

Relax! A Failure is NOT an Emergency.

Arjen Lentz, Open Query (arjen@openquery.com)

Modern geeks have a life. The trick is to keep it! I have a daughter, and I want to spend my free time with her, uninterrupted. For a child, LOVE is spelled T I M E. Work has no business invading free time. I don't want to be woken up in the middle of the night by a blipping mobile, or distracted across a weekend by feeling any need to check my email.

One of the most important aspects with this topic is to identify what kind of problem we're actually dealing with. We're technical people and keywords such as high availability, redundancy and failover definitely sound technical. But I've come to the conclusion that it's actually primarily a business issue. It's a matter of strategic focus, and if the company doesn't have this focus, your technical work is unlikely to have the desired improved lifestyle outcome. So we'll cover that first.

Good Business

Many people regard the existence and occurrence of "*emergency situations*" as normal and inevitable. I have chosen to challenge this and I've been shaping my work environment accordingly. I own the company I work in so I have a fair amount of influence; but it's not just for me, it applies to all my co-workers also. We do not have night/weekend duties, or an emergency contact number. Lots of other companies offer emergency services, with corresponding higher cost structures. We just don't offer such services, that's in our offering, our clients know it, and it's proving to be positive differentiator! Why? Offering the same as the competition is senseless, and doing less can be an improvement: we have lower costs, a different focus, and our clients like both. It also enhances our pool of talent as people who want to keep their free time free would, given a choice, simply not apply to jobs that involve emergency duties.

Do you know what do these factors have in common: *stress, anxiety, lack of sleep*? They all affect a person's mental and physical health. If there's too much work, there's a human resource problem and *you* spending more time isn't going to fix it as that's neither temporary nor sustainable: you will burn out. It won't be temporary because the company will grow and your workload is only likely to increase further. Sysadmins love automation, but that still takes time to research, test and implement - so in the end it's still a time/resource problem that needs to be addressed through company policy.

What it comes down to here is that a inadequate resourcing (people, systems, equipment) at each stage of a company's growth actually *causes* emergencies (triggering all the above-mentioned nasties). So you see, without having covered anything particularly technical yet, addressing the problem actually make sound business and commercial sense both for companies as well as their clients and service providers (that is, upstream and downstream).

SLAs and Commercial Risk

I generally ask a new client a) how much downtime they can afford (practically or legally) and b) how much time system recovery takes. These simple two questions result in surprising, interesting and shocking insights. Many have of course never tested recovery, so they just don't know. Clearly, something to work on then.

When they have tried and do know, and the recovery time is longer than their allowed downtime, *I simply suggest to them that they don't in fact have a backup*. It often takes them a while to acknowledge that worrying fact (quite understandably, really).

Through such questions I also get to hear a lot about SANs that can't/won't fail, "enterprise architecture", and Service Level Agreements (SLAs) that ensure the company's continuity. Except they don't. One Google rule is "everything will fail" and there's no small print excluding SANs.

SLAs won't save your business: they contain negative incentives and threats designed to encourage a service provider to deliver a certain level of service, and to a great extent these things primarily help the pre-sales cycle. But the financial penalties do not compensate for your business damage, and neither will a lawsuit after the fact. If you want your business to survive, you won't want to rely on an SLA to manage your commercial risk. You'll want to build your business in such a way that it can cope.

Technicalities

If recovery is not an option in terms of time required using currently present technology, then either different more suitable technology is called for, or an online solution. Yes that's right, if you can't afford being offline, make sure you stay online! 99.999% uptime (the famous five 9s) is not 100%, in fact it's more than 5 minutes short of that. Can your business afford that? If not, then asking for or aiming for five 9s is not just costly, but silly.

Think of the powergrid: providing reasonable reliability naturally costs, but beyond that... ensuring near-continuous power is very costly and strictly speaking not necessary at all. Were you to redesign it now, you might consider that it's not an absolute choice between being up or down. Some systems are more important than others, and in any case having some systems be temporarily unavailable may not affect the rest. That is, if you architect the whole with that in mind.

In your business, it may also be acceptable for a service to be (for instance) slower for a short period of time, without interrupting service altogether. This allows you to not only handle failure scenarios better, but also regular maintenance and upgrades.

Application servers will fail, so we have more of them (also for load). Disks will fail, so we use RAID. Database servers will fail, so we use replication. Replication can fail, so we set up monitoring in a master-master setup with automatic failover for the applications and the slaves.

Note that none of this actually prevents things from breaking, but rather prevents a breakage from having to wake you up. You can, in fact, stroll in to work the next morning, notice that there's been a failure, do a post-mortum to find out what happened, and fix things up.

Think about what a superb arrangement this is! If it were a real emergency, you wouldn't have time to do that all-important post-mortum since your primary task would be to get the system back online! That's how the business focus and resulting technical decisions affect everything.

Your Objective

How do you test or prove resilience of an infrastructure? Why, you pull the plug on each component. Yea, literally, pull it. It's a wonderful thing to do! Even better, make you boss/client do it, and allow them to enjoy the experience of not actually being in an emergency when something fails.

On the “not quite dead” front, you can tell a Linux firewall to drop a portion of the packets, and emulate even more tricky behaviour. Excellent for testing replication and monitoring resilience, and whether failover actually works.

The most costly solutions (note the ones with an “*enterprise*” sticker!) often integrate everything for control purposes. While very resilient, they don’t prevent the possibility of failure so that’s something you’d still have to deal with. So you have to look across all your systems, and on the basis that each component *will* fail. It’s not an if, but merely a when. You should be comfortable with that, not stressed.

Using a some cloud trickery can be handy, particularly with fast-growing or highly variable workloads. Just remember that the hardware behind any virtualisation will, at some point, still fail.

Conclusion

Some or even much of this story might seem rather obvious. But in my experience, companies often focus on one aspect while ignoring others, blindly go with trusted brands or Enterprise labels, or make another misjudgment on one topic, undermining the resilience of their overall architecture. And if a flaw is caught during implementation, does anyone check that everything still makes sense after the one fix? Some changes can appear small but in fact be quite fundamental.

Focusing on the prevention of emergencies is good for your health. Literally. And also that of the company you work for, and your clients. It is simply sound economic and business strategy. Never should it be seen as a luxury or something that’s practically unattainable and thus might as well be disregarded. It’s something worthwhile to focus on, as all the other aspects come together with this one.

In closing, an educational analogy: the Swedes have a goal of 0 road deaths. When a Swedish representative visiting Australia was interviewed about this a few years back, his Australian counterpart challenged this number, stating that 0 was simply unrealistic. The Swede, unfazed (probably not the first time this had been put to him) responded:

“We regard any road toll count greater than 0 as unacceptable, so we have a clear aim for 0. It’s not an easy task and we might not reach it any time soon, if ever. But if not 0, what number would you regard as acceptable?”

Our systems are (generally) not quite that lethal, but they are nevertheless important in different ways. Knowing and acknowledging that everything will fail, we can put appropriate architecture in place that can handle this fact. *My* company is growing during a recession, I sleep well at night, and have a great time playing with my daughter. How about you?