Realtime
publishers

Protecting Data with a Unified Platform

The Essentials Series

sponsored by

Acronis

Dan Sullivan

Introduction to Realtime Publishers

by Don Jones, Series Editor

For several years now, Realtime has produced dozens and dozens of high-quality books that just happen to be delivered in electronic format—at no cost to you, the reader. We've made this unique publishing model work through the generous support and cooperation of our sponsors, who agree to bear each book's production expenses for the benefit of our readers.

Although we've always offered our publications to you for free, don't think for a moment that quality is anything less than our top priority. My job is to make sure that our books are as good as—and in most cases better than—any printed book that would cost you $40 or more. Our electronic publishing model offers several advantages over printed books: You receive chapters literally as fast as our authors produce them (hence the "realtime" aspect of our model), and we can update chapters to reflect the latest changes in technology.

I want to point out that our books are by no means paid advertisements or white papers. We're an independent publishing company, and an important aspect of my job is to make sure that our authors are free to voice their expertise and opinions without reservation or restriction. We maintain complete editorial control of our publications, and I'm proud that we've produced so many quality books over the past years.

I want to extend an invitation to visit us at http://nexus.realtimepublishers.com, especially if you've received this publication from a friend or colleague. We have a wide variety of additional books on a range of topics, and you're sure to find something that's of interest to you—and it won't cost you a thing. We hope you'll continue to come to Realtime for your educational needs far into the future.

Until then, enjoy.

Don Jones

Copyright Statement

Realtime
publishers

Preserving Data Availability

Businesses run on data. Remove data from today's business, and you remove a critical part of a complicated operation. Part of data protection is ensuring data is available when it is needed, where it is needed, and in the format it is needed. A backup copy of your financial ledger is useful, but if it takes hours to restore that ledger to a server running your financials applications, then your operations that depend on that data are essentially blocked for that span of time. This article, the third in this series, considers:

- The need for data availability
- Methods for ensuring data availability
- The need for adaptability in data protection strategies

Let's begin with business concerns about data availability and some of the risks that underlie those concerns.

Need for Ensuring Data Availability

It is hard to imagine running your business for very long or very effectively when your IT systems are down. Business continuity is one of the chief drivers behind ensuring data availability. Not all systems or types of data are equally important or time-sensitive, and these factors should be taken into consideration when assessing your business continuity requirements.

Varying Impact of Data Loss

Consider, for example, if your department-level data mart were unavailable for two days. Managers might not have up-to-date reports on product sales, sales personnel performance, or other summary reports. The lack of timely reports may be unwelcome but the information will eventually be available when the system is restored. Decisions that might have been made during the outage can be made afterwards with marginal adverse effect on the overall outcome. In use cases such as this, data availability is important in the long run but lack of data for a day or two probably will not have substantial adverse effects.

Next consider a customer-facing Web application that is down because data was lost from the application database. For systems with high volumes of transactions, the longer it takes to restore the data, the more revenue is lost. These types of systems measure recovery time in minutes not days. In order to ensure continuous business operations, you need to be able to rapidly restore data and applications.

Direct business concerns, such as revenue, are just one motivating factor for ensuring business continuity. Some industries and governments are subject to regulations that require data to be preserved in accessible forms. Publically-traded companies in the United States, for instance, are subject to the Sarbanes-Oxley Act (SOX), which requires chief financial officers (CFOs) to attest to the validity of financial reports. This requirement in turn necessitates accurate data. Data loss in a financial management system could significantly impair a company's ability to produce the necessary reports to demonstrate compliance with regulations. The root cause of these kinds of scenarios can vary, although the net effect is the same: loss of data availability.

Threats to Data Availability

Continuous access to your data requires a secure, reliable hardware and software environment. This setup can be undermined by a number of potential problems.

Hardware failures are relatively common. In many instances, you might not be aware of a hardware failure because systems detect and compensate for errors. A desktop OS might detect a failed data block on a disk drive and avoid writing data to that block. A failed disk in a RAID array can be compensated for by recovery mechanisms built in the disk array. More catastrophic failures, such as the failure of multiple components at one time, can lead to data loss. In spite of best engineering efforts, there might be no way for the storage device to recover lost data in such a scenario.

