

[Save Article Instructions](#)[Close](#)

## CDP delivers granular recovery

BY DAVE SIMPSON

The first generation of continuous data protection (CDP) products did not live up to their exalted expectations, to say the least. And there were a number of reasons for that.

For one, most of the early pioneers were relatively unknown startups, and most users were loathe to trust something as important as data protection, let alone *continuous* data protection, to startups. Second, most of the large storage vendors and other trusted backup suppliers were not yet in the market (although almost all of them are now, often via acquisitions or OEM partnerships). Third, the early CDP software and appliances were very expensive. Fourth, most users simply did not need the primary feature that early CDP vendors touted: recovery to any point in time.

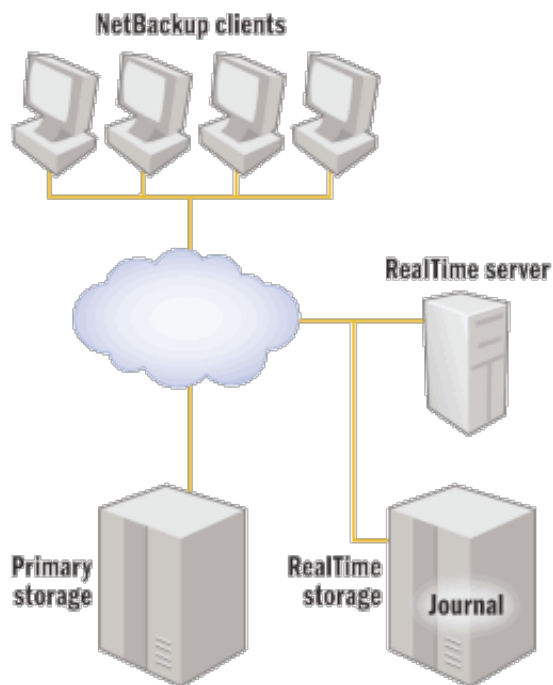
And lastly, there was considerable confusion among end users as to what continuous data protection actually was. End-user surveys (including InfoStor's) during the first generation of CDP revealed a seemingly fast adoption rate for CDP. Upon closer examination, however, it turned out that many users considered weekly full backups with daily incremental backups and recovery from tape to qualify as continuous data protection, thus falsely jacking up the apparent adoption rate statistics.

First, let's clear up what CDP—or at least true CDP—is. The most commonly cited definition comes from the Storage Networking Industry Association's Data Management Forum:

*Continuous Data Protection (CDP) is a methodology that continuously captures or tracks data modifications and stores changes independent of the primary data, enabling recovery points from any point in the past. CDP systems may be block-, file-, or application-based and can provide fine granularities of restorable objects to infinitely variable recovery points. According to this definition, all CDP solutions incorporate three fundamental attributes:*

1. *Data changes are continuously captured or tracked*
2. *All data changes are stored in a separate location from the primary storage*
3. *Recovery point objectives are arbitrary and need not be defined in advance of the actual recovery*

### How RealTime CDP works



**RealTime uses a dedicated server/storage architecture with a journal of application changes to provide continuous protection.**

The SNIA DMF (which includes a CDP Special Interest Group) also states that, although there are a variety of ways to implement disk-based CDP, the key benefits are faster data retrieval, enhanced data protection (through granular recovery and near-instant restore), increased business continuity, and lower overall cost and complexity vs. traditional backup/recovery methods. In addition, CDP essentially eliminates the backup window, providing so-called “window-less backup,” and minimizes data loss in the event of a failure.

### Trend toward integration

The key trend in the CDP space over the last year or so has been a movement away from standalone products to CDP as part of traditional backup/recovery suites and/or integrated into appliances that can be configured with other functionality, such as snapshots, replication, etc.

“The pure play CDP products have pretty much disappeared, and CDP functionality has been integrated into broader solutions, such as appliances from vendors such as FalconStor, InMage, DataCore, etc. that combine CDP with technologies such as snapshots and replication,” says Eric Burgener, an analyst with the Taneja Group research and consulting firm. “And backup vendors, such as Symantec and Atempo and others, have integrated CDP for data capture into their backup products.”

Symantec, for example, offers two CDP products: the Windows-only Continuous Protection Server for its Backup Exec software, and Veritas NetBackup RealTime Protection for NetBackup software. RealTime is block-level CDP, based in part on technology that Symantec got via its Revivio acquisition, that can be used as a standalone product but is typically used in integration with NetBackup and its application-specific agents.

“One of the reasons behind the failure of first-generation CDP was that it was sold as a standalone product,” says Matt Kixmoeller, Symantec’s vice president of product management for NetBackup.

It’s important to note that not all of the backup software vendors have fully integrated CDP into their software, but virtually all backup vendors have some level of CDP as an option.

