WHAT MAKE A COMPLEX SOCIO-TECHNICAL SYSTEMS BRITTLE: EVIDENCES FROM AN EVENT ANALYSIS

Luigi Macchi[1], Florence Magnin[1] and Jean Paries [1]

[1] Dédale S.A.S., 15, Place de la Nation, Paris, France

lmacchi@dedale.net

Abstract

A considerable operation was planned, prepared and executed in 2014 by the IT division of a major European corporation. The operation consisted in replacing two cores of the two data centres to simplify the architecture of the network, to improve its exploitation and to allow future technological evolutions. Despite months of preparation, several rehearsals in mock-ups, and despite trained, competent and experienced personnel, the replacement did not go as expected: part of the corporate network was slowed down, some communication lines were shut down, and back-up processes were disturbed. To cope with these problems, a crisis management unit was put in place and several days were needed before the IT division was able to return to its normal functioning. The relevance of the operation, the fact it was brought forward by the division management and the fact the solution of the problems required the participation of multiple actors contributed to make this event a major corporate crisis. An internal "accident" analysis was quickly performed by the IT division personnel. Four independent dysfunctions were spotted and a number of technical failures and human errors were identified as contributing factors to the event.

Unfortunately this event was not an isolate case. In the months preceding it, as well as in the following period, a number of other smaller unwanted events occurred. This situation shook both the team in charge of the operation and the whole IT division. It undermined the image of the division *vis-à-vis* of the clients, and the confidence they had in their ability and competences. The self-confidence of the IT division personnel was tackled as well. Overall, people's mind was deeply marked.

Despite the performed accident analyses the management of the IT division was still unsatisfied with the understanding of the "*deep*", root causes of the events. It was thus decided to take a closer look at the human and organisational factors which could have contributed to the "accident" and crisis. Since the internal accident analysis mainly focused on what happened during the operation, the management expressed the will to expand the analysis to the preparatory phase as well as to the crisis management one.

This paper presents the analysis process and the results of the investigation performed by the authors.

1 A WEAKENING CONTEXT

To understand the reasons behind the accident and to explain why few human errors led to a major organisational crisis, what happened during the planned operation has to be set in the overall organisational and operational context.

One of the leading objectives of the IT division is to improve and maximise its performance and to provide its clients (which are other divisions of the same corporation) the best and more effective technological solutions for their operational needs. This objective, coupled with the inherent tendency of the IT domain to look for and develop new technologies at a high pace, pushes the IT division to move towards a virtualisation and mutualisation of its services and products. The virtualisation of services, e.g. to move some services into *clouds,* and the mutualisation of products, i.e. to develop and implement shared IT products for the different clients, aims at reducing costs on facilities, power, cooling, hardware, administration and maintenance.

In conjunction with what can be considered a general trend of the IT industry, other contextual aspects deserve to be considered. First of all, the IT division is somehow subordinate to the other divisions of the corporation and it has to do its best to meet their expectations and respect the constraints they put forward. For the sake of the above mentioned mutualisation the IT division's operations tend to be more and more multi-clients (i.e. one operation concerns simultaneously multiple divisions of the corporation). In addition, since each client has different activities, the IT division has to deal with multiple specific, and sometimes conflicting, requirements. Finally the IT division has a recent record of multiple organisational changes including transfers and turnover of personnel (both at managerial and operational level).

All these aspects, as described in the results section of this paper, played a role in reducing the ability of the

IT division to cope with the unexpected situation during the replacement of the cores of the two data centres.

2 UNDERSTANDING AN ACCIDENT BY UNDERSTANDING ORGANISATIONAL BRITTLENESS

A theoretical framework was needed to understand the reasons why a mayor crisis was triggered during the execution of the planned operation. This framework must provide a specific set of *lenses* to collect information and data, and to make sense of them. A first and traditional framework option would have consisted in looking into the technological failures and the 'human errors' which occurred during the replacing of the cores of the two data centres. To a certain extent this was the option taken by the IT division in preparing their internal accident report.

