It's all about Application Availability!

By Garth Tucker and Randy Bucking

Are you like me? Are you sick to death of "expert" DR planning consultants popping up like Whack-A-Moles at the Ex and plowing the same "**Chapter 13 from the *Backup and Recovery Guide***" ground over and over and over and over again without ever discussing our everyday, real needs and what our long term goals should be?

Have you heard all the doom and gloom, worst case, sky-is-falling scenarios rehashed so many times that you can mouth the words before they are spoken, but can't translate this to your real needs?

In reality, how often have you walked into your machine room and found a smoking hole where a server used to reside? There's no doubt that hot-site agreements are an important piece of the puzzle, but they're not the whole story. You should stop and consider the source of the information you're being given before signing on the dotted line. After all, are you likely to declare if you lose two drives in a RAID set? No. It's unquestionably a disaster, but not likely serious enough to pull the trigger on your DR hot-site agreement. What does this mean? Essentially, if we spend our budgets buying big-ticket hot-site agreements right away, we may not have the budget to implement the first steps required to provide for continuity. This also ignores the more common types of disaster, such as someone making an application or network change that affects the user's ability to access mission critical applications, or a disk failure that stops us from getting to our data. A hot-site agreement will likely be part of our solution eventually, but before we get there, let's take care of the foundation before building the roof. A more comprehensive view of how we deal with ensuring our systems are available to the business continuously is required. So, as **Monty Python** says, "Now for something completely different!"

A survey of 450 Fortune 1000 companies, conducted by the Strategic Research Division of Find/SVP, found that downtime costs U.S. business over $4 billion per year.

Business Continuity

The term business continuity is pretty self-explanatory. Simply stated we ensure that our application service level agreements with the business are met, and that business functions perform as designed without interruption. Easier said than done? Continuous operations provide access to information and applications even during planned system maintenance. Unreliable systems are not the only disrupter of information availability. Systems also undergo planned downtime to install upgrades and perform routine maintenance. Planned administrative activities account for 95 percent of downtime in IBM iSeries environments.

Nights... Weekends... Holidays...

These previously comfortable windows for IT maintenance are rapidly being eliminated by the 24x7 requirements of e-Business, ERP, virtual users and global operations operating in different time zones. Meanwhile, maintenance task loads keep increasing. Operations staff struggle to schedule system downtime when it least affects user productivity and business profitability. In order to ensure the highest level of business continuity, we need to start with simplifying our IT structure and management. Common sense dictates that fewer systems equates to less overhead to manage, thus making our management job easier and our applications more accessible.

Our IBM AS400/iSeries/i5 systems come out of their wrappers pretty simple from the start and if you look inside your system, there are all sorts of tools for us to run our business, from TCP/IP servers to system management tools. However, due to intrusion from other platforms, it's been deemed that man cannot live by AS400/iSeries/i5 alone and we all have other platforms in our machine rooms breathing our primary platform's air. So in order to begin simplification of these other platforms, we need to start with storage consolidation.

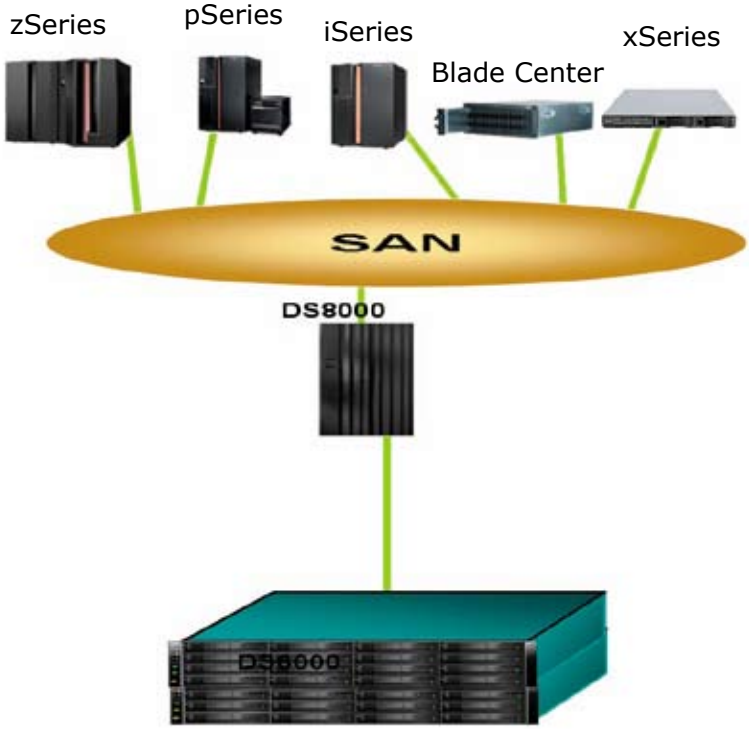
SAN and NAS technology for many platforms bring features such as virtualized storage, fast disk access, improved disk protection, storage optimization, concurrent maintenance and so forth. Everyone has their pile of local disk in the closet: SCSI, Fiber Channel, desktop drives, etc. (Don't deny it.) Do you KNOW how much local disk you actually have? Do you know what's on these drives? And more importantly... is there sensitive data on these drives?

