

# Why 'Plan B' often works out badly

Posted: Friday, March 18 2011 at 05:51 am CT by Bob Sullivan

Engineers used to talk about guarding against the "single point of failure" when designing critical systems like aircraft control systems or nuclear power plants. But rarely does one mistake or event cause a catastrophe. As we've seen in Japan, disaster is usually as function of multiple mistakes and a string of bad luck, often called an "event cascade" or "propagating failures."

In Japan's case, early reports indicate an earthquake knocked out power to the nuclear plant's cooling system, then the tsunami knocked out the backup generators. The third tier of protection - backup batteries -- were only designed to provide a few hours coverage - enough to get the generators repaired. But the backup backup plan didn't account for the time it would take to complete generator repairs under duress, such as when Japan's infrastructure had been decimated by an earthquake.

Separately, the failure of backup systems isn't enough to create a disaster; but taken together, the results can be catastrophic.

"If you add up probabilities independently, everything looks good. But in this case, there is a high degree of correlation and you can't treat these as independent variables," said Bruce Schneier, a risk management expert.

Defending against and preparing for such event cascades is a problem that vexes all kinds of systems designers, from airplane engineers to anti-terrorism planners. There's a simple reason, according to Peter Neumann, principal scientist at the Computer Science Lab at SRI International, a not-for-profit research institute. Emergency drills and stress tests aside, Neumann said, there is no good way to simulate a real emergency and its unpredictable consequences. Making matters worse is the ever-increasing interconnectedness of systems, which leads to cascading failures, and the fact that preventative maintenance is a dying art.

"People just wait to fix things when they are broken," he said.

History is replete with stories of failed backups -- in fact, it's fair to say nearly all modern disasters involve a Plan B gone bad. Neumann keeps a running list of such events, which includes a long series of power outages (and backup power failures) that shut down airports, including Reagan National in Washington D.C.; failed upgrades that felled transit systems like San Francisco's Bay Area Rapid Transit; and backup mismanagement that delayed the first Space Shuttle launch.

There's a simple reason backups work well in theory but often fail when they encounter real-life trouble, Neumann said.

"It's impossible to simulate all the real things that can go wrong. You just can't do it," he said. "The idea that you can test for unforeseen circumstances is ridiculous. When unforeseen circumstances arise, you realize your test cases are incomplete. In general you can't test for worst case emergencies. You can't anticipate everything."

Emergency tests -- like fire drills -- can easily take on an air of artificiality. Think about the last time you lined up to exit a school or office building during a faux fire. Did that really make you better equipped to escape during a real fire?

Those who run critical systems have a hard time simulating the pressures and emotional reactions that come with real crisis. Even if they do, sometimes it's functionally not possible to fully simulate a disaster in progress, says M. E. Kabay, an expert at risk management who teaches at Norwich University.

"It is exceedingly difficult to test a production system unless you have a completely parallel system, and often, you can't. Then, what are we supposed to do, shut off the cooling system at a nuclear power plant to run a test? It's not easy," he said. "Very few people will agree to have their electricity turned off so we can test a response to a breach of coolant. And provoking a critical system that is unstable (like a nuclear plant) is itself unconscionable."

## Why they don't work

Plan Bs can fail dozens of ways, but they often fall into three groups:

**\*Synchronization failure.** It's harder than it looks to keep the backup system in the exact same state as the production system. Think about all the software patches that are installed on your computer; is the software on backup computer completely identical?

**\*Bad fallback plans.** Many failures occur when a system is being upgraded. Risk managers stress the need to be ready to fall back to the system when it worked before, but sometimes, that's not possible. The New York City public library once lost thousands of records this way, as did the Dutch criminal system, Neumann said. In the latter case, criminals actually went free.

**\*Not in working condition.** Backup power generators can sit idle for years. They might be full of fuel, but are they full of lubricant? Are gaskets dry and prone to cracking? Can they really handle a long-term full power load? Hospitals struggle to keep backup generators in working order. More than 100 hospital deaths during Hurricane Katrina have been blamed on the failure of backup power generators; many hospitals simply hadn't planned for 15 feet of water. Even when generators worked, they couldn't power air conditioners to fight off triple-digit temperatures.

It's human nature to let backup systems that are rarely needed degrade over time. In fact, it's built into our DNA, says Kabay.

