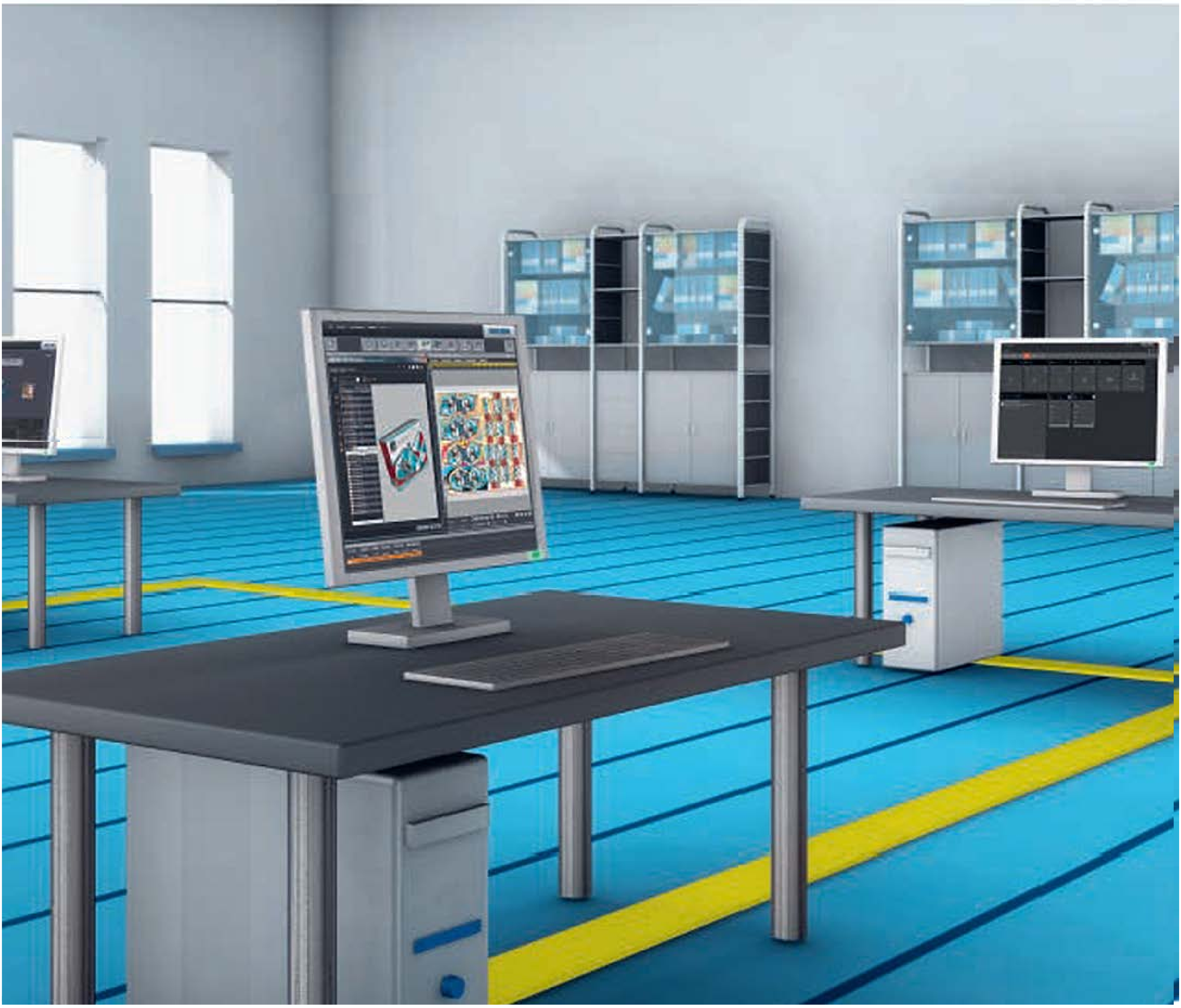
Service

Data protection for print shops.



