Brought to you by:

WIZ

Google Cloud Security Foundations

for dummies

A Wiley Brand

Build faster and more securely

Secure all of your cloud environments

Full coverage without agents

Wiz Special Edition

Steve Suehring

Hey readers,

Before you dive into this book, here’s what you need to know about Wiz and how we secure cloud environments for our customers.

We started out in 2020 with a simple idea: to democratize security, so that companies of all sizes and using any cloud service could make their security effective. In this book, we look at best practices for securing your Google Cloud environments.

Google Cloud is a comprehensive cloud platform, often used for data services such as BigQuery and Vertex AI, and also their compute and container services. Google secures its infrastructures, but it’s up to users to protect their own data. With the explosion of AI, it’s becoming very simple to create complete pipelines, unfortunately it’s more complex to have visibility and thus ensure that they are secure. This can lead to data breaches and leaks. So, it’s crucial to understand your cloud environment, understand its context, and prioritize your risks. We’ll cover all these points in this book and give you helpful tips and best practices along the way.

If you are interested in learning more about how Wiz can enhance your cloud security processes, check out wiz.io/demo or scan the QR code below to see the product in action.

Cheers,

Assaf Rappaport

CEO and Co-Founder, Wiz.io

Google Cloud Security Foundations

Wiz Special Edition

by Steve Suehring

for dummies®
A Wiley Brand

Google Cloud Security Foundations For Dummies®, Wiz Special Edition

Published by
John Wiley & Sons, Inc.
111 River St.
Hoboken, NJ 07030-5774
www.wiley.com

LIMIT OF LIABILITY/DISCLAIMER OF WARRANTY: THE PUBLISHER AND THE AUTHOR MAKE NO REPRESENTATIONS OR WARRANTIES WITH RESPECT TO THE ACCURACY OR COMPLETENESS OF THE CONTENTS OF THIS WORK AND SPECIFICALLY DISCLAIM ALL WARRANTIES, INCLUDING WITHOUT LIMITATION WARRANTIES OF FITNESS FOR A PARTICULAR PURPOSE. NO WARRANTY MAY BE CREATED OR EXTENDED BY SALES OR PROMOTIONAL MATERIALS. THE ADVICE AND STRATEGIES CONTAINED HEREIN MAY NOT BE SUITABLE FOR EVERY SITUATION. THIS WORK IS SOLD WITH THE UNDERSTANDING THAT THE PUBLISHER IS NOT ENGAGED IN RENDERING LEGAL, ACCOUNTING, OR OTHER PROFESSIONAL SERVICES. IF PROFESSIONAL ASSISTANCE IS REQUIRED, THE SERVICES OF A COMPETENT PROFESSIONAL PERSON SHOULD BE SOUGHT. NEITHER THE PUBLISHER NOR THE AUTHOR SHALL BE LIABLE FOR DAMAGES ARISING HEREFROM. THE FACT THAT AN ORGANIZATION OR WEBSITE IS REFERRED TO IN THIS WORK AS A CITATION AND/OR A POTENTIAL SOURCE OF FURTHER INFORMATION DOES NOT MEAN THAT THE AUTHOR OR THE PUBLISHER ENDORSES THE INFORMATION THE ORGANIZATION OR WEBSITE MAY PROVIDE OR RECOMMENDATIONS IT MAY MAKE. FURTHER, READERS SHOULD BE AWARE THAT INTERNET WEBSITES LISTED IN THIS WORK MAY HAVE CHANGED OR DISAPPEARED BETWEEN WHEN THIS WORK WAS WRITTEN AND WHEN IT IS READ.

For general information on our other products and services, or how to create a custom *For Dummies* book for your business or organization, please contact our Business Development Department in the U.S. at 877-409-4177, contact info@dummies.biz, or visit www.wiley.com/go/custompub. For information about licensing the *For Dummies* brand for products or services, contact BrandedRights&Licenses@Wiley.com.

ISBN 978-1-394-24585-7 (pbk); ISBN 978-1-394-24586-4 (ePDF); ISBN 978-1-394-28311-8 (ePUB)

Publisher's Acknowledgments

Some of the people who helped bring this book to market include the following:

Project Manager: Jennifer Bingham

Acquisitions Editor: Traci Martin

Editorial Manager: Rev Mengle

Client Account Manager: Cynthia Tweed

Content Refinement Specialist: Tamilmani Varadharaj

Introduction

A few companies dominate the cloud computing space. Amazon with Amazon Web Services (AWS), Google with Google Cloud, and Microsoft with Azure. Other players exist in the space but the mindshare seems to be captured by those big three. AWS is the current market leader but Google Cloud is a strong contender. Notably, Google is the thought leader in large scale containerized deployments because of Kubernetes.

Google faces challenges unlike the other two companies in creating, deploying, and maintaining their own services. Where Amazon began as a bookstore and Microsoft as a software company, Google began as a search engine. Its entire business model depends on being available, fast, and accurate with responses to searches. It's natural then to expect that Google would get cloud services correct.

Google Cloud is an extension or byproduct of its own service offerings, but made available to customers. Customers of Google Cloud can deploy on consumer-level versions of many of the same technologies that Google operates to deliver the main `google.com` search results and other popular and widely used services from the company. To be fair, AWS was created out of the same idea: that organizations might want to use the same tech that Amazon uses to deliver shopping, video, audio, and more.

What's unique about Google Cloud is how transparent Google is when sharing how it operate its services. Google provides significant and extensive architectural guidance for cloud deployments and playing nice with others. That architectural guidance is freely available but is also sometimes difficult to sort out. This book helps to sort out some of the technologies available with Google Cloud with a specific focus on security-related technologies and how those offerings can be used to help secure cloud-based applications.

Foolish Assumptions

This book is intended for chief information security officers (CISOs), DevSecOps, DevOps, and cloud security operations staff. These individuals are familiar with the underlying concepts in computer security and even familiar with cloud-related security.

The speed with which cloud providers roll out new technologies means that keeping up with the latest from each provider can be difficult. As you'll find while reading this book, the technologies exist but would benefit from having more automation and unification.

Icons Used in This Book

Within the book you will see the following icons. These icons are meant to share tips and other important information.

TIP

This icon shares some additional thoughts on the current subject that you might find helpful.

REMEMBER

This icon is meant to call out a particularly important topic so that you might commit it to memory or remember where to find it later.

WARNING

The Warning icon denotes something that might go wrong if not heeded. Like opening up a compute node to the world with known vulnerabilities. Don't do that.

TECHNICAL STUFF

This icon shares information that is even more technical than the technical information being shared elsewhere in the book.

Beyond the Book

Google has done an excellent job of creating content that is helpful for deploying and securing workloads on Google Cloud. Start at `https://cloud.google.com/security` to see the highest level view of Google Cloud Security. Other URLs are shared throughout the book where appropriate.

IN THIS CHAPTER

- » Understanding cloud computing
- » Examining Google Cloud security

Chapter 1 Introducing Google Cloud and Cloud Security

By now, the cloud is a familiar term not just to meteorologists but to everyone involved in computing. People tend to use the term the cloud to mean just about anything that's not running on a server in the building. But the cloud is more than just technology and refers to a shift in how computing problems are solved.

This chapter sets a baseline of cloud computing so that we all understand how we got here. Once that's complete, the chapter focuses on Google and its Google Cloud offering as being one of the leaders in the cloud provider space.

Defining Cloud Computing

Cloud computing represents a paradigm shift in the delivery of software and application solutions, but what exactly is the cloud? It's easy enough to type "What is the cloud" into your search engine of choice and come up with answers containing words like *on-demand* and *compute resources* and *as-a-service*. These definitions are true but incomplete.

REMEMBER

Somewhere within the 2.5 billion results of a Google search for *cloud security* should hopefully be a single source of truth on security in the cloud, but maybe not. Just as verifying compliance and data integrity and having visibility are important regardless of the location of the application and data, securing a computing workload in the cloud presents unique challenges, not the least of which is the complexity or sheer number of moving parts involved.

