Australian Government
Australian Signals Directorate

ASD AUSTRALIAN SIGNALS DIRECTORATE
ACSC Australian Cyber Security Centre

Gateway security guidance package: Gateway security principles

First published: July 2022

Key gateway concepts

In order to contextualise this guidance, the concept of a gateway, and where gateway capabilities need to be deployed must be understood by decision makers, architects, engineers and operators.

What is a gateway?

A gateway is a boundary system that separates different security domains and acts as the first line of defence to provide security capabilities for an organisation and its systems. In doing so, a gateway provides policy enforcement mechanisms for data flows by only permitting data to flow between different security domains in accordance with an organisation's security policy. A common example of a gateway data flow is between an organisation's internal network and the internet.

A gateway is comprised of a collection of physical and logical components that enforce a security policy to manage access to, and transfer of, data from one security domain to another. In a cloud service provider model, or a managed service provider model, a gateway may be abstracted to a set of security services and capabilities that are exposed to consumers.

Factors that heavily influence an organisation's gateway design include:

- business and operational requirements and strategies
- technical capabilities
- confidentiality, integrity, availability and privacy requirements
- threat modelling, risk appetite and risk management strategies
- integrations between self-hosted and cloud services
- sourcing and service delivery models
- location and number of data centres and office locations
- availability of staff to design and sustain gateway capabilities.

A gateway should be deployed as a combination of physical and logical components and capabilities that operate collectively as a single integrated solution. There are a number of architectural approaches to gateway design an organisation can consider:

- monolithic - provides all gateway security functions through one centrally managed system (for example a Secure Internet Gateway)
- disaggregated - provides service specific gateway functions through discrete but interoperable systems, which do not share a common control plane
- hybrid gateways - provides all required gateway services through a mixture of central and disaggregated service offerings and control planes.

Regardless of the architectural model deployed, it is critical that gateway services and cybersecurity defence capabilities evolve to support an organisation's changing business and risk management needs.

What is a security domain?

The *Information security manual* (ISM) defines a security domain as:

A system or collection of systems operating under a consistent security policy that defines the classification, releasability and special handling caveats for data processed within the domain.

As ICT footprints evolve, an organisation may need to evaluate multiple factors when defining its security domains. A security domain may be a collection of ICT services that have a degree of commonality so as to justify the collection being treated as a single homogeneous group in terms of security. Factors may include:

- purpose
- ownership/sovereignty (Australian or foreign owned/controlled)
- consistent administrative and security control
- consistent security and operational visibility
- value (security or otherwise)
- threats and risks to which they are exposed (risk profile)
- interdependency on other systems (and perhaps data)
- data classification, business impact level, information management markers and other caveats
- ability to perform cybersecurity incident response (e.g. data spill)
- encryption.

For example, an organisation's OFFICIAL and PROTECTED networks are separate security domains, as are the PROTECTED networks of two different organisations. Multiple government tenancies within a shared service (e.g. managed service providers [MSP] or cloud service providers [CSP]) are also separate security domains.

An organisation will usually have at least two (an inside/trusted and an outside/untrusted) security domains within its operating model. Subject to the organisation's risk-based approach, just as security domains can span different data centres, a security domain can span other segmentation and segregation approaches, such as those offered by CSPs. Note, the security policy for a security domain is inclusive of security governance, personal security, physical security and information security as outlined in the *Protective Security Policy Framework* (PSPF).

Each organisation needs to determine if a managed service or cloud service is an extension of its security domain, or if it needs to be considered a separate security domain. This is done through a detailed analysis of its business requirements, security policy, compliance requirements, risk tolerances, consumer guides and existing enforcement capabilities. It should not be assumed that it is automatically one or the other. Non-corporate Commonwealth entities (NCEs) should consider ICT and protective security requirements outlined in the PSPF when changing how they define their security domain boundaries.

Multiple networks can be joined together to create a single security domain by using appropriate security platforms. For example, Virtual Private Networks can join a number of office locations together with the resulting network being considered the security domain, provided appropriate controls are used. Note, the security domain does not need to correspond to a specific network topology, provided controls isolate the organisation's security domain from the underlying transport network.

The controls implemented in a gateway should be designed to reduce or eliminate the attack surface associated with the flow of data entering and leaving a security domain.

An organisation should consider using modern and traditional security approaches to control and inspect traffic traversing security domains via the use of technologies such as proxies, web application firewalls, Software Defined Networks, host-based firewalls, behavioural analytics and Application Programming Interface (API)-based business logic.

An organisation will need to carefully assess a gateway's system boundary prior to the gateways Infosec Registered Assessors Program (IRAP) assessment. As per the ISM, at higher security classifications (i.e. SECRET and TOP SECRET), a Cross Domain Solution (CDS) is required as part of the gateway design.

Consumers and service providers may have different perspectives on how a security domain is defined (see Figure 1). However, as each organisation is responsible for ensuring the protection of its own systems and data, it is expected to determine its own security domains. Where a service provider offers multi-tenancy gateway services, the service provider may retain administrative and policy control of the system. In cases where there is no clear delineation of where the security domain boundary exists, it is recommended that an organisation consider treating these environments as separate security domains until it is satisfied that there is equivalent security policy enforcement and operational visibility.

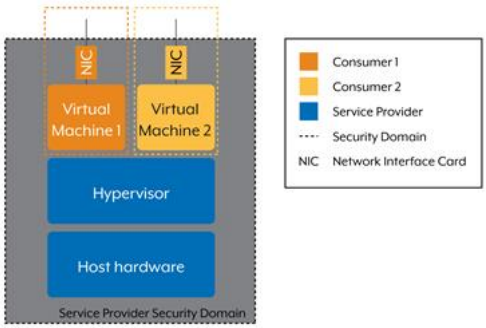

Figure 1: Example multiple security domain perspective in a multi-tenant environment

The Australian Signals Directorate (ASD)'s CDS publications are:

- ASD, *Introduction to Cross Domain Solutions*
- ASD, *Fundamentals of Cross Domain Solutions*

What is a policy enforcement point?

A policy enforcement point (PEP) may be a hardware or software component, security device, integrated appliance, tool, function or application. PEPs may perform a combination of blocking, inspecting, filtering, and validating incoming and outgoing data to mitigate identified ICT risks, in accordance with an organisation's security policy. PEPs may be applied, synchronously or asynchronously, at different levels of the Open Systems Interconnection (OSI) model that collectively enable a defence-in-depth approach. A PEP can be thought of as one or more gateway capabilities used to enforce an organisation's security policy as data traverses between security domains. Multiple PEPs can be chained together to achieve the full set of security policy enforcement expected of a gateway (see Figure 2).

