Availability: sharing information on causes and effects of IT failures

Summary

This paper is about sharing of information about the performance or other characteristics (typically non functional[1]) of software components, capabilities, or services as observed by users of IT systems in practice.

Information sharing is particularly important for organisations, either where software is being used under circumstances they were not designed to meet, or where a software product or service is used by a large number of customers. This is particularly relevant as open source and third-party services increasingly become part of any organisation's architecture. The benefits to organisations who share this information include the potential to avoid outages and to take proactive action to avoid incidents and outages, and therefore avoid customers being negatively impacted through outages.

Information sharing requires both organisational structures to facilitate collaboration and a shared terminology. Terminology is needed to describe known vulnerabilities (before outages are experienced), causes of operational service outages (which enable an understanding of the user impact) and metrics for user impact. In this paper we[2] recommend additional work, to enhance and complement the structures and terminology in place for cyber security, for use in sharing information on causes of outages due to IT causes other than cyber-attacks.

Cyber Security Information sharing

The benefits of information sharing have been identified and recognised within the realm of cyber security for some time, and there are a set of active standards

[1] Non-functional requirements are the criteria that define how a system should behave, rather than what it is supposed to do. https://www.perforce.com/blog/alm/what-are-non-functional-requirements-examples

[2] In this text and others in the series of ITLF Availability Papers, "we" refers to the Availability/Service Resilience Working Group of the IT Leaders Forum of the BCS – the Chartered Institute for IT and co-opted colleagues who have provided additional insight.

and frameworks, a common terminology, and organisations and communities in place to support the sharing of information about cyber security threats, actors and scenarios, as included in the advance information on a proposed UK Government Bill[3]. There is a significant overlap between cyber security and operational resilience[4], and indeed, in the event of an outage, both cyber security and resilience have to be considered.

Existing Codes of Practice

The UK Government has published a Software Security and Governance Codes of Practice and associated documents.

- Software Security Code of Practice[5]:
- NCSC Code of Practice[6]:
- NCSC Implémentation guidance [7]:
- Cyber Governance Code of Practice[8]:

These rely on suppliers providing data on their processes for design and testing of software, as for a Kite Mark for products. The guidelines are not tied to outcomes such as performance or characteristics observed by users in practice.

Organisations promoting information sharing

In addition, there are a number of organisations which facilitate that information sharing. These include:

- **FIRST[9]** FIRST brings together a variety of computer security incident response teams from government, commercial, and educational organizations. FIRST aims to foster cooperation and coordination in incident prevention, to stimulate rapid reaction to incidents, and to

[3] https://www.gov.uk/government/publications/cyber-security-and-resilience-bill-policy-statement/cyber-security-and-resilience-bill-policy-statement

[4] Resiliency describes the system's ability to withstand and recover from disruptions, including failures and attacks; it focuses on the speed and effectiveness of response and recovery after a disruption or outage.

[5] https://www.gov.uk/government/publications/software-security-code-of-practice

[6] https://www.ncsc.gov.uk/section/software-security-code-of-practice

[7] https://www.ncsc.gov.uk/collection/software-security-code-of-practice-implementation-guidance

[8] https://www.gov.uk/government/publications/cyber-governance-code-of-practice/cyber-governance-code-of-practice

[9] https://www.first.org/

promote information sharing among members and the community at large. Currently FIRST has more than 700 members, spread over Africa, the Americas, Asia, Europe and Oceania.

- **CERT**[10] The CERT Division of the Carnegie Mellon Software Engineering Division is a leader in cybersecurity. They study problems that have widespread cybersecurity implications and develop advanced methods and tools to counter large-scale, sophisticated cyber threats. They collate and publish reports of vulnerabilities and provide extensive education and training.

- **NCSC**[11] The National Cyber Security Centre has as it s mission Making the UK the safest place to live and work online”. We support the most critical organisations in the UK, the wider public sector, industry, SMEs as well as the general public. When incidents do occur, we provide effective incident response to minimise harm to the UK, help with recovery, and learn lessons for the future”.

Information sharing within retail banking

Categories of failure

The recent analysis of the sources of failure in Nine Banks[12] was made more challenging by a lack of common terminology across the banks. There were a number of factors highlighted, for example:

- Failures due to introducing changes were nearly half of the reported outages
- “Incidents” (using the ITIL terminology[13]) are failures for which the cause is not the introduction of change, but no other terminology appeared to be in use to describe the root cause.
- 3rd party software was responsible for less than a fifth of the outages but the scope of the outages tended to be wider as many organisations use common components, eg Crowdstrike.

[10] https://insights.sei.cmu.edu/divisions/cert/

[11] https://www.ncsc.gov.uk/

[12] https://www.bcs.org/media/rxdmjr5h/availability-6-report-nine-banks-data-roundtable-220425.pdf

[13] https://www.axelos.com/certifications/itil-service-management/

Need for a common terminology

The analysis of the Nine Banks data identified three aspects where a common terminology is needed:

- Sharing on vulnerabilities – build on existing terminology
- Sharing on causes of incidents – lacking a common terminology
- Sharing on impacts to users – use existing terminology

Any terminology should build on and complement the terminology already in place for cyber security information sharing.

Sharing on vulnerabilities

FIRST[14] has a Common Vulnerability Scoring System, which is in wide use.

The CERT Division of the Carnegie Mellon Software Engineering Division is a leader in cybersecurity. They collate and publish reports of vulnerabilities, see[15].

NCSC collates information on incidents supplied by organisations – see[16].

Sharing on causes of incidents

The list below of possible causes of failure incidents can be extended from that outlined in *Resilience of Services*[17] to include:

1. Third-party and Provider Issues

Outages can also originate from failures at cloud, colocation, or telecommunications providers, which are increasingly responsible for a significant share of incidents as organizations rely more on external services.

[14] https://www.first.org/cvss/

[15] https://www.kb.cert.org/vuls/report/

[16] https://www.ncsc.gov.uk/collection/vulnerability-management/understanding-vulnerabilities

[17] https://londonpublishingpartnership.co.uk/books/resilience-of-services-reducing-the-impact-of-it-failures/

Both Commercial Off the Shelf Software and Open-Source software may contain weaknesses which cause outages across multiple organisations. These can be internal to the third-party software or system, or due to incompatibilities with external and internal (eg legacy) software and systems.[18]

2. Cyberattacks and Security Failures

Cyberattacks, including ransomware and distributed denial-of-service (DDoS) attacks, have become a regular and growing cause of outages. These incidents can result in lengthy downtime and data loss. Data from the Uptime Institute[19] and the Nine Banks analysis share both suggest that less than 10% of outages are due to cyber-attacks, though their impact on loss of data integrity may be systematically more far reaching.

Security lapses, such as unpatched vulnerabilities or weak access controls, can enable such attacks.

3. Capacity and Demand Surges

Unexpected spikes in user demand or insufficient system capacity can overwhelm infrastructure, leading to performance degradation, system thrashing or outright outages.

4. Hardware Failures

Failures of physical components—servers, storage devices, networking hardware—are a common cause, especially as equipment ages or if preventive maintenance is lacking.

Critical services often use duplicate / fallback hardware.

5. Human Error

Human mistakes, such as typing errors, misconfigurations, accidental shutdowns, or incorrect deployment procedures, are a persistent and significant cause of downtime. These errors are often the hardest to detect and fix, and they can have prolonged impacts.

[18] https://www.ncsc.gov.uk/collection/vulnerability-management/understanding-vulnerabilities

[19] https://uptimeinstitute.com/about-ui

Management failures, such as insufficient training or inadequate procedures for staff and contractors, contribute to the risk of outages. Badly designed interfaces and erroneous data in the system can be caused by end user error and lead to shutdown.

6. Software Bugs and Configuration Errors

Software bugs, faulty updates, and configuration or change management errors are major contributors to outages. Inadequate testing, rushed deployments, and poor change management practices can introduce critical faults.

According to a 2023 survey[20], configuration or change management issues accounted for 64% of IT system and software-related outages, making this the most common root cause in that category. This compares with 43% in the Nine Banks analysis[21].

7. Network Problems

Network issues have become increasingly common and are now often cited as the top cause of IT service outages. These issues include configuration errors, firmware bugs, corrupted routing tables, and failures at third-party network providers.

The complexity of modern, software-defined networks increases the risk of misconfiguration. These errors are liable to cause cascading failures.

7. Power Failures

Power outages remain a leading cause of major technology outages, particularly in data centers. Failures in uninterruptible power supplies (UPS), transfer switches, or generators can bring entire systems offline.

As grid reliability is variable globally, the risk of power-related outages may be hard to quantify, especially if backup systems are inadequately maintained.

8. Cooling System Failures

[20] https://datacenter.uptimeinstitute.com/rs/711-RIA-145/images/AnnualOutageAnalysis2023.03092023.pdf

[21] As [12]

In data centers, failures in cooling systems can cause overheating, which in turn can trigger shutdowns to protect hardware. Failures are also seen in laptops and smart phones subject to exposure to heat.

Commentary

Further work needs to be done to build a comprehensive categorisation of causes of failure, which aligns with the objectives of the information sharing, and enables banks or other organisations to take appropriate action.

Sharing information on impacts to users

The NIS framework[22] is included in the Policy Statement[23] - with a UK Government bill due to be introduced later this year. It suggests data on user impact of outages be collated under the headings below.

Four metrics for user impact

Parameter	Metric
Availability	Your service was unavailable for a number of affected users for a duration of 60 minutes (lost user hours).
Integrity, authenticity or confidentiality	The incident resulted in a loss of integrity, authenticity or confidentiality of the data your service stores or transmits, or the related services you offer or make available via your systems.
Risk	The incident created a risk to public safety, to public security or of loss of life.
Material damage	The incident caused material damage to at least one user.

[22] https://ico.org.uk/for-organisations/the-guide-to-nis/what-is-nis/

[23] As [3]

Recommendations

Information sharing requires both organisational structures to facilitate collaboration and a shared terminology.

Terminology is needed to describe known vulnerabilities (before outages are experienced), causes of operational service outages (which enable an understanding of the user impact) and metrics for user impact.

This paper recommends using existing terminology of vulnerabilities and for user impacts wherever they are specified. Additional work can enhance and complement the terminology in place for cyber security, for use in sharing information on causes of outages due to IT causes other than cyber-attacks.

A separate paper[24] proposes organisational structures for information sharing, under the heading *Cyber Security and Resilience*.

[24] https://www.bcs.org/membership-and-registrations/member-communities/bcs-it-leaders-forum/papers/