A very brief history of cloud computing

Throughout the 1990s, deploying application software for an organization, especially customized line-of-business applications, required significant planning. Waterfall methodologies were used to plan, analyze, and design what was going to be created and then implement the resulting software. Whether the software met the requirements or vision of the stakeholder was often an afterthought. Projects gained a momentum of their own, and change was feared and rejected.

Part of the planning process was attempting to size hardware appropriately for the projected workload of the software. This included predicting not only initial utilization but also planning for growth and usage over the following 12 to 18 months or more. This level of planning helped determine the initial need for hardware and the need for additional disk space. Planning to this level also enabled service-level agreements (SLAs) to be planned and hardware purchases adjusted accordingly.

If estimates for utilization were too low, additional hardware needed to be ordered, built, and deployed. If estimates for utilization were too high, hardware sat idle and unused. Scaling up and down with physical hardware was a manual and expensive process.

TECHNICAL STUFF

In the late 1990s and into the 2000s, virtualization of server workloads grew in popularity. As part of this shift, organizations purchased large physical servers with multiple processors, large amounts of memory, and deployed storage area networks (SANs). The physical servers were then logically divided into many virtual servers. From the perspective of developers and users, virtual servers appeared the same, using the same operating systems and the same software as they would use on physical servers.

The seeds that grew into cloud computing were planted with the shift to virtualization. Although physical hardware remained the backbone of virtualization, the consequences of incorrect estimates became less important and scaling an application to meet demand became easier. Testing software prior to launch became easier because a virtual machine used for testing could be decommissioned immediately after the test thereby making the hardware resources available for other uses.

Off-premises hosting solutions also existed in various forms. Organizations could co-locate their own hardware at data centers and Internet providers in order to gain regional coverage to meet demand as well as providing disaster recovery and redundancy. Virtual private servers and hardware-based private servers were also viable options throughout this time if an organization needed to host an application off-site in order to avoid incurring up front hardware costs.

Companies like Google and Amazon were using combinations of hardware and virtualization at scale in order to meet demand. The computing services used internally at these and similar companies could be used by others and for some of the same advantages as would be gained with a virtualized infrastructure.

Moving from physical to virtual was a layer of abstraction that decoupled hardware from the workloads that ran on it. Moving from virtualization to cloud is another level of abstraction, further decoupling a server-based infrastructure into an infrastructure where compute, memory, and disk are distinct, and where fundamental components like networking and security are handled virtually as well.

REMEMBER

But the shift to the cloud is about more than just moving work off-site. The shift to the cloud is about on-demand resources billed on a granular pay-per-use basis. For example, compute resources are frequently billed not only by the number of processor cycles or CPU cores but per minute as well. A single-core processor may cost fractions of a penny per minute but a large multicore multiprocessor compute node might cost a dollar or more per minute.

Just as hardware could be purchased from multiple vendors, so too can cloud services be purchased from one or more vendors. There are numerous big-name companies in the cloud space including Amazon with Amazon Web Services (AWS), Google with

Google Cloud Platform (GCP), and Microsoft with Azure. Other big names include IBM and Oracle.

A cloud-first paradigm changed how software is developed, tested, delivered, and secured. Software and data that could be physically protected within the walls of an organization is now hosted in one or more regions, sometimes subject to multiple regional regulatory requirements. Where an operations engineer could point to the server that held the data, they now point at a map on a screen. This additional layer of abstraction presents challenges of its own for security.

When developing software, avoiding a big-bang deployment means shifting responsibility for testing earlier in the software development lifecycle (SDLC). By necessity, moving testing earlier in the SDLC means that environments for testing need to be built earlier. The shift-left mentality is part of the DevOps or DevSecOps movement that combines elements of development, security, and operations to deploy software faster and more accurately than before.

Defining Google Cloud

Google Cloud is one of a few major cloud providers, competing with Amazon Web Services (AWS) and Microsoft Azure. Google Cloud works like other cloud providers, with per-resource and per-minute billing for services. Like other providers, Google Cloud operates globally, with data centers in 39 regions.

Google Cloud offers numerous services to meet the varied computing-related needs of its clientele. Standard services provide compute power, database, storage, application-level, big data, and others. Specialized or industry-centric services are also available like those aimed at healthcare and life sciences.

REMEMBER

You may be familiar with the varied acronyms connected with *as-a-service* like infrastructure-as-a-service (IaaS), software-as-a-service (SaaS), and platform-as-a-service (PaaS). The products and services found on Google Cloud encompass all of those. Just as the responsibilities for operating an application depend on the as-a-service model, so too does the responsibility for security depend on the model chosen.

Introducing Google Cloud Security

Security in Google Cloud is like that of other cloud providers, where certain elements are secured by Google and others remain to be secured by you, the client. For example, physical security is something that you need to care about when running your own data center. On one level, you need to control access to the building and to the area within the building that physically holds the equipment. You also need to provide redundancy at the hardware level in case of disaster.

TIP

When using Google Cloud, you're no longer responsible for providing and maintaining physical security for the hardware itself, Google does that. Google secures its own data centers and ensures that unauthorized people are kept out. Google also has physical controls in place like fire suppression, but it's still up to you to determine how to provide for redundancy in case of natural disaster. Google will work to restore service at a data center as soon as possible but mission-critical workloads must still be shifted by you with minimal downtime to another region until the data center is recovered.

The principles of shared responsibility and shared fate are helpful when understanding the models used for security in the cloud. *Shared responsibility* is the term used to describe the most basic of cloud security: Google Cloud is responsible for securing the data center and related infrastructure and you are responsible for securing the application being deployed on Google Cloud. Figure 1-1 illustrates the shared responsibility model of cloud computing.

Along with shared responsibility is shared fate (for more on that, turn to Chapter 3).

Deployment of an application in the cloud shifts the risk associated with securing physical assets, but protecting the application is a risk that remains. Mitigating the risk is increasingly difficult because of the constant escalation of cyberattacks. In many ways, protecting against known risks in the physical realm is straightforward. But the increasing sophistication of cyberattacks means that yesterday's protections may not work today.

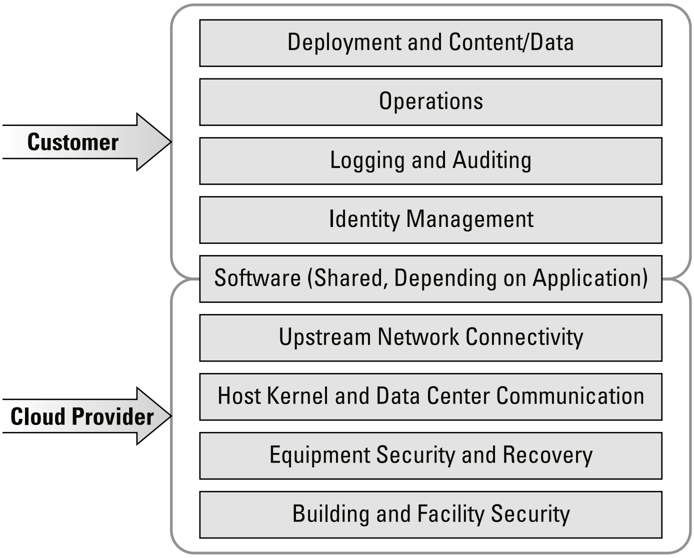
Customer
Cloud Provider
Deployment and Content/Data
Operations
Logging and Auditing
Identity Management
Software (Shared, Depending on Application)
Upstream Network Connectivity
Host Kernel and Data Center Communication
Equipment Security and Recovery
Building and Facility Security

FIGURE 1-1: Sharing responsibility for cloud security between the client and the cloud provider.

Recalling the history of cloud computing, Figure 1-2 shows how responsibility for some of the key components of hosting on-premises compares to cloud computing and how responsibility is now shared in the cloud.