### True CDP vs. near CDP

According to the Taneja Group’s Burgener, the other key trend is a movement away from so-called true CDP (fine-level granularity of recovery from any point in time) toward what is sometimes referred to as near CDP (or by some industry wags as “snapshots on steroids”), which usually refers to solutions that offer granular, time-based snapshot functionality, but typically without the ability to roll back to any point in time.

“What’s driving the movement toward near CDP is that users really want to recover from application-consistent recovery points,” says Burgener, noting that a lot of CDP vendors offer an application recovery kit that integrates with APIs in specific applications, allowing users to automatically create application-consistent recovery points and then use the CDP product to snap to disk.

Applications such as Exchange and SQL Server are commonly supported by most CDP vendors, but some vendors also support applications such as SharePoint, SAP, Blackberry Server, etc., and databases from Microsoft, Oracle, IBM or Sybase. And breadth of application support (via application-specific markers) is one of the ways that CDP vendors try to differentiate their products.

Despite the general trend toward near CDP, one area where true CDP is particularly advantageous, according to Burgener, is in data forensics for root cause analysis, where users need to figure out exactly what happened when, say, a database became corrupted.

“But generally, users want to be able to recover from transaction-consistent points—as opposed to crash-consistent points—and near CDP is sufficient for that,” says Burgener. “Vendors don’t have to throw out the true CDP functionality, but in reality, most users don’t need it.”

In fact, the true CDP vs. near CDP discussion seems to be fading away because many vendors offer the full spectrum of recovery granularity.

“The argument is irrelevant because you can do both,” says Fadi Albatal, FalconStor’s director of marketing. “It just depends on the customer’s environment. Many companies will never need true CDP; they just need to be able to recover from the latest available snapshots.”

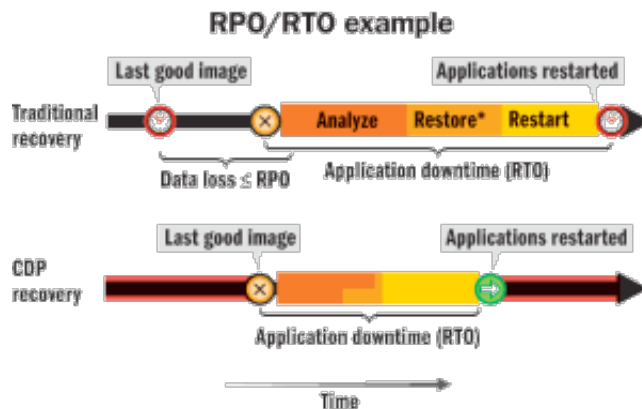
Albatal adds, “It all depends on your SLA [service level agreement], RPO [recovery point objective], and RTO [recovery time objective] requirements. It’s a matter of what is acceptable risk and/or data loss.”

As do other vendors, FalconStor offers a hybrid solution that combines true CDP (continuous journaling), snapshots, replication and other data protection technologies. “You have to have multiple levels of recovery—both locally and remotely,” says Albatal.

Adding a little controversy, NetApp argues that snapshots are sufficient for most companies’ recovery requirements. The argument is simple: “Snapshots can meet most companies’ RTO and RPO requirements,” says David Chapa, director of backup/recovery solutions marketing. “CDP provides finer granularity, but how granular does a company need to get? It just depends on the importance of the data and how granular you need the recovery point to be.”

Although NetApp acquired CDP vendor Topio, the company has not yet introduced a product based on Topio’s technology.

“CDP is more granular than snapshots,” says John Ferraro, president and CEO at InMage Systems. “True CDP continuously captures changes at the block level, and when you recover or failover, you can go back to any point in time. With snapshots, you can’t get near-zero RPO, and it takes up more disk space.”



**This chart shows an example of the recovery point objective (RPO) and recovery time objective (RTO) differences between a traditional recovery environment and a CDP recovery environment.**

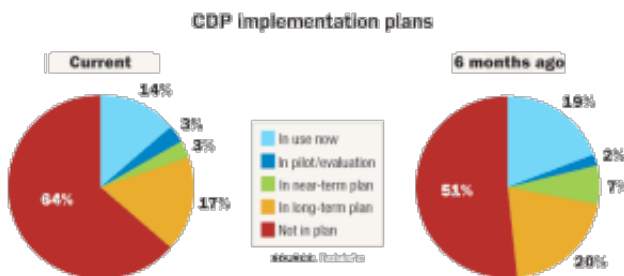
Although InMage is among the true CDP vendors, the company positions its flagship DR-Scout product as a disaster recovery solution rather than CDP technology. “We don’t sell CDP; we focus on disaster recovery,” says Ferraro.

He adds that another key trend in the CDP space (and all other areas of the storage industry) is a tight coupling with virtual server environments, noting that 70% to 80% of InMage’s DR-Scout customers have virtualized their servers.

**Variations on a theme**

A few years ago, the CDP market had only a half-dozen or so vendors. But the market is getting crowded, with more than two dozen CDP players (see vendor listing). And since some vendors use the term CDP loosely, that list could probably be stretched to three dozen or more.