It boils down to security holes, time consuming backup windows, endless support and management headaches, and most of all — a loss of control over the data for which you're responsible. Well, do something about it! Sure, you have a mix

of AS400/iSeries/i5/AIX/*Nix and MS environments; all with different file systems and different capacity, different throughput and different security needs; possibly in different physical locations.

How Will a SAN Work For You?

You have a network infrastructure in place now; servers need to talk to each other and the clients they serve. Why not get more ROI out of that network by placing all that local disk and RAID onto a SAN? Another industry buzzword and a lot of investment right? Well take a look at your staffing support and downtime costs when data isn't available and think again. A SAN doesn't care about file system type, OS, or hardware platforms, it's a *virtual* physical disk pool that appears local to the host and allows you to stash all that scattered data in one place. Now you have control, security, accountability, and availability — and it's becoming easier to leverage the cost of a redundant storage pool against disparate legacy disks.



FICON, FCIP, iSCSI, FC and even CIFS/NFS can all be transported over increasingly cost effective Gigabit Ethernet, FC and even WAN connections (read: remote branch office sites or DR sites) all using secure, portable, common and *available*... IP. Why not? You're already using IP for most of your communications today. Now you provision, backup, audit and secure everything in one place. Failed disk? Failed power supply? Your data is still available. Out of disk space on a certain host? Grow it on the fly. Redundant network links will keep it available. It's portable and allows you to develop best practice storage policies to keep the lights on 24/7 — secure, safe, and sound.

Information Lifecycle Management (ILM)

To quote the Storage Networking Industry Association (SNIA) — http://www.snia.org/home) from October 2004, "ILM is comprised of the policies, processes, practices, and tools used to align the business value of information with the most appropriate and cost-effective IT infrastructure, from the time information is conceived through its final disposition." What this means is that we need to identify the important data and put it on the right media at the right time until we are ready for final disposal.

To Implement ILM:

First we would start with **Planning and Assessment** to identify, evaluate and predict trends. This involves data categorization which allows us to establish policies for our data retention and management.

Next we move on to **Active Data Management** which virtualizes our physical storage. This is accomplished by pooling of our storage by class of service, policy based file allocation/placement, and policy based file migration/simplification and consolidation of file systems. With high-end data mining, CRM and ERP resources sharing the same pool of storage with current email, VOIP voicemail and 2 year old email archives, it's prudent to identify performance, security, availability and accountability (have you heard these before?) policies to those unique servers and the users who need access to them.

Identify the data, active and inactive (archive). Identify each type's users and then current and future capacity requirements. Add in security for those users and the data. Do you see what you're building? Finally identify backup, snapshot, DR, audit or other processes, it isn't hard to see the overall picture once you've mapped it out like this and it's easy to build on and implement. Most companies stuck in the disaster prepared mode haven't accounted for this, let alone have a strategy in place to manage their active data before a disaster strikes. Do you?

Finally, inactive data needs to be assessed and dealt with; this can include email, database and file archives allowing for retention/destruction or long term retention for compliance. Business never stands still. Over time you need to:

- Introduce new products to expand your markets and stay competitive
- Reorganize administrative processes to increase productivity and accountability
- Change production processes to improve efficiency
- Change your sales and marketing channels to address new opportunities and be more competitive

Change never ends. And, as your business evolves, your business databases and applications must evolve with them. These changes inevitably lead to changes in your databases. New fields must be added, existing fields must be deleted, or other schema changes become necessary.

While database changes may be essential, so is maintaining your operations. You cannot afford to stop your business to restructure your databases to accommodate an upgraded application. What is the answer?

The 7 Tiers of Disaster Recovery

First, where did this standard come from? In 1992, the SHARE user group, in conjunction with IBM, defined a set of Disaster Recovery tier levels to address the need to properly describe and quantify various different methodologies for successful mission-critical computer systems' Disaster Recovery implementations. Whew! That was a mouthful...