However, the fact that the IT division had been suffering recurrent incidents led to the presumption that this event was not an out-of-the-ordinary situation but rather the symptom of some dysfunctional mechanisms characterising the organisation. On the basis of this presumption, it was decided to adopt a systemic and resilience oriented perspective on the event investigation. This implied achieving a description of the event at the level of the socio-technical system as a whole, rather than looking only at local failures and errors. It also implied to understand the normal functioning of the IT division. Special attention was therefore placed on understanding the interactions between humans, technology, and organisation within the IT division as well as between the IT division and other relevant organizations and stakeholders.

Adopting this perspective the research question which steered the event investigation became: *What made the IT division so brittle that it was unable to anticipate and manage the unexpected events that happened during the operation of replacing the cores of its network?* And the analysis framework included two main concepts: the notion of *migration towards the boundaries of acceptable performance*; and the notion of *capacity of manoeuvre. .*

The analysis of the event hence started with the acknowledgment of the natural tendency of organisations to migrate towards the boundaries of some kind of acceptable performance (Rasmussen, 1997) under pressures for achieving the objective of functioning *faster, better* and *cheaper*. By the very first interactions with the personnel (both at managerial and operational level) working in the IT division, it appeared evident that the organisation in question was not an exception to this general rule. By being pushed towards their functional boundaries, organisations are at risk of exhausting their ability of remaining in control of operations, i.e. to absorb disturbances and to stretch their functioning in case of sudden increase of demands. Woods and Branlat (2010) describe this ability as a *potential to gracefully extend* the functioning of an organisation over the boundaries of the acceptable performance. Such potential goes under the name of *Capacity of Manoeuvre (CfM)*.

The *CfM* of an organisation is influenced by two main factors. On one hand there is the way an organisation allocates its resources. Two main strategies are possible. The first one consists in striving for maximising performance efficiency in normal situations. This means that the organisation decides to reduce slack resources and buffers on the idea that, for normal operations, they represent unnecessary costs and excesses. The lean and total quality management perspectives are good examples of this management approach. The second one consists in allocating resources to cover for peak of demands and to deal with and overcome unexpected situations. It has to be pointed out how these resources will show they value only when disrupting events occur. The second factor influencing the *CfM* is related to the fact that an organisation exists and operates in a network of other organisations. Each of them, either intentionally or unintentionally, constrains or expands the *CfM* of the targeted organisation by their operational modes and by setting demands and requirements on the organisations they are related to. Despite the reasons behind the available *CfM* of an organisation at any given moment, it should be noted that the more an organisation is brittle the less it has capacity for adapting to surprises and this will require high amount of energy and resources.

3 DATA COLLECTION AND ANALYSIS

The event investigation was based on two main sources of information. The first one consisted in the review of the available documentation, i.e. the internal accident analysis report developed in the aftermath of the event as well as the minutes of the accident management meetings, an accident analysis report concerning another event which occurred a couple of months after the above mentioned one. The organisational structure of the IT division was as well part of the collected information. The second source of information consisted in the data collected during in-depth interviews and focus groups to gain knowledge about the three main phases of the operation i.e. its preparation (lasted almost a year), its execution and the crisis

management. Sixteen in-depth interviews were conducted with:

- Representatives from the IT division directly involved in the event at both operational and managerial level.
- Representatives from client organisations (other divisions of the same corporation) which had been affected by the accident
- Representatives from the technical supply and support organisations

On the first round of interviews the topic discussed with the interviewees ranged from leadership and communication issues, to decision making and competence management, from professional culture to fatigue and working rhythms.

During a preliminary analysis of those data, some main topics were identified and a second round of data collection was organised. This was conducted in the form of two focus groups with a total of eleven representatives of the personnel from the IT division not directly involved in the accident. The scope of these focus groups was to understand the "normal" functioning of the IT division and explore the hypothesis that the human and organisational factors contributing to the event were not so unique and exceptional (as thought by the IT division management) but rather recurrent aspects of the overall organisational functioning.

The analysis was conducted by exploiting the theoretical framework to make sense of the information collected and to structure it in cluster of factors weakening the IT division and eroding its *CfM*. The preliminary results of the analysis were exposed and discussed in feedback meetings with the IT division personnel. This allowed to achieve a more solid and shared understanding of the reasons why the IT division was not able to anticipate the unexpected events that happened during the operation of replacing the cores of its network.