Lauren Whitehouse, an analyst with the Enterprise Strategy Group (ESG), covers the CDP space and closely tracks vendors such as CA, Double-Take, InMage, EMC, IBM, BakBone, FalconStor, Symantec, Atempo, and SteelEye.



But Whitehouse notes that there are a number of different approaches that may fall under the general CDP umbrella. For example, some replication solutions offer the capability to perform continuous data replication (CDR), or the ability to maintain historical replication data (e.g., multiple remote replica versions at a disaster recovery location). This capability protects against situations where a corruption is replicated and allows IT organizations to roll back to a known consistency point in time, although these solutions are not as granular as true CDP. Examples of CDR vendors include Double-Take (which also has a true CDP product, in addition to CDR functionality in its replication software) and CommVault with its Continuous Data Replicator product.

As examples of vendors with “kinda CDP” solutions that essentially use snapshots to provide multiple recovery points, Whitehouse cites examples such as Symantec’s Backup Exec, Asigra’s Televaulting, Microsoft’s Data Protection Manager (DPM) and Iron Mountain (via the technology it acquired from LiveVault).

Lastly, Whitehouse notes that there is yet another CDP subcategory that includes solutions that provide save-on-write functionality for desktops/laptops that enable multiple recovery points. Examples include IBM (Tivoli CDP for Files), Atempo, and Yosemite.

In fact, according to a report from theTaneja Group (“Next Generation Data Protection Emerging Markets Forecast, 2007 – 2011”), low-end CDP products targeted at laptops and remote offices made up about half of all CDP revenues in 2007.

Yet another variation on the general theme is CDP functionality for image-based backup and disaster recovery software. For example, UltraBac Software recently introduced a CDP option—dubbed Continuous Image Protection (CIP)—for its UltraBac 8.3 and UBDR 4.0 software. CIP continuously backs up changed blocks and enables users to restore from any point in time (vs. the last full, incremental, or differential snapshot).

“The other advantage of image-based CDP is that the software is application-agnostic,” says Morgan Edwards, UltraBac’s president and CEO.

In a recent survey of Fortune 1000 storage professionals, conducted by TheInfoPro research firm, only 14% of the respondents said that they were currently using CDP (which is actually down from 19% in TheInfoPro’s survey conducted six months previously). Another 3% have CDP in pilot or evaluation stages; 3% have it in their near-term plan, and 17% have CDP in their long-term plan. However, well over half (64%) do not have CDP in their implementation plans (see figure, “CDP implementation plans”).

## Ready to spend

InfoPro also queried current users of CDP regarding their spending plans on the technology: 44% of the respondents plan to spend more on CDP over the coming year, 38% will spend about the same, and 19% plan to decrease spending on CDP.

Not surprisingly, EMC is the most prevalent CDP vendor among TheInfoPro's Fortune 1000 users. But when asked what vendors were in their future CDP implementation plans (including those in the pilot or evaluation phase), survey respondents cited a number of other vendors, including Symantec, IBM, FalconStor, CommVault, InMage and Hitachi Data Systems (which OEMs Asempra's CDP technology).

## Additional uses for CDP copies

There is a wide range of additional uses for point-in-time copies, including:

- *Surgical recovery of production data*

Recovery of a file or database table from an earlier point-in-time. The required point-in-time can be selected and the file or database table copied or extracted and moved back to the production server without affecting other data on the production volume.

- *Compliance analysis*

Using time-based views, an image can be mounted at a specific point-in-time and compliance tools can be run against the image without impacting the production volume. This can also be useful to access data at an earlier point-in-time before it may have been changed on the production server.

- *Data warehouse seeding*

The biggest cost when seeding or updating a data warehouse is copying the data from the production environment into the data warehouse environment. Data can be exported from the point-in-time copy into the data warehouse without affecting the production server.

- *End of period operations*

At the end of each period (e.g., week, month, quarter, etc.), a point-in-time image of the production data can be used to exactly represent the data as it existed at the end of the period. The image can be used for archiving a copy to disk or tape with the assurance that the archived data is an exact representation of the production data.

- *Cloning of an application environment for development testing*

It is common for users to have a development and test environment where updates to production software are developed and tested before they are deployed into production. A point-in-time image of the production data can be used to build the development and test environment so that the real production data can be used during the development and Q&A cycles, allowing users to expose any issues with the updates before they are pushed to production.

**Source:** Storage Networking Industry Association (SNIA) Data Management Forum (DMF)

---

To access this Article, go to:

[http://www.infostor.com/infostor-2/en-us/index/articles/display.articles.infostor.volume-13.issue-1.special\\_report.CDP\\_delivers\\_granular\\_recovery.html](http://www.infostor.com/infostor-2/en-us/index/articles/display.articles.infostor.volume-13.issue-1.special_report.CDP_delivers_granular_recovery.html)