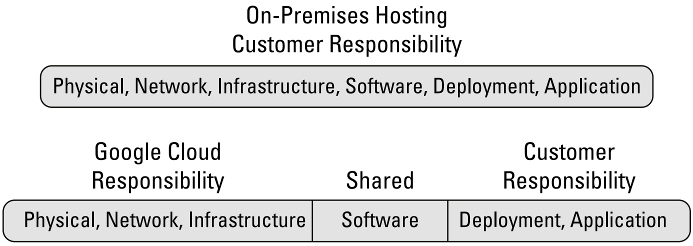

FIGURE 1-2: Sharing responsibility for security with the cloud.

The shared responsibility model only tells part of the story when it comes to transforming into a digital-first, cloud-first organization. Governance and compliance touch every aspect of infrastructure, network, application, and data security. Ensuring that an organization meets regulatory requirements and audits emphasizes the need for visibility across the entire software supply chain. This is illustrated in Figure 1-3.

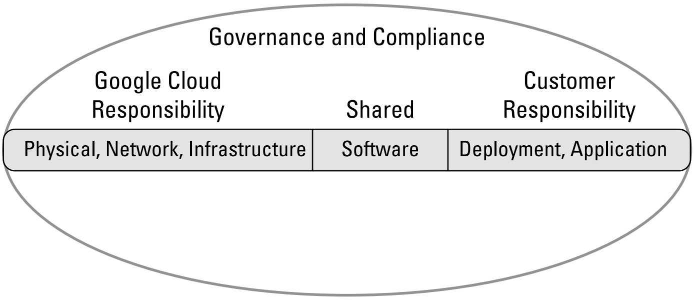
Governance and Compliance
Google Cloud Responsibility
Shared
Customer Responsibility
Physical, Network, Infrastructure
Software
Deployment, Application

FIGURE 1-3: Governance and compliance encompass all aspects of security in the cloud.

Of course, increasing visibility and ensuring compliance comes at a cost. Simply adding logging and tracking anomalies doesn't solve the problem and often merely increases noise. This prevents security and operations staff from focusing on the threats that matter most.

A cloud-native approach is needed to provide secure workloads in the cloud. The various control points that exist for on-premises deployments don't exist in the same form or at all in a cloud-based deployment. The speed with which software is developed and deployed to production has increased greatly in the last several years. Technology has kept up with the demand for rapid deployment by emphasizing containerization, repeatability, and code as configuration.

IN THIS CHAPTER

- » Examining defense in depth
- » Keeping things secure
- » Examining Shared Fate

Chapter 2 Architecting Cloud Security Solutions

At its most basic, computer security is about maintaining confidentiality, integrity, and availability. By applying processes, tools, training, and techniques, security professionals can begin to protect the assets of an organization.

The challenge, the thing that keeps the security professionals up at night, is that security needs to be correct every time but an attacker needs to be correct only once. It's that one single successful attempt that can enable an attacker to escalate the problem into a catastrophic outage.

There is no single solution for security. But you can do things to delay an attacker and reduce the impact of a successful attack. When architecting cloud solutions and migrating workloads to the cloud, choosing cloud-native security creates a better starting point for compliance and, ultimately, for security.

But cloud-native security solutions are only part of the story. Organizations frequently deploy workloads to multiple cloud providers. Although there are some integrations available related to security, a higher level cloud-native application protection platform (CNAPP) provides that higher level abstraction necessary to obtain a holistic view of cloud security posture and maintain compliance.

This chapter examines several key architectural components related to Google Cloud that help to enhance security. The goal is to relate security concepts that are familiar to their solutions found in Google Cloud. Google frequently takes the security concept to the next level when solving the problem.

Defense in Depth at Scale

Defense in depth describes the multiple layers of security that are necessary to decrease the risk of a successful attack and mitigate the effects of both successful and unsuccessful attacks. When deployed on a single machine, protecting data is the same logical challenge as when an application is deployed across a few machines. However, when an application is decoupled and deployed across multiple services, maybe even across multiple regions across multiple cloud providers, the complexity increases.

This section looks at security from the perspective of scaled security on Google Cloud.

Zero trust

Google Cloud implements a zero-trust policy for all services. Just because a connection is initiated from a known IP address or from an internal IP address doesn't mean that the connection should have any additional privileges.

REMEMBER

Zero trust is a foundation on which defense-in-depth is built. Rather than authenticating and authorizing requests at the network perimeter, a policy of zero trust authenticates and authorizes every request regardless of source.

Google Cloud carries zero trust to bare metal by integrating the Titan security module at the hardware level. In this way, Google can ensure and attest to a verifiable chain of execution for cloud workloads.

Zero trust on Google Cloud operates on three principles:

- **All network traffic is untrusted.** All requests and traffic flows throughout the network need to be validated when encountered and denied by default unless explicitly allowed.
- **Use least-privilege.** Grant only the minimum permissions on the minimum objects necessary to fulfill a request and then only for the duration needed.
- **Monitor everything.** Monitor for anomalies and track trends and current active requests for access.

Using zero trust, an organization spends less time pursuing attackers because even successful attacks can't leverage their position to gain further access. Zero trust also provides a consistent model on which applications can be architected. Developers don't need to be concerned with where their application is deployed because the process for identity and access management will be consistent.

Multilayered security

A key characteristics of defense-in-depth is multilayered security. Multilayered security creates barriers to successful attacks. Imagine a perimeter fence followed by guard dogs, followed by locked doors controlled by biometric authentication. All the while, the attacker is monitored by multiple systems. This is a multilayered security approach from any movie script but it mirrors what is done in a computing environment. The perimeter fence is the external firewall and guard dogs provided through monitoring and alerting. Locked doors are the credentials needed to authenticate into the target of interest for the attacker.

REMEMBER

Google Cloud data centers take a multilayered approach at the physical level, minus the guard dogs. Perimeter fencing, biometrics, laser intrusion detection, and other protections exist to protect systems from attack.

In a cloud-native multilayered security approach, deny-by-default and zero trust are the base. Together, a firewall plus virtual private cloud (VPC) networking prevent access from crossing boundaries. Key-based authentication through services like Google Cloud Key Management Service (KMS) and customer-supplied encryption keys (CSEK) help to secure data at rest while transport layer security (TLS) encrypts data while in transit, to, from, and within the network.

Even if the attacker can gain access to a system within the customer VPC, there is no guarantee that they can escalate or leverage themselves into another system. Further, an established chain of trust for cloud deployments can be created using Artifact Registry and Binary Authorization in Google Cloud. These cloud-native systems assume zero trust but also assume that scaling and repeatability will be needed.

Secure by default

Secure by default describes the overall posture of security architecture with Google Cloud. Rather than assuming trust, Google Cloud uses configurations that enhance security in their default states. A good example is encrypted-everything while in transit to and from Google Cloud. Customers can choose to add more protections through IPSec tunnels and other means for enhancing protection through encryption of data while in transit. Google also encrypts communications between virtual machines (VMs) within a VPC network.

TECHNICAL STUFF

All requests that go to Google Cloud traverse Google Front End (GFE). GFE not only provides distributed denial of service (DDOS) protections but also provides routing services and termination of HTTP(S), TLS, and TCP traffic. GFE is distributed worldwide.

Data at rest is also encrypted by default within the Google Cloud Platform. The Advanced Encryption Standard (AES) is frequently used for this purpose.

Maintaining compliance

Google Cloud is certified compliant with numerous regulations and many of the services can be used in a way that helps you to achieve or maintain compliance with industry-specific regulations. For example, the Health Insurance Portability and Accountability Act (HIPAA) has specific requirements for data protection. You can create a Business Associate Agreement (BAA) with Google that ensures the services are compliant with HIPAA.

TIP

Not all services within Google Cloud are covered as part of HIPAA. See https://cloud.google.com/security/compliance/hipaa#covered-products for more information on covered services relevant to HIPAA.