Foreword

This white paper focuses on the unique issues that print shop owners face in keeping their production running in the event of a system outage. As printers continue to evolve their operations into more fully automated workflows, they must be increasingly concerned with two concepts that often go hand in hand:

Business continuity

The plans designed to keep print production running through various unexpected interruptions such as system crashes, power failures, software and hardware upgrades/migrations, natural disasters, illness or departure of key employees, etc.

Disaster recovery

The plans used to resume print production after some kind of disruptive event such as a tornado or fire, or something like a computer hardware/software failure or a virus infestation. A popular term being used for a single system now is Bare Metal Restore, i. e., bringing a failed system back to life on a new system. This means you need all the system state data (users, device configuration, security policies, etc.). It is human nature to ignore disaster recovery and business continuity planning, because many people consider disasters and accidents to be highly unlikely. In the first half of 2017 alone, 46 disasters were declared in the USA.[1]

You may be surprised by some of the following statistics:

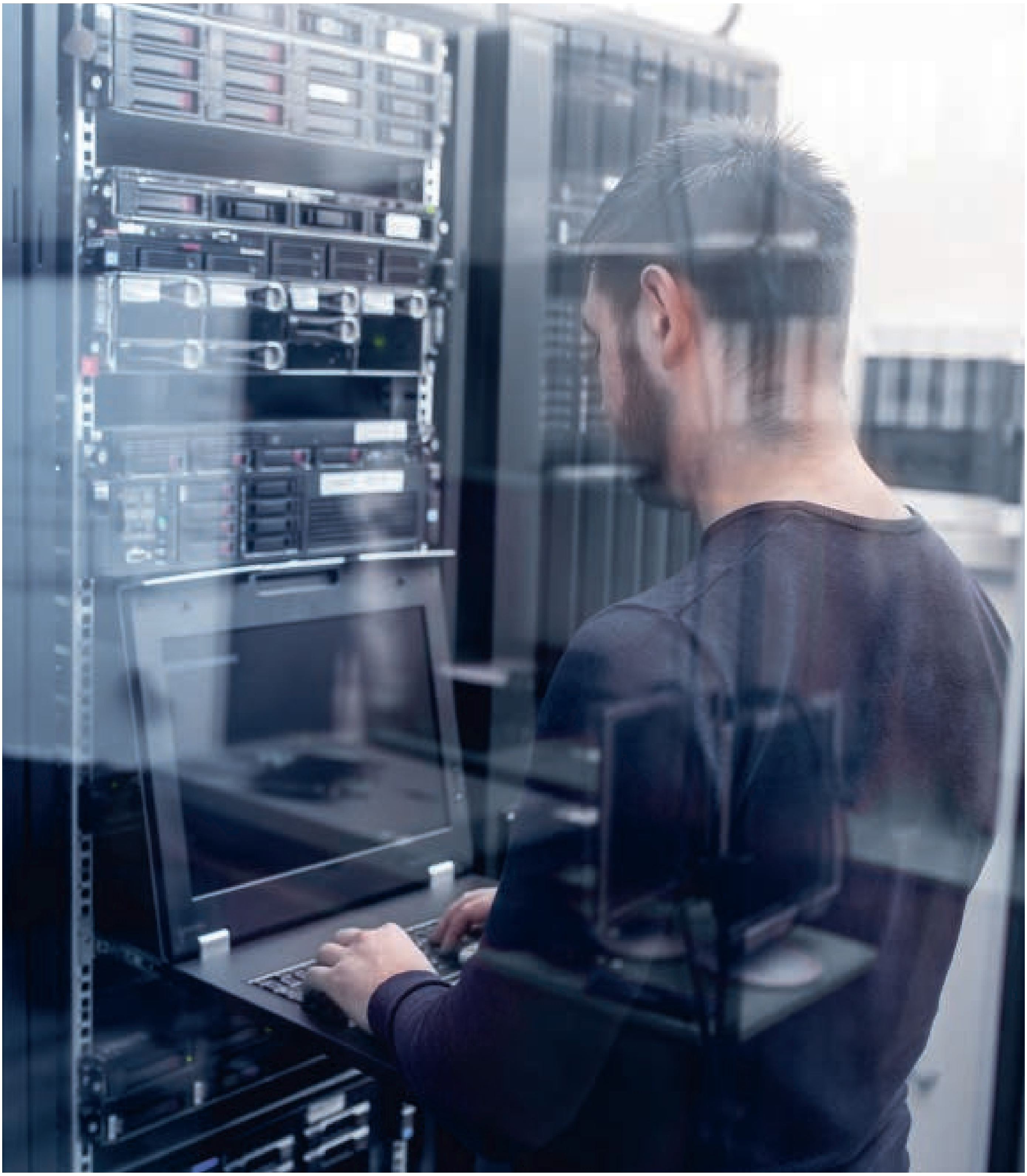
- 10 % of hard drives fail every year – approximately 130,000 a week in the US.[2]
- 6 % of all PCs will experience data loss in any given year. That is approximately 10 per week in the United States.[2]
- 30 % of businesses that have a major disaster go out of business within a year. 60 % fail within five years.[2]
- 34 % of companies fail to test their tape backups. 77 % of those that test it have found tape backup failures.[2]

