the Availability Digest

Replicating Windows and Linux Environments with Double-Take
August 2009

Double-Take (www.doubletake.com) is a unidirectional asynchronous data-replication engine that replicates file changes in Windows and Linux environments over unlimited distances. It intercepts file and directory changes and replicates these changes in real time to one or more target systems, where they are applied to the target file systems. Likewise, multiple source servers[1] can replicate their files to a single target system, where an image of each source file system is maintained.

For data writes, Double-Take replicates only byte changes. Entire files need not be replicated, thus minimizing network load and replication latency. Either upon command or automatically, Double-Take will compare a target file system to its source and will correct any errors that have occurred.

With Double-Take's mirroring capability, files can be migrated to other systems by copying the files first and then keeping them updated via replication. Users are unaffected during the migration process.

Should a source system failure occur, Double-Take can automatically fail over to the target system and fall back if desired when the source system is returned to service.

Any file can be included in the list of files to be replicated. Thus, an entire server can be replicated to the target system. Should the source server fail, the entire operational server can be quickly restored onto the target system. Also, entire servers can be migrated from one system to another without affecting the users.

The Double-Take Data Replication Engine

The Double-Take replication engine installs as a software component between the operating system and the file system in Windows and Linux environments.

Double-Take intercepts file-system modification calls and queues them to be replicated to the target system. File-system modification calls include not only data updates but also other calls such as file creates and deletes, directory operations, file-parameter changes, user authorizations, and so on.

---

[1] In this article, "system" is used to denote a physical system, and "server" describes a logical server running on a physical system.

Double-Take Source Side

Figure 1 shows a simplified diagram of the Double-Take source-system software component.[2] As file-system calls are issued to the operating system by an application, these calls are normally passed to the file system for execution. However, Double-Take inserts itself into this flow to accomplish replication.

The Double-Take source side contains the following components:

- File-System Interface (FSI)
- Source Replication Manager (SRM)
- Source Communication Manager (SCM)
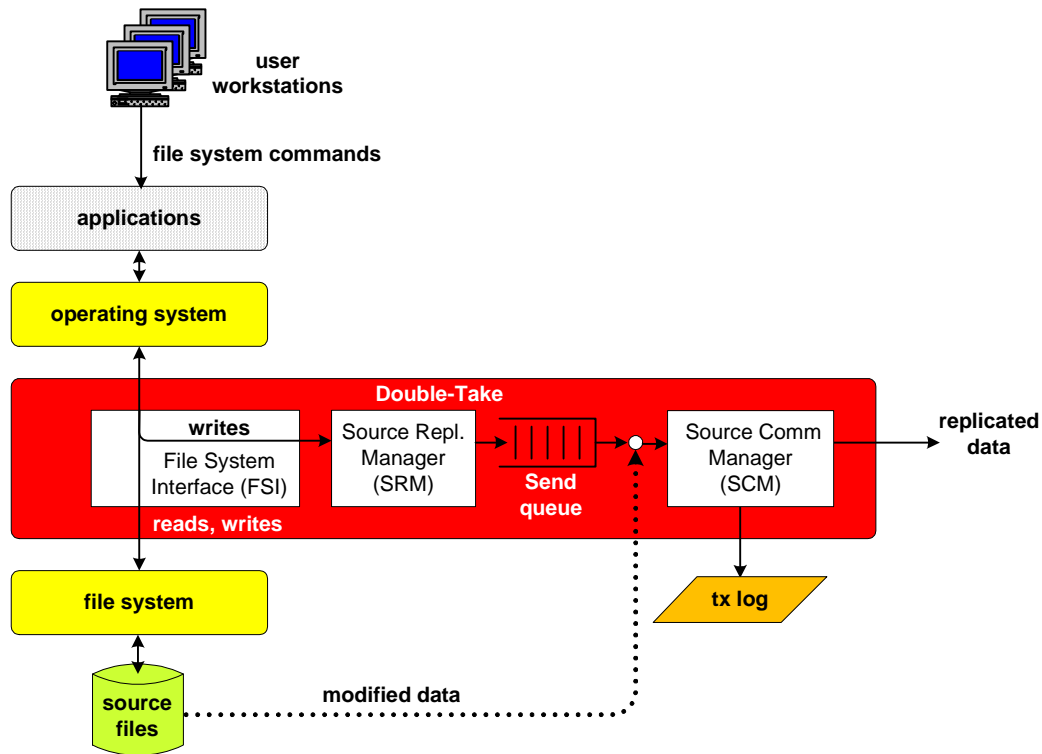
Double-Take Source Server Software Component
Figure 1