An organisation should note that implementing the same controls and capabilities in multiple places does not necessarily constitute defence-in-depth. For example, an implementation (either physical or virtual) of a firewall or router with access control lists is an example of a PEP where differing policies are enforced. Each interface is a different security zone and a policy is applied between them.

A PEP capability may be implemented in a number of different locations, such as a gateway (e.g. CDS, application proxies client or server agents), group policy or both. Overall, a gateway consists of a number of PEP capabilities that enforce an organisation's security policy.

When controls are inadequately enforced, a malicious actor may:

- gain unauthorised access to steal, copy or interfere with sensitive data
- establish covert channels into or out of systems
- compromise the integrity of trusted systems or data (such as audit logs)
- bypass security-enforcing functions or policy enforcement points
- interrupt the availability of critical systems or services
- move laterally to compromise internal system.

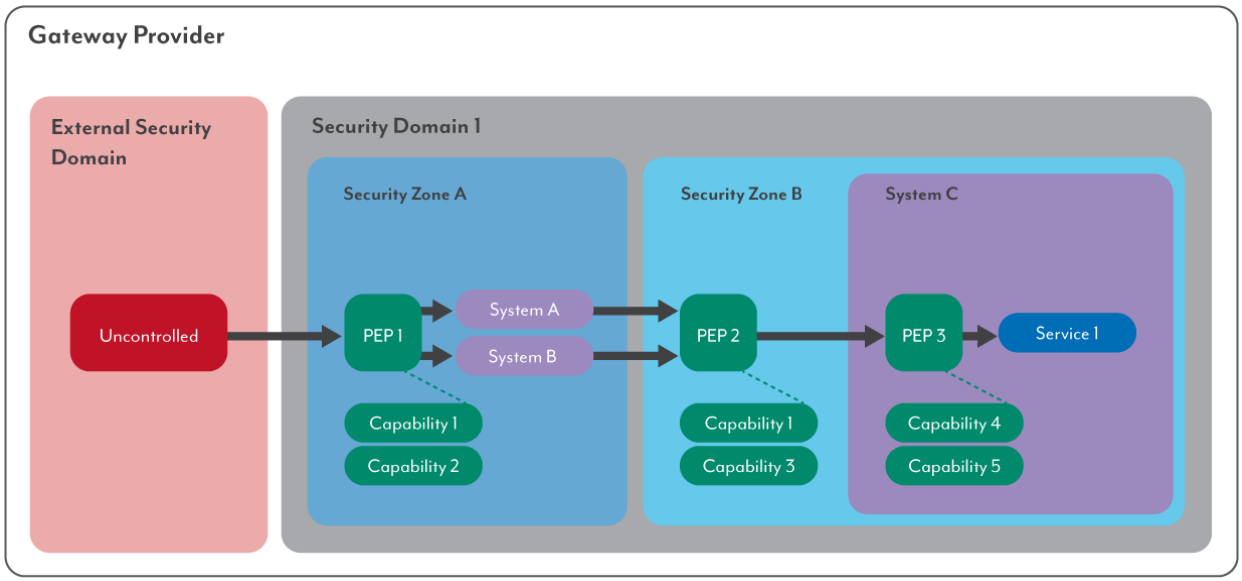

Figure 2: Chaining PEPs

Note, the concept of a PEP has been taken from the National Institute of Standards and Technology (NIST)'s Special Publication 800-207, *Zero Trust Architecture*.

Gateway security principles

ASD has designed a number of governance-related gateway security principles that an organisation should be aware of and consider when implementing or consuming a gateway.

These principles should be applied when designing, procuring, operating, maintaining and disposing of a gateway.

Risk cannot be outsourced

An organisation owns its security risk, even if the organisation outsources or transfers, in full or in part, the implementation responsibilities. The Department of Home Affairs' *Protective Security Guidance for Executives* states that the procurement of goods and services does not transfer the operational risk from the Commonwealth. Security controls are intended to reduce an organisation's risk, but an organisation cannot eliminate all risks. The extent of the organisation's security implementation responsibilities will vary significantly depending on the type of deployment methodologies, services and contractual terms involved, but the organisation will always own the risk. An organisation inherits risk from dependent and underlying systems, inclusive of dependent services across multiple providers (e.g. MSPs, hyper-converged and multi-vendor cloud deployments). As outlined in the PSPF, each organisation is responsible for maintaining the confidentiality, integrity and availability of all official information.

Each organisation will have different responsibilities, threats, risks, legislative requirements, and obligations within its service delivery ecosystem. Actions should not be taken, whether contractual or otherwise, that prevent an organisation from meeting its responsibilities and obligations. Whether a supplier, service provider or other organisation, each will have unique systems and responsibilities, and risks should be wholly understood across the entire ecosystem for visibility, detection and prevention capabilities.

An organisation that enters into a contract that is not consistent with the ISM and the guidance provided in this document may be at risk of not fulfilling the requirements of the PSPF and the *Public Governance, Performance and Accountability Act 2013*.

Security management is continuous

Tactics, techniques, mitigations, technologies and good practices evolve over time, so processes should be developed to remain current with the evolving threat environment. Gateways should form part of an organisation's defence activities, proactively implementing a range of security measures to strengthen a network or system to make it more robust against attack.

It is also better practice to re-visit a security domain's overall risk assessment periodically for it to remain relevant and meaningful. Malicious actors that intend to cause harm or gain access to sensitive information are not static and will continually evolve their tactics, techniques and procedures (TTPs) to achieve success.

An organisation must continually evolve its gateway defences, and in some cases reassess the fitness of IT systems for their intended purpose. An example of this may include programmatically monitoring information sources, such as the ISM or software release notes changes, for operational team review and implementation.

Risk is continuously managed

Threat modelling complements the risk-based approach to cybersecurity outlined within the ISM and the PSPF.

The MITRE ATT&CK framework is a useful tool as part of an organisation's threat modelling process for its security domains, as it provides a knowledge base of malicious actor behaviour, taxonomy for adversarial actions across their life cycle, information on how to detect attacks, and mitigation strategies. Threat modelling aids in identifying an organisation's mitigation coverage of known tactics, which then can be mapped back to ISM controls and implementations. More information on threat modelling can be found in the *Gateway security guidance package: Gateway operations and management* document.

