Lessons learned from the 2003 northeastern blackout

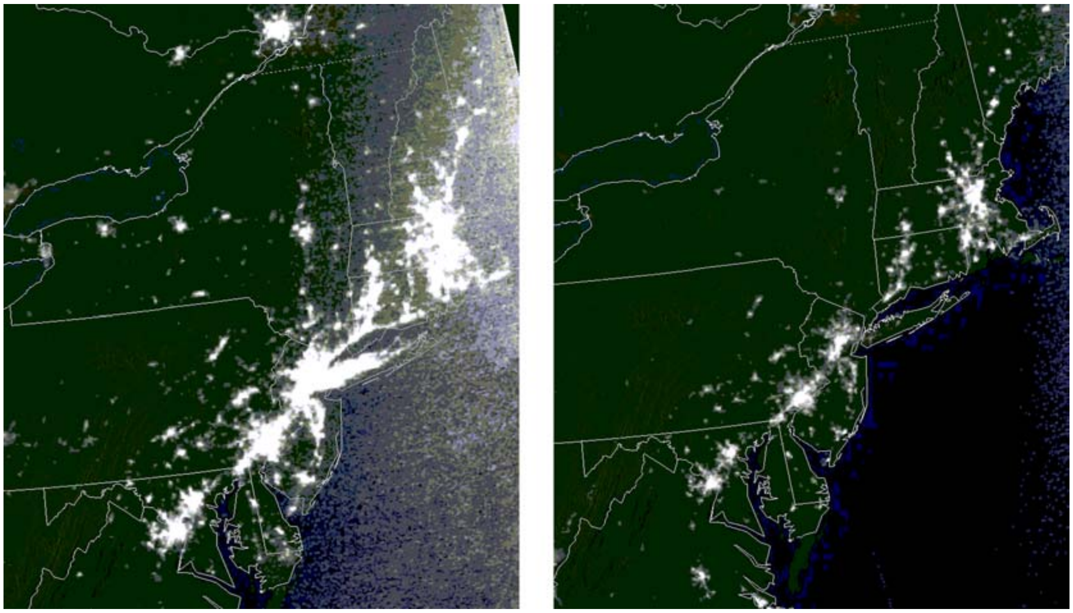
Article

Ron LaPedis has been with the HP NonStop Enterprise Division for 23 years and is the senior product manager for platform security and business continuity products. He has been a Certified Business Continuity Professional since 1990 and a Certified Information Systems Security Professional since 2000. In addition, he is a member of the FBI's InfraGuard program, a public/private partnership dedicated to protecting the nation's infrastructure. He has published many articles and has taught and consulted in the platform security and business continuity fields around the world. Ron is a visiting scholar at the East China College of Computer Technology in Shanghai, and is also a licensed amateur (ham) radio operator, an instructor, and a volunteer examiner.

From 16:10 (4:10 p.m.) on Thursday, August 14, 2003, to 21:03 (9:03 p.m.) on Friday, August 15, 2003, major portions of the northeastern United States and the Toronto area of Canada were hit by a power outage. It spanned 93,000 square miles and affected 60 million people. Millions had no drinking water due to pump failures, mobile phone cell sites stopped working because they didn't have backup power, and cordless phones and voice-over-IP systems failed. Of course, no subways, electric trolleys, or streetcars were operating, and many high rises needed to be shut down because the water, escalators, elevators, fire alarms, and sprinklers in them were inoperative. Also, many buildings had no air conditioning, and the outside temperature was around 92° F (33° C). Luckily, it wasn't winter and it wasn't snowing or raining.

Not everyone was affected

Photos of the blackout area were taken by the National Oceanic and Atmospheric Administration (NOAA). In the figure, the satellite image on the left was taken 20 hours before the blackout, while the one on the right was taken seven hours after the blackout began. Almost immediately after the blackout, a satellite image showing the area in almost total blackness appeared on the Internet and was reproduced in newspapers and magazines. However, this image is a hoax. (To view the false image, go to www.snopes.com/photos/blackout.asp.) In the genuine NOAA image, the pockets of light in the blackout area are evidence that not everyone in the region was affected. This is because there are communities with local power generation capability and residences and businesses with uninterruptible power supplies.



Before

After

Before and after images of the blackout area (photo courtesy of NOAA).