The magic number appears to be 10 days, meaning printers that are not able to resume full production within ten days are unlikely to survive. Bottom line, if you plan properly for business continuity and disaster recovery, a major disaster can be an inconvenience instead of a catastrophe.

[1] https://www.fema.gov/disasters/grid/year
[2] http://www.bostoncomputing.net/consultation/databackup/statistics/

Content

After severe data loss, printers that are not able to resume full production within ten days are unlikely to survive.



Traditional backup/restore.

Since the advent of time (which in UNIX is January 1, 1970) there has been a need to backup computer systems.

Even with paper tape systems, there were Paper Tape utilities that let you make copies of punched tapes in case they got lost or damaged. As those paper tape systems moved to floppy disk-based systems, there were utilities that made an exact duplicate of the floppy disk data. As hard disk drives emerged, software became available that made a copy of the hard disk drive on several floppy disks or magnetic tape.

While there have been many approaches used over the years, many printers tend to follow the same strategies that they followed in the 1990s:

- They rely heavily on RAID storage to protect their data.
- They keep multiple copies of important files duplicated on different servers, on CDs/DVDs, and, more recently, removable USB drives.
- They purchase backup software for particular machines that they deem to be mission-critical and back up to DLT, AIT and LTO tape drives, sometimes tape libraries with one or more drives and a robotic arm.

Unfortunately, these strategies often provide a false sense of security and peace of mind. In actuality, it is not uncommon to find printers with no backup at all (especially of any laptop data).

Pros and cons of common data backup techniques.

Common backup strategies often provide a false sense of security and peace of mind. It is not uncommon to find printers with no backup at all.

