Bad things happen to good information systems. That is how life is; everything is moving along swimmingly and then, KA-POW, nothing is moving at all. It is impossible to prevent all bad things from happening; all that can be done is to devise ways to rebound when they do occur. Some organizations are content to wait until something goes wrong before figuring out what to do. This may be fine for small businesses with little information, long lead times for their transactions and extensive insurance policies. Any organization with a lot of data in use all of the time and that must be available shortly following a disruption must plan for recovery in advance of the aforementioned bad things.

**PLANS AND PLANNING**

If it were only clear which bad things were going to happen, this would all be much easier. But it is in the nature of bad things not to let on; that is one of the things that make them bad. This necessitates planning for many more bad things than actually will occur and probably in a greater degree of detail than will be necessary at the time. But to quote former US president Dwight D. Eisenhower, “The plan is nothing. Planning is everything.”¹ In information systems terms, Eisenhower’s dictum means that consideration of needs and acquisition of necessary resources are more important by far than a neatly printed emergency manual.

Organizations must devise emergency response plans for the immediate period following an incident, with the emphasis on preserving human life and safety and only secondarily on information resources. A crisis management plan guides management in making and executing decisions to minimize the effect on an organization until operations return to normal. A business continuity plan prepares organizations to carry on vital (and ultimately all) operations under adverse circumstances.

**DATA LOSS AND DOWNTIME**

The speed and volatility of modern business create some confusion regarding disruptions to and recovery of information systems, as well as the recovery of the organizations that depend on them. For what exactly are plans needed? Total destruction? Inaccessibility? Application mishaps? Only long interruptions or short ones, too? Will one plan adequately address all exigencies?

A disaster recovery plan, as applied to information systems, is intended for response to a catastrophic event that destroys all or most of a data center, renders it inoperable or impossible to reach. This is the so-called “smoking hole” scenario. Of course it is a plan for extreme circumstances, but there have been too many floods, hurricanes, toxic spills, terrorist attacks, fires and airplanes crashing into data centers to discount such events on the basis of rarity. They are credible threats and must be addressed. Briefly stated, organizations need an alternate data center with the right equipment, current data and a network to reach them. They need a set of processes for transitioning and carrying on operations in the alternate data center. Oh, yes, and they need all these things at a price that management considers prudent and affordable.

Questions arise in trying to make the “smoking hole” plan apply to lesser disruptions, such as failures of equipment, software or network services. Is an organization that is prepared for disasters ipso facto ready to deal with interruptions of service? Or, put another way, are service availability and disaster recovery the same or different concerns and can one plan suffice to deal with both? Is a service availability plan the same as a plan for recovering from disasters?

If there is anything positive to be said for a disaster, it is that, as with the sight of the gallows, it wonderfully concentrates the mind. There are no wherefores and maybes; a smoking hole is a powerful inducement to action. The same cannot be said of system failures. A virus, for example,
may cause a service interruption, but it is not disastrous in a physical sense. System failures may go on for some time before anyone realizes which failure caused a disruption. It may be necessary to diagnose what has caused an outage in order to fix it; the same cannot be said of a physical disaster.

More central to the discussion is that the responses necessarily differ, or at least they do most of the time. The operative principal of disaster recovery is to go where the disaster is not. For service interruptions, it generally makes more sense to stay in one place and fix the problem. But one critical element unites them: in both cases, it is essential to have current data. This leads to the determination of how timely the data must be or, viewed differently, how much data can an organization acceptably do without?

This question leads backward to requirements and forward to solutions. There are some business activities that require that not a bit of data be lost. In financial services, millions may be made or lost in seconds, so loss of the data generated in those seconds is unacceptable. Lives are at stake in hospitals and the military, so these industries have similar needs. But, for most organizations, and all organizations some of the time, a little loss—minutes to hours or even days—is tolerable. Similar considerations apply to the determination of acceptable downtime.2

RISK AND AFFORDABILITY

The more downtime and data loss approach zero, the higher the cost will be.

It may be fairly stated that the more downtime and data loss approach zero, the higher the cost will be. The cost is based on having alternate locations with backup equipment and on capturing the data in multiple locations. It also stems from the disk and tape storage required to hold all the data, the network to transport them and the repository in which to store them.

Who, then, is to make the decision regarding risk tolerance and affordability? Business managers are supposed to set the limits of acceptability, but they are often so swayed by the cost of reducing data loss that they understate their needs. Business continuity and disaster recovery managers should respond to business drivers; they are often in no position to contradict the stated needs of business leaders, even if they fear that their organizations are underprepared.

What is needed is a programmatic approach to managing all outages, whether they are caused by disasters or lesser events. At issue are not the causes but the effects, such that disaster recovery and service availability merge, albeit incompletely. No disaster will cause seconds of downtime and no operational problem would be allowed to continue for weeks. In the middle, though, it is possible to evaluate the ramifications—in financial, operational and reputational terms—of outages of various durations with data losses of various magnitudes. Business managers should not be asked how much loss their functions can tolerate, but rather how much money will be lost in seconds, minutes, hours and days of downtime. How badly will operations be disrupted? How much effect will there be on customer and public confidence? If the impact falls within certain ranges, the decision on funding continuity of service, regardless of the cause, should be made impartially and systematically for the business.

In short, bad things can be made better by careful and skillful analysis beforehand. There need to be plans for both disasters and service outages. Very short-term and very long-term disruptions should be seen as extremes that must be planned for separately. If these are called disaster recovery and service availability plans and are kept in two different drawers of the same desk, no harm done. Outages that fall in between—for more than minutes but less than weeks—are the ones most likely to be faced. In this case, the disaster recovery and service availability plans had better say the same things.

ENDNOTES

1 Dwight David “Ike” Eisenhower was the leader of Allied forces in Europe in World War II and served as President of the United States (1953-61). I have seen this quote in various forms, but the gist of it is always the same, putting the emphasis on planning over the product of the process.

2 Astute readers of this column will recall, from previous columns in this space, the terms “recovery point objective” and “recovery time objective” and consider them in this paragraph.