Security breaches pose another kind of threat to data availability. Attackers are often motivated to steal data for their own profit or other motivation. Another risk is compromise of the integrity of your data. This compromise could be intentional, for example, in the case of a disgruntled employee willfully deleting data. It could also be unintentional, such as in cases in which malware designed to steal information or gain control of a device inadvertently corrupts data on the device.

Your own software could also be the root cause of a data loss. A bug in a program could overwrite correct data with erroneous data or even delete data that should not have been deleted. Storage devices are subject to a number of threats that can result in data loss and compromise your businesses operations.

Methods for Ensuring Data Availability

Like other successful operations in IT, ensuring data availability starts with planning and sound practices. The next consideration is the selection of a good implementation option for your requirements.

Plan to Maintain Data Anytime, Anywhere

At the risk of sounding like we are thinking too big, imagine what is required to maintain all of your data anytime regardless of where the data resides. This idea might sound like overreach but consider how quickly data storage practices have changed. Many businesses are using devices that run one or more OSs using one of several virtualization platforms. Employees are storing business data on their personal devices. Sometime those might be laptops, but they could be smartphones or tablets running yet another set of OSs/platforms.

As a matter of design principle, you should solve the data protection problems you have today but also look ahead to plan for changes. For example, you might not have a large number of tablet users today, but that could change within a year. Similarly, you might have limited use for cloud storage currently, but in several months, that could change if vendors drop their prices or offer innovative storage options.

Understand Sound Practices

Sound practices can vary from general principles to very targeted, application-specific procedures. For the subject of data availability in constantly changing environments, sound practices will tend toward the former more generalized end of the spectrum.

Unified View of Backup

Try to maintain a unified view of backups. A single application can depend upon data from multiple sources running on different platforms. If you have to check multiple backup platforms to assess the state of backups, you will need more time and training to get the job done. A consolidated view of backups can make it easier to track the backups you need, identify the data you need to restore, and perform restore operations in a timely and efficient manner.

Plan for Recovery

Recovery operations should be planned in advance of adverse events. These plans should be well understood by systems administrators and documentation should be continuously maintained. The goal is to have easily executed procedures in place so that recovery operations are not delayed while an inexperienced systems administrator learns the details of the restore procedure.

Know Your Time to Recovery and Recovery Point Objectives

Also, plan around acceptable standards for time to recovery and points of recovery. Time to recovery is a term for the span of time that passes between a failure and the time needed to recover. Backups saved to disk might be recovered in minutes; backups to tape that are kept onsite could be used to restore data in a matter of hours; archive tapes stored offsite might have time to recovery measured in days. Understand your recovery point objective, which is a measure of how much data loss you can tolerate in terms of time between a data loss and the last backup. Nightly backups can lead to a one-day loss of data if the failure occurs just prior to the start of the next nightly backup. The best plans and practices require execution with appropriate backup applications and implementation strategies.