RAID storage

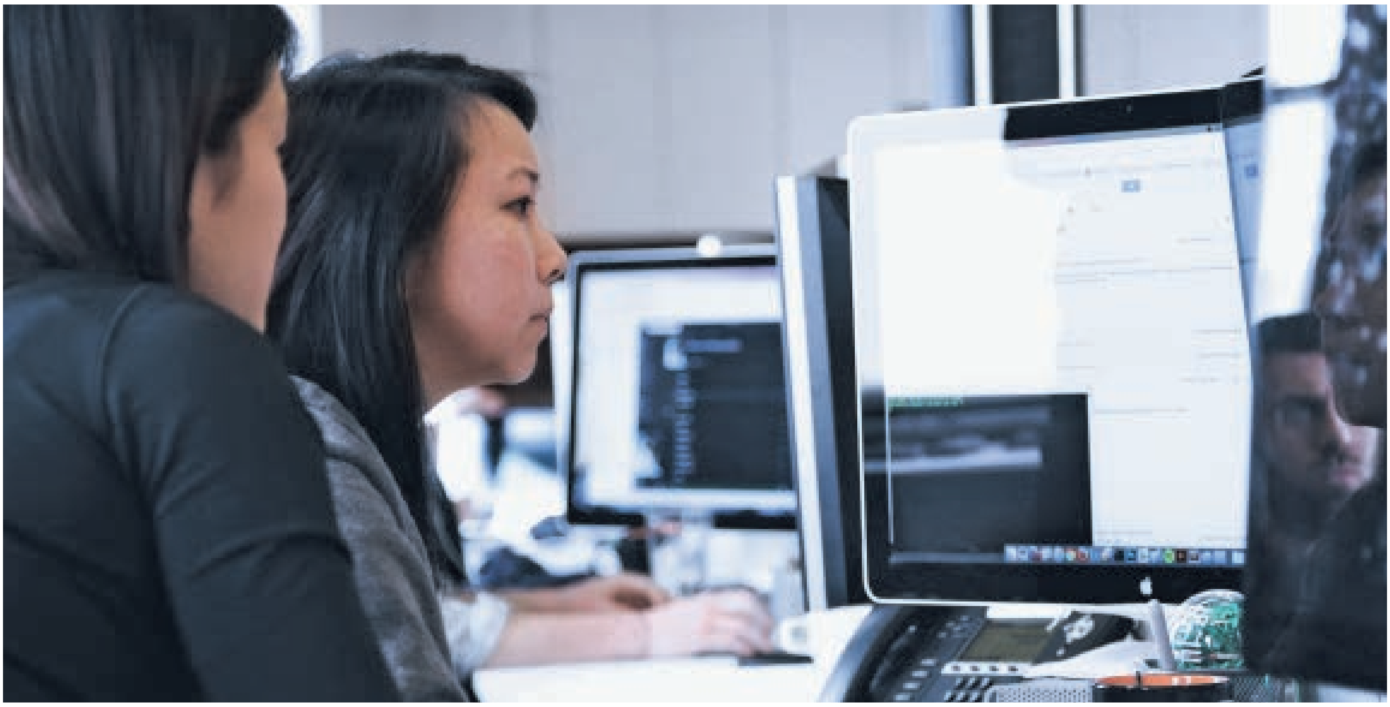
There is no question that all production servers should have hardware RAID storage. It is a proven technology and protects against the high possibility of a disk drive failure. But, hardware RAID storage is not a backup or a suitable replacement for having regular backups. Generally, hardware RAID (only use hardware RAID, not software RAID) is only available for server-class hardware, not workstations. The problem with hardware RAID is that when it fails (e. g., a controller or a multiple drive failure), it can be a lengthy and complicated recovery process. Since it is usually reliable, people are not familiar with the recovery process and move very slowly because of fear of losing data.

Local backup software

This has been a mainstay of backup for the past 25 years. There can be a huge price differential in software: from less than 100 to 5.000 US dollar depending on how robust the features are. Backup software is usually licensed per user or per workstation. There is typically a cost to upgrade to new versions and, sometimes, it is required because of new operating systems (i. e., moving from Windows XP to Windows 7, or moving from 32-bit to 64-bit). Potential problems are cross-platform (Windows, Mac, Linux) environments, where some of the software only works with Windows, for example. The bigger problem is that print applications are so finely tuned that introducing a third party backup software can negatively impact performance as it consumes hardware resources, or can affect reliability due to conflicts. For example, a print application could require a hotfix for the operating system that is incompatible with the backup software.

Backing up to USB drives

A popular and relatively inexpensive solution. The downside: USB drives are generally slow when it comes to performance, and the MTBF (Mean Time Between Failures) may be too low for some commercial business applications. The other possible concern is that USB drives can be easily stolen because people know how to use them, and they are desirable.



Backing up to CDs/DVDs

In the past, this has been popular, but its popularity has decreased with the advent of USB drives. It is fairly inexpensive, especially since it is common to find DVD burners in many modern computers. However, the most common factors driving people to USB drives include slow speed and burn failures. In addition, CDs/DVDs can be easily lost or stolen, you need a person to manage them, and also a place to store them.

Backing up to tape

In the past, it was common for print shops to utilize tape drives or tape libraries for backup – whether DLT, LTO or AIT. Sometimes this approach is used in conjunction with a service for off-site storage of the tapes. The biggest issue for users is the uncertainty in dealing with tape. The tape will not read when the drive is replaced, the process is typically slow, the read and write waiting time is also fairly long. In addition, as this hardware begins to become obsolete, the parts and service become harder to find and more expensive. Also, migration to different media is very slow and, in some cases, impossible. As with any removable media, tape cartridges can be easily lost or stolen, someone must manage the tapes, and they need a place to be stored.