Techniques in the MITRE ATT&CK framework relate to actual real-world threats based upon threat intelligence and cybersecurity incident reporting. The mitigation strategies can be mapped to other technical publications, including the ISM. While security risks are discussed in the frameworks, these should not be considered an exhaustive list for a specific activity or technology. As such, the ISM's cybersecurity guidelines provide an important input into each organisation's risk identification and risk treatment activities, but cannot represent the full extent of such activities.

While cybersecurity guidance can assist with risk identification and risk treatment activities, an organisation will still need to undertake its own risk analysis and risk evaluation activities due to the unique nature of each system, its operating environment and the organisations risk tolerances. As per the PGPA Act, any risk-based decisions used to design a gateway must be in the context of that organisation's system of risk oversight and management.

The invisible cannot be protected

An organisation should have visibility of all data entering and exiting its security domain(s). Visibility can be provided either through gateways within the given security domain, or a combination of gateway capabilities within the security domain combined with appropriate and trusted visibility sources provided by the external security domain. If the latter approach is taken, it must be aligned to the organisation's risk-based approach. This allows organisations to balance the investment in gateway capabilities with the risks associated with trusting the external security domain for a subset of the visibility required.

PEPs cannot provide adequate protection where they are unable to apply security policy to data transfers between security domains. The use of encryption can result in security policy enforcement blind spots. An organisation must understand and appropriately defend information holdings, paths and interfaces across its security domains.

Choosing not to decrypt content poses a security risk of encrypted data moving between security domains. In addition, encryption could mask information at a higher classification being allowed to pass to a security domain of lower classification, which could result in a data spill. Where gateway decryption, inspection and policy controls are not feasible or achieving necessary security outcomes, substitute policy enforcement point security controls, such as network segmentation and endpoint control capabilities on client or server endpoints, are required to appropriately manage the risk.

Gateways protect organisations and staff

A gateway should be positioned appropriately to prevent the flow of information between boundaries and prevent the bypass of organisational security policies. A gateway should have enough context to permit the data flow, but also to enforce security policy controls as required.

A gateway will block traffic by default, and only allow traffic flows that are expressly permitted. Gateway policy should protect an organisation from malicious actions. A gateway should implement security controls to protect staff from malicious or inappropriate content, and play an important role in preventing security policy breaches through the inadvertent or deliberate actions by staff.

Commonwealth entities have specific obligations

Commonwealth entities will undertake necessary measures to identify and defend against cyberthreats including monitoring, cybersecurity incident response and dynamic cyber defence capabilities. These include products and services for Commonwealth entities provided by ASD. Just like any product or service, these will have different requirements which may influence architecture designs. A Commonwealth entity needs to consider its obligations, including strategies outlined in the PSPF, PGPA Act, and ISM.

Plan for security flaws

An organisation's ICT systems should be resilient to the failure of security controls. When designing systems, the concept of defence-in-depth should be considered, particularly where control failures have occurred in the past. An organisation should design systems where a control failure (or other flaw) of a component does not result in catastrophic system failure or compromise.

Prior to implementing new systems, organisations should undertake table-top exercises and subsequently develop cybersecurity incident response plans (CIRPs) for data spills, particularly when data is no longer within a security domain controlled by the organisation.

Control design and placement should anticipate and allow for the subversion of critical components to the extent reasonably practicable, noting that the organisation may have to plan for a number of malicious actors.

All systems are susceptible to bugs and security issues. Plans should be made in advance on how to deal with them. Pre-emptively building cyber resilience and response capabilities into systems can make it easier to deal with events as they arise; for example, being able to apply patches quickly.

Gateways may provide organisations options to quickly implement compensating controls where a vulnerability is detected through either restricting access to a vulnerable service or monitoring incoming until remediation can occur.

Balance business and security

Critical to effective security is ensuring that business and security objectives are balanced to empower business units to function both efficiently and effectively, while managing risk through the implementation of commensurate security controls. Each security domain will have unique systems, security requirements and business objectives. This is particularly pertinent for security domains associated with development, sandboxing, research or dedicated 'dirty' networks supporting an organisation's cybersecurity operations (e.g. malware analysis).

An organisation should demonstrate consistency in its approach to risk. Security policy exemptions and inconsistent architectural and security capabilities in different systems introduce risk to an organisation. An organisation's gateway should apply security policy consistently.

An organisation should consider what additional governance mechanisms and oversight are needed where multiple gateways are deployed, or where security policies are enforced through different technology stacks.

Placement of gateways

An organisation's architects and engineers should be aware of industry trends to consider gateways as a suite of capabilities, rather than a specific network architecture. This allows organisations more flexibility in architecting solutions, while still retaining focus on the suite of capabilities that will best reduce risk in the specific environment in which each organisation operates.

Gateways should still be the only network path for information to be transferred into and out of an organisation's security domain. There are a number of potential architectures, and there may be multiple gateways deployed in an organisation (e.g. internet, voice, cloud service, suppliers and partners, etc).

Gateways should be designed around business requirements for availability and performance considerations. Not all services need high-availability (HA), but some business functions may require higher gateway resilience or performance than others. It may be more effective to design multiple gateways to cater for different business requirements and services.

All design decisions for a system (or gateway) should be documented and the rationale of these decisions captured. These decisions should reflect the principles and considerations presented in this document, and should summarise the architecture and controls that have been implemented within the gateway. These decisions are effectively the gateway design and should be reviewed regularly by the system owner and the authorising officer that is responsible for acceptance of the security architecture and the residual security risks. An organisation should leverage the expertise of enterprise architects, and engineering and operations teams, and the advice of experts (such as consultants, portfolio organisations, or the IRAP community).

Cloud

The following cloud related sections of this document should be read in conjunction with the *Cloud assessment and authorisation* and *Cloud assessment and authorisation FAQ* publications. These publications provide an introduction into several key cloud concepts, including the shared responsibility model, the trade-off between cloud computing and self-managed systems, IRAP assessments to cloud gateway deployments, and other considerations. The following guidance is also potentially applicable to many other outsourcing scenarios, including most managed service provider offerings.

