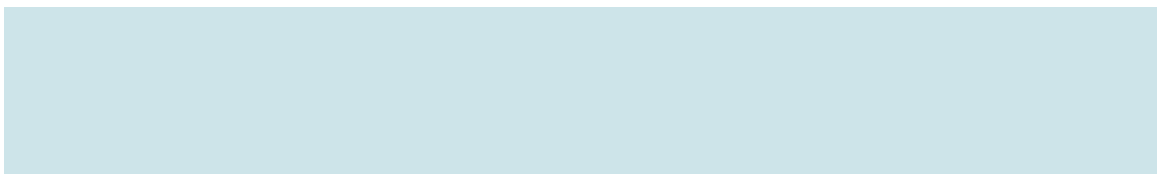




Continuous data protection: addressing timely data recovery and much more

Carl Greiner: carl.greiner@ovum.com

October 2005





Continuous data protection: addressing timely data recovery and much more

Continuous data protection (CDP) with near-instantaneous data recovery capability will be implemented in all major data centres over the next 24 months. It will augment traditional solutions and effectively become a part of the information lifecycle and business continuity solution sets.

The evolution of back-up and recovery continues

Traditional back-up and recovery technologies are tape-based solutions and primarily address data centre disaster scenarios with recovery time measured in days. Such solutions continue to meet over 50% of enterprise disaster recovery (DR) requirements. However, new business continuity (BC) and compliance requirements beyond traditional DR are driving IT organisations (ITOs) to evaluate more timely solutions that can reduce the potential outages and data loss to hours, minutes or seconds, depending on the application, the recovery situation and its business criticality. Fortunately, disk-based storage has seen a rapid acceptance of good enough, cost-effective, high-capacity disks (both SATA and fibre-attached large capacity solutions) as an enabler for more stringent recovery and secondary requirements. Moreover, varying requirements for corporate risk mitigation demand different solutions, and often require multiple solutions within and between applications (for example, depending on user or line of business). In addition, CxOs realise that there is an enormous risk of losing critical data, and the inability to recover in a timely manner transcends the loss of a data centre. Therefore, one size no longer applies, and solution sets must address a multitude of requirements and be readily adaptable to changing business requirements.

A wide range of solutions

The ultimate solution requires multiple data centres with synchronous replication between at least two, with asynchronous capability from these two sites to a third in another region. While such solutions represent the most costly and the ultimate in data centre availability and BC or DR requirements, alternative solutions are required to minimise or eliminate daily and weekly back-up windows, corruption issues, and local hardware and software outages. As a result, there has been wide acceptance of mirroring, snapshots and clones, providing point-in-time data views (dictated by predetermined frequencies; for example, several times daily) that can be recalled if there are problems. However, data consistency evaluations can often be very time consuming or require complex scripts if automated. Virtual tape libraries (tape



emulation to disk) have been well accepted as an intermediary device to tape, and are non-intrusive to the existing data centre environment and back-up software, while mitigating many of the personnel, time and reliability issues of tape. Yet such solutions represent a point-in-time instance of the data (often with incremental write updates) and also require consistency exercises to ensure the integrity of the data. While all the current solutions answer certain requirements, none truly allow for the rollback to the time of the error and/or the last sync point without extensive people evaluation time, which elongates the recovery time and business exposure.

CDP to the rescue

The Storage Networking Industry Association's (SNIA) CDP Special Interest Group's working definition for continuous data protection is '*a methodology that continuously captures or tracks data modifications and stores changes independently 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.*'

More simply stated, the technology provides the ability to access or recreate data to the exact data state as it existed at any previous point in time in seconds or minutes, thus initially transforming local application recovery. One should note that the above definition does not currently address data centre disaster recovery; however, CDP could evolve and be included in the DR continuum. CDP for the most part is currently targeting local data corruption problems such as database corruption, virus or accidental deletion issues with quick restore requirements. The objective is to get the application back online quickly in a consistent form and minimise the loss of data. Mirroring will not help because the mirror copy will also be corrupted. Snapshots can help, but determining what snapshot to use and what updates to apply can be time consuming and problematic. Also, if corruption is not discovered immediately, the rotation of storage supporting the daily snapshots may result in corrupted snapshots as well. Tape is always an alternative; however, the process could take hours/days depending on data centre procedures and back-up implementation.

CDP offers the ability to capture every write and sync point from the application, allowing for a fine-grained data recovery solution in a very timely and controlled manner. Initially, CDP will best address the local back-up and recovery of data associated with a particular set of critical applications with the following general characteristics:

- they have a high number of writes, thus data changes frequently
- they need to run continuously (24x7x365) and have a major business impact when down
- the amount of data (database) is large, thus making activities like traditional back-up difficult and time consuming
- they are normally large transactional systems



- the application has a recovery time objective (RTO) of seconds/minutes (near zero) and a recovery point objective (RPO or amount of data lost) of zero or near zero (last completed transaction).

CDP is disk-based and captures all data writes and maintains a journal for all historical data states, allowing for an 'any point in time' recovery versus the more traditional approach of a scheduled point in time. CDP will not initially replace traditional methods for storing data to disk or tape on a scheduled basis for back-up or long term archiving. However, it will be used to minimise or eliminate local back-ups for critical applications, with off-site back-up being taken at a point in time from the CDP device (or snapshot), utilising existing back-up software. Over time, this could change as CDP becomes an integral part of back-up and recovery and replication solutions, particularly when considering the gradual merging of back-up and archive requirements as information lifecycle management (ILM) becomes more instantiated within ITOs.

Types of CDP solutions

The key to CDP is the constant capture of all data changes as they happen and are time-stamped, allowing for the ability to roll back to any point in time. It is important to note that this is different to periodic snapshots, which are often referred to as near CDP since there is a data exposure between the snapshots and less granular recovery points. In addition, disk capacity requirements for periodic snapshots (particularly full volume) greatly exceed true continuous solutions. CDP systems can be either block (volume) or file based. Block-based solutions operate at the block level of logical devices. As data blocks are written to primary storage, copies of the writes are stored and their timestamp and location managed by some form of metadata manager. Block-level data capture works across all data types (structured, semi-structured and unstructured). Application-level integration is through APIs (such as Oracle and SQL Server). Such integration is required to create consistency of data. For example, a given database transaction may consist of multiple writes, thus information on transaction status must be traced to understand what has to be backed out to achieve a consistent data state. Block-level CDP solutions are targeting the data centres and database application recovery for the obvious reasons of granularity and flexibility across multiple platforms and application environments.

File-based solutions operate in a similar manner to block-based ones, but only at the physical file level. File-based recovery CDP solutions can achieve more granular recovery with applications (such as Exchange or SQL Server) since they can recover at the file level versus the whole volume. File-based solutions are normally platform-centric and application-centric, since there is no common file-level solution across all environments.

CDP implementations differ, yet operate in a similar manner. Some products are software based, while others are appliances. The solutions may operate at the application, file system or volume level; some are agent-less, while others require host-based agents and/or drivers. Some are in the data path (in-band) or outside it



(out of band). The degree of recovery granularity and scalability can vary significantly depending on the solution and the underlying architecture, requiring ITOs to have a solid understanding of their requirements versus solution capabilities. The vast majority of data centre deployments are targeting block-based multi-platform and application solutions due to the need for flexibility and the desire to share the solution, where appropriate, across a diverse environment.

Data consistency and federated environments

There are many federated environments within a data centre, meaning that multiple applications and/or databases are interrelated, running on single or multiple heterogeneous servers and storage arrays. Recovering a federated environment requires consistency between the various components that make up a given environment, thus creating application-consistent replicas. Such recovery points could be used when the ITO does not know when the corruption occurred by determining the right recovery point and initiating a traditional restore and roll forward. Therefore, a robust data centre CDP solution must be able to address the following scenarios and provide consistent recovery points:

- multiple databases, single host, single arrays
- multiple databases, single host, multiple arrays
- multiple databases, multiple hosts, single arrays
- multiple databases, multiple hosts, multiple array.

As with any evolving technology, the delivery of functionality is iterative, with the more complex solutions often lagging behind the more simple straightforward functionality. Since data centres are risk averse, their implementations of the multiple examples listed above (the last three) will normally follow in the deployments of more straightforward and simpler single host/single array ones. This is so that operational procedures can be properly adjusted to fully support the technologies before progressing onto the more complex requirements. Therefore, we believe most vendor solutions will roll out such functionality incrementally.

Where to start

ITOs must have a solid understanding of the risk the corporation is willing to accept by line of business and application. Business executives must establish acceptable tolerance levels for different types of situations – for example, data centre/building disaster, database corruption, viruses, infrastructure outages, hardware and software issues, and out-of-compliance ramifications – as they relate to each and every application. It is only through such a business impact and risk analysis that true expectations are set and the proper solutions deployed. Not all applications require the highest availability solution for all potential business impacts. It is only with business buy-in and agreement that availability and compliance expectations can meet reality. To this end, successful ITOs have establish availability matrices that demonstrate where current applications are, as they relate to basic recovery from the



primary causes of outages, giving business executives a true picture of their risks. In addition, alternative high availability solutions (like CDP) should be indicated with a relative cost increase versus current cost. The primary objective for ITOs should be to attempt to achieve the availability service levels appropriate to a single data centre without a secondary site for disaster. Once a second site is deemed a requirement by the business, the ITO should begin to cost out alternative options by application requirements.

Notwithstanding the above requirement for success, there are a number of ITO-initiated proof points that can be undertaken prior to full corporate buy-in for CDP. Situations that should initially be considered include the following:

- applications with shrinking or no back-up windows
- applications prone to corruption and/or deletions; for example, web-based applications, Exchange, database and software development environments
- large data stores/databases with high transaction rates
- unprotected applications.

As in all such initiatives, ITOs must have baseline costs for their current environment so that proof points can be given as to the additional capabilities delivered, business impact and at what cost. Without such incremental proof points most efforts rarely see fruition.



Client re-use disclaimer

- This is a verbatim reproduction of independent material that has previously been published by Ovum within the last 6 months
- Ovum operates under an Independence Charter. For full details please see www.ovum.com/about/charter.asp
- Ovum may have been paid by the client for the right to re-use the material
- Ovum may have a deal with the client to supply research or consultancy. However, no other relationship exists between the 2 companies (e.g. shareholdings, loans, non-executive directorships etc)
- Ovum does not endorse companies or their products
- While we take every care to ensure the accuracy of the information contained in this material, the facts estimates and opinions stated are based on information and sources which, while we believe them to be reliable, are not guaranteed. In particular, it should not be relied upon as the sole source of reference in relation to the subject matter. No liability can be accepted by Ovum Limited, its directors or employees for any loss occasioned to any person or entity acting or failing to act as a result of anything contained in or omitted from the content of this material, or our conclusions as stated
- This material is the copyright of Ovum Ltd.