Backing up to hard disk

This has become increasingly popular as the cost of drives has reduced, and the drive capacity has increased. Right or wrong, people feel more comfortable with disk drives – whether it be SCSI attached storage to a server, a network appliance or an NAS (Network Attached Storage)/SAN (Storage Area Network). The problem has been if you lose the storage, you have lost everything, and it cannot be physically taken off-site. In addition, storage demand has grown exponentially in the past decade. This is particularly true for printers who deal with high resolution, 4-color, imposed data. You must look seriously at deduplication technologies that can reduce twenty times the storage requirements.

Backing up to the cloud

Another popular approach is to replicate data off-site to the cloud, i. e., renting storage on the Internet. This offers the benefit of off-site protection, and the maintenance costs of the hardware become the responsibility of the provider. The downside: you have recurring fees and are limited by the Internet bandwidth capability, which is particularly limiting when it comes to restoring the data. There has also been given a lot of attention to potential security issues that can ensue with loss of control.

It is important to look at backup from the perspective of data protection and the high availability of mission-critical systems.

Choosing a new approach.

The problem in the traditional approaches to backup is that no method is all encompassing and none address the challenges printers face in the highly competitive environment of 2017, where it is important to look at backup from the perspective of data protection and the high availability of their mission-critical systems.

Situations that complicate the transition to a more modern approach:

- Printers do not realize the total cost of ownership (TCO) in their existing approach. A piece of backup software may cost a couple of hundred dollars. But you have to look at your investment in the target device (e.g. a tape drive), the media itself (e. g., the tape cartridges), the manpower to administrate both on-site and off-site, and the cost of any software upgrades, training and/or support contracts.
- Printers do not always have on-site IT staff to manage procurement, installation decisions, troubleshoot problems, and handle upgrades.
- Printers often do not really know how effective their backup is because they have never had to recover seperate. A common confusion is that often backup/restore is confused with archive/retrieve functions. They were designed to provide solutions to two very different needs. Backup/restore is to recover from an unexpected event like a hardware failure, for example. Archive/retrieve is intended to retrieve a job that needs to be reprinted or to move jobs waiting for approval to near-line storage until they are approved.

Data protection encompasses the processes a printer follows to ensure that data is securely backed up and recoverable.

Good data protection also means maintaining data integrity. In developing data protection plans, a printer needs to understand the threshold for the following two parameters:

Recovery Time Objective (RTO)
This is the maximum amount of downtime before a printer is negatively impacted in all or part of their print production processes. Any outage that exceeds the RTO would trigger restoring backup data.

Recovery Point Objective (RPO)
This is the maximum amount of data a printer is prepared to lose. For example, you initiate a backup of a mission-critical system each night at midnight. Then at noon the next day, that mission-critical system fails. Recovery can bring you back to the point in time of your last successful backup, which in this case would be midnight. So, effectively, you would have lost any new files or revisions that occurred in the past 12 hours.

Together, a printer's RTO and RPO influence how they determine the level of data protection and high availability they implement in their infrastructure. The smaller the RTO and the RPO you want, the larger the investment you need to make in your infrastructure.

High availability deals with the uptime of your mission-critical production systems.

A “mission-critical” application is one that is essential to the printer, and if it became unavailable for any significant amount of time, it could negatively impact business, i. e., missed production schedules, lost customers or reputation, frustrated employees, lost revenue, inability to accept new jobs, complete jobs or close/bill finished jobs, etc. In addition, it includes any computer that contains sensitive information that is considered private, as well as any computer that is accessed directly by print buyers or that faces the Internet. Finally, any component that is core to the infrastructure, e.g., DNS servers, domain servers, DHCP servers, or any similar server that provides services to the entire plant. This includes any computers running a central database. A “mission-critical” application has some basic principles that should be adhered to:

- The server should perform a dedicated role, not be a catch-all for installing other applications.
- The server should be a server-class computer not a workstation-class computer.
- The server should run a server-class operating system, not a desktop operating system.
- The server should have N+1 power and cooling, and the server should be protected by an uninterruptable power supply (UPS) with battery backup and surge protection. The UPS should, of course, be configured for graceful shutdown of the system in the event of a power failure.
- The storage should be hardware RAID protected.
- The server should be secured with the latest critical security updates from the OS vendor, as well as protected from viruses, worms and Trojan horses or other malware, spyware, or adware with some anti-virus solution.
- Routine system maintenance should be scheduled and performed like backup, disk cleanup, check disks, defragmentation, monitor and clear logs, etc.
- The server administrative accounts must be password protected.
- The server should not be used for general web surfing or e-mail access. These tasks should be done on a desktop computer.
- The hardware and software should be maintained under service contract.

All of the above are good business practices and good computing practices, but they do not necessarily create a high availability environment. High availability identifies and eliminates single points of failure wherever it can. Over the years this has been referred to as fault tolerance or failover. A strong focus is placed on this in the area of virtualization.

High availability is measured by how often a system will fail and then how long it will be down for. High availability is usually expressed as a percentage of uptime in a given year. In a given year, the number of minutes of unscheduled downtime is calculated for a system. The total unscheduled downtime is then divided by the total number of minutes in a year (525,600), producing a percentage of downtime. This is then subtracted from 100 % to create the uptime. This is what is meant when you hear the term “four nines”, i. e., 99.99 % uptime. Each additional “nine” of availability provides an improvement in having your systems up and in production. The following chart depicts how this is calculated:

Percent uptime	Minutes downtime	Hours downtime
99.99+ %	70	1.2
99.99 %	105	1.75
99.9 %	1,051	17.5
99.5 %	5,256	87.6
99 %	10,752	175.2
95 %	52,560	876

The more mission-critical your system is, the more important this is. A mission-critical system is a system upon which your print shop depends on it being available. This means accessibility and working at a performance level that makes it usable. Of course, there are two kinds of downtime: planned and unplanned. Planned downtime is for routine system maintenance like software updates, backups, virus scans, disk defragmentation, etc.

A data protection analysis will prepare you for the possible “what-if” scenarios and identify potential gaps that need to be corrected.

Summary.

As print shops face growing competition and expand into new products and services for their print buyers, keeping production up and running and ensuring their customers' data is protected is essential.

Downtime equates to lost revenue due to missed deadlines and pulled jobs, which can lead to a poor reputation with customers – concerned that you may experience future downtime. Additionally, downtime can create increased operating costs as a result of paying employees who cannot work during downtime and paying for the resources necessary to recover.

Every business, not just printers, encounters the problem of balancing new technology and streamlining production costs, yet still guaranteeing their customers' jobs are safe and agreed schedules are met.

Imagine if your print production was down for 8 hours or 24 hours or even longer. How much would that cost you? How long would it take you to recover?

What if you spent four hours on a complex new job for a new customer and the data was suddenly lost due to a mistake or a failure? How much would that cost?

If you ever experienced a loss on your personal computer at home, imagine the magnitude of the time and loss increased by a factor of 10 or 100.

A data protection analysis will prepare you for the possible "what-if" scenarios and identify potential gaps that need to be corrected, identify who would be responsible, define your RTO/RPO, and, most importantly, show your customers that protection of their job data and schedules is a priority. Heidelberg can help printers with services designed to protect production. A data protection analysis can help you formulate a business continuity plan and/or a disaster recovery plan. There are also products that work with your workflow to make the transition easier and successful.

Heidelberger Druckmaschinen AG
Kurfuersten-Anlage 52 – 60
69115 Heidelberg
Germany
Phone +49 6221 92-00
Fax +49 6221 92-6999
heidelberg.com

Trademarks

Heidelberg, the Heidelberg logotype and Prinect are registered trademarks of Heidelberger Druckmaschinen AG in the U.S. and other countries. All other trademarks are property of their respective owners.

Subject to technical modifications and other changes.

Liability for contents

The contents of this brochure have been prepared with great care. No warranty or liability is accepted for the correctness, completeness, or accuracy of the information. This brochure does not constitute a contractual offer and is solely for the purpose of providing (non-binding) information.