"From a biological and evolutionary standpoint, if you spend time looking at things that are not a threat you decrease your evolutionary fitness," he said. "A baboon looking around for nonexistent lions is not going to succeed from an evolutionary standpoint. ... Ignoring things is an inevitable response to habitual success."

At the same time, building redundancy into systems makes them far more complex, adding to maintenance headaches.

"Designing fault tolerant mechanics can more than double the complexity of a system," Neumann said, "and that can make the likelihood of failure much greater." It also adds to the likelihood that a backup system will be neglected by busy engineers.

Bureaucracy can also keep engineers from fully testing backup systems, or fully synching them up with online systems. At one point, NASA's rigid code verification process for the space shuttle meant each programmer could generate only three to five lines of code per day, Neumann said. Such processes make it tempting to skip effort to mirror minor changes between online and backup systems

"That's an example of where you still need to go back through the process," Neumann said. "But often don't."

Then there's the key problem of interconnectedness, which makes circumstances ripe for an event cascade. The more systems are integrated, the more a problem in one can spread to another. The classic example is the Morris worm, which took down much of the Internet in 1988, but dozens of bugs and attacks have since spread to millions of computers because of the Web interconnectedness. An even better example -- the cascading failure of the U.S. power grid in the Northeast coast during 2003.

There's still a lot of conjecture around the reasons for the failure of the Fukushima cooling system, but Neumann has sympathy for planners who faced tough decisions when they designed it. While an earthquake followed by a tsunami was a predictable one-two punch, it's doubtful engineers tested their design against a magnitude 9.0 earthquake and the subsequent wall of water it would generate.

"You come up with a worst case scenario and you design the system around that flood," he said. "They clearly hadn't designed for a flood this size."

#### **Safety costs money; tradeoffs everywhere**

One ugly reality of safe system design -- even for life-critical systems like mass transit or nuclear power plants -- is cost. It's easy to say Japanese designers should have spent more on cooling system backups, Kabay said, but most people misunderstand the tricky cost-benefit analysis routinely conducted at such plants. Safety engineers don't have infinite budgets. Every day, they make educated guesses -- in other words, they place bets.

"Many people don't realize that risk management is a balancing act. Somebody had to make a decision at some point about where the cutoff would be. Some group had to decide as best they could that the probability of events beyond a certain threshold had dropped below the level that they could cope with," he said.

Hypothetically speaking, he said, an engineer could raise generators off the floor 10 feet to protect against a flood likely to occur every 50 years, or they could raise them 25 feet to protect against a flood that might occur every 100 years. If a plant has an expected life of 50 years, engineers would probably choose the lower structure and the cost savings. "The cost-benefit analysis said we could make it more resistant to a once in a century event, but that will triple the cost, they'd settle on protecting from a 1 in 50 year event and saving the money."

One terrible irony of risk management is the better you do, the more your techniques will come under attack, Kabay said. The longer we go without a dangerous nuclear event, the more safety engineers are accused of overspending.

"The better precautionary measures do, the less effective they appear," Kabay said. "...There is an exceptional psychological tendency to narrow your functional view and forget the earlier conditions we have improved." That's why funding for preventative measures against major disasters tend to vacillate over a half-generation. The recent memory of a bridge collapse leads to tougher civil engineering laws; a distant memory leads to accusations of overkill and overbuilding. "Many people start thinking 'we're wasting money here, we've been wasting all this money on backup systems we never need.'"

And then there's the fundamental problem of what Kabay calls a "disjunction" between the people who decide how much money should be spent on safety measures, and the people who suffer the consequences of those choices. Often, a detached group of distant stockholders wants to save money, but it's the neighbors who will suffer if there's a radioactivity leak.

"Many times the managers who make the decisions know they won't be around when there's consequences," he said. The only way to fix the disjunction problem is with regulations and laws designed to fix consequences back on the decision-makers -- through fines, criminal liability -- so they share in the risk.

In a world of just-in-time manufacturing and corporate penny-pinching, this is easier said than done, warned Neumann. It's hard to get companies to spend money on Plan B when they are cutting things so close on plan A.

"Preventive maintenance is fundamental, but it is a dying art," he said. Airlines often don't do preventive maintenance until flight checks spot problems, he said. And power companies rarely reserve spare generation power for critical incidents.

"Most companies just ignore things until they get burned."