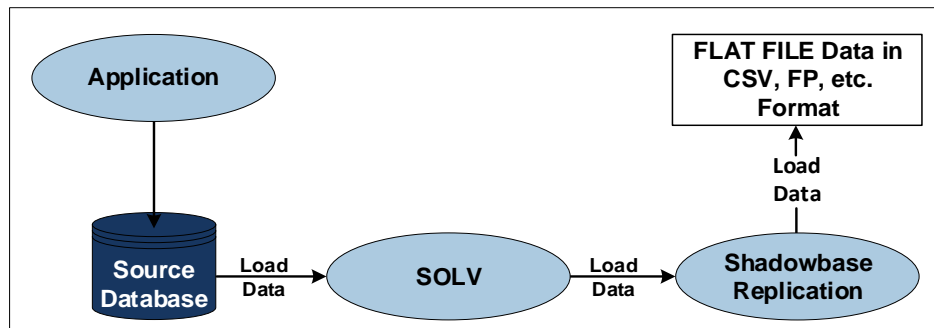


Figure 4 – SOLV Unloading with HPE Shadowbase ETL Output

For example, Figure 5 depicts a typical scenario where an application is actively modifying the source database, and SOLV is unloading that source data into a target flat file in CSV or fixed-position format along with any changes made to the source for the data of interest. The application change data is merged with the load data.

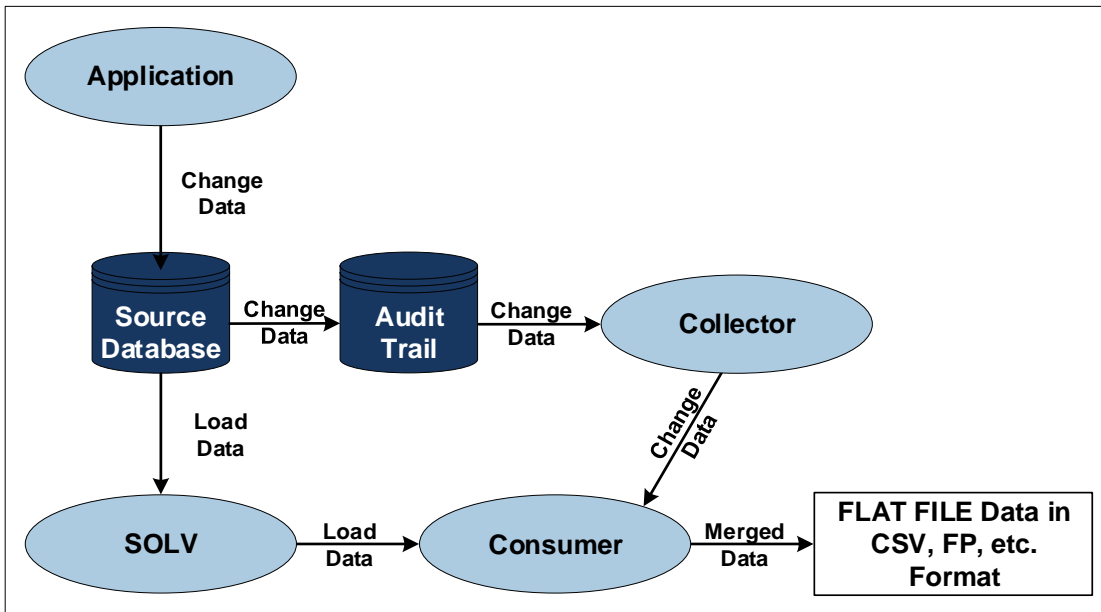


Figure 5 – SOLV Unloading Integrated with HPE Shadowbase Replication and ETL Output

HPE Shadowbase SOLV Manager (SOLVMGR)

In addition to generating flat files for output, HPE Shadowbase replication can also process flat files of input data and inject their contained events into the replication engine for subsequent application into a target database or another target flat file.



Figure 6 depicts a typical HPE Shadowbase ETL scenario for injecting ETL data into the replication engine. In this figure, the HPE Shadowbase SOLVMGR (also known as the Shadowbase *File Chaser*) is reading data to inject into the Shadowbase replication engine from a variety of formats, including a flat file, an application change data log file, a middleware queue, or other application-generated file. The SOLVMGR (or in some cases a Shadowbase user exit directly built into the Shadowbase replication engine), will read the events as they are generated and inject them into the replication stream for subsequent processing by the engine.

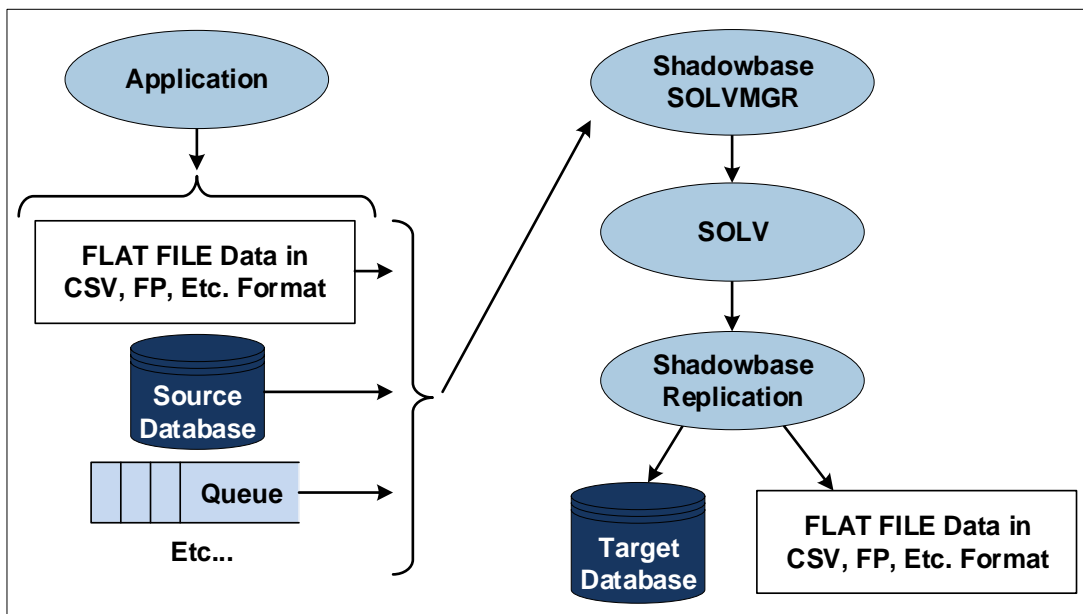


Figure 6 – HPE Shadowbase SOLVMGR Injecting Flat File Data into the Replication Engine

This SOLVMGR is a component of the SOLV loading product that can monitor source objects (sets of flat files, application log files, queues, etc.) and *chase* the data insertions into these objects, typically at end of file (EOF), extracting them for Shadowbase software to process (often referred to as *EOF chasing*). The Manager will read and prepare that data and inject it into the engine for processing.

Summary

HPE Shadowbase Online Loader and the HPE Shadowbase ETL Toolkit are utility products designed to perform initial database loads, including any subsequent target database resynchronization, providing online and offline loading (or unloading) of a source database into a target database, which can be open for reading and updating while the load (or copy) occurs. The SOLV utility plays an important role in eliminating planned downtime via zero downtime migration (ZDM) and also provides an option for replicating non-audited HPE NonStop Enscribe and SQL data. The ETL Toolkit is used to extract database changes or initial load data into flat files for subsequent ETL loading into a data warehouse using comma-separated value (CSV), fixed-position, and tab-delimited formats. The ETL Toolkit supports a variety of features and databases.

International Partner Information

Global

Hewlett Packard Enterprise

6280 America Center Drive
San Jose, CA 95002
USA
Tel: +1.800.607.3567
www.hpe.com

Japan

High Availability Systems Co. Ltd

MS Shibaura Bldg.
4-13-23 Shibaura
Minato-ku, Tokyo 108-0023
Japan
Tel: +81 3 5730 8870
Fax: +81 3 5730 8629
www.ha-sys.co.jp

Gravic, Inc. Contact Information

17 General Warren Blvd.
Malvern, PA 19355-1245
USA
Tel: +1.610.647.6250
Fax: +1.610.647.7958
www.shadowbasesoftware.com
Email Sales: shadowbase@gravic.com
Email Support: sbsupport@gravic.com



Hewlett Packard Enterprise Business Partner Information

Hewlett Packard Enterprise directly sells and supports Shadowbase Solutions under the name **HPE Shadowbase**. For more information, please contact your local HPE account team or [visit our website](#).

Copyright and Trademark Information

This document is Copyright © 2019, 2020 by Gravic, Inc. Gravic, Shadowbase and Total Replication Solutions are registered trademarks of Gravic, Inc. All other brand and product names are the trademarks or registered trademarks of their respective owners. Specifications subject to change without notice.