File-System Interface (FSI)

The FSI is given a set of files for which replication is desired. This set can be specified by name, by wild cards, by directory, or by volume. The FSI intercepts file-system calls directed to the file system from the operating system. If the call is a read-only call that makes no modification to any file, it is passed directly to the local file system for execution. However, if it is a call that modifies a file that is on Double-Take's replication list, the call is intercepted. These commands include writing data, creating files, deleting files, creating, renaming, or deleting a directory, changing file or directory attributes, changing file ownership or permissions, and any other operation that modifies the file system.

---

[2] The following descriptions of the source-side and target-side Double-Take components are taken from the Double-Take patent, D. E. Beeler, Jr., Real Time Backup System, Patent No. 5,819,020; October 6, 1998.

The FSI will pass the file-system modification call to the local file system so that the call can be executed on the source system. In addition, FSI will pass the parameters of the call to the Source Replication Manager for replication to the target.

Source Replication Manager (SRM)

The role of the SRM is to control the flow of updates to the Source Communication Manager (SCM). If the SCM is currently idle, the SRM will pass the command as a *replication packet* to SCM for replication. If the SCM is currently busy, the SRM will place the replication packet in a Send queue for later replication.

A replication packet contains all the parameters describing the command to be replicated. These parameters include the command, the file name, and if data is being changed, the offset, the number of bytes changed, and the new data.

The Send queue is memory resident unless it grows too large, in which case it overflows to disk. The size of the memory-resident queue is configurable. Its size can be up to one gigabyte for 32-bit systems and is unlimited for 64-bit systems.

Source Communication Manager (SCM)

The SCM receives write requests from the SRM or from the Send queue. It will check to see if any other replication packets in the Send queue contain overlapping or contiguous updates to the same file. If the updates can be packaged into a single update that will fit within a communication block, they are aggregated and sent as a single change.

The SCM logs every command that it replicates into a temporary memory-resident transaction log. When the target system acknowledges that it has made the change, the log entry is removed. If the target system does not respond positively, either the change was lost in transmission or the target system was unable to apply it. In this case, the replication is retried. If the failure persists, the log entry is marked as an error and is kept in the log file. The Double-Take replication engine is rebooted, and an attempt is made to resynchronize the target file system and to continue replication.

The replication channel can be any IP communication link. Source and target systems can be separated by arbitrary distances by using WAN channels. The SCM can be directed to use only a certain portion of the bandwidth so as to not lock out other application activity that is being carried over the network.

As an option, the SCM can operate in synchronous mode. In this mode, the change to the source file system is held up until the target has acknowledged success. If the target system cannot make the change, the change is aborted and is not made to either system.

As further options, the SCM can encrypt and/or compress messages being transmitted. Four levels of compression are supported, including no compression.

