



Protecting your data from corruption

How to build a more resilient business by
improving data management



pwc

Introduction

The topic of data increasingly finds itself in front of boards of directors. This is driven by a combination of the exponential growth in the amount of data held by firms, and how these datasets are used to support decision making, coupled with increasing regulatory interest in its proper use and management. Those firms which address these challenges head-on will be able to build trust with customers and regulators, while creating a strong foundation for innovation and growth.

This whitepaper equips senior business leaders with an understanding of the risks around data integrity, to help improve their partnership with technology leaders both within their firm and at third parties.

The introduction of the General Data Protection Regulation (GDPR) in May 2018 was a turning point, with a raft of new rules being applied and the prospect of significant penalties for personal data breaches. The rising prominence of data has even spawned three new senior roles at many firms, with Chief Information Security Officers (CISO), Chief Data Officers and more recently Chief Information Security and Resilience Officers (CISRO) now recognised as distinct from Chief Technology Officers (CTO) or Chief Information Officers (CIO). Furthermore, there is a common misconception that the need for transformation to deliver, for example, better customer insights, is a technology issue. In fact it is a data issue, and its solution needs data-centric principles at its heart to deliver on business and customer needs.

Why is this important?

Through our work across different industries, we see many firms running services on systems which cannot be effectively recovered in the event of a mass data corruption incident. Even where recovery is possible, the timeframe to recover would likely be weeks or months. For high volume transaction systems, it's likely to be uneconomical to attempt to recover lost or corrupted data from, say, a 24-hour period.

However, there are actions which can be taken to manage this risk, both to contain a potential wide-spread impact, and to reduce the likelihood of an incident occurring. Understanding how and where critical data is replicated across systems is vital to manage the risk of the spread of malware or corrupt data, and therefore the ability to recover business services. Firms will need to update their data governance frameworks and solutions to enable recovery at pace, and broadly migrate data onto more resilient technology. **This is a systemic problem for industries to solve – not just one firm.**



Historically firms have struggled to get their technology and business-facing departments to speak the same language. Businesses have struggled to articulate their requirements clearly, while technologists have struggled to explain the implications of investment decisions and the risks associated with them.

The growing importance of data

Data presents an opportunity for businesses to gain competitive advantage. Organisations with mature data processes are able to leverage real-time access to immense quantities of data with high levels of integrity. However, many tech-enabled business processes have been built without a true focus on data or its quality.

Most firms do not even have data corruption as a critical risk event which could lead to them ceasing trading. There are examples of where this has been catastrophic, such as the case of Magnolia which admitted in 2009¹ that it could not recover user files from a corrupted database and ceased trading, or Travelex, which went into administration following ‘a recent ransomware attack and being acutely affected by COVID-19’. Cyber attacks are often directed at financially vulnerable organisations and depending on

how severe the wider financial risks are can threaten the survival of the company with the former Travelex Group being a recent example².

While firms invest in a package of prevention, detection and response controls to defend against these types of catastrophic incidents, it is only a matter of time before these controls fail to spot and mitigate the impact. As the UK financial services regulators suggested in their 2018 discussion paper on operational resilience, ‘boards and senior management should assume that individual systems and processes that support business services will be disrupted, and increase the focus on back-up plans, responses and recovery options’³.



Boards and senior management should assume that individual systems and processes that support business services will be disrupted, and increase the focus on back-up plans, responses and recovery options.

Data protection

GDPR has required wide-scale privacy changes in all regulated organisations, and regulators have gained unprecedented powers to impose fines. The regulation defines the concept of a ‘personal data breach’ which means a breach of security leading to the accidental or unlawful destruction, loss, alteration, unauthorised disclosure of, or access to, personal data transmitted, stored or otherwise processed.

However, a firm does not have to experience a breach to come under regulatory scrutiny. A firm could be censured, for instance, for not being able to respond to subject rights requests within the required calendar-month time limits. An extension of this time period due to corrupt data or unavailability of systems will prove difficult to justify. Moreover, the accuracy of personal data is critical not only in the eyes of the regulation but also in the eyes of internal and external business users who rely on personal data – e.g. call centre staff responding to a customer query or marketing teams who are launching new campaigns. The considerations that need to be made around data corruption are wider than just security.

