

goitil.co.uk

Complimentary Report

PM1

Top tips for dealing with major system failure's

Author: Rob Sloggett
Co-Author: Greg O'Hare
Publication date: 1st June 2006
Version no: 1.0
Document Name: CR_PM1

Introduction

When it comes to computer systems, regardless of their size, two things are inevitable.

Firstly, as time goes on, they will need upgrading. As hardware becomes faster and memory becomes cheaper, the software houses are spending large amounts of time and money to ensure that the application you bought 2 years ago is now out of date, and out of date means “unsupported”. The only way to keep on top is to upgrade!

Secondly, at some point in time your computer system WILL fail. It maybe 10 minutes after you ran your last backup on your hard drive holding all of your data or in the case of my “sat nav” system, the day before a business trip of 500 miles to a place I had never been to before.

Fortunately system failures can range from “inconvenient” to “catastrophic” and most individuals can manage somewhere between 50 – 75% along this range based on life experience alone.

But what happens when your system failure end up close to the catastrophic end?

Anything between anger, blind panic or indecisiveness can rear its head and this is where this report comes to hand.

Written in consultation between two experienced problem managers, this team have handled everything from an outbreak of blaster potentially affecting 1200 PC's throughout the UK, to planned upgrades that have not gone according to plan finishing with a flood and a fire in the main computer room.

The top tips in the next section are designed to provide you with additional skills to help you manage a major computer system failure where experience and blind luck are just not enough.

<p>The complimentary reports provided by goitil are based on the experiences of ITIL qualified service management professionals. They illustrate best practice and practical implementation of the ITIL framework in real life business situations and are offered as complimentary to the ITIL theory to assist you in implementing it in real life situations.</p>
--

Our top 10 tips:

1. Information is key

The last thing you want to be doing in the middle of a major system failure is scrabbling around for information. Worse still is sat that knowing that all the information you need is saved on the hard drive of the system that has just failed.

To manage a major system failure you usually need the following things:

- Key contact names and number
- Copies of any support agreements
- Technical information and if possible, system diagrams

These items alone will not guarantee you a fix, but they do provide you with all of the raw information you need to get the right people together and to make the right decisions.

Make sure all of this information is held in a paper format, ideally in a file that can be accessed easily. Assign an owner and if anything changes make sure it is updated.

2. Deadlines help focus the mind

One of the first questions we always ask at the start of a major system failure is “when do we need to get the service back?”

Now there are always two times’, the time the customer would like it back (which is normally in the next hour!) and the time the customer actually needs it back.

Working to an unrealistic and unnecessary deadline causes corners to be cut and poor decisions to be made. This can result in only a partial success which just comes back to catch you out later on.

Deadlines should be driven around key business activities and if flexibility exists always take the latest window. Experience has told us that if a technical person tells you how long it will take to fix, always multiply it by 3.

3. When to have a plan b

Having clearly understood deadlines also allows you to re-plan when things are not going well. If you have been working on a strategy for 12 hours and you only have 4 hours left, it may be prudent to start considering other options. It is not normally efficient to approaching a problem from various angles. Firstly, resource and costs do not normally allow it; secondly, it pays to have all of your knowledge working in the same direction.

BUT occasions do occur when no matter how long you have been working on it or what resource has been thrown at it, the break through has not been made.

The hardest (and sometimes bravest) decision is to ask all of the support people to down tools for a short period of time and to start consider contingency plans or other avenues of investigation. Normally they will want to continue to explore their current theory and that’s fine, but at some point you have got to start putting other eggs into the basket and getting the framework of these into place.

4. If needed, allow one to direct whilst another one writes

I had only been in the job for 6 weeks, when I received a phone call from my boss telling me our main computer room was on fire. This resulted in multiple systems failing and having to juggle a lot of deadlines and support people.

The approach was quite simple, he directed the operations and I recorded everything, no matter how trivial or insignificant.

From the times that people arrived on site, to decision we made (and who signed them off) down to things we needed to check and key deadlines, it was my responsibility to capture everything and then remind my boss of any key points or key times.

This is only really relevant for big system failures or when you lose a few systems at the same time, but by clearly define the roles of the “manager” and the “scribe/reminder” the problem can be managed much tighter with key issues not being missed.