The visibility, detection, prevention, and protection security capabilities provided by PEPs across data flows traversing security boundaries remain important to the overall security posture of a system hosted in the cloud. Infrastructure as a Service (IaaS), Platform as a Service (PaaS), and Software as a Service (SaaS) all have data flows and, therefore, it is equally applicable to all of them. When considering security capabilities across data flows, a range of interrelated themes need to be considered, including:

- security boundaries with reference to the system's architecture, components and dependencies
- inbound and outbound interfaces
- tenancy segmentation and segregation
- the classification and sensitivities of data held within the system
- the business context of the service
- visibility and monitoring requirements.

Cloud consumers need to consider all aspects of the CSP, its cloud services and a cloud consumer's own systems to make an informed decision about its use, and not rely on a single factor to determine suitability.

Some cloud gateways may only offer limited or proprietary logging. Organisations should understand what telemetry is available and how to implement it into their security monitoring capability. In cases where high confidence is required, an organisation may implement non-native capabilities to instantiate additional logging or telemetry collection to complement default cloud capabilities.

Shared responsibility model

In providing cloud PEP solutions, there are shared responsibilities between the different parties providing the complete cloud solution. One party may be predominantly responsible for the necessary security capabilities, or different aspects may be shared between parties. In all cases, the cloud consumer must retain visibility across data flows and have the ability to further restrict or filter data flows as the consumer sees fit. For example, filtering data flows containing files based on a YARA rule, and HTTP requests based on HTTP headers.

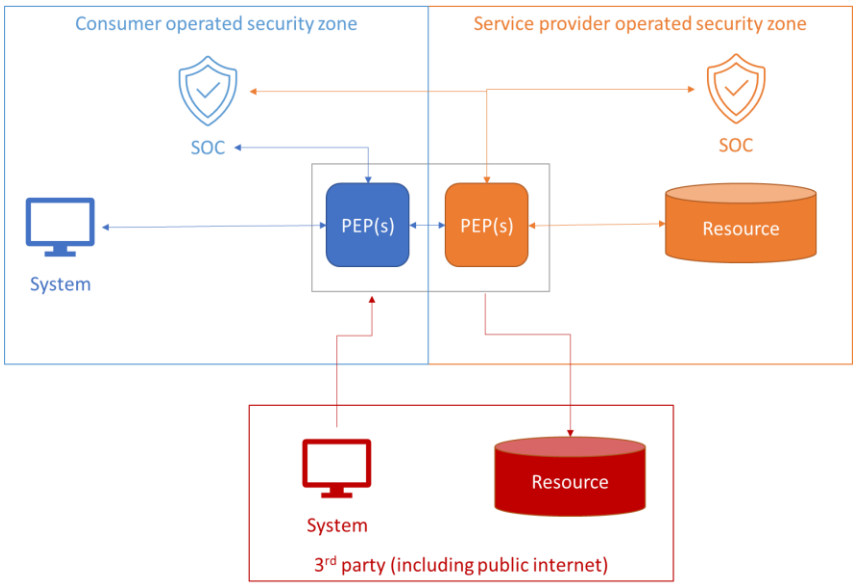

Figure 3: PEPs and security domains

It is important to understand that PEPs can be effectively layered together across data flows. For example, one cloud service may use the security capabilities provided by another service offered from the same CSP, or the cloud consumer may deploy its own additional capabilities inline between the client and the cloud service resource. When layering policy enforcement point capabilities, it is important that these do not lower the security baseline provided by the CSP and introduce new weaknesses. The connection between the two security domains needs to also be appropriately secured. Cloud consumers should also consider wider aspects when evaluating the desired architecture, including any impacts on user experience or availability.

Further information

- YARA, *Documentation*

Tenancy considerations

It is important that services exposing interfaces for clients to consume also provide mechanisms for a client's PEPs to appropriately restrict traffic to only the expected tenant or instance. For example, by providing:

- a unique tenant domain or URL path
- a combination of IP address and port, with a unique HTTP header on all requests
- a Software Defined Networking (SDN) construct including unique virtual network interface, or
- a virtual network route that can be enforced by a PEP on the client side.

If any system component can make a request outbound to an unauthorised or uncontrolled tenancy (even if the client has a valid signed payload), this could enable an outbound command-and-control path that the consumer would be unable to identify or constrain.

Inbound and outbound data flows need to be protected by PEP capabilities. Particularly for SaaS and PaaS with multi-tenant services, CSPs need to consider how they will enable each consumer's own security policy to be applied across both outbound and inbound data flows. A CSP should support one or more methods. As an indicative example, if the responsibility for some of the security controls falls on the consumer to implement, then that service could support using a consumer-provided proxy via either a SDN construct (including a Gateway Load Balancer construct as referred to by some CSPs) or application construct (including API based event bus for policy enforcement).

Cloud native and cloud aware capabilities

The increased prevalence of APIs and other architectural changes across cloud native offerings brings with it opportunities to introduce additional ways to deliver security capabilities across data flows. This provides the building blocks for a range of potential benefits, including improved security posture or improved user experience, for both CSPs and cloud consumers to leverage.

Service native integrations

Cloud native gateways should adopt cloud services, instead of pointing existing monolithic architectures to the cloud.

Service native integrations could provide stronger security capabilities, as they can be service-and application-aware, such gateways and services have two-way integrations that provide the best outcomes. Some benefits could include:

- Continuous protection instead of point-in-time protection. For example, a traditional gateway may only apply signature-based detection against payloads at the time data enters the security domain. On the other hand, a service native integration could scan all the data stored within the service when a new signature is added.
- Enhanced user experience. For example, instead of potentially silently dropping an inbound email at the gateway, a user could be provided an application interface to gain visibility across blocked emails and have the option to take appropriate streamlined action against false positives if required. This could include a render safe preview of filtered or higher-risk emails.
- Providing a broader and more integrated view of security, as the controls are applied within the context of the services fulfilling business objectives.

