ITLF Service Resilience Paper 4

The consequences of shortcomings in resilience

The Context

Shortcomings in service resilience result in service outages that have consequences for the organisations providing these services, their customers and the wider economy. These consequences include corruption of or illegitimate access to data, revenue loss, and costs to customers.

A central issue in analysing these consequences is that the supply chain is complex. Some issues are directly experienced by organisations which provide services. Other consequences are those experienced by the customers or users of the service. Although the direct consequences are readily observable and measurable by service providers, effects on customers and the wider economy are outside of the profit and loss or other performance metrics used by service providers.

Ordinarily, such effects provide a motive for regulation. For example, environmental regulation is justified by the claim that the effects of waste and pollution are imposed on people who are not producing the waste and are not the polluters. This regulation is enforced through penalties imposed by governments on the polluting organisation. This logic follows from the inability to specify a direct connection between harm to one individual from the actions of the polluter.

In the case of service resilience shortcomings there is, in principle, a more direct connection between the cause of the harm and its measurable result. So, it might be argued that individuals or organisations could be recompensed for harm through private dispute resolution, e.g. through claims of damages resulting from failure to maintain the terms of a service contract.

This ‘damager pays’ principle, however, confronts the problem that damages can be widely dispersed, with each individual suffering relatively small damages. The aggregated damage may be large, but no one represents the ‘aggregate’ and therefore action does not necessarily follow. In other contexts, such as product safety regulation, the state takes responsibility for representing the aggregate public interest – in recognition of the scale of the aggregate problem.

Digital services are not in general subject to regulation which has teeth through fines charged for breaches. In the UK, an exception is the regulatory regime for Registered Data Service Providers, under which fines have been levied for loss of customer data.
In seeking to quantify the magnitude of the damage to the UK economy from service outages and data breaches, a first step is to understand the categories of consequence.

Categories of Consequence

Software-enabled services are subject to a variety of types of failure, each of which may have consequences of differing magnitude depending on the nature of the incident and the context in which it occurs. We use two categories to describe consequences, namely they can be direct to the service provider, or they can indirectly affect the customer.

The direct consequences experienced by service providers who are selling their services will be revenue loss. The size of loss will depend upon the length of an outage and the terms of service agreement with customers. Conspicuously, many service contracts are based on a 'best effort' principle in which the outage only is seen as incurring damages proportional to the duration of the service period. Thus, a 24-hour outage might be taken as a diminution of 1 day of service being billed monthly. If this principle governs, the service provider will suffer relatively small losses of revenue, even if its customers' monthly service bills are reduced by a day. This is current practice for many utilities.

However, context matters. Services often include content as well as connectivity. For example, an online retailer offers connectivity to their online store as a service (for free) but this connectivity leads to direct revenue in the purchases made by customers. An outage not only affect 'the number of people in the store' but also the revenue generated from sales.

As online advertising has become a major method for finding and buying offline services, outages that affect connectivity mean that outages of online services cause revenue loss to the suppliers of offline services.

Services are also part of larger business processes. For example, an inventory system provides a service to a company who may in turn be taking orders or making delivery promises based upon the state of their inventory. An inventory system outage can therefore cause direct losses by blocking orders that would have normally occurred.

Service outages have the capacity to generate further losses by interrupting business processes that are interdependent. The losses stemming from this interdependency can be difficult to unravel or fully measure. This is because there is a complex chain of consequences.

One of the key potential consequences of digital system failures is data corruption or loss. Due to interdependent systems, data loss or corruption can have a domino effect causing the subsystems that rely upon the data to degrade or fail. Again, unravelling the cause and consequences of this type of failure can be time consuming and thus costly, and sometimes unsuccessful.

As noted earlier, service outages also have effects that are 'indirect' from the viewpoint of service suppliers but very direct from the view of their customers. An outage makes it impossible to use the service and may put users' data at risk. Assessing the scale of damage to service users resulting from service outages is difficult because of the lack of systematic data. However, it is possible to construct a series of estimates underpinned by a combination of data and assumptions.

Quantification of consequences

A fundamental concept needed to quantify the losses to customers due to service outages or data breaches is the concept of 'opportunity cost.' Opportunity cost arises when an opportunity is blocked and has two components.

One component is the foregone value of the foregone service to the user. These costs are nearly impossible to quantify. They vary from the potential for serious health consequences when someone is excluded from using an online service, e.g. to book a GP appointment, to the merely inconvenient when someone is unable to access an online ticketing system. In some – and increasingly few - cases the affected user will be able to utilise a different service to achieve the same purpose.

The second component of opportunity cost is the time wasted that would have been spent using the service but is lost trying to access it. This time could have been used for some other purpose. Economists regard the opportunity cost of lost time as approximating to the individual wage rate. This uses the logic that if the time were not wasted it could be used to work more hours. At first glance this seems rather unrealistic as most people can't immediately use time saved to earn more money nor do they necessarily self-report 'time off' doing their own business on company time. However, the logic is that the wage one accepts is the amount required to attract an individual to work where their alternative is leisure.