Traditional risk and recovery techniques do not work for data corruption

In general most organisations' technical recovery capability, effort and investment have been about recovering IT services to a different site, should the primary site fail. This principle assumes that there is a problem with the technology hardware or the utilities driving it, rather than the data that underpins it, and therefore a firm would be able to reboot systems elsewhere. This does not account for those incidents where the data, rather than the hardware, has been compromised.

The exponential use of IT services and their data has precipitated massive advances for organisations across all sectors in the products and services they can offer and the manner in which they can offer them. Connectivity of technology has enabled firms to enter new markets and benefit from resources that have historically been outside their reach. Such new opportunities also fundamentally change the risk profiles of firms and need to be managed accordingly.

The regulatory perspective – insight from the Financial services industry

Within financial services, UK regulators are consulting on policy changes which mandate firms to identify the important business services they deliver to end users, to map out how they are delivered (and the underlying resource dependencies), to set tolerances for disruption, thinking about the impact on end users, the market and the firm itself, and to test their ability to remain within those tolerances.

However, the regulators say that a firm is not required to remain within its impact tolerance 'if this would put the firm in breach of another regulatory obligation, conflict with the proper exercise of a discretion granted to it under any rule or regulation, or result in increased risk of harm to its clients or the soundness, stability or resilience of the UK financial system or the orderly operation of the financial markets'⁴. This could include a scenario where data has been compromised and there is insufficient confidence that the issue has been fully resolved before attempting to reboot.

At a Treasury Select Committee hearing in 2019⁵, Lyndon Nelson (Deputy CEO of Prudential Regulation Authority, PRA) referred to a data integrity issue as 'probably the one that [the PRA] would fear the most'.

Information security is a key theme within the existing EBA Guidelines on ICT and security risk management. Beyond requirements for monitoring and testing, for example, there is a need to 'define and implement data and ICT systems back-up and restoration procedures, [...] in line with business recovery requirements and the criticality of the data'.

These requirements also form part of the current draft of the European Digital Operational Resilience Act⁶ which looks to harmonise expectations across the whole financial services industry and will make a more coherent connection to the NIS Directive⁷. That means approximately 22,000 firms in scope of the proposal would be required to set out information security objectives as part of the ICT risk management framework and to establish an information security policy.

The movement of data between organisations means it presents a risk to the overall financial system and not just to individual firms. The European Systemic Risk Board published a paper on systemic cyber risk in February 2020 which set out why it is fundamentally different from other sources of operational risk. This is primarily because of the speed and scale of its spread as well as the potential intent of threat actors. There is acknowledgement that a cyber incident can evolve into a systemic crisis when trust in the financial system is eroded, so this topic looks set to remain high on regulatory agendas.

Traditional risk and recovery techniques do not work for data corruption

However, firms' approaches to managing the risks have not kept pace. All too often in a major cyber or technical outage the board instinctively instructs CIOs to use their IT Disaster Recovery (ITDR) capability, only to be told it's the 'wrong type of disaster' and then suffer an extensive loss.


Of the 'wrong' types of disaster, massive data corruption caused by a deliberate cyber attack or technical failure has the potential to cause significant impacts on customers, financial markets and firms themselves.

This type of failure is difficult to recover from because it often destroys the recovery site's data at the same time as the corruption is copied over to it. Its impact is also significant as it has the potential to not only destroy the data IT services processes and uses, but the services themselves. This is because the software components that make them up are also corrupted, disabling them en masse.

Depending on the organisation's back-up technology, the corruption can spread there too, further limiting recovery.

We can illustrate how contamination can spread, and how it can be controlled, using an analogy of a water supply to a city. In the diagram below the city's water supply is delivered via two reservoirs which are linked together. A network of water pipes feeds into each building. This arrangement has served the city for over 100 years without any major issues arising.




Key:  = dam with sanitisation plant



- The city has two reservoirs supplying it, both with sanitisation plants cleaning the water before it gets to the pipeline.
- Both reservoirs are fed with rivers and groundwater originating in the mountain range.
- The city has not had any issues on water contamination or pollution but recent industrial accidents and terrorist acts in the region have led to concerns in the City Council.