Lessons learned

Because of the events of September 11, 2001, the *Interagency Paper on Sound Practices to Strengthen the Resilience of the U.S. Financial System* (see www.sec.gov/news/studies/34-47638.htm) was adopted on April 7, 2003, by the Federal Reserve Board, the Treasury Department, and the SEC. Due to intense lobbying by local governments and by many financial institutions, the distance requirement for backup sites was watered down from the requirements in the draft versions of the paper. It appears that some cities and companies are still trading away uptime for short-term profit, and many found out too late that their backup sites were located in the same outage zone as their primary sites.

So what lessons did we learn from the blackout? First, we learned that the United States electrical grid is fragile and that such an event could happen again. On that particular day in August 2003, three of the FirstEnergy's high-voltage lines in Ohio sagged into unkempt trees and “tripped” off. Because FirstEnergy’s computerized alarm failed silently, control room operators didn’t know they were relying on outdated information and, trusting their systems, discounted phone calls warning them about worsening conditions on their grid.

After the alarm function crashed in FirstEnergy’s control center, unprocessed events began to queue up, and within 30 minutes the EMS server hosting the alarm process failed as well. When the backup server kicked in, it also choked on the message queue and quickly failed. By the time FirstEnergy operators figured out what was going on and restarted the necessary systems, hours had passed, and it was too late. It took experts in the power industry almost a year to perform a root-cause analysis of the incident, and the conclusion is not pretty.

Second, we learned that this kind of thing can’t only happen in the United States, because similar outages occurred in the United Kingdom and in Italy in the months following the U.S. blackout. Widespread outages have also occurred in Australia and New Zealand in recent years.

The Root cause

Although al-Qaida claimed responsibility for the blackout, it was something much more insidious that brought it about—a minor software error that led to a major outage.

Like many power companies, FirstEnergy’s control center relies on the General Electric XA/21 EMS system. This system, which doesn’t run on the Microsoft® Windows® operating system, has some 1 million lines of code making up the Alarm and Event Processing Routine, written in the C and C++ languages. Mike Unum, manager of commercial solutions at GE Energy, said that his team “spent a considerable amount of time analyzing that, trying to understand if it was a software problem, or if—like some had speculated—something different had happened.”

After several weeks of working late into the evening and early hours of the morning, engineers were able to reproduce the Ohio alarm crash in GE Energy’s Florida laboratory. In the end, they had to slow down the system, injecting deliberate delays in the code while feeding alarm inputs to the program. What they discovered was what is known as a *race condition.* “There were a couple of processes that were in contention for a common data structure, and through a software coding error in one of the application processes, they were both able to get write access to a data structure at the same time,” says Unum. “And that corruption lead to the alarm event application getting into an infinite loop and spinning.”

Peter Neumann, principal scientist at SRI International and moderator of the Risks Digest, says that the real root problem is that makers of critical systems aren’t availing themselves of a large body of academic research into how to make software bulletproof.

“We keep having these things happen again and again, and we’re not learning from our mistakes,” says Neumann. “There are many possible problems that can cause massive failures, but they require a certain discipline in the development of software, and in its operation and administration, that we don't seem to find. . . . If you go way back to the AT&T collapse of 1990, that was a little software flaw that propagated across the AT&T network. If you go 10 years before that you have the ARPAnet collapse.

“Whether it’s a race condition, or a bug in a recovery process, as in the AT&T case, there’s this idea that you can build things that need to be totally robust without really thinking through the design and implementation and all of the things that might go wrong,” Neumann says.

Even the best-laid plans . . .

The New York Stock Exchange (NYSE) ran on generators overnight until regaining power on the Friday of the failure at 06:00, but the American Stock Exchange (ASE) wasn't so lucky. Although it had backup power to its computer systems, it didn't have backup power for its cooling systems. It only opened 15 minutes before closing time on Friday to exercise expiring securities options. It seems the lack of cooling under emergency power possibly could have been discovered earlier if the ASE had fully exercised its plan. How many of you test your generators by starting them to see if they will fire up, but then don't actually run all of your operations on them?

A not obvious side-effect of owning high-availability systems is that unless you have your standard operating procedures documented, no one may remember how to restart the system. In fact during the 2003 blackout, an HP NonStop system customer called the HP Global Customer Support Center (GCSC) to ask for assistance in restarting their system and applications.