4 RESULTS OF THE EVENT ANALYSIS: WHAT MADE THE IT DIVISION BRITTLE

The reasons for the weakening of the IT division could be clustered in four layers ranging from a macro to a micro perspective. Each layer reduced the potential the organisation has to cope with unexpected situations and it contributed to the erosion of its *CfM*.

The four layers are:

1. Context and strategic choices (e.g. tendency of sharing applications for different clients)
2. Relationship between the IT division and the clients (e.g. submission of the IT division to clients' requirements)
3. Structural and organisational choices (e.g. lack of a precise methodology for managing changes)
4. Professional culture (e.g. step backwards perceived as a failure)

It can be considered that what happened during the replacement of the cores of the two data centres results from the normal ordinary functioning of the IT division rather than from some out-of-the-ordinary causes.

4.1 Context and Strategic Choices

The outer layer is related to the organisational context of the corporate and to the strategic choices made over the years**.** The choice to go for a mutualisation and virtualisation of IT applications has the drawback of making the technology so complex that it becomes virtually impossible for both managers and operators to fully know all the details of the technical system itself, of its functioning, and to identify all the risks which could be encountered in changing some parts of the system. As stated in one of the interviews: "*It is impossible to model everything. We run tests for the most of part of the functional aspects, and for what concerns the specificities we keep our fingers crossed*". The choice of mutualising applications among multiple clients has another effect: one problem impacts multiple stakeholders and a broader section of the network at the same time. Thus, when an incident occurs both the pressure to solve it and the perception of its criticality increase.

The effort to increase and optimise performance of the network pushes operators of the IT division to accept risks they would not accept under different conditions and contexts. This effect is amplified in those cases, as the operation of core replacing, where the projects are supported by the hierarchy of the organisation since operators perceive a higher pressure for accomplish their missions. Simultaneously to the aim of mutualising applications, the effort to optimise performance leads to the creation and deployment of more and more specific and dedicated applications to answer clients' needs. The proliferation of applications makes it challenging to assess and identify the sources of problems when they appear.

4.2 Relations between the IT Division and the Clients

The second layer comprehends the factors related to the relations of the IT division with respect to its clients and to the overall corporate.

Due to the fact that between the IT division and the other divisions of the corporate there is a supplier-client type of relation, the IT division is bound to satisfy the requirements and expectations of the clients and is subject to the pressures they put on it. This has been the case for the three phases (preparation, execution and crisis management) of each activity the IT division carries out. For example, during the preparation phase the IT division is limited by its clients in the selection of the time and planning for the operation. According to one of the interviewee "*[the clients] are always winners in the negotiation; finally there is not a real negotiation*". During the execution phase the IT division is pressured by the clients for respecting the planning of the operations. This pressure can sometimes be implicit and "simply" perceived by the IT personnel as one of the interviewee expressed: "*We had already postponed the check 5 or 6 times, and we did not want to delay it once more*". Even in the case of a crisis management, the IT division is affected by the pressure of its clients. This is the case, for example, when a solution to a problem has to be found and different conflicting logics exist. While the IT personnel would prefer to address a crisis by thoroughly understanding the reasons behind a problem, the clients are keener in finding a quick fix to the issue. As in the case of the event here discussed, the strategy for coping with the crisis was to a certain extent decided by the clients. This fact was perceived by most of the IT operators as unfair and it exacerbated tensions and frustrations.

The relation with the other divisions of the corporate plays as well a role in the process for demanding the authorisation to execute the operations. Since the other divisions are often critical toward the execution of operations, the IT division has to spend time and effort in communicating with them about the upcoming operations and the risks associated to them. To facilitate the task of obtaining their approval, the IT division tends to minimise, in its communication, the impacts the operations could and would have on the core activities of the corporation. The drawback of doing so is that the IT division creates idealistic expectations on the way operations will go and on the complete absence of risks; as an interviewee reported: "*We anticipate their refusal so we include less information and we state the operation will have no impact at all*".