Traditional risk and recovery techniques do not work for data corruption



Key:  = contaminated water

- The City Council thought that having a backup water reservoir and protecting both reservoirs with sanitisation facilities would ensure clean water to the city.
- However, a further risk assessment analysis showed that a single source of pollution at one of the sources has the potential to contaminate both reservoirs serving the city. Existing sanitisation plants may not be equipped to cleanse the water sufficiently for human consumption which could lead to public health issues.
- The council now understands that this risk mitigation strategy would not be sufficient to manage the impact of a range of water contamination events.

 = water tower  = contamination detection device

- The City Council quickly initiated a remediation programme, keeping the reservoir structure unchanged but installing sensors from the reservoirs to the city, in order to detect any contamination that may potentially evade sanitisation.
- Fully understanding that no preventive measure can eliminate these risks, they also added local water supplies in critical buildings, such as hospitals, to enable these buildings to continue functioning if a water contamination event were to materialise.
- The City Council is now feeling more confident in preventing and responding to threats related to water contamination or pollution.



In most instances, the only course of action is to recover from the off-site or off-line back-up on the assumption that it survived, that it works, and that it is not completely overwritten with corrupt data. However, recovering from back-up can take weeks or even months, with the IT infrastructure affected having to be largely re-built, and even then some data will have been lost permanently. This means an extended period of disruption to a service.

As highlighted previously, in our experience, very few organisations are tracking data corruption as a major risk and even fewer are acting on it. Consequently, organisations' IT divisions are generally not considering the potential causes, scope and impact of widespread data corruption as they build and update their IT estates. Configuring data replication solutions without considering the threat of data corruption can make a disruptive incident much worse.

For organisations with a high number of transactions and or uptime expectations, this level of interruption is likely to be very damaging. In a few cases, organisations are looking at "third data" vaulting solutions, where data is copied away from the IT estate with enough time lag to allow IT to cut it off if data corruption occurs.

Unfortunately, recovering effectively – particularly where there are large volumes of data and complex interrelated IT services – is complicated and time consuming with the limitations of the recovery being fatal to the organisation. To address this, organisations have to properly consider the architectural approach and changes required in their environment to make recovery from a vault viable.

What can organisations do to manage this risk?

The first and most important step is to assess the risk and the technical factors that exacerbate the impact. The technical solution is more challenging. In the shorter term, this means changing the way IT estates replicate their data for both sharing as well as resilience and recovery purposes, so the probability and impact of data corruption is reduced to a manageable level.

The first step is to decide what data should be copied and when. Data being used for the purposes of transactions may still need to be replicated from a resilience perspective as it is generated, but the programmes (“executables”) that make up the IT service may not need to be copied as frequently as they do not change often.

The second and more important step is to select the technologies that actually perform the copy, and potentially use different ones for each data type. The rationale here is that some technologies have a higher risk and impact than others with respect to data corruption. Also, through limiting the scope of what a tool copies and when, if data corruption happens it will be in a limited area. This may still cause a major incident, but it will not totally disable the IT estate.

While this might sound straightforward, it can be very complex in larger environments and the level of maturity in terms of understanding the environment may not be sufficient to make the changes.

Additionally, given the legacy infrastructure and state of larger organisations’ IT estate, the technical options may be limited as well as expensive. In the longer term this issue could potentially be mitigated as part of a major re-development of the IT estate, such as moving to the cloud if the capability is part of the target design. Given the proliferation and importance of IT services, their complexity, data sharing and replication within an IT estate, the risk and impact of massive data corruption has to be managed more effectively than it is today.

Finally, the business must be an active stakeholder with the final word through the process of defining the problem, identifying and funding the solution. The business must make an informed decision on the residual risk.

The approach in financial services

From our work with clients we can see that some banks have already identified and assessed the potential impacts which could be caused to clients, markets and themselves if a core dataset was damaged. This has highlighted the need to invest in additional recovery capability.

A smaller subset of banks have implemented or are in the process of implementing such capability in a variety of methods, using vendor bespoke solutions (on premise) and also leveraging cloud services.