Quantification – reported studies

RTI/NIST study

The most systematic and directly related effort at quantifying losses from software failure was conducted under the direction of Gregory Tassey for the US National Institute of Standards and Technology (NIST) in 2002 by RTI (Research Triangle Institute), an independent consulting company[1]. The study, hereafter called the RTI/NIST study, was based upon a questionnaire survey of firms in the US transport equipment and in financial services industries. The aim of the study was to estimate the savings available by making feasible improvements in the software testing infrastructure. To derive this estimate, survey respondents were asked about the incidence and repair costs of 'bugs' (the 2002 term for software faults) as well as the potential cost savings by making feasible improvements in the testing infrastructure.

The conceptual framework of the RTI/NIST study was an attempt to construct a series of counter-factual scenarios in which there were fewer software faults. It was recognised that a fault-free scenario was infeasible. Respondents were queried as to cost reductions possible by less faulty software in the following categories: major failure costs, minor failure costs, purchase decision costs, installation costs, maintenance costs and redundant system costs. While our concern is primarily with the first of these, each cost category is affected by failures in operational software. In both transport equipment and financial services sectors, 60% of respondents indicated that they had experienced major failures.

The RTI/NIST study can be extrapolated to estimate the total costs of software errors as US$59 billion in 2002 (or, adjusting for inflation, $97 billion in 2022). This estimate for the US economy (excluding the public sector) is based upon taking the two sectors as representative of the economy and scaling up according to the number of employees.

The RTI/NIST study recognises several limitations. Three are of particular note:

- "Quantifying the impact of inadequate testing on mission critical software was beyond the scope of this report. Mission critical software refers to software where there is extremely high cost to failure, such as loss of life. Including software failures associated with airbags or antilock brakes would increase the national impact estimates."

- "...the costs of software errors and bugs to residential households is not included in the national cost estimates. As the use of computers in residential households to facilitate transactions and provide services and entertainment increases, software bugs and errors will increasingly affect household production and leisure. Whereas these software problems do not directly affect economic metrics such as GDP, they do affect social welfare and continue to limit the adoption of new computer applications."
- In addition, the scaling up process is based on calculating the cost consequences per employee of the software 'bugs' in the surveyed firms, and then using these costs per employee to scale to other service and manufacturing sectors of the US economy. As the authors of the report note, this involves an important assumption about constancy of costs across sectors.

Much has changed over the past 20 years.

The failure of software systems used in households now has an impact on GDP in several different ways. These include loss of productive time, loss of the value of services that would have been provided, and costs of identifying alternatives. In 2002, the distribution of software-based systems was much more uneven. Then, the finance, insurance and real estate sectors had a much more dominant share of IT systems and investments than other sectors. The method of using total employees for the scaling up is therefore more likely to be appropriate today than it was then.

In summary, a 2002 estimate of US$59 billion (or $97 billion adjusting for inflation to 2022) appears to be a sound estimate for the US economy.

Tricentis and CISQ estimates

More recent estimates have been made by business consultancy organisations for inclusion in white paper-type reports.

One estimate comes from Tricentis, an Austin, Texas based company that provides software testing solutions[2]. As a part of their effort to raise awareness of software quality issues they performed a systematic review of English language press accounts of operational software failure (as well as other types of failure). The accounts reported a mix of actual losses, repair costs and the assets affected as if they were totally written off. They aggregated these to reach a global estimate.

Tricentis' 2018 report was based upon 606 reported failures affecting 314 companies in 2017 although only 80 of these had an estimate of the loss or description of the assets affected. The report estimates that, for 2017, the total global loss and assets affected in the English language press amounted to US$1.72 trillion[3]. This total reflects press reports and illustrates a limitation of the earlier RTI/NIST report. Several incidents reported in the press resulted in hundreds of millions of actual losses – these 'outliers' contribute to real world totals but are likely to be missed by the sampling method employed in the RTI/NIST study.

The major shortcoming to the Tricentis study is that a preponderant share of the total comes from the 'assets affected' category. These are not properly losses. For example, the software problems with the F-35 fighter jet program are reported to have added $1.7 billion to the cost of the $400 billion programme. The number recorded as part of the above total is $400 billion not $1.7 billion. For this reason, the Tricentis estimate appears to be a substantial over-estimate of directly incurred costs of operational software failure.