Implementation Options

Businesses have several options when planning their backup and recovery implementation. When cost is a primary consideration, disk to tape backups might be appropriate because tape is the lowest-cost backup media. A disadvantage of tape is the time to recover is longer than that of recovering from disk. Disk to disk backup is an option with good recovery time but at a higher cost than tape. You can combine the benefits of fast recovery with lower cost by backing up to disk and then to tape. Tapes could be kept for long periods of time while the higher-cost backup disk would be used only for more recently backed up data.
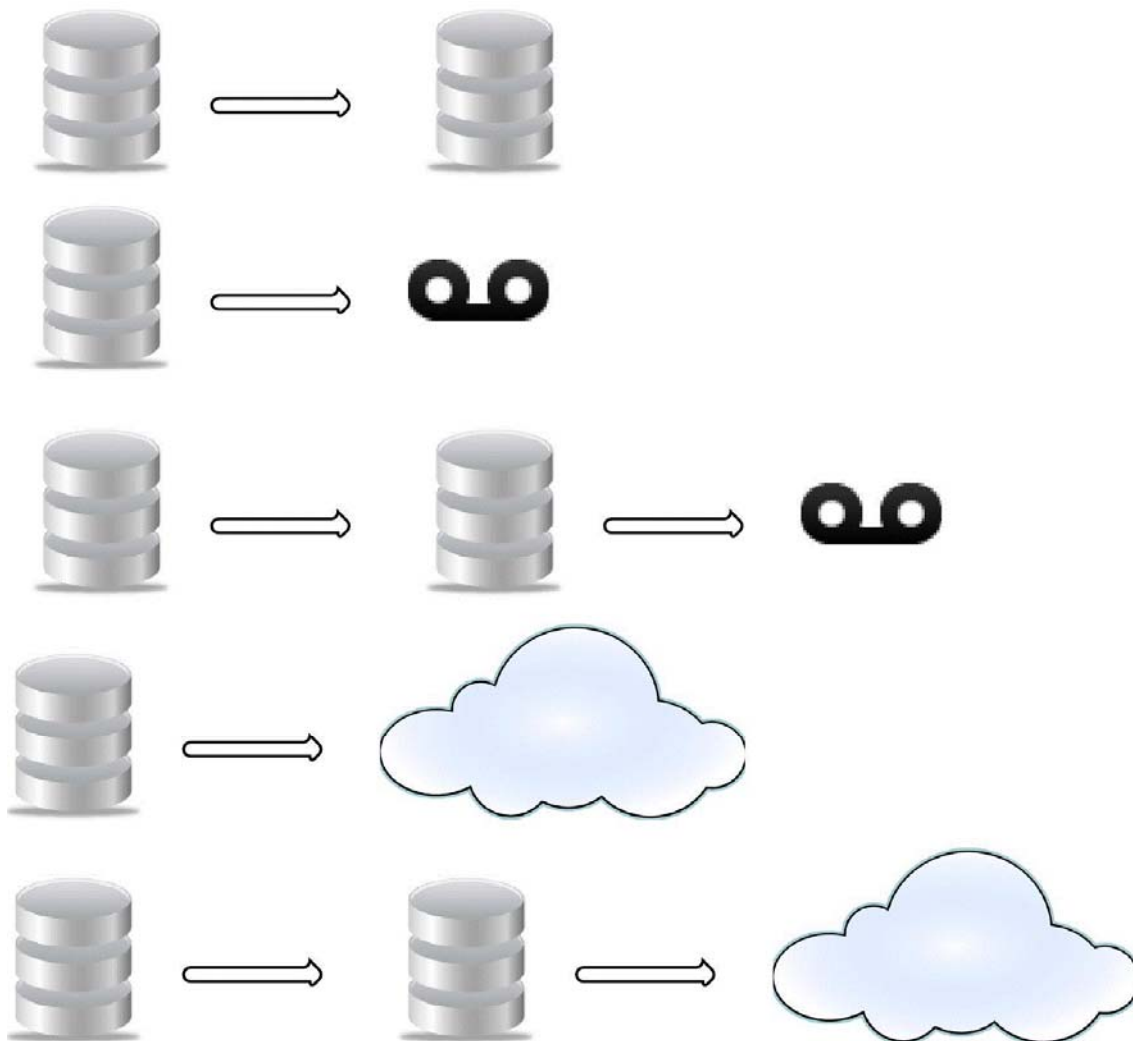

Figure 1: Backups can be performed as (a) disk to disk (b) disk to tape (c) disk to disk to tape (d) disk to cloud or (e) disk to disk to cloud.

Data can also be backed up to public cloud providers directly from disk or in a disk to disk to cloud model. This option allows for growth in your backup storage without significant capital investments. In cases where you still want to have backups on--site for immediate recovery, you could use the disk to disk to cloud model.

Regardless of the implementation model you choose, consider how your backup system will function in a changing environment. Tapes and disk were the primary backup media options in the not too distant past. Public cloud providers are now offering a third cost-effective option. Cloud can also help with data consolidation in highly distributed environments. Unified backup solutions are particularly important to help manage the diverse IT platforms that will depend on the implementation option you choose.

Summary

Businesses depend on their data to enable normal business operations and must plan for potential disruptions caused by human error, hardware failures, security breaches, and software errors. Plan for data recovery operations. This task includes having a backup and recovery plan in place and well documented, assessing your needs for recovery time objectives and recovery point in time objectives, and selecting a backup implementation scheme that fits your needs.

Realtime
publishers