With respect to the relation between the IT division and the clients a third and final factor making the IT division brittle exists. In order to comply with the other divisions' stringent requirements in terms of the availability of their applications, the IT division is sometime pushed to trade safety for efficiency in the way to execute its operations. An illustrative example of this situation is that the IT division decided to change the cores of both its datacentres simultaneously rather than doing it in two times. The second option would have been more cautionary (at least they would have been sure that one of the datacentre was operative) but it would have implied that some clients would have been bothered two times and not just one.

4.3 Structural and Organisational Choices

The third layer contributing to make the IT division fragile in the face of unexpected situations comprehends multiple factors related to the structural and organisational choices. A first set of factors in this layer is related to the process and approach for conducting changes and performing operations. To make sure that all operations on the network are authorised by relevant stakeholders, the IT division put in place a dedicated organisational system. This system raises issues with respect to its effectiveness. Sometimes it is not the most appropriate representative of a client organisation who grants the authorisation and therefore some constraints can be not considered at this stage. In addition, a technical perspective on the operation is often lacking, and the organisational system does not allow a clear differentiation and identification between operations at high or low stake. Finally, by the fact that the documentation of the organisational system is normally compiled by highly specialised IT personnel and read by non-technical personnel, it is often prepared in a technical jargon difficult to be understood by the receivers.

The IT division does not impose on its personnel a strict methodology for preparing operations. For this reason there is a rather big variety in practices. This includes, for example, the way in which operations are classified (as an IT project with all the administrative aspects related to it, or not). As a side effect of this situation, there are uncertainties in the way operators allocate their time to preparing operations, in the way managers follow that preparation, in the way validation milestones are implemented, and in the way risk analysis is performed. These uncertainties are experienced by the personnel of the IT division which declared, for example, that "*Up to today, I still do not know if I should have been part of the management team of the project or not...*" or that "*[For that operation], we did a risk analysis by rule of thumb"*

Sometimes the IT division is pressured to perform multi-site operations (mainly to accommodate clients' constraints). As in the case of the accident occurred, this can complicate the construction of a shared

understanding of what is going on at the sites. In addition, this increases the temporal pressure for synchronising operations and can lead (as it has been the case) to misunderstanding and incoherent decisions.

4.4 Professional Culture

The specific professional culture which could be observed at the IT division also played a role in making the socio-technical system fragile. The personnel is highly committed to their work. This positive and desirable trait has led in the past as well as during the execution of the operation of replacement of the data centres to some side effects. For most of them it is difficult to step back from an operation or an activity. It appears to be difficult to "*say no*" to challenges and work demands. The commitment of the personnel for example resulted in a problematic shifts' change during the event in analysis. Some operators did not want to leave the operation site because they wanted to help and know how things developed, as well as others did not respect the instruction to take a day off from work because they cared about the result of the work and "*wanted to know where they were"* with respect to the resolution of the problem.

For each operation the possibility to withdraw from its execution and the reestablishment of the main functionality of the system is considered. As a matter of fact the personnel, even when it prepares for it, does not really conceive the withdrawal as a realistic possibility. It is perceived as a failure and therefore it is postponed as much as possible. The withdrawal is often used as an argument with the clients for obtaining the authorisation to perform operations in the sense that "*in case of problems we can at any time withdraw and you [the clients] will not be affected".*

Another cultural trait characterising the IT division is the gap between the managerial and the operational level. This is evident in the different degree of technical knowledge the two groups have. Since it is typical of the IT domain a rapidly evolving technology, managers often possess an out-of-date technological competence which limits their possibility to follow up operations in their preparation and execution. Even in the case of incidents they have limited visibility on the technological problems, and their focus to solve an accident (i.e. to resolve it quickly) is different from the one shared by technicians (i.e. first understand the problem then fix it).

5 DISCUSSION AND CONCLUSION

The complexity of the work performed by the IT division and the level of uncertainties associated with it make it de facto impossible to foresee all the potential problems emerging during operations and prevent their occurrence. The analysis of the accident which occurred during the replacement of the cores of the data centres allows identifying a number of factors which contributed to erode the Capacity for Manoeuvre of the organisation. For the sake of presenting the results of this analysis, those aspects have been clustered in four main layers of weakening factors. The aspects and their effect on the IT division are summarised in the following table:

Table 1. *Summary of impacts on the brittleness of the IT division*

Observations	Impacts on the brittleness
Virtualisation	Increased difficulty to foresee barriers for identified risks Increased difficulty to adapt to hazards
Mutualisation	Increased possibility for problems to widely spread in the network Increased possibility for problems to affect multiple clients
Optimisation of performance	Increased risks acceptance for meeting performance demands Increased difficulty to assess the source of problems
Relation with clients	Reduced preparation time Increased time pressure which could lead to errors
Organisational system for authorising changes does not allow an optimal preparation of operations	Increased difficulty to assess and validate technological solutions
Difficulty to build a shared vision of the situations in case of multi-site operations	Increased possibility for misunderstandings and incoherent decisions

Management with out-of-date technical competences	Increased difficulty to follow operations' preparation and execution Increased difficulty to take over crisis management
Withdrawal from operations considered as a failure	Contingency plans insufficiently prepared
Commitment to work	Increased difficulty to hand over tasks execution
Multiple strategies for problem solving	Increased difficulty to build a shared strategy to solve problems

The traditional approaches for preventing human errors and addressing organisational factors seem to have reached their limits in supporting organisations to further improve safety level. The concepts of human and organisational reliability fall short in ensuring accidents prevention. The same appears to be true for the safety management approaches aiming at constraining performance and reducing internal and external variability. The Resilience Engineering community has been advocating for a change in safety management approaches and practices for more than a decade (e.g. Hollnagel, 2004; Hollnagel et al, 2006; Hollnagel et al, 2008; Dekker, 2011).

The framework exploited in this event analysis acknowledges that complex socio-technical systems operating in a network of organisations are pushed towards the borders of a space of acceptable performance by multiple conflicting pressures. This approach offered a practical perspective for identifying, highlighting and acknowledging the elements which reduce the IT division's capability to effectively cope with surprises. The results of this analysis provide content for the management of the IT division to rethink their safety management approach and to evolve in their safety management practices. While the possibility to work and invest time and resources in improving operators' reliability and enhancing safety culture should not be neglected, it seems that this will not be enough for making the IT division less brittle. At least three lines of work could be conceived for expanding the *CfM* of the IT division - or at least for limiting its erosion. The first two concerns accidents prevention. One consists in conducting a reflection on the elements which make the system brittle i.e. by identifying and revealing the impacts of the strategic choices (e.g. mutualisation of applications) on the activities of the IT division. The other consists in reducing the possibility for vagaries. In the specific case of the IT division this would mean for example to define, in collaboration with the clients, a set of maintenance slots during which operations could take place. This would have the effect of facilitating the obtaining of authorisation for operations, of easing the communication about risks and of reducing the tensions between the IT division and its clients. Finally, the third line of work concerns how to deal with and recover from unexpected events. Improving risk understanding (for example by classifying operations according to their sensitivity) and accident management (for example by creating buffers in the system to reduce the impact of accident) represent a possible strategy for expanding the capacity for manoeuvre of the organisation.

REFERENCES

Dekker, S.W.A., 2011. Drift into Failure: From Hunting Broken Components to Understanding Complex Systems. Ashgate Publishing Co., Farnham, UK.

Hollnagel, E. (2004). Barriers and accident prevention. Aldershot, UK: Ashgate.

Hollnagel, E., Woods, D. and Leveson, N. (Eds.). 2006. Resilience Engineering: Concepts and Precepts. Aldershot, UK: Ashgate

Hollnagel, E., Nemeth, C. P. & Dekker, S. W. A. (Eds.) (2008). Resilience Engineering Perspectives, Volume 1: Remaining Sensitive to the Possibility of Failure. Aldershot, UK: Ashgate.

Rasmussen, J. (1997): Risk Management in a Dynamic Society: A Modelling Problem. Safety Science, 27(2/3):183-213.

Woods, D. D., & Branlat, M. (2010a). Basic Patterns in How Adaptive Systems Fail. In E. Hollnagel, J. Pariès, D. D. Woods & J. Wreathall (Eds.), Resilience Engineering in Practice (pp. 127-144). Farnham, UK: Ashgate.