The Seven Tiers of Disaster Recovery solutions offer a simple methodology of how to define your current service level, the current risk, and the target service level and target environment. Let's look at each of the levels and define where we are and where we need to be. As you may notice, there is a difference in the numbering assigned; IBM does not recognize Level 0 as a tier versus Baker et al, but we will discuss Level 0 as it's pertinent to some shops. IBM splits Level 6 into Levels 6 and 7, so in effect; we sorta/kinda/in-a-round-about-way have 8 Tiers. Confused yet?

How Are These Tiers Defined?

Tier 0

Businesses with a Tier 0 Disaster Recovery solution have no Disaster Recovery Plan. There is no saved information, no documentation, no backup hardware, and no contingency plan.

Typical recovery time: The length of recovery time in this instance is unpredictable. It may not be possible to recover at all. Many of you just finished reading that and thought, "They're dead when disaster strikes..." Well, that's true in a worst-case, smoking hole in the floor scenario. However, if the disaster was a network disruption would you still think the same way? They would experience an outage, but would not be forced to close their business down and apply for food stamps.

This is an unacceptable level to be at, but is probably more common than a lot of us would believe. At a minimum, they need to start backing up their systems and get a basic roadmap to follow in an emergency.

Tier 1

Businesses that use Tier 1 Disaster Recovery solutions send their back up data to an off-site facility. Depending on how often backups are made, they are prepared to accept several days to weeks of data loss, but their backups are secure off-site. However, this Tier lacks the systems on which to restore data.

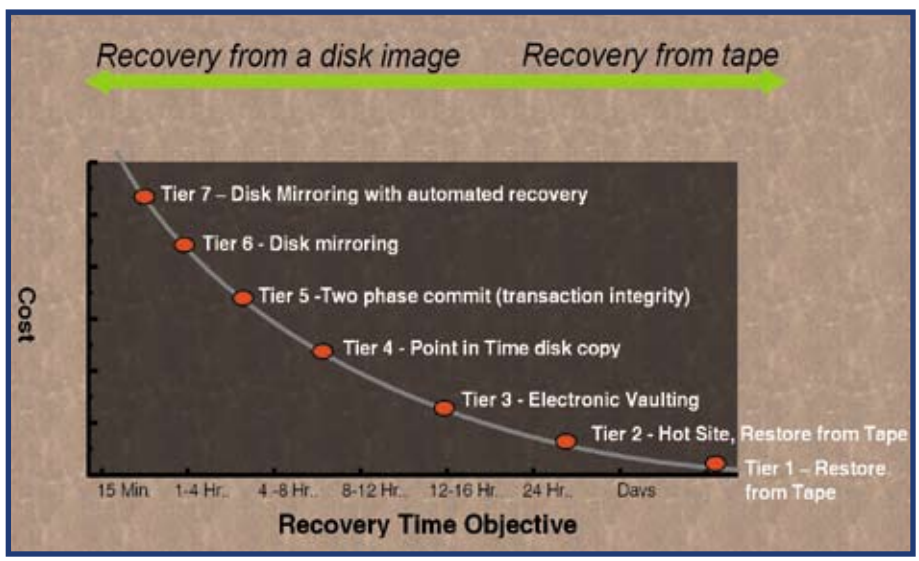
Examples of Tier 1 Disaster Recovery solutions:

- Pickup Truck Access Method (PTAM)
- Disk Subsystem or Tape-based mirroring to locations without processors

In my perambulations through many machine rooms throughout the city this is where I see about 90% of SMALL (< 20 employees) companies. They are usually here because of one of the two following reasons.

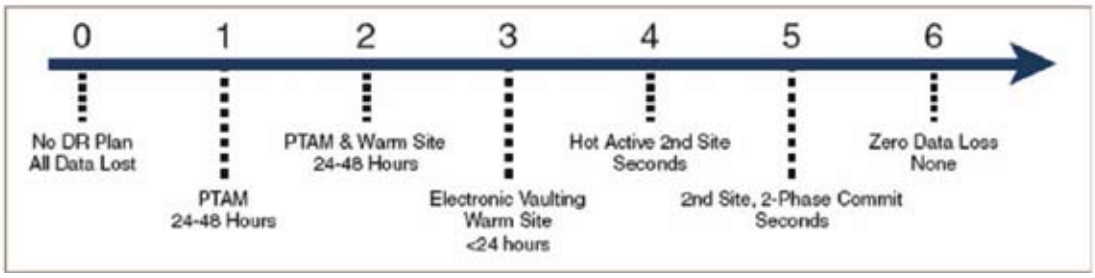
1. Inertia: They have been doing things this way for 20 years and it's worked for them so far. However, "worked" is a relative term. In most cases they have never had to test the theory and so they just believe that it works. In the past 6 months, I have attended two different sites after multiple disk failures in the same RAID set. In the case of one, they were dead. Their backups were not setup correctly and thus they lost a significant portion of their IFS. The second site was luckier in that we were able to rebuild MOST of their data and they ONLY lost 2 days worth of transactions that had hard copies from which to re-enter their changes.