Despite its informal methodology, the Tricentis report from 2018 has become a basis for a further report from CISQ in the same year[4]. At p.15 of the CISQ report, the US$1.7 trillion (taken to be a global sum of actual losses) is translated into a US loss of $1.275 trillion on the somewhat dubious basis that 75% of the world's English speakers are Americans[5]. A 2020 CISQ report raises the estimate to $1.56 trillion based on the assumption that the growth has been 22% over the two years (the rate is not substantiated). Returning to the Tricentis study, it appears that actual costs might be more in the range of $20 billion for the 80 companies reporting. Taking the average of these losses and attributing the same average for the other 234 companies would suggest a cost to the economy of $51 billion.

Estimates for the UK economy

Using the RTI/NIST inflation adjusted estimate of $97 billion as an upper bound and the very approximate $51 billion derived from the Tricentis study as a lower bound, we now attempt to use this to estimate costs for the UK.

The US labour force was about 148 million in 2020. In the same year, the UK labour force was 33 million. So, UK employment is 23% of the US. Employing this share to attributed costs of US failure ($51-97 billion in 2020) yields an estimated range of UK costs of US$12-22 billion or £10-18 billion.

In terms of GDP, the 2020 UK was $2.7 trillion or 12.9% of the US GDP in that year ($20.94 trillion). This provides another, lower, estimate of the range of costs of operational software failure: $6.6-12.5 billion or £5.4-10.2 billion.

We have critically considered these estimates. We conclude that the current (2024) costs for the UK are likely to be in the range of £9-15 billion. For purposes of discussion, we suggest using £12 billion as a conservative estimate. This is conservative because it largely neglects the opportunity costs imposed on users; the figure is dominated by the costs incurred directly by the organisations providing software services to remediate operational software after failure.

Extensions and Further Implications

The above quantification is necessarily crude because of the absence of systematic report of service outages attributable to software failure. The 22-year-old US estimate in the RTI/NIST study remains the most systematic, in that a representative group of companies were willing to disclose their actual experience. Although this produces a credible estimate of £12 billion per annum to the UK economy, it is important to note what is missing.

What is missing is that an important share of costs to the economy are external to the company that directly experiences the service outage. We suggested that one means of assessing these costs is to consider the value of time lost to the users during a service outage. The costs incurred by users are due to failing to complete tasks requiring the downed service and the time lost in order to accomplish the task by other means. There is no systematic data reflecting actual experience, so the scale of these costs can only be estimated by introducing a series of assumptions concerning the aggregated amount of lost time and the average 'opportunity cost' of this lost time.

The UK population is approximately 68 million individuals, with approximately 28 million households. 97% of the UK population use the internet, which we can imagine is 97% of UK households. Arranged by decile the average gross income implies an estimate of the hourly contribution to gross income of from £7.40 for the 2.8 million in the lowest decile to £98 for the 2.8 million in the highest decile. Using these decile weights to create a total for a one-hour loss implies that if the effects of outages were spread evenly over the entire population, the cost of a one-hour outage to the entire population would be £888 million. The following table translates these assumptions into an estimate of user costs of lost hours per annum using four different levels of lost time.

Average Hours Lost per Annum per Household	Total opportunity cost (£ billion)
5	4.4
10	8.9
20	17.8
40	35.5

These numbers constitute a significant percentage of UK total GDP, ranging from 0.2-1.56%. In effect these losses can be understood as a drag or leakage on national income and productivity that is similar in proportion to the difference between positive growth and recession in recent years. Of course, these losses cannot be eliminated at a stroke, but their reduction would provide a significant boost to the UK economy.

As emphasised throughout, more accurate estimates of loss and hence a better gauge of the possible gains from improving resilience would be highly desirable. This will require systematic data gathering on the actual experience of companies providing digital services and their customers.

---

[1] RTI/NIST, The Economic Impacts of Inadequate Infrastructure for Software Testing, Final Report May 2002 (Prepared for Gregory Tassey by RTI Health, Social, and Economics Research, Research Triangle Park, NC 27709. Available at https://lara.epfl.ch/w/_media/misc/rti02economicimpactsinadequateinfrastructuresoftwaretesting.pdf

[2] The developer of the Tricentis study and author of the software failure reports for 2017 and 2018 Chelsea Frischknecht generously provided the underlying dataset which reveals ambiguities in how costs were measured primarily because of the 'assets affected' category which dominated the reported total.

[3] Tricentis, Software Fail Watch 5th Edition, https://www.scribd.com/document/427481278/Software-Fails-Watch . This link is to an online archive Scribd as the Tricentis site no longer carries the report. The link may be a summary of the original report but definitively states the US$1.72 trillion (actually US$1,715,430,778,504) as the loss from software failures.

[4] www.it-cisq.org/wp-content/uploads/sites/6/2023/09/The-Cost-of-Poor-Quality-Software-in-the-US-2018-Report.pdf

[5] For example, there are 125 million English speakers in India. There are 67 million people in the UK and over 30 million people in Australia and New Zealand (the vast majority of whom are English speakers). The 2021 US population is 331 million.