Also when you come to trace your steps later on (to prevent it happening again or to review what worked and what did not), you will be surprised the small details that are recorded.

5. Conference calls

Most system failures normally involve 4 groups of people:

- Those controlling it
- Those trying to fix it
- Those who are directly affected
- Those who are indirectly affected but just want to be involved

The most time consuming part of managing a system failure is normally around the areas gathering information, making decision and keeping people updated. Depending on the size of the groups involved, the most efficient way to do this is to utilise one of the many telephone conference call services advertised on the internet. With rates of around 5ppm billed directly to the caller bill, these can be a low cost method of updating large groups quickly as well as discussing possible options without people feeling left out.

Be aware that these non face to face meetings need a strong person in the chair and the expectations of meeting need to be outlined at the start. Due to people’s attention span tending to drift quickly, the chair should take the opportunity to recap and clarify points on a regular basis.

6. Don’t change anything without ...

Recording it and making sure, everyone who needs to agree to it.

It is so easy in the heat of a major system failure to stumble across a possible fix and to make changes in quick succession with no real rationale of why you have done it. Worse still, if the changes do not work (and notice the emphasis on “changes” as it normally results in several), it is very difficult without referring back to records, to remember exactly what was changed and why. This is especially relevant if several different individuals have provided the support over a prolonged period (as at the point of regression, some of them may be catching up on their sleep).

Agreement to change to important for two reasons;

Firstly, most changes come with a risk. The risk that you may extend the outage or the risk that it may do more damage than the current situation you find yourself in are two which spring to mind very quickly. Therefore, it is essential that in controlling the system failure, you set up a small group of people who will validate any decisions and where necessary question the reason for them. These do not have to be people from the end user community, but it is beneficial to have someone who understands the business and can explain the impact if the change causes further issues.

Secondly, some changes need a level of technical understanding. A change to “system x” to resolve a problem may seem straight forward but in doing so may cause a new problem for “system y”. It is essential that all changes are considered within the scope of all of your computer systems, not just the one that is broke.

7. Trade off full service for vital business functionality

The common approach when dealing with a system failure is normally to want to restore the full service but when deadlines are short or the availability of technical support is limited consideration should be given to restoring vital business functionality only.

Vital business functionality as the name suggests, looks at restoring only the key aspects of a system that the business needs to continue its operations to maintain a state of profitability or customer satisfaction. It is not normal for companies to have this documented for each system (although that would be beneficial), but generally is discussed near the start of a system failure to establish requirements early on. This gives the technical support the option of restoring a partial service to keep the business going or to try to fix the full problem.

8. 18 hours is enough!

Whether it is the support staff or the person managing the system failure, productivity and accuracy start to get impaired at around 12 hours into the fault. Once you hit 18 hours, your effectiveness is significantly diminished.

Strangely, this does not seem to be influenced by when you last woke up from a nights sleep, so whether the problem starts at 9AM or 9PM the 12 & 18 hour rule seems to be the same (the difference is in your personal recovery time after getting the next good nights sleep!)

Whenever a system failure is drawn out and due to its impact on you business you are working into the late hours, you should always aim to have a handover point somewhere between the 12–18 hour mark. Don't forget, that this handover should include a full briefing of all the decisions that have been made and any courses of action that are currently being considered along with a summary of key milestones that are due to appear.

9. Blame has no place in restoring the service

It is so easy in the heat of dealing with a system failure to start looking for someone to blame. This is normally linked with the question why?

Why did it happen?

Why did we do it at that time?

Why did we listen to that advice?

Why didn't we take a backup?

All of these whys are great questions and actually add to the prevention next time, but during the failure they can encourage negative feeling or create barriers.

No matter how much you want to pin the blame on someone or something, it should always be avoided, your energy is best invested in finding out why the failure has happened and what you need to do to fix it.

10. Restoration is only half of the job

Once you have returned your service back to a useable state, it may be time to take a deep breath and pat both yourself and any people who have assisted you firmly on the back. But good problem management does not stop there.

The next steps involve fully understanding and documenting both the reason for the failure and what actions were taken to restore the service back to normal. Following this, actions to prevent it happening again need to be identified and where costs allow, implemented.

Further information on this final step can be found in report PM2 – Major system failure reviews and preventing future failures.