Replication latency can be measured in seconds unless the communication channel becomes heavily loaded. Alternatively, replication can be scheduled at intervals to reduce replication load if the potential consequence of data loss due to source-system failure is found to be acceptable.
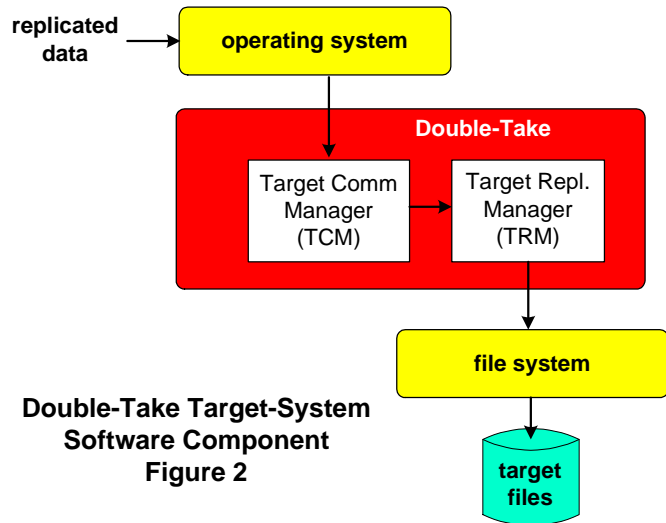
Double-Take Target Side

As shown in Figure 2, the Double-Take target-system software component comprises the following modules:

- Target Communication Manager (TCM)
- Target Replication Manager (TRM)

Target Communication Manager (TCM)

As in the source system, Double-Take positions itself between the operating system and the file system at the target side. Replication packets received by the operating system are passed to the TCM. The TCM passes these to the Target Replication Manager for application to the target file system.

The replication packets are sequenced-numbered to ensure that changes are applied in the correct order to maintain database integrity. Should a replication packet be missing in the replication stream, TCM pauses the passing of replication packets to TRM, thus pausing replication until the missing packet has been received. Unacknowledged replication packets are retransmitted by the source SCM to effect packet recovery.

Target Replication Manager (TRM)

The TRM receives replication packets from the TCM and executes the file-system commands contained in those packets.

While replication is taking place, the target file cannot be opened for query or writes. By turning off Double-Take file locking, a target application can open a file. However, while it has the file opened, replication is paused. There is no mechanism to prevent an application on the target system from writing to a file, and in most cases this would corrupt the file.

System Replication and Recovery

Double-Take is given a list of files (by volume, by directory, by wild cards, or by name) to be replicated. These files can include any file, not only data files. Operating system files, application executables, registries, DLLs, patches, service packs, event logs, system settings, and other information can all be replicated. Therefore, an entire server can be replicated.

Should the source system fail, the operating system, applications, data files, configurations, and all other information are already resident on the target system. All that needs to be done to recover operations to the target system is to start up the operating system and applications and to move the users. Recovery can be done in minutes.

Since Double-Take supports a single target system holding the replicates of several source servers, a single target system can provide a consolidated backup for many servers.

Server Restoration

By simply reversing replication, Double-Take can restore a system that has gone down. In this way, an original source system can be returned to full service when it has been repaired. Once the original source file system has been synchronized with the current operational file system, users can be returned to the original system to restore the normal system configuration.

The restoration of the file system is accomplished via Double-Take's file mirroring capability that copies files from one system to another. If the failed system has not been down long, the restoration of files can optionally be sped up by replicating only file blocks that have changed. To do this, rather than replicating a file block, Double-Take will calculate a checksum for the source file block and send it to the target system. If the checksum for that block at the target is the same, that block need not be replicated. Only if the checksums are different are the blocks replicated. The size of the replication block can be adjusted to optimize mirroring speed when checksum replication is used.

Replication to a file continues during the mirroring process. However, the target system will queue the replicated changes until the mirror has completed. The queued changes are then applied to the file to bring it into synchronization.

Server Migration

Using the same technique as system restoration, Double-Take can be used to migrate servers, for instance, from an old hardware platform to a newer, more powerful platform. Migration is accomplished with no impact on the user. The entire file set is copied from the old system to the new system while the old system remains active. The new server is tested, and users are then moved over to it.

The old server can be kept operational for a while to ensure that the new server is operating satisfactorily. If problems arise, users can be moved back to the old server until the problems are resolved.

Replicating Virtual Environments

Double-Take is at home in virtual environments as well as physical environments. It supports Microsoft's Hyper-V and VMware's ESX hypervisors and can replicate these environments over arbitrary distances.