2. Budget: They have not been given the money by the business to implement a better solution. The key here is to prove to the business the folly of not recognizing the danger. Have IBM or your IBM Business



The IBM Seven Tiers of Disaster Recovery

Partner help you make your case and find cost effective solutions.



The Seven Tiers of Disaster Recovery Service (Baker et al, 1997)

Tier 2

Businesses using Tier 2 Disaster Recovery solutions make regular backups on tape. This is combined with an off-site facility and infrastructure (hot-site) in which to restore systems from those tapes in the event of a disaster. This tier solution will still result in the need to recreate several hours to days worth of data, but it is less unpredictable in recovery time. Tier 2 Disaster Recovery solutions:

- PTAM with Hot Site available
- IBM BRMS or Tivoli Storage Manager

A lot of shops that I go into are using some form of this tier and are comfortable enough. However, they may not have gone back to the business recently to ask whether losing a few hours or a days worth of data is acceptable. Most business modules need their eyes opened to the possibilities and they depend on us IT geeks to provide them with this information. If we haven't educated them, they cannot make an informed decision on the level of protection they require. They are responsible to fund the availability of the applications and servers, so let them make the call and accept the responsibility, otherwise if the worst happens; you're going to be left holding the bag.

Tier 3

Tier 3 solutions utilize components of Tier 2. Additionally, some mission-critical data is electronically vaulted. This electronically vaulted data is typically more current than that which is shipped via PTAM. As a result there is less data recreation or loss after a disaster occurs. Tier 3 Disaster Recovery solutions:

- Electronic vaulting of data
- IBM BRMS or Tivoli Storage Manager with Disaster Recovery Manager

This tier is differentiated from Tier 2 by the use of a low level form of HA and management agent software that allows us to automate a recovery, but still requires us to take an informed role in the recovery.

Tier 4

Tier 4 solutions are used by businesses that require both greater data currency and faster recovery than users of lower tiers. Rather than relying largely on shipping tape, Tier 4 solutions incorporate more disk-based solutions. Several hours of data loss is still possible, but it is easier to make such point-in-time (PIT) copies with greater frequency than data that can be replicated through tape-based solutions. Tier 4 Disaster Recovery solutions:

- Batch/Online Database Shadowing and Journaling
- PPRC-XD
- FlashCopy
- FlashCopy Manager
- Peer-to-Peer Virtual Tape Server
- Asynchronous Cascading PPRC
- IBM Tivoli Storage Manager / Disaster Recovery Manager
- iSeries IASPs with FlashCopy

The primary HA functions that SAN and NAS allow for are disk clustering, remote mirroring, and copy services such as Flashcopy. In recent years the iSeries added similar functionality to its disk arrays with Independent ASPs which provides storage that can be dynamically switched between systems. This can be used to keep applications available during planned and unplanned outages by allowing a backup system to access the data in the event of the primary being unavailable. When correctly coupled with Cross-site mirroring (geographic mirroring) these can be very powerful tools for application availability.

Tier 5

Tier 5 solutions are used by businesses with a requirement for consistency of data between production and recovery data centers. There is little to no data loss in such solutions. The presence of this functionality is entirely dependent on the application in use. Tier 5 Disaster Recovery solutions:

- Software,
- Two-phase commit

Tiers 6/7

Tier 6 Disaster Recovery solutions maintain the highest levels of data currency. These are used by businesses with little or no tolerance for data loss with a need to restore data to applications rapidly and have no dependence on the applications to provide data consistency. Tier 6 Disaster Recovery solutions:

- PPRC
- XRC
- GDPS/PPRC Storage Manager
- Peer-to-Peer VTS
- Asynchronous Cascading PPRC
- PPRC Migration Manager
- eRCMF (pSeries)
- GeoRM (pSeries)
- AIX Logical Volume Mirroring
- IASPs with PPRC (iSeries)

Tier 7 solutions include all the major components being used for a Tier 6 solution with the additional integration of automation. This allows Tier 7 solutions to ensure consistency of data above that which is granted by Tier 6 solutions. The recovery of the applications is automated, allowing for restoration of systems and applications much faster and more reliably than would be possible through manual Disaster Recovery procedures. Tier 7 Disaster Recovery solutions:

- GDPS (Geographically Dispersed Parallel Sysplex)/PPRC
- GDPS/XRC (Extended Remote Copy)
- GDPS/PPRC with Open LUN Management
- GDPS/PPRC with HyperSwap
- HACMP/XD (High Availability Cluster Multi-Processing)
- ESS support of GDS for MSCS
- iSeries High Availability Business Partner software

Tiers 6 and 7 provide us the most secure methods of Disaster Recovery in that all of our data is replicated elsewhere and in the event that our building is not available, our business processes and applications will continue at the remote site. These Tiers are the ideal; however, they come with a hefty price point as well. How many of us can afford these solutions? Probably 10-20% of us can afford the price tag.

HA offerings allow you to retain access to your critical data and applications even during system or system-component failures. It is achieved through fault tolerance and other availability management strategies. Today, we rely so heavily on information systems that a computer outage leaves tens, hundreds, or even thousands of well-paid employees with little or nothing to do, and leaves customers without the means to buy from you or to receive service.

Consumers expect to be able to visit your Web site at their convenience, not yours. People in time zones around the world expect to be treated equally. Businesses running multiple shifts expect to be able to interact with your systems whenever their workers are on the job. And, when competitors are only a click away, the potential loss of customer loyalty and revenue due to system downtime is enormous.

Ninety eight percent of respondents in an International Data Corporation (IDC) survey stated that unscheduled system downtime significantly affects their business. Uninterrupted business operations are imperative for today's enterprise. ERP solutions, eBusiness, business-intelligence software and people networked together in virtual organizations are all elements of many modern and successful businesses. These new solutions create a new requirement for 24-hour-a-day, 365-day-a-year data and application availability. No exceptions.

Disaster Recovery Defined – Unplanned downtime results from a variety of events: disk crashes, power outages, hardware failures, software failures, lightning strikes, fires and so on. Many organizations maintain backup systems and copies of critical data at secondary facilities. When a disaster occurs, operations shift to the recovery site.

Do you need a Disaster Recovery Plan?

- Could your business survive a significant loss of data or application usage during a peak business period? **No?**
- Is your company in a highly regulated industry such as finance, banking or pharmaceuticals? **Yes?**
- Is the value of transactions processed each hour particularly high? **Yes?**

Regardless of your answers to these questions, the RIGHT answer will always be a resounding, **YES!** We must have a plan in place to help us recover from even the most innocuous of problems and there must be a level of accountability for each step taken to recover. Losing access to critical information for even short periods is no longer acceptable. You cannot wait a day or more to fix a crashed system and reload lost data. Business information must be available whenever and wherever you need it.

Is it a Disaster or a Fault? Hardware, software and networks occasionally break down. Fault tolerance is required to keep the business running when they do. However, fault tolerance will not suffice in the event of a catastrophic failure, such as a fire or flood. A multiple system strategy that employs geographically distributed systems is required to protect operations in these situations.

And the Right Answer for Application Availability Is?

It depends on your answers to the questions raised in the planning process. There is no silver bullet that will tell you where you should be. Roll up your sleeves and find out what the thresholds are from the business and make them accountable for the decisions after they are presented with all the options and costs associated. TUG

Garth Tucker *is an IBM iSeries/i5 Specialist. IBM Certified in: Technical Solutions Design and Implementation, Sales, Linux and Windows Integration for iSeries with OS/400 versions V4R3 through V5R3 as well as being CompTIA Linux+ Certified. He has many years of experience with AS400, iSeries, and i5; and helped write the Technical Overviews of OS400; V4R4, V4R5 and V5R1 with the IBM/ITSO. Garth has presented sessions at COMMON, TUG MoMs, and TEC.*

Randy Bucking *is a Cisco network specialist with several Professional certifications in Network Design, Security, VOIP, Storage, and LAN/WAN routing and multi-layer switching. He has worked within Microsoft, Sun and various Unix environments and has authored technology reviews, new product and technology testing, and deployment scenarios in industry publications. Randy embraces bleeding edge technologies to allow customers cost effective approaches to solving problems and building solutions.*

Further Reading:

- A white paper on Business Continuity Solutions by **John Ghrist** appeared in *iSeriesNetwork magazine* and was republished on their Web site: http://www.i3tg.com/images/i3iSeriesNetwork.pdf. It discusses some of the same ideas and was also a partial inspiration for this article.
- A global view of application availability can be outlined by looking at *Business Continuity, Information Lifecycle Management and the 7 Tiers of Disaster Recovery*.