Beginning with a cloud-native deployment saves significant time attempting to audit and certify each layer. For example, the lower level Google Cloud infrastructure is covered within the BAA already. Auditors with experience in cloud-related deployments may use the certification and compliance by Google as evidence of the same for your applications when deployed on Google Cloud.

There are numerous compliance standards that Google is prepared for when using Google Cloud services. See https://cloud.google.com/security/compliance/offerings for more information on the compliance-related solutions found in Google Cloud.

Implementing defense in depth

GFE terminates connections at the ingress point to Google Cloud. These connections are then proxied by GFE to the appropriate Google Cloud service where the customer application is hosted. Traffic within the network is encrypted and authenticated, thus ensuring that the traffic flow is authorized to traverse the network within Google Cloud itself. All of this is done invisibly to the customer. You needn't do anything extra to obtain this level of protection.

Shifting left

Shifting left refers to concepts surrounding the DevOps and DevSecOps movement. A few of those concepts are particularly relevant:

- Test early and often
- Manage configuration as code
- Repeatable and automated deployments

In a traditional waterfall software development life cycle (SDLC), testing would be held until after development was complete. In an agile development-based SDLC, testing is done more often because of the iterative nature of development. However, even with agile methods, testing is traditionally separate.

REMEMBER

When shifting left, testing becomes more integral to the development process. A developer commits and pushes their code which then triggers a build process that in turn triggers a deployment process. Tests are executed in an automated manner, requiring as little intervention as possible from testers, developers, or operations staff.

The need to shift deployments earlier and then to execute those deployments earlier means that instructions for creating images and containers necessarily are created earlier. These instructions, in the form of configuration files, can change during the lifetime of the project. Just as it's necessary to trace code changes, so too is it necessary to track changes to configuration files. Managing configuration as code is then part of the shift left process.

Repeatability and automation are facilitated by shifting left. Deployments can be created on demand and automatically through build solutions like Google Cloud Build. The pipeline of a continuously integrated and continuously deployed application begins when these elements are shifted left.

TIP

Security needs to be shifted left to meet development earlier in the SDLC as well. Google Cloud includes several technologies to facilitate the integration of security at the earliest stages of development. The goal is to establish a chain of trust from development to deployment. Rather than trying to bolt on security later, integrating development with cloud-native tools alleviates pain points before they are ever felt.

Google identifies five phases of the software supply-chain, illustrated in Figure 2-1.

These phases have unique characteristics — services within Google Cloud help with each phase:

- Secure baseline (physical infrastructure)
- Source code (Cloud Code)
- Constructed artifact (Code Build)
- Deployed artifact (Artifact Registry)
- Promoted artifact (Binary Authorization)

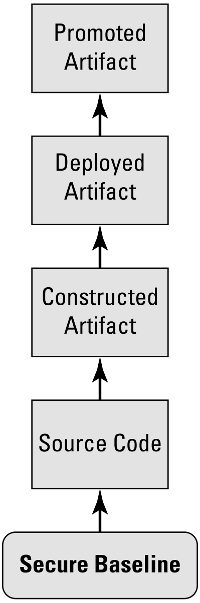

FIGURE 2-1: The five phases highlighted by Google.

At the base of a secure SDLC is the underlying hardware on which code is created, built, and deployed. Google works with hardware vendors to ensure security of the hardware foundation of Google Cloud, including the Titan security chip. Beyond those elements, architecting the configuration for that hardware is frequently application-dependent. Shielded VMs provide an environment that is free from rootkits or other untrusted software that might be injected into the boot process of the machine.

A declarative infrastructure or managing infrastructure as code includes configurations for underlying services like the continuous integration server and other basic services needed for the code-deployment process. Ultimately, this leads to the capability to deploy every part of the process through automated means and, in fact, prevents direct interactions. Instead, configuration changes even to underlying infrastructure components are pushed through a configuration management cycle and toolset. This reduces an attacker's ability to escalate privilege or a rogue (or inexperienced) administrator from making changes to the process.

At the next level is the code itself. This includes creating the code while assuring the security of the code-creation process along with repository storage security. Automated testing and scanning of the code are key elements of this aspect of security. Google Cloud includes several services that facilitate automation at this level. In addition, Open Source Insights examines dependency chains for vulnerability and licensing issues that might encumber the application that you're developing.

Code commits can be digitally signed by the creator and static code analysis performed to help provide assurances that the code at the lowest levels is free from basic defects, has not been tampered with, and meets organizational requirements for formatting and code structure.

TIP

Google Cloud also includes tools to scan for sensitive information and secrets.

Beyond lower level infrastructure and code commit is artifact creation and deployment. Google Cloud services are particularly adept at managing and securing artifacts to ensure a secure software deployment. The end of the process is Binary Authorization, which provides attestations of the provenance of the artifacts being deployed.

More information on shifting left with Google Cloud, including other best practices, can be found at `https://cloud.google.com/static/files/shifting-left-on-security.pdf`.

Impactful visibility

Google Cloud Monitoring is a central service within Google Cloud to which metrics can be sent to gain insights into performance of cloud-based workloads. There are numerous areas where events are observable throughout a typical Google Cloud deployment. Many services include logging of some form. Collecting of logs to a central location for real-time and archival analysis is also a central theme found in cloud-native deployments.

Beyond Cloud Monitoring are several other services that have more specific focuses toward security. For example, Security Command Center monitors for misconfigurations, threats, and vulnerabilities as well as helping to maintain compliance. Some of the services found within Security Command Center include:

- **Cloud Asset Inventory:** Discovery, monitoring, and analysis, including analysis of identity and access management policies affecting objects.
- **Security Health Analytics:** Identify misconfigurations and connect to the underlying benchmark or standard related to the correct configuration.
- **Web Security Scanner:** Automatic detection and scanning of web applications deployed within your cloud organization.
- **Event Threat Detection:** Log analysis to help identify potentially malicious activity.
- **Container Threat Detection:** Monitor activity in container workloads.

In addition to Security Command Center, Access Transparency and Access Approval enables an audit trail of access by Google Cloud personnel into data. Access can be traced back to a support ticket that was opened with Google.

Understanding BeyondProd

BeyondProd is the name given to the cloud-native architecture developed and implemented by Google. BeyondProd is used to secure the entirety of Google infrastructure, including the cloud offerings.

BeyondProd incorporates many of the security best practices described in this chapter but goes a step further to describe how all of the individual pieces fit together to create an integrated solution.

Containerized core

Code runs as microservices that run inside of containers. The use of containers means that portability and decoupling is necessary simply as part of the architecture. Because containers can be rapidly and repeatedly deployed, the capability to scale up and down is inherent. If a given microservice needs to service more requests, additional nodes can be deployed.

Containers have the advantage of being able to be deployed where needed, whether regionally or across cloud providers.

TECHNICAL STUFF

The container orchestration at Google is called Borg. Borg served as the basis for Kubernetes.

Trust

BeyondProd is zero trust. The root of trust begins at the hardware level within the BeyondProd network. This trust begins at the Titan chip, continues through firmware and to the application level.

Specific authentication and authorization is required for access within the BeyondProd network. Code that is untrusted would not have access to authenticate and thus could not initiate a call to another service. Escalation or using a compromised system as a launching point for other attacks is not possible.

Application Layer Transport Security (ALTS) authenticates and encrypts remote procedure calls (RPC) while also providing integrity and identity management for those calls. The key architectural takeaway from ALTS is that identity is bound at the service level rather than at the host level. This level of identity management decouples trust from hosts both upstream and downstream and enables scaling.

Code must have followed a known provenance chain of trust in order to be executed in the BeyondProd environment. This is achieved with the Binary Authorization service in both the BeyondProd architecture and Google Cloud.