Unfortunately, no matter how well your plan is written and tested, things can still go wrong. Backup generators at the 1,946-room Marriott Marquis in New York City failed, despite being tested weekly and having been serviced only three weeks before the blackout. This left the hotel without water, elevators, fire alarms, or sprinklers. Guests had to be evacuated and sleep under the stars. Staff with flashlights climbed 47 flights of stairs to retrieve prescription medicines and other guest items.

Critical infrastructure checkpoints

Think about how much you rely on continuous electric power, even when it may not be obvious. Following any disaster, phone lines and cellular towers may be damaged or simply overwhelmed with volume, making it difficult to get calls through to an area. Chances are high that voice-over-IP systems will be down, and even if your company's main switchboard has battery backup, it may only last for a few hours. Assuming that the cellular infrastructure is up, what happens when the battery in your phone, laptop, or personal digital assistant (PDA) finally runs out of power? What infrastructure does your recovery plan rely on that may not be available?

Here are some key points to keep in mind that will help in your preparation:

- Know your recovery requirements and the risks you face.
- Make sure that you have enough backup power and that all critical components are connected to it.
- Critical areas should have at least one corded, directly connected telephone.
- If your plan requires employees to travel by air to a backup site, do you have an alternate travel plan? Although air traffic control centers and many airports had backup power for some areas, operations were shut down during the blackout because the outage crippled security-screening facilities, bag handling, reservations, and other airline operations.
- You may not be able to rely on mobile (cellular) phones or voice-over-IP systems in a disaster. If your plan relies on them, consider changing it now.
- Amateur radio operators can help with people and property safety. Start a club at work (HP has a very active ham radio operator club). In many countries, a beginner's license doesn't even require learning Morse code, just a basic understanding of electricity and radio safety.

Be prepared

The next time you're in the restroom at work, check out the fixtures. Are the toilets automatically flushed and do the sinks and soap and paper towel dispensers rely on electricity to work? Unless they are battery operated, you may need to evacuate your building for safety reasons, even if the computer room and work areas have backup power. Electronic gadgets may be nice, but only if they are working.

During the blackout, many employees who could not get home had to sleep in their offices. You may want to think about having a three-day supply of emergency food and water on hand for each of your workers in case this kind of event happens to your company. In fact, a program to give employees a discount on bundled emergency kits for home use might be a good idea, too. It's hard to concentrate on recovering key business functions if your employees are hungry and thirsty, or worried about their families.

The basic kit recommended by the U.S. Federal Emergency Management Agency (FEMA) includes a battery-powered radio with extra batteries; nonperishable food and drinking water; a first-aid kit; and soap, water, bleach, and other sanitation supplies. These items are useful for dealing with all types of disasters, not just those related to blackouts.

You'll also want to think about including in your kit a small amount of cash, a flashlight with fresh batteries, and any items that are essential to your health, such as prescription drugs. Being a professional business continuity planner and living in "earthquake central" (and therefore, tending to lean toward the side of paranoia), I have bright orange emergency supply backpacks in my house and in the family cars.

Your family should establish an out-of-state contact person to act as command central in the event of a disaster or other emergency. Make sure this person knows that he or she is the emergency contact and has essential phone numbers and e-mail addresses on hand.

Family members should carry the family contact's phone number and e-mail address with them at all times and plan on calling or e-mailing that person if they have problems getting in touch with each other directly. Once your employees know their families are safe, your recovery should go more smoothly.

In summary

Rehearse (exercise) your plan and then rehearse it again. Make your rehearsal as realistic as possible, and run as much of your operations as you can in backup mode. If you need to, have a contingency plan in case your normal plan doesn't quite work. How far should you go in your planning? What are your risks and what are the cost/loss tradeoffs? You should plan for the 100-year flood, and if planning for the 1,000-year flood doesn't cost that much more, maybe you should plan for that, too.

Spreading backup systems around geographically for security and continuity purposes needs to be done across hundreds if not thousands of miles. Building a backup facility across the street or across the river may not be enough. Companies will need to rethink what it means to create truly redundant business operations.

And finally, whereas security analysts have painted a fearful picture of what would happen if terrorists combined a major physical attack with a well-timed and well-executed cyber assault, a similarly nasty one-two punch by Mother Nature may be just as likely to occur without any human intervention at all.

By the way, did I mention you should rehearse your plan?

For more information, go to www.hp.com/go/nonstop.
05/2004