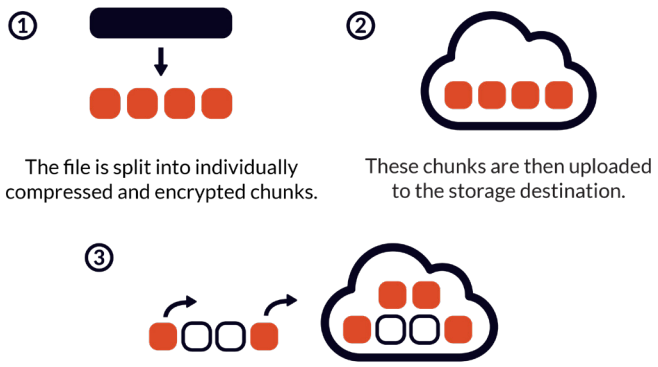
How Does Comet Backup's Chunking Algorithm Work?

Comet's chunking algorithm is the groundbreaking technology that drives Comet Backup. Departing from traditional fixed offset chunking methods and intricate full + differential chain backups, Comet Backup has revolutionized data backup systems. This innovative approach intelligently segments data based on content, dynamically adapting to changes within the dataset.

The result: reduced storage requirements, faster backup times, and unparalleled efficiency as only modified portions are backed up. Comet Backup eliminates the need for repetitive full backups, ensuring optimal resource utilization and unwavering data integrity.

Backing up

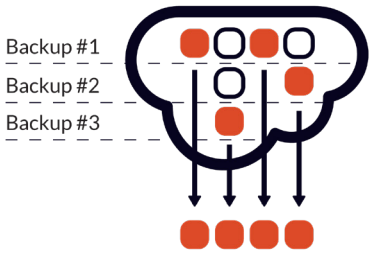


The file is split into individually compressed and encrypted chunks.

These chunks are then uploaded to the storage destination.

When the backup runs again, it detected what chunks already exist and only uploads changed data (chunks).

Restoring



When it comes to restoring, the software is super-fast because it directly downloads only the chunks required for the job.

Comet's chunking algorithm allows for

- ✓ Incremental forever backups
- ✓ Client-side compression and deduplication
- ✓ Fast backup and restoration performance
- ✓ No full reuploads—ever!

What this means for you

Your oldest backup restores just as quickly as your most recent one or any backup in between, without any risk of errors or data loss.

Every backup is immutable, hash-checked, and audited, eliminating the possibility of errors or data loss due to merging.

Comet Backup

How Chunking Compares To Other Backup Algorithms

Feature	Chunking	Always full (e.g. tape rotation)	File versions	Full + diff chain (with periodic reupload)	Full + diff chain (with server-side synthetic rollup)
Backup	✓	✓	✓	✓	✓
Incremental	✓	✗	✓ File-level only	✓	✓
Incremental forever	✓	✗	✓	✗	✓With high IOPS load
Supports cloud storage	✓	✓	✓	✓	✗ Requires complete redownload/reupload
Supports compression	✓	✓	✓	✓	✓
Supports encryption	✓	✓	✓	✓	✗ With server trust
Storage efficiency	✓ High	✗✗ Very low	✗ Low	⚠ Must store multiple full archives	✓ High
Restore performance	✓ Fast, constant	✓Fast, constant	✓Fast, constant	✗ Fast head, slow tail	✗ Fast head, slow tail
Retention cleanup performance	✓ Good	✓✓ Best (overwrite)	✓Good	⚠ Fast, but requires additional storage	✗ Slow, requires rollup
Duplicate file consolidation	✓Perfect	⚠ With long-range solid compression	✗	⚠ With extra scan	⚠ With extra scan
Renamed file detection	✓Perfect	✗N/A	✗	⚠ With extra scan	⚠ With extra scan
Block deduplication	✓	⚠ With long-range solid compression	✗	✗	✗
Renamed file detection	✓Perfect	✗ N/A	✗	⚠ With extra scan	⚠ With extra scan