Enforcement and isolation

Workloads are isolated within the BeyondProd network architecture. Some of this isolation is achieved as a byproduct of other protections such as least privilege and architected network layout. Enforced policies within the network also ensure policies that are set at the administrative level for compliance are maintained at all times.

TECHNICAL STUFF

The gVisor kernel is a tool to separate workloads at the kernel level. Exposure to the host kernel is thereby limited, thus reducing the attack surface where escalation or boundary crossing could potentially occur.

Edge protection

The perimeter of the network needs to be protected, as always. With cloud workloads, deny-by-default firewall policies remain best practice. The standard protection of blocking unauthorized connections is available but also DoS attack mitigation is available on cloud-native firewalls.

Google Front End (GFE) also provides protection at the edge for the BeyondProd network. Cloud Front End is the equivalent that you will encounter in Google Cloud.

Repeatable change control

BeyondProd uses blue-green deployment to avoid affecting existing traffic flows thereby eliminating downtime. Service migration is an automated process, meaning that connections are automatically drained and pointed toward new workloads as needed.

Building and deploying requires a two-person code review but is otherwise automated. For critical updates, such as vulnerabilities in the underlying operating system or in dependencies, live migration of virtual machines (VMs) can be used to prevent downtime.

TIP

Live migration isn't available for confidential VMs, for VMs that have graphics processing units (GPUs) attached, Tensor Processing Unit (TPU), spot, or preemptible VMs.

Service mesh

BeyondProd uses a service mesh to separate and share infrastructure security features. The effect is that security-related functions don't need to be integrated or even included at the application level. Rather, the infrastructure assumes the responsibility for incorporating common security requirements.

Common issues that were managed at the application level like management of identity and data access are handled at the service-level in a cloud-native architecture like BeyondProd. TLS termination is another layer of security that doesn't need to be handled by the application.

Instead, cloud-native architectures perform these functions at a different level. This change promotes consistency, without the need to determine whether a given development team implemented identity correctly or took shortcuts around encryption. Policies can be enforced and verified at common points throughout the infrastructure.

Within Google Cloud, the Anthos Service Mesh provides the fully managed experience.

TECHNICAL STUFF

Anthos Service Mesh is the Google Cloud implementation of the Istio open source project.

Understanding Shared Fate

The shared responsibility model has long been the foundation of cloud architecture. The cloud provider and the cloud customer are each responsible for certain areas of application security. For example, the physical layer and physical security are solely the responsibility of the cloud provider but securing the application code or ensuring that the application doesn't contain vulnerabilities has long been the responsibility of the cloud customer. Some areas have overlap, such as updating the underlying operating system.

Shared responsibility is clear for simple cloud deployments and workloads. However, in practice when more complex needs drive the solution, shared responsibility becomes more challenging. That's where shared fate comes in.

REMEMBER

Shared fate goes beyond shared responsibility and emphasizes best practices for security, viewing security as the problem of both provider and customer. As complexity increases, the number of security options and configurations increases. This makes it more difficult to maintain, validate, and verify compliance and security. With shared fate, Google Cloud takes a more active role in ensuring that the end-to-end experience is more secure by default.

Blueprinting solutions

As part of shared fate, Google wants to help secure your usage of the cloud. To achieve that goal, several blueprints have been created that provide a strong architectural foundation. A few such blueprints include:

- Security foundations
- Secured data warehouse
- Vertex AI Platform notebook security
- Anthos
- Secured serverless

The Security Foundations blueprint, which can be found at `https://cloud.google.com/architecture/security-foundations` is aimed at CISO, practitioners, and compliance officers. An important element of the security blueprints provided by Google is that they not only include hands-on how-to solutions but also include some of the reasoning behind the decisions and guidance.

Using assured workloads

Google offers solutions that help you meet compliance and regulatory requirements around data storage and access. For example, data residency provides a means to ensure data is stored in specific regions and only in those regions. Data can't be moved to unapproved regions.

Data sovereignty is another means of control over data access within Google Cloud. Personnel-based access controls can be used with Assured Workloads. These access controls along with ownership controls ensure that Google personnel must meet certain requirements such as physical location in order to access Google Cloud customer data.

TECHNICAL STUFF

Google also manages FIPS-140-2 compliant encryption keys and you can add your own keys through customer-managed encryption keys (CMEK).

Visibility and alerting for security

Security Command Center Premium provides a single point for aggregating security-relevant information gathered from several sources. Event Threat Detection and Security Health Analytics provide insights by gathering information from your cloud deployments. For example, Security Command Center might report that there is a misconfiguration within a container that could lead to a security issue.

The full list of sources for security information available with Security Command Center Premium is:

- Rapid Vulnerability Detection
- Security Health Analytics
- Web Security Scanner
- IAM Recommender
- VM Manager
- Policy Controller
- Sensitive Data Protection
- Anomaly Detection
- Container Threat Detection
- Event Threat Detection
- Forseti Security
- Virtual Machine Threat Detection

The single location of Security Command Center saves significant time compared to gathering information from each separate source. Security Command Center can be integrated with an already existing security and event management (SIEM) or security orchestration, automation, and response (SOAR) system. For example, cloud logs can be ingested into the SIEM or response automation integrated between systems.

Security Health Analytics

Security Health Analytics is worth highlighting separately. Security Health Analytics manages vulnerability scanning for several Google Cloud services, including:

- Cloud DNS
- Cloud Key Management Service
- Cloud Monitoring and Cloud Logging
- Cloud SQL
- Cloud Storage

- Compute Engine
- Google Kubernetes Engine
- Identity and Access Management (IAM)

Security Health Analytics operates in three modes:

- **Batch scan:** A once-daily run for all projects or organizations that are enrolled.
- **Real-time scan:** Certain detectors execute a scan when a change detected.
- **Mixed mode scan:** Certain resource types may not reflect a detectable change, in which case a combination of batch and real-time is used.

Google has leveraged the expertise that comes from running a huge Internet-dependent operation and a series of applications that all depend on security.

The concept of shared fate helps frame the commitment that Google demonstrates to sharing the experience of securing workloads on the Internet. Even with shared fate and with the number of tools available in Google Cloud, additional integration may be needed.

A CNAPP becomes integral to success, going beyond shared fate and the tools available in Google Cloud. Using a CNAPP, integration within a single cloud and across clouds becomes possible. An agentless CNAPP also alleviates the need to worry whether a given workload is being monitored properly.

IN THIS CHAPTER

- Securing lower levels
- Sharing responsibility
- Protecting the Application Layer
- Keeping the software supply chain safe
- Delivering data security

Chapter 3 Examining Cloud Security Layers

The multilayered approach to computer security isn't new. Rather than being a turtle, with a hard outer shell but a soft underside, security professionals know that to be successful, security has to be deployed in multiple places and through multiple methods.

Google approaches security from a layered perspective, understanding that applications will now be deployed on hardware that is owned, operated, and secured by Google on behalf of its customers who use the hardware to solve computing problems.

There is also an understanding of shared responsibility, where both Google and the customer need to take positive steps to secure the applications and workloads deployed in the cloud. Google takes shared responsibility a step further into shared fate. Shared fate emphasizes the integral role that Google can take in helping to secure applications and data for their customers.

This chapter examines layers of security with a Google Cloud perspective.

Securing Data at the Lowest Level

Google Cloud security provides a multilayered approach to security. Core infrastructure at Google Cloud is secured by Google. Google is responsible for the data center in a Google Cloud-native application. The security of its infrastructure provides the foundation, but other services exist that customers can interact with to enhance their own application security for Google Cloud-deployed applications.

REMEMBER

The infrastructure behind Google Cloud is the same as the infrastructure that delivers the Google services the world uses every day. The physical data centers that you would secure in a non-cloud deployment are secured by Google when your application is deployed to Google Cloud. Limiting access to the facilities and controlling and auditing what happens at those facilities is something Google handles through cameras, various forms of identification, and other physical control mechanisms.