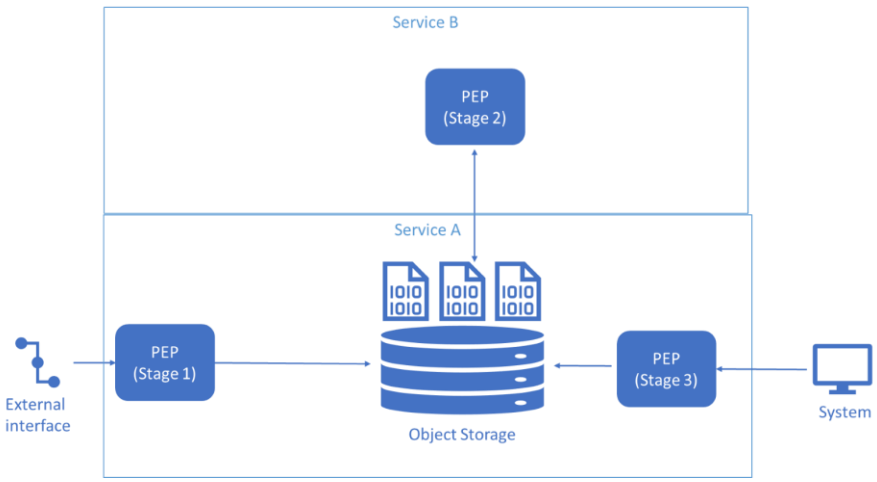

Figure 4: Abstract data flow pattern

Figure 4 presents a pattern that could be applicable to use in cases where:

- Service A is an email service, business cloud storage service, or big data lake ingress
- Service B applies additional security capabilities, including those that would be traditionally applied in line outside of the application context.

Additionally, an organisation could apply multiple services (in the location of service B in the Figure 4) and collect metrics in relation to the operations of each service to assist with comparing each service's value. For example, processing time, reported false positive rate or error rate.

Figure 4 presents an abstract data flow pattern in which:

- The initial security capabilities applied in the stage 1 PEP focus on protecting the external interface, but not the payload object that is stored directly in the object storage. For example, PEP1 could be a WAF/CDN/HTTP reverse proxy (configured not to look at http body).
- Asynchronously service B picks up the object for processing via the stage 2 PEP and stores its findings alongside the object. For example, PEP2 could be a file sanitizer or AV.
- The stage 3 PEP enforces that no consumer system can read the object until the service B processing is complete and the findings are in line with the organisation's security policy. For example, PEP could be part of the authorisation system.

Cross-organisation collaboration within the one cloud service

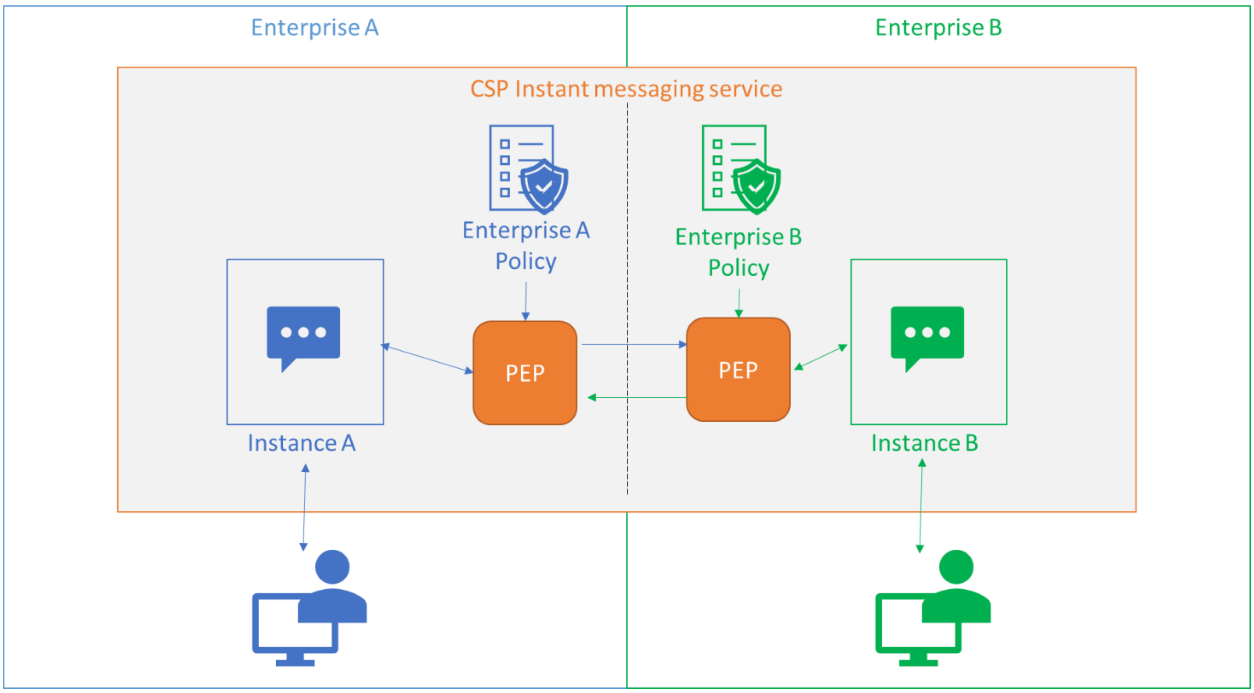

Figure 5: Cross-organisation collaboration within the one cloud service

When using the same underlying cloud service, and data flows occur between two different security domains (in this example two different organisations), that both organisation's security policies must be enforced. This results in both organisations having visibility and control of the data flows. Notably, the Enterprise B PEP could apply its own

additional capabilities across the data flows. For example, additional message filtering or content conversion security capabilities via an API integration. Additionally, it is recommended that relevant application logs and context related to the specific data flows traversing both security domains are made continuously available to both participating organisations.

