Backup Software Technology

Backup Technology Whitepaper

Backup Software Technology

All the Dirty Secrets: How Backup Software Actually Works

Introduction

After reading this document you will be able to size up any backup software and have an expectation for how it will perform in your environment. Most backup technology was created before the Internet was a boom and the average business just turned off their lights at the end of the day. Consistent backups are simple when no one is using a computer system. There is little need for high performance backup when you can take a server down for 12 hours to perform the nightly backup.

Does your business shutdown access to all of its computing systems at 5 PM every day? Strange as it may sound that's the assumption backup applications are built around.

What software are you using for your server backups?

- Did you know Legato Networker was first released in 1990 when the largest hard disk readily available was only 270 MB?
- Did you know GNU Tar was written in 1986 when a gigantic hard disk was a whopping 70 MB?
- Did you know the author of rsync Andrew Tridgell did not intend rsync to be used as a backup application? Instead he developed rsync for general purpose file transfer and mirroring over slow WAN links. Tridgell explains in his PhD thesis paper that while the rsync algorithm would be highly useful for optimizing tape backup applications the rsync application itself is not intended to be used for backups.

Contents

- Why Are Server Backups So Painful?
- Categories of Backup Software
- File Based Backup Technology
- Block Based Backup Technology
- Bare-Metal Restore
- Computing Deltas - File Attributes
- Computing Deltas - Check Sums
- Computing Deltas - near-Continuous (CDP)
- near-Continuous V.S. True CDP
- Backup Method - Full Backup
- Backup Method - Incremental Backup
- Backup Method - Differential
- Backup Method - Virtual Full Backup