Google also ensures security of the underlying hardware. From designing its own circuit boards to working closely with component and chip vendors, Google employs the means to verify and validate that the hardware on which their services run is secure.

Google secures the boot process as well. The custom Titan chip, which is found in Google Pixel devices, ensures that there is a chain of trust beginning at the root of the process tree. Where other hardware is involved, Google has a lockable firmware chip or other microcontroller thereby providing verifiable integrity at the hardware level for everything in the data center.

Google further secures the low-level infrastructure by automating and centralizing processes for:

- Machine life cycle issues
- Certificate rotation
- Diagnosis of hardware problems
- Controlling network access to verified machines
- Software version integrity and verification

REMEMBER

Google has built an end-to-end chain of trust for hardware. This level of provenance isn't normally achieved within an on-premises deployment at an organization but helps Google attest to the lowest levels of security as the physical hardware layer turns to the virtual software layer.

Sharing Responsibility at the Network Layer

With an application hosted entirely on-premises, you're responsible for the security of traffic at the ingress and egress points for every data center. With Google Cloud, protection of ingress and egress is primarily Google's responsibility. A service called Google Front End and its cloud counterpart Cloud Front End provide reverse proxy services to virtual machines (VMs) that need to be exposed externally.

TECHNICAL STUFF

As it relates to scaled and distributed denial of service (DoS) attacks, Google mitigates their effects through a centralized DoS service. This service receives reports from within Google's network from services such as GFE and load balancers. When necessary, the service applies remediation for attacks.

Encrypting traffic

Though a given for web communication, encryption at the network level isn't necessarily a given for internal communication today. For instance, an HTTPS request coming from an external Internet user reaches the web application firewall inside of an organization. From that point, the encryption is offloaded and communication travels in an unencrypted manner back to the web server. This is illustrated in Figure 3-1.

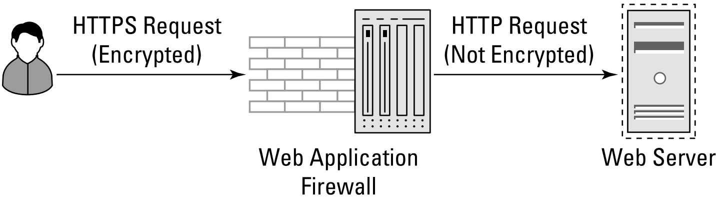

FIGURE 3-1: Network traffic traveling unencrypted on an internal network.

Within the cloud, end-to-end encryption is the standard. Rather than encryption ending at the perimeter, data is encrypted at all points and through all networks.

TECHNICAL STUFF

Transport Layer Security (TLS) is the standard for encrypting data while in transit. Within the cloud, end-to-end encryption is also the standard. Google Cloud accomplishes this through virtual network encryption. Google also uses Application Layer Transport Security (ALTS) for encryption along with authentication and data integrity. More information on the technical details of encryption in transit is available at `https://cloud.google.com/docs/security/encryption-in-transit#service_integrity_encryption`.

Leveraging a virtual private cloud

Just as with an on-premises deployment, Google Cloud consists of local area network (LAN) components and wide area network (WAN) components. The WAN components are Internet-related but also may be deployments of hybrid network solutions between a customer and Google. In addition, a virtual private cloud (VPC) can be created by a Google Cloud customer. A VPC provides a virtual or logical version of a physical network implementation.

There are several use cases that VPCs facilitate, including:

- Private network space, globally, and within and across regions that does not traverse the public Internet
- Connectivity to on-premises networks through the use of virtual private networks or Cloud Interconnect
- Advanced subnetting and network topologies based on organizational need

Network segmentation helps to provide a means for logical control of traffic flows within and between VPCs. As a customer, you can configure routing and segmentation to prevent traffic from passing into and out of secure areas of the implementation.

Using VPC Service Controls

VPC Service Controls have elements of firewall, exfiltration prevention, and access control. Isolation can be achieved through VPC Service Controls by setting service perimeters. In this way, you limit the ability for an attacker to leverage a compromised system in order to escalate their attack.

TIP

VPC Service Controls have a dry-run mode that merely logs requests that would otherwise have violated the control policy. Logs can be analyzed to further refine the VPC Service Controls. Using dry-run mode is particularly helpful for a development scenario to determine which traffic flows would fail if the project was deployed to production.

Implementing firewall controls

From a security standpoint, traffic from, to, and within VPCs can be controlled through several means. A firewall provides a cloud-native approach for preventing traffic from traversing the network. Firewall rules are managed in a centralized manner just as with an on-premises deployment.

VPC firewall rules control ingress and egress connections for both IPv4 and IPv6 in a stateful manner. Standard protocols are supported and connection tracking is available. Limitations apply for VPC firewall rules depending on the type of VM in use.

- **Standard shared-core machine:** Limited to 130,000 stateful connections.
- **1 to 8 virtual CPUs (vCPUs):** Limited to 130,000 stateful connections per vCPU.
- **More than 8 vCPUs:** Limited to 1,040,000 stateful connections in total.

TIP

Firewall rules default to least privilege for ingress. There is an implied deny rule that doesn't appear within the Google Cloud console for ingress. There is an implied allow rule for egress as well.

Google Cloud also implicitly blocks certain traffic such as DHCP offers and acknowledgements unless that communication is with the Google Cloud metadata server. In addition, by default, outbound TCP port 25 is also blocked to prevent spam. This limitation can be removed for certain customers and also does not apply for internal traffic. Figure 3-2 demonstrates this configuration.

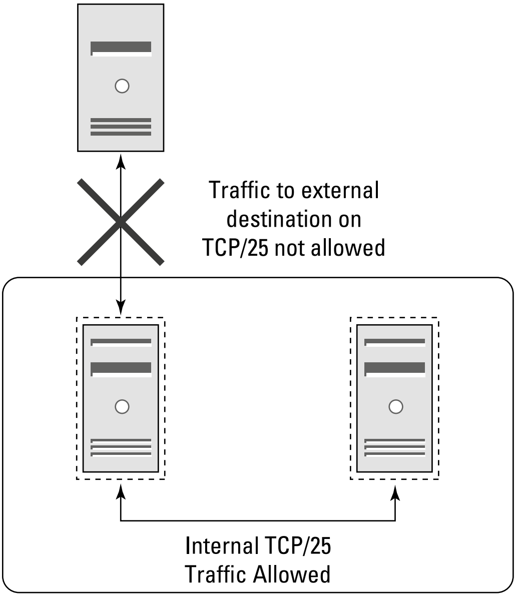
Traffic to external destination on TCP/25 not allowed
Internal TCP/25 Traffic Allowed

FIGURE 3-2: Google Cloud blocks outbound port 25 by default to prevent spam.

Firewall protection can be applied in multiple ways with a VPC in Google Cloud. See `https://cloud.google.com/vpc/docs/vpc#firewall_rules` and `https://cloud.google.com/firewall/docs/firewalls` for more information.

Applying bandwidth control

Monitoring and controlling egress bandwidth are helpful activities that can be used to mitigate attacks resulting in exfiltration and situations where a machine has been taken over and is being used for unauthorized purposes.

Certain controls on network bandwidth are available and under the customer's control with Google Cloud. The type of VM deployed within Google Cloud is a controlling factor for communication flows within a VPC. The use of Tier 1 networking capability is also a factor, notably for egress to external destinations. With Tier 1 networking, 25 Gbps is the maximum egress available, whereas 7 Gbps is the maximum available if Tier 1 isn't available.

TECHNICAL STUFF

Per-VM, per-flow, and per-project egress bandwidth is subject to limitations set by Google. You should be aware of these limitations and size the implementation accordingly.

Google Cloud doesn't impose limitations on ingress bandwidth within a VPC. The limiting factor is typically the VM's capacity and capability. Traffic from outside a VPC has restrictions on a per-VM basis of 20Gbps.