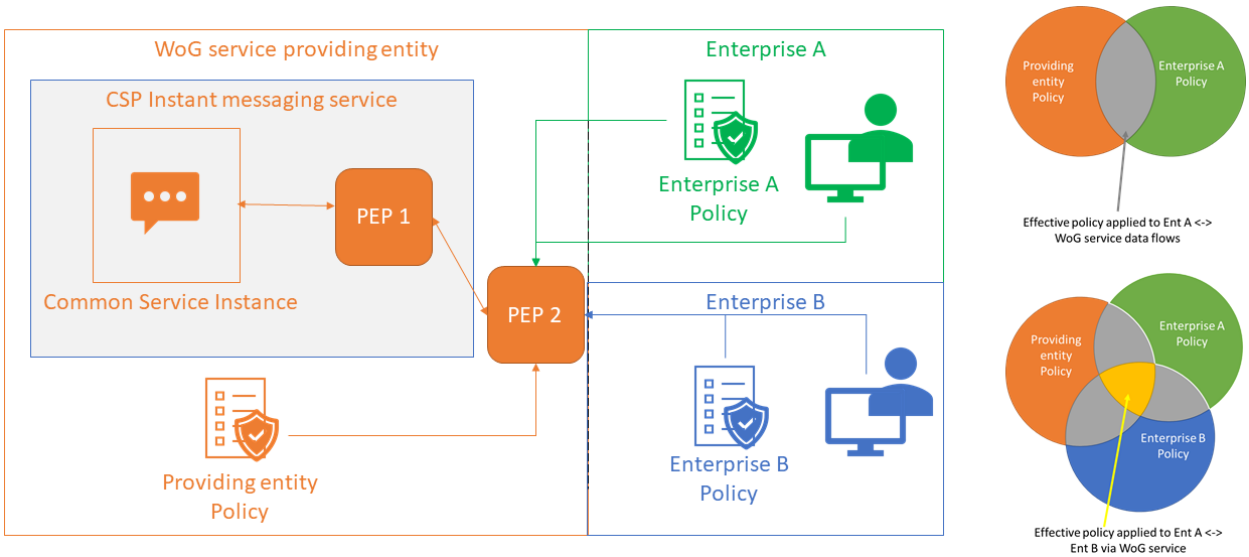

Figure 6: Shared common security domain

Figure 6 shows a similar example, but where two organisations are relying on a shared common security domain, for example as found in some Whole of Australian Government (WoAG) services. In this example, the providing organisation has decided to deploy an additional PEP that is capable of enforcing the security policy of all organisations at the same time (as shown in the Venn diagrams on the right of Figure 6). It would be expected that the relevant information from the providing organisation is shared with the consuming organisations, and that the consuming organisations would have direct control over their own security policy.

Each organisation is responsible for its own risk assessment and granting authorisation to operate (ATO) of the common service - in effect, authorising the organisation's data to be transmitted, shared or processed by the shared service, and ensuring that the common service and PEP are in line with its own risk tolerances. For simplicity, additional PEPs within enterprise A and B have been omitted, but it is noted that they may be deployed by the consuming organisation.

Segmentation and segregation

Network segmentation involves partitioning a network into smaller networks, while network segregation involves developing and enforcing a ruleset for controlling the communications between specific hosts and services.

The evolution of cloud service offerings has lowered some of the barriers to macro-and micro-segmentation. Segmentation and segregation design considerations that are applied to cloud native workloads should also be considered for the security capability implementations; for example, the applicability of ephemeral microVMs for data processing.

Additionally, different security zones can have different security controls applied against them. This can be very helpful when trying to enable specific business scenarios.

Further information

- ASD, *Implementing network segmentation and segregation*

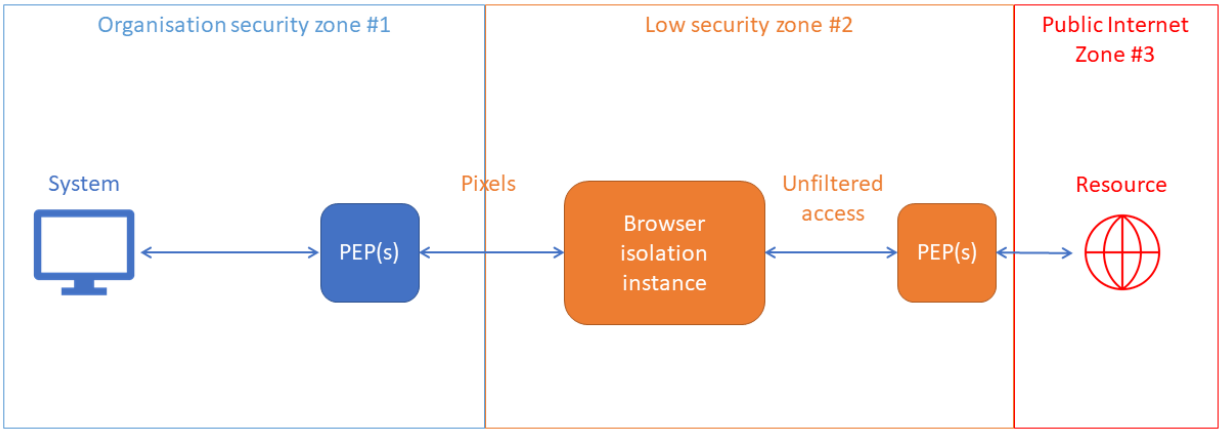

Figure 7: Security zones

In Figure 7:

- the PEP in security zone #1 applies all the appropriate protections to protect the enterprise data held within zone #1
- the PEP in security zone #2 provides logging security capabilities but does not filter or break TLS between the virtual browser and the public internet. This can be done, as there is no sensitive enterprise data within the lower security zone to protect.

This same pattern can be applied to sandbox or training environments that hold no sensitive enterprise data and to a lesser extent to particular development scenarios, as long as data flow from the low security zone to the enterprise security zone has appropriate security capabilities applied. For example, any code from the lower security zone would need to be treated as untrusted. See 'Non-persistent virtualised sandboxed environment' strategy within the *Strategies to mitigate cybersecurity incidents: Mitigation details* for related information.

Further information

- ASD, *Implementing network segmentation and segregation*

Gateway feeds for SOC and cybersecurity incident response teams

Not all organisations have the skills, experience and budget to operate a 24x7x365 Security Operations Centre (SOC), but that does not mean that they should not collect and make use of logs and other system telemetry generated by gateways.

An organisation should prioritise the centralisation and protection of system logs. This should be done prior to implementing other systems, as this data is essential to undertake cybersecurity incident response (i.e. prioritise building a scalable logging solution prior to investing in a Security Information and Event Management [SIEM]).

Gateway logs and telemetry may come from a variety of sources:

- systems and applications (e.g. authentications, operational logs, headers, geo-location)
- network flows (e.g. IPFIX/Netflow/JFLOW/SYSFLOW)
- traffic payload artefacts (e.g. packet captures of decrypted content, remote object inspection [ICAP]).

An organisation's cybersecurity incident response teams should have access to logs, telemetry and other artefacts from gateways that provide authentication, traffic inspection, and logging of network traffic traversing its gateways.

High-value logs from a gateway include traffic to and from authentication systems, DNS servers, web proxies, mail relays, load balancers, IPS/IDS, remote access solutions, sandbox and ICAP services, and other security services. Network flow telemetry from gateways can enhance security capabilities.

Operational teams, software developers, suppliers, service providers and security teams need to work together to identify the logs and telemetry that may indicate that a system/application is operating correctly (identifying standard activities) or when systems are operating out of specification (i.e. potentially compromised), and when policy enforcement features are not operating correctly. A SOC should have operational visibility into planned system changes to reduce the risks of false positives that may be associated with authorised changes to a system. Mature organisations can gain additional insights into malicious actor TTPs through the payload analysis of an attack. This may be achieved through packet capture, or the results of sandbox detonation.

Complete packet capture and storage of all gateway traffic is impractical; however, organisations should have the capability to decrypt and store network traffic, and test the capability to conduct targeted data capture of higher risk sessions and services, through random sampling or in response to a cybersecurity incident or investigation. Processes should be developed to provide captured data to the organisation's security operations centre (typically a cybersecurity incident response team) in a format that facilitates forensic analysis. This data could be in the form of a Packet Capture (PCAP) file, or streamed data.

During a cybersecurity incident, the availability of logs becomes more important to cybersecurity incident response teams. Gateways, and systems deployed in gateway DMZs, will typically have the ability to tune systems to increase the verbosity of system logs, providing additional visibility. An organisation should develop and test procedures to increase and decrease log verbosity in response to different scenarios.

An organisation's SOC will, of course, be interested in more than just gateway events. An organisation should also collect and analyse internal network flow telemetry and endpoint event logs (servers and workstations), behavioural analytics of processes, endpoint security related logs, and events from any internal security capabilities.

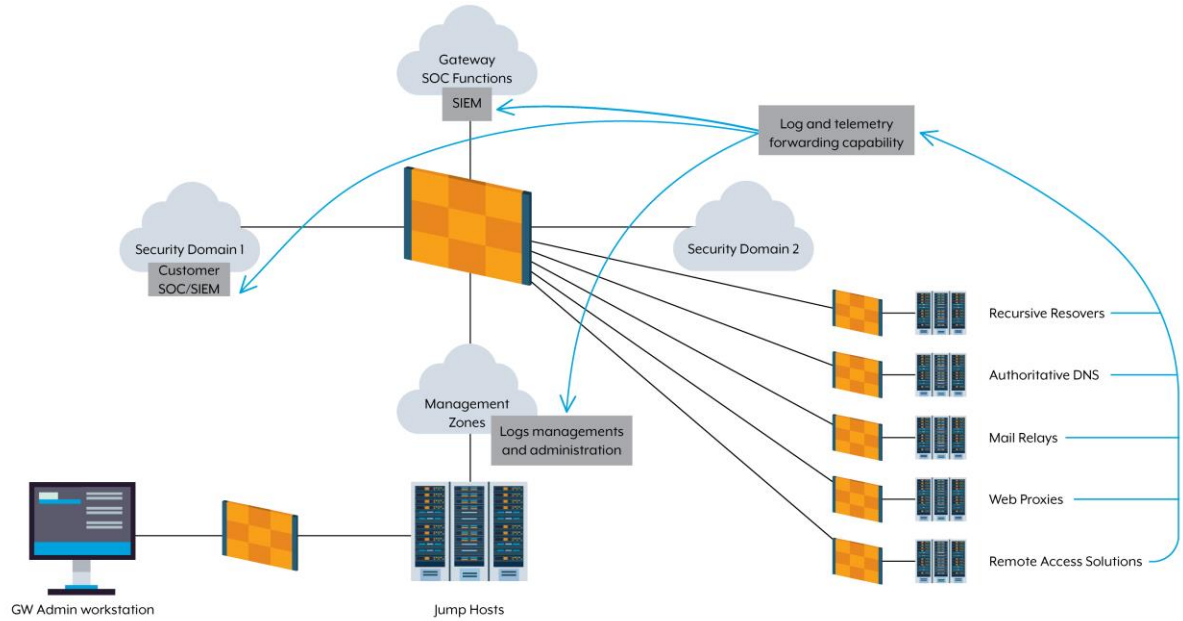

Figure 8: Logs and telemetry transfer from gateway systems

Cybersecurity incident response

Cybersecurity incident response activities involve a range of stakeholders, including senior executive, application business owners, security teams (cybersecurity incident response, forensic analysis), ICT support teams, media teams, and key service provider contacts. An organisation's CIRPs prioritise the active maintenance of contact details and the roles of key personnel involved in the cybersecurity incident response process, storing this information offline in accordance with the organisation's business continuity plans.

A system's CIRP should include components for foreseeable events. Data spills occur and a CIRP should document the processes and responsibilities of various stakeholders (e.g. what to do, and who performs which functions, who is in control of the cybersecurity incident, and how that impacts on services). An example may be to document the process to manage a data spill into a cloud environment. 'war-gaming' and post-cybersecurity incident review activities should identify if additional controls, processes and training could prevent the spill.

An organisation should not look at a CIRP as just another project deliverable. A CIRP should accurately reflect what an organisation will do when crises occur, and the cybersecurity incident response context should have a wider context than just IT operations.

An organisation should conduct 'war-gaming' cybersecurity incident response exercises to refine and demonstrate that the plan would be effective in the event that the plan needed to be executed.

CIRPs should specifically address the fact that cybersecurity incident managers and participants need to be empowered to perform their function during a crisis. An organisation should document how cybersecurity incident management activities will be implemented, including people, processes, operations and management functions. Changing these plans during a cybersecurity incident can result in sub-optimal outcomes.

An organisation may not have the expertise or resources to sustain an extensive cybersecurity incident response capability. Business leaders should evaluate the benefits of having third party expertise on retainer to assist with the management of a significant cybersecurity incident.

Control failures can and do occur. Software is imperfect, even in security products. An organisation should regularly war-game its CIRPs. This helps test assumptions before a cybersecurity incident, and helps teams gain an understanding of why various datasets are useful.

Gateways can help an organisation respond to vulnerabilities; for example, as a temporary measure, a network intrusion prevention system or web application firewall may be able to actively block attacks against vulnerable systems that require patching, and firewalls can ingest cyberthreat intelligence (CTI) to automatically block sources of malicious traffic.

A service provider's response to the questions below can indicate their ability to handle cybersecurity incidents.

Timely service provider support

- Is the service provider readily contactable and responsive to requests for support, and is the maximum acceptable response time captured in the service-level agreement (SLA)?
- Is the support provided locally, or from a foreign country, or from several foreign countries using an approach that 'follows the sun'?
- Which mechanism does the service provider use to obtain a real-time understanding of the security posture and configuration of the service provider's services?

Vendor's CIRP

- Does the service provider have a CIRP that specifies how to detect and respond to cybersecurity incidents in a way that is similar to cybersecurity incident handling procedures detailed in the ISM?
- Can the organisation review the service provider's CIRP?

Training of service provider's employees

- Which qualifications, certifications and regular information security awareness training do the service provider's employees require to know how to use the service provider's systems in a secure manner, and to identify potential cybersecurity incidents?

Notification of cybersecurity incidents

- Will the service provider notify the organisation via secure communications of cybersecurity incidents that are more serious than an agreed threshold, especially in cases where the service provider might be liable?
- Will the service provider automatically notify law enforcement or other authorities, who may confiscate computing equipment used to store or process an organisation's data?

Extent of service provider support

- To what extent will the service provider assist the organisation with investigations if there is a security breach, such as an unauthorised disclosure of the organisation's data, or if there is a need to perform legal electronic discovery of evidence?

Access to logs

- How does an organisation obtain access to time-synchronised audit logs and other logs to perform a forensic investigation, and how are the logs created and stored to be suitable evidence for a court of law?

Cybersecurity incident compensation

- How will the service provider adequately compensate a consumer if the service provider's actions, faulty software or hardware contributed to a security breach?

Data spills

- If data that an organisation considers too sensitive to be stored in a location is accidentally placed in that location (referred to as a data spill), how can the spilled data be deleted using forensic sanitisation techniques?
- Is the relevant portion of physical storage media zeroed whenever data is deleted? If not, how long does it take for deleted data to be overwritten by consumers as part of normal operation, noting that some service providers have significant spare unused storage capacity?
- Can the spilled data be forensically deleted from the service provider's backup media?
- Where else is the spilled data stored, and can it be forensically deleted?

Further information

- ASD, *Cybersecurity incident response planning: Executive guidance*
- ASD, *Cybersecurity incident response planning: Practitioner guidance*
- ASD, *Cyber Incident Management Arrangements for Australian Governments*

Architect for maintenance

It is common for vulnerabilities to be exploited within hours of a security patch becoming available. An organisation should adopt architectures and business processes to ensure patches and mitigations can be deployed with very short notice. An organisation should generally prioritise unplanned outages over the risk of system compromise. Where patching is not feasible, gateways may provide organisations options for compensating controls for vulnerable systems.

An organisation should align change management processes to support emergency patching processes. As SecDevOps processes are more widely adopted in an organisation's supply chain, an organisation should consider if its existing change management processes are introducing barriers for operational teams. For example, cloud service providers are not beholden to a consumer's change management processes, but they do proactively implement processes to patch infrastructure in order to protect their consumers, consistent with the service providers shared responsibilities model.

Further information

- ASD, *Patching applications and operating systems*

Cyberthreat intelligence

A gateway should use CTI to proactively respond to threats and vulnerabilities that are relevant to the consumer(s) of the gateway. A gateway should allow an organisation to both derive and consume CTI from its operation, and generate data that allows security analysts to produce CTI. The *Gateway security guidance package: Gateway technology guides* document provides more information about how a gateway can be used to generate and consume

CTI. For the purposes of this gateway guidance, CTI is: data about current or potential threats and malicious actors that has been added to or analysed in order to derive contextual meaning to one or more organisations. The products of the intelligence process should provide meaning and context in order to help an organisation make data driven decisions (i.e. 'what does this mean to me?').

An organisation may use CTI in a range of activities depending on their context. This might include automated blocking of URLs, file hashes, search activities, or even simply a confirmation of whether similar CTI have or have not been observed.

Importantly, CTI can be more than an Indicator of Compromise (IoC) (e.g. a file hash, IP address or email in isolation does not constitute a CTI) as this provides no additional context to the IoC. An example of CTI may be advice to look for a particular file in a specific location, using a specific method, on a certain date, because a malicious actor was attempting to compromise systems in this specific way. In the context of CTI, intelligence may be considered as applying a hypothesis to an artefact (as opposed to conducting a forensic activity to conclusively establish facts).

CTI often includes a confidence indicator to assist organisations decide how to prioritise and take action on the CTI. There are levels of confidence in CTI, and, therefore, there are uncertainties in the conclusions that can be drawn from CTI. An organisation that consumes CTI has a number of decisions to make, such as: How will CTI be used in a gateway environment? What is the impact of relying on a piece of evidence?

Gateways should generate and consume CTI. Mature gateways may take automated response based on high confidence CTI received. This may include actively checking if an IOC has been observed, or taking action to block an attacker's IP or email address.

Some CTI, especially IOCs, may be highly perishable and will only be relevant and actionable for a short time. Advice is likely to have a short life (less than a day, often only hours). Advice should inform organisations of what they need to be actioning as a priority. As gateway teams (and SOC/cybersecurity incident response teams) have a limited number of resources, advice should include a priority as well as environmental contexts to assist with the triage of the CTI. Organisations should make security decisions based on CTI (e.g. by changing security settings, adding new security capabilities, changes in business processes, adjusting training and policy, or by making architectural changes).

An organisation should consider what resources are needed to react to the CTI they develop and share (context and importance should be part of the product).

Priority services for security visibility

While there are many services that are provided by a gateway provider, ASD recommends government organisations prioritise uplifting the security capabilities in the following five services:

- recursive resolvers
- web proxies
- mail relays
- reverse proxies
- remote access.

The *Gateway security guidance package: Gateway technology guides* document contains service-specific advice on different types of gateway services (e.g. mail gateways, DNS, web proxies, and remote access).

Most vulnerabilities are either logical faults in the application layer of these protocols, or are payloads transferred via the application layer of these protocols. The *Gateway security guidance package: Gateway technology guides* document discusses mitigation capabilities and strategies. Vulnerabilities frequently arise from misconfiguration or insecure default configurations in these services. References to hardening advice for network protocols is provided where available.

Contact details

If you have any questions regarding this guidance you can write to us or call us on 1300 CYBER1 (1300 292 371).

Disclaimer

The material in this guide is of a general nature and should not be regarded as legal advice or relied on for assistance in any particular circumstance or emergency situation. In any important matter, you should seek appropriate independent professional advice in relation to your own circumstances.

The Commonwealth accepts no responsibility or liability for any damage, loss or expense incurred as a result of the reliance on information contained in this guide.

For more information, or to report a cybersecurity incident, contact us:

cyber.gov.au | 1300 CYBER1 (1300 292 371)

Australian Government

Australian Signals Directorate