Fewer still have tested the end-to-end capability to assess the time it would take to recover from targeted cyber attacks. For example, recovery from a ransomware attack on a critical service can take weeks, which falls outside their newly defined impact tolerances.

To manage risk, emphasis has been placed on introducing new or enhanced prevention, detection and response (containment) controls, until a time when a more efficient recovery solution or substitutable capability can be implemented.

What questions should you ask?

For the business owner to take more active accountability for managing the risk of data corruption they need to understand what good data looks like, what their current exposure is, and what contingencies are in place in case an incident occurs. This seems simple enough but it can sometimes feel like business and technology areas are speaking different languages.

Here we suggest some questions that can help to overcome confusion on technical language and improve collaboration and understanding.



Business to technologist: What is the risk and impact of data corruption within our organisation? How do we report it? If it is reported as low, what evidence do you have to support that position? What are you doing about it?



Business to technologist: We have already agreed on recovery objectives for my applications. Are these objectives still achievable in a data corruption scenario for a single system? What about data corruption across a service chain or a similar, wide-scale corruption? If not, what would be appropriate?



Business to technologist: In the event of a data corruption, what would happen to our service? What in terms of recovery is the best you can do? Can you get us up and running in a reasonable period?

If the answer is yes, how confident are you about what you can do in such a situation? Have you tested these capabilities in realistic situations?



Business to technologist: What would be the state of my business in the event of a data corruption? Do I even have a business at that point?



Technologist to business: What interim measures could we take today to ensure continuity of service, while longer term measures are put in place, as any investment will take time? Do you have any other sources for the data that would not be impacted by such an event (e.g. external and authoritative data sources, paper records etc.)



Board to both technologist and business: Have you clearly articulated the answers to the questions above to us, briefing us on the risk and asking for investment to increase resilience where needed?

Conclusion

Despite the growing importance of data there has been a historic underinvestment in processing, transferring and storing it in a way which maintains data integrity and quality over time. Added to this, we see a misconception in many firms on their current level of resilience to operational incidents which can impact the integrity or availability of the data. This is usually predicated on a belief that traditional disaster recovery techniques can be deployed where data integrity has been compromised, whereas in reality such issues will bring greater disruption and entail more substantive responses.

To address this, business and technology leaders need to establish the minimum resilience standards and agree on the right way to meet them.

Taking decisive action to build greater resilience into the storage, transfer and use of data provides an opportunity for firms to get ahead. Preparedness to respond to issues that materialise, through the use of scenario testing, will build confidence and corporate memory which can be useful in real time.

Firms may increasingly use their credentials on this topic to build reputations for strong data stewardship and gain competitive advantage.

Contacts



Chris Oxborough
Partner, Digital Resilience Lead

T: +44 (0) 7818 510537
E: chris.oxborough@pwc.com



Matt Burns
Director, Cyber Security & Resilience

T: +44 (0) 7711 562536
E: matt.burns@pwc.com



James Cooke
Director, Digital Resilience

T: +44 (0) 7813 784338
E: james.r.cooke@uk.pwc.com



Stella Nun
Director, Crisis and Resilience

T: 44 (0) 7932 144627
E: stella.nunn@pwc.com



Cem Osken
Senior Manager, Digital Resilience

T: +44 (0) 7483 423041
E: cem.osken@pwc.com



Adam Stage
Senior Manager, Operational Resilience

T: +44 (0) 7483 422845
E: adam.stage@pwc.com

¹ <https://www.ft.com/content/bb4a1fea-0d0d-11de-a555-0000779fd2ac>

² <https://www.theguardian.com/business/2020/aug/06/travelex-falls-into-administration-shedding-1300-job>

³ <https://www.bankofengland.co.uk/-/media/boe/files/prudential-regulation/discussion-paper/2018/dp118.pdf?la=en&hash=4238F3B14D839EBE6BEFBD6B5E5634FB95197D8A>

⁴ FCA SYSC 15A.2.10G <https://www.fca.org.uk/publication/policy/ps21-3-operational-resilience.pdf>

⁶ <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A52020PC0595>

⁷ Within the UK, financial services are not currently considered 'essential services' under the [NIS Directive](#), but the principles are embedded in the PRA and FCA approaches to resilience.