TIP

Various other factors can limit bandwidth capabilities for a VPC or a VM. See `https://cloud.google.com/compute/docs/network-bandwidth` for more information.

Securing the Application Layer

Up to this point, security for a cloud deployment on Google Cloud has primarily been the responsibility of Google. This includes default VPC security, firewall rules, transit encryption, and session limitation. As the customer, you can configure behavior of those fundamental elements in order to effect the configuration changes that you need for a project.

Application security is primarily the responsibility of the customer. That's true whether deploying a traditional application, an application programming interface (API), or another component that requires customized software. Visibility into performance of an application enables rapid response for security-related issues.

Scanning and testing

Part of a cloud-native deployment is scanning for known classes of vulnerabilities. Web Security Scanner, part of Google Cloud, is able to scan for security vulnerabilities for applications deployed on Compute Engine, App Engine, and Google Kubernetes Engine (GKE). Web Security Scanner can be run on-demand and automatically.

Security Health Analytics provides vulnerability assessment for several services, including:

- Cloud DNS
- Cloud Key Management Service (Cloud KMS)

- Cloud Monitoring and Cloud Logging
- Cloud SQL
- Cloud Storage
- Compute Engine
- Google Kubernetes Engine
- Identity Access Management (IAM)

Like Web Security Scanner, Security Health Analytics can be executed in multiple modes, including:

- **Batch scan:** Scans run once per day.
- **Real-time scan:** Scans are kicked off when a configuration change is detected.
- **Mixed-mode scan:** Combination of batch and real-time because not all services detect changes instantly and any that aren't detected would be covered by the batch run.

Numerous classes of vulnerabilities can be detected by Security Health Analytics. For example, API keys that are being used too broadly or that are unrestricted can be detected by Security Health Analytics. Having Secure Boot disabled or using a default service account are other examples. Classes covering SQL, DNS, containers, pub/sub pattern, and others are also available by default. Security Health Analytics detectors correspond with various compliance standards, including:

- Payment Card Industry (PCI) Data Security Standard
- National Institute of Standards and Technology (NIST) 800-53
- International Organization for Standardization (ISO) 27001
- Open Web Application Security Project (OWASP) Top Ten
- Center for Information Security (CIS) Google Cloud Computing Foundations Benchmark

Adding WAAP

Web App and API protection (WAAP) solutions are another standard means to protect an application. WAAP solutions combine best practices for common types of attacks.

REMEMBER

Cloud Load Balancing automatically defends against DoS attacks by scaling the application to meet demand. For example, if an attacker begins to send enough requests to hold connections open and exhaust resources, Cloud Load Balancing can add more capacity by deploying additional nodes to meet the demand. Cloud Load Balancing should be combined with robust monitoring, alerting, and cost controls to ensure that a DoS attack doesn't result in extreme scaling and added costs.

Cloud Armor provides layer 7 (application level) protections that are essential for a web-related application. For example, you might block requests that contain potentially malicious HTTP request headers or query strings that indicate malicious activity or bot scans. Geographical blocking is also available with Cloud Armor.

reCAPTCHA Enterprise can also be deployed to prevent malicious users or bots attempting to register, authenticate, or complete web forms repeatedly.

Although Cloud Load Balancing and Cloud Armor can be used to protect APIs, tools within Apigee API Management can also be deployed for additional protection. Apigee API Management includes API Gateway which can be used to throttle or limit calls to APIs as well as implementing authentication through OAuth and key validation, among other techniques to protect an API from malicious use.

Securing the Software Supply Chain

A modern software development life cycle (SDLC) includes steps for coding, building, testing, and deploying software to a production environment. Figure 3-3 illustrates a typical SDLC.

SDLC basics

The SDLC represents a software supply chain, with a combination of processes and tools that combine with the code produced by developers and managed by the DevSecOps and project teams. Each of the phases in the SDLC introduces an opportunity for an attacker to inject malicious code.

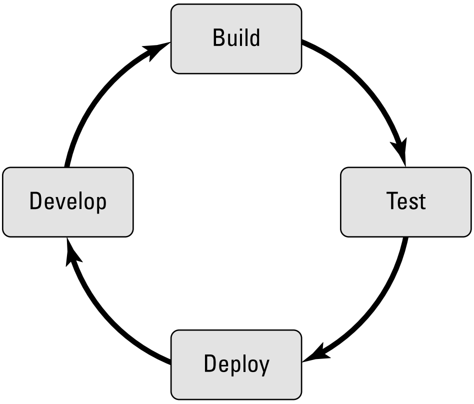

FIGURE 3-3: Developing software through an SDLC.

During the development stage, an attacker could change the source code itself if they gain access to the code repository in which the source is stored or can escalate privileges elsewhere in the early development process, such as gaining access to a developer machine.

When code reaches the build stage, which it may after every commit to a certain branch, an attacker has another opportunity to inject malicious code or otherwise alter the code to affect integrity. An attacker may leverage their access to cause the build process to use a compromised version of a compile library. This type of attack may not be noticed by code analysis tools looking solely at the source code itself.

During testing, an attacker can cause certain tests to not run or to appear to have run successfully, thus hiding the change that was made to inject a malicious binary program into the build. The test system or systems themselves represent another attack surface to be potentially exploited. Notably, test systems may not be secured in the same way as a production system thus enabling an attacker to successfully exploit relaxed network or authentication controls on that system.

Finally, an attacker that can place themselves in the deployment process can also inject a malicious binary at production deployment time.

Vulnerabilities can also exist in a nontargeted manner as well. Failure to update frameworks or other software involved in the development or build process can leave the resulting software vulnerable to exploits at later stages. Downstream dependencies on vulnerable libraries was noted as a particular problem with the log4j issue in 2021.

Leveraging Google Cloud security in the software supply chain

The entire software supply chain needs to be secured so that a chain of trust is established. Google Cloud provides several tools that help facilitate end-to-end secure software delivery:

- **Open Source Insights:** Public dataset based on dependencies within many popular open source software packages.
- **Cloud Build:** Build tools that can be integrated as part of the chain of trust.
- **Artifact Registry:** Manage container images and packages and automatically detect vulnerabilities in container images.
- **Binary Authorization:** Binary Authorization provides oversight throughout the build process.

It would be reasonable to assume that your code might not be vulnerable to any particular issue and to discover that a library included in your code also depends on an upstream library, which also depends on another library. It's that upstream analysis where vulnerabilities may be found. To help solve this problem at the code level, Google utilizes the Open Source Insights dataset to identify security-related issues within dependencies. The Open Source Insights dataset helps to create a dependency graph of your code that uses open source software but then carries the analysis a step further by also analyzing those packages for their dependencies.

Beyond security vulnerabilities, Open Source Insights also examines the licenses of the open source software found in the dataset. This feature can help keep an organization from incurring liability for unlicensed usage, abide by the terms of the licenses, and not be encumbered by license terms that can't be met.

The chain of trust begins at development time. As development transitions to build, Cloud Build is used to provide the next link in the chain. Cloud Build integrates natively with Google Cloud

Source repositories and with popular source code repository providers like GitHub.

Cloud Build obtains source code from Cloud Code or another repository source and builds the software as specified, producing one or more artifacts. Those artifacts are then analyzed with Artifact Analysis and eventually go to Artifact Registry. This process is depicted in Figure 3-4.

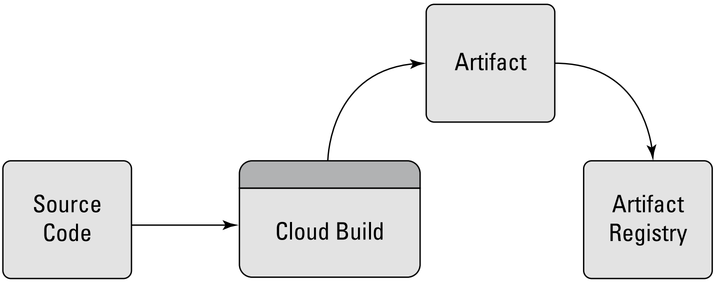

FIGURE 3-4: Establishing a chain of trust from source code to artifact creation.

Artifact Analysis enables decisions to be made on next steps for deployment based on metadata found within the artifact configuration. Artifact Analysis can perform scans within containers, looking at operating system level vulnerabilities within a container as well as Go, Java, Python, and node.js packages.

Artifact Registry is managed storage for container images and packages related to your deployment. From Artifact Registry, deployment to Google Kubernetes Engine, Cloud Run, Compute Engine, and App Engine become possible and fully integrated parts of the Continuous Integration/Continuous Deployment (CI/CD) regimen.

Cloud Deploy provides automation of the deployment architecture and just prior to deployment, Binary Authorization acts as a final gatekeeper to provide digitally signed attestations of the images created as part of the CI/CD process in Google Cloud.

TIP

You may be more familiar with another name for attestations: digital signing. Creating an attestation is also known as signing an image.

Binary Authorization uses policies to determine whether a container image is authorized to be deployed to production. As part of this process, rules are applied based on organizational need through policies. For organizations moving towards more automated, cloud-native processes, Binary Authorization can be configured to work in a number of modes including a dry-run mode where container images are not prevented from deployment.

TECHNICAL STUFF

The Kritis specification, part of Grafeas, is the foundation on which Binary Authorization is built. More information on Kritis can be found at `https://github.com/grafeas/kritis/blob/master/docs/binary-authorization.md`.

Providing Data Security

Envision a scenario where a supersecret account exists and can be accessed through a specially crafted GET request or by adding admin=1 to the address bar. Assuming no one ever finds out about the admin=1 trick, the application is secure. But once someone finds out, then this bit of security-through-obscurity is forever compromised.

Securing your data is a shared responsibility between the customer and Google Cloud. Google Cloud will provide security at the physical level, making sure that no one walks away with the server that contains your customer database. But securing the application itself is the responsibility of the data owner.

Google Cloud provides tools to encrypt data while in-transit and at-rest. Isolation-related technologies like a VPC prevent many types of attacks from occurring as well. Even with a secure-by-default posture, vulnerabilities, whether intended or unintended, that are written directly into the application can effectively undo even the best protections.

Like the scenario at the beginning of the section, proper management of secrets and private keys is the responsibility of the customer. Storing sensitive data in publicly accessible storage locations is something that Google Cloud would allow, of course not knowing that the data is sensitive. The customer has the responsibility for securing data whether that data is hosted on-premises or in the cloud.

Ensuring that data backups are completed and securely stored is also the responsibility of the customer. Google Cloud provides tools for backing up data. Ensuring redundancy at the data level is a basic requirement regardless of where that data is physically located.

Google offers a service aimed toward the enterprise customer called Sensitive Data Protection. Sensitive Data Protection includes the former Data Loss Prevention and contains components to discover and inspect potentially sensitive data. Protections such as automatically de-identifying the data stored in Google Cloud Storage, redaction of sensitive data from chat logs, and other protections are available with Sensitive Data Protection.

Further protections are available through the Google Cloud Confidential Computing platform. Defined by the Confidential Computing Consortium, the Confidential Computing platform combines hardware elements found at the CPU level with confidential storage space to create a trusted execution

Data is encrypted by default while at rest when stored in Google Cloud. Google Cloud Key Management Service (KMS) provides finer grained control over how keys are used. By default, Google controls the keys but enables significant customization of key-related infrastructure depending on your needs, including the following:

- **Default encryption:** AES-256-based encryption generated and managed by Google.
- **Cloud KMS with software-based keys:** Keys generated by Google and you can encrypt and sign using a key under your control.
- **Cloud KMS with hardware-generated keys:** Also known as Cloud HSM, these keys are stored in Google's hardware security modules (HSMs).
- **Cloud KMS with key import:** Import keys that you have generated and that are managed outside of the Google Cloud platform.
- **Cloud KMS with external key manager (EKM):** Use keys that are stored outside of Google Cloud but only used by Google Cloud for encrypting data at rest.

Customer-managed encryption keys (CMEK) are keys that you create using tools in Google Cloud.

TIP

See `https://cloud.google.com/docs/security/key-management-deep-dive` for more information on Cloud KMS.

A different solution than Cloud KMS or CMEK is customer-supplied encryption keys (CSEK). CSEK can be used to encrypt data on Google Cloud Storage and with Compute Engine. Customer-supplied keys are not stored other than in transient memory but rather used through derived keys.

Chapter 4
Ten Strategies for CNAPP Success

When deploying application workloads to the cloud, there are numerous problems that organizations need to solve. As the number of moving parts increases, so does complexity. The problem initially seems quite difficult to solve as organizations shift left. This chapter provides the decision-maker and decision-influencer with ten key considerations when considering application security in the cloud.

Consolidate and Validate

Multicloud deployments often involve the use of different tools. Indeed, each cloud provider offers its own native tools, as is the case with Google Cloud and Security Command Center. However, it's best to opt for a multicloud platform, or Cloud-Native Application Protection Platform (CNAPP), which integrates with the native tools and gives a complete view of the security of all your workloads.

Establish a Chain of Trust

Establishing a chain of trust is the fundamental step to prevent untrusted code from being deployed to production. Using a CNAPP can prevent untrusted code from running across cloud providers. Managing the software bill of materials (SBOM) and ensuring integrity of an image as the last step prior to deployment is vital to ensuring integrity of the overall application as it is deployed to production.

Looking specifically at Google Cloud, tools like Cloud Build, Artifact Registry, and related services provide the end-to-end chain of trust for cloud applications that ends with Binary Authorization as a gatekeeper before being deployed. In the case of a multicloud environment, a CNAPP can provide a holistic approach to validate artifact integrity prior to their deployments. Establishing a continuous integration and continuous deployment (CI/CD) pipeline is a prerequisite to achieving this goal and a CNAPP will integrate well with the shift-left processes of CI/CD.

Use Cloud-First Security

Given enough time, you can adapt older technologies to work against simple cloud applications. But when placed into a zero-trust environment, legacy tools may be cumbersome to run. CNAPP is the current state-of-the-art solution for securing cloud-based and hybrid workloads. CNAPPs are built for cloud rather than merely providing plugins for the cloud.

Prefer Agentless

Needing to install an agent to perform scanning of workloads introduces operationalization challenges. Instead, a CNAPP that can deep-scan an entire environment like Google Cloud should be preferred over one that requires agent software to be installed. Agentless delivery also means that visibility is naturally increased. You no longer need to worry about whether the agent software is supported by a given platform or technology or if it has been deployed everywhere.

Reduce Noise

A security tool should help reduce unnecessary alerts and other noise and false positives. Doing so can be challenging in any environment and even more challenging in a multicloud scenario across cloud providers. A CNAPP should help by coordinating and consolidating to reduce noise and provide focus to security-related efforts. Personnel can then understand critical risks based on context and prioritize their remediation efforts on actual issues rather than notice-level warnings.

Reduce Compliance Effort

Technologies like Cloud Key Management, Cloud IDS, Secrets Manager, and BeyondCorp are just a handful of the tools in Google Cloud that can be used to verify and validate cloud-native applications. Coordinating all of these tools into coherent output is a characteristic of a successful CNAPP. A CNAPP should also assist with license verification and other compliance-related activities.

Decrease Remediation Efforts

When a security issue is found, a CNAPP should know how to remediate the issue immediately and without intervention. Providing a means to feed the notification backwards to DevOps helps relieve the burden of chasing security problems to their source. The shift-left mentality that incorporates automation should extend to the security tooling used within the deployment and across the organization.

Reduce Costs

Don’t continue to pay for siloed security tools that aren’t optimal and don’t work when connected to a cloud environment. On the contrary, deploying a CNAPP reduces complexity by consolidating tools. Overall, providing a unified platform improves the security posