



## Cohesity SnapTree™

### Executive Summary:

The 24/7 nature of today's enterprise requires that Recovery Time Objectives (RTO) and Recovery Point Objectives (RPO) are reduced to a minimum. Cohesity Data Protect is built on patented Cohesity SnapTree technology to deliver a robust data protection strategy that delivers (i) Protection from accidentally deleted files, application crashes, data corruption (ii) Data retention for long periods of time to satisfy compliance and regulatory requirements (iii) Low RPO & RTO with a distributed web scale platform for secondary storage infrastructure.

**Challenges with current Snapshot technologies:** Data protection in legacy storage solutions leverage copy-on-write/redirect-on-write snapshot technology to create copies of data. These snapshots of a file system at a particular given point in time form a chain, tracking the changes made to a set of data and form the basis for organizing and storing copies of data. Every time a change is captured, a new link is added to the chain. As these chains grow with each and every snapshot, the time it takes to retrieve data on a given request grows because the system must re-link the chain to access that data. This process is highly inefficient and is unable to cope with the growing demands of achieving shorter RPO and RTO windows.

### Cohesity SnapTree™ Technology

Cohesity has brought innovation to data protection through the SnapTree technology. SnapTree is a 'Distributed-Redirect-on-Write' (DROW) snapshot mechanism that provides speed and scalability in addition to the inherent benefits of RoW snapshot. The design is optimized for write performance so any changes are redirected to new blocks. Additionally, all nodes participate in this process thereby leveraging the scalability elements of the Cohesity cluster. **SnapTree** is available as part of **Cohesity DataPlatform**.

### Comparison between Traditional snapshots and Cohesity SnapTree

	Traditional snapshots	Cohesity SnapTree
Read/write performance	Degrades as the number of traverses to reconstruct data increases linearly with time.	No change as the number of traverses to reconstruct data from any snapshot is fixed.
Number of snapshots	Limited by performance impact of frequent snapshots particularly with write intensive workloads. High data change rate will mean high IO (reads/writes) overhead on any changes/updates after snapshots are taken.	Unlimited because data change does not result in accesses needed to retrieve the data. Moreover, the snapshots are distributed across all nodes in a cluster, hence the term DROW snapshots.
RPO	Order of hours to days.	Order of minutes.
RTO	Since snapshots are not fully hydrated, RTO times are high with traditional snapshots.	SnapTree provides fully hydrated snapshots to recover quickly to a certain point in time.
Cloning for test/dev purpose	Several vendors consolidate changes to build out periodic full copies. This can take significant time due to traversing the entire snapshot chain.	Instantaneous with SnapTree.
Storage consumption	Inefficient due to periodic fulls required to consolidate data changes over time.	Efficient by eliminating the need to do periodic fulls.

## Cohesity SnapTree Benefits

- Achieve faster backup and restore with reduction in RTO and RPO windows by over 90%
- Obtain pay as you grow scaling that ensures maximum data reduction while backups are readily available for instant recovery and clone capability
- Lower Total cost of ownership (TCO) by limiting storage growth from multiple data copies

One Platform.  
Infinite Possibilities.



Cohesity's patented SnapTree technology allows businesses to take a large number of snapshots at any time interval with uncapped retention policies, without ever affecting performance or consuming additional space. Cohesity's DROW implementation keeps track of changes by writing the changed data to new blocks.

SnapTree creates a tree of pointers that limits the number of traverses it takes to retrieve blocks of data, regardless of the number of snapshots that have been taken. SnapTree uses a B+ tree data structure (Fig 1) such that access to any point in the tree takes a fixed number no matter how many snapshots there are, without having to rebuild any chain linkage.

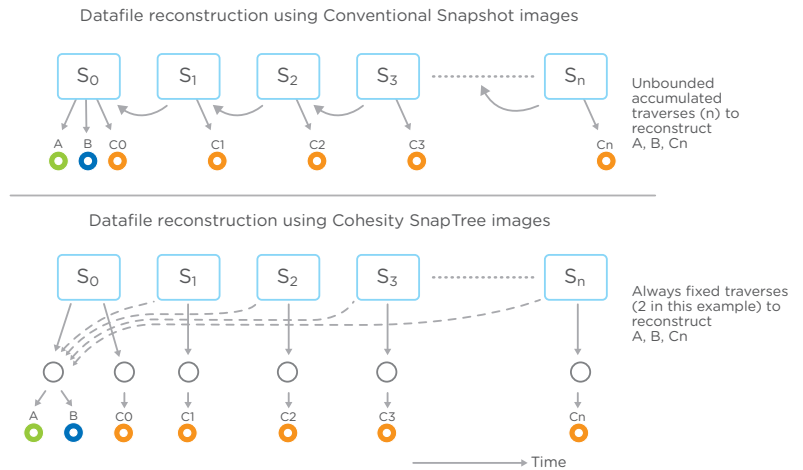


Fig 1. Cohesity SnapTree technology helps create snapshots without incurring the recovery penalty of traversing the entire snapshot chain seen in traditional snapshot architecture.

Because SnapTree is implemented on a distributed file system (Fig 2), every node sees the same nested structure of the chain with a fixed depth independent of where the actual data is stored in the cluster. Keeping the snapshots fully hydrated improves the recovery times of any snapshot from t<sub>0</sub> to t<sub>n</sub> because it does not incur the time penalty of traversing the entire chain of changes. Each of these snapshot clones is fully hydrated so that businesses can achieve fast RTO and near-continuous RPO objectives. SnapTree is available as part of Cohesity DataPlatform.

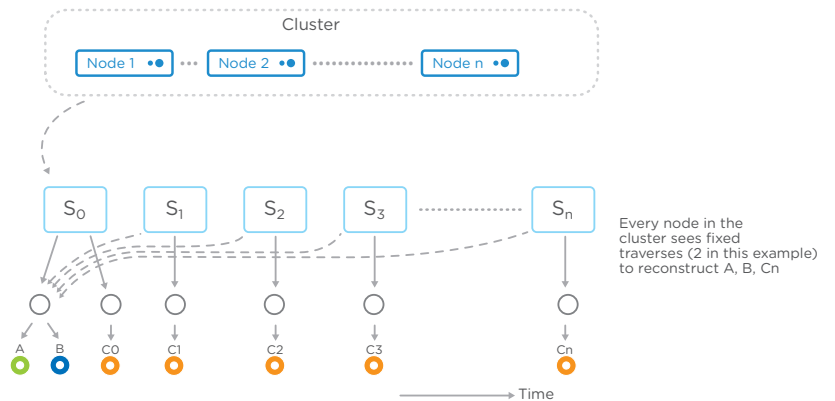


Fig 2. Distributed-Redirect-on-Write' (DROW) snapshot ensures that every node sees the same nested structure of the SnapTree