For virtual environments that use a host operating system, Double-Take can be used in one of two ways. It can either replicate individual virtual machines to different physical hosts, or it can replicate the host operating system with all of its virtual machines to another physical system.

If the hypervisor is a bare-metal hypervisor like ESX, the virtual machines themselves can be individually replicated. In addition, Double-Take has been integrated with ESX at the kernel level so that the entire physical machine with all of its guest operating systems can be replicated.

Using Double-Take, replication can be configured between physical and virtual environments. In addition to replication to another physical server (P2P), a physical server can be replicated to a virtual machine (P2V). Virtual machines can be replicated to other virtual machines (V2V) or to physical servers (V2P).

Replicating Cluster Environments.

Clusters have two major problems that can be solved with Double-Take's GeoCluster product:

- The shared disk volumes represent a single point of failure. Even if they are redundant, they must be collocated and are subject to a site disaster. With Double-Take, the nodal databases can be replicated to other nodes in the cluster or to nodes in another cluster that can take over should the shared database fail.

- Because the nodes in the cluster need to be close to the shared database, they are subject to a data-site disaster. Using Double-Take, a disaster-recovery cluster can be established thousands of miles away.

GeoCluster prevents a failover to a node with a suspect file system. Failover is only allowed to nodes with a known good file system.

Replicating Cloud Environments

Compute clouds and storage clouds may be the wave of the future, but today they are inherently unreliable.[3] It is imperative that any critical applications running in the cloud be backed up, either by running separate instances in the cloud with separate and geographically-distributed file systems or by backing up the processing capability and data storage on another system.

Double-Take can provide a cloud backup facility via its data-replication capabilities. It currently supports Amazon's EC2 (Elastic Computing Cloud) compute cloud.

System Management

The Double-Take Application Manager (DTAM) is used to manage the Double-Take replication environment:

- It ensures that the application environment is configured properly for replication.
- It automatically establishes application data sets and configuration settings on the target system.
- It controls the entire application protection process from configuration through failover, data restoration, and fallback.
- It maintains persistence of configuration settings.

The Double-Take Dashboard monitors the health of the protected servers, the mirror status, and the event logs on one screen.

Double-Take will forward replication statistics and events via SNMP for unified monitoring and control with enterprise management facilities.

Double-Take can send event notifications via email. It can be configured to email different recipients for each server in the replication network.

Operating-System Support

Replication must stay within the same operating system for any source/target combination. Cross-operating system replication is not supported. Though the operating systems may be different versions on the source and target systems, the Double-Take version must be the same on both systems.

---

[3] The Fragile Cloud, *Availability Digest*; June 2009.

Windows

Double-Take can replicate between physical or virtual machines running Windows Server 2008 Enterprise, HPC, Standard, Web, Small Business Server, Enterprise x64 or Standard x64. It also supports Windows Server 2003 or 2003 R2 Enterprise, Standard, Web, Small Business Server, Enterprise x64 or Standard x64. All Windows 2003 operating systems require Service Pack 1 or later.

In virtual environments, Microsoft's Hyper-V or VMware's ESX 3.0.1 are supported.

Replication can be done between any Windows environments, though replicating to an earlier version of Windows may create problems.

Linux

Within a Linux environment, replication may be between Red Hat Enterprise Linux and/or CentOS. Replication can also be done between SUSE Linux Enterprise environments.

Summary

Double-Take, via its file-system replication capabilities, provides many options for disaster-recovery sites and migrations with no planned downtime. It supports Windows and Linux servers. Replication can be between physical systems, virtual machines, and cloud facilities.

Headquartered in Southborough, Massachusetts, Double-Take is a Nasdaq-listed public company (DBTK) with almost 20,000 customers in a variety of industries. Installations range from two servers to several hundred servers. Many industries are active users of Double-Take, including financial (over 1,000 institutions), sporting, health care (over 1,100 hospitals), hospitality, realty, educational (over 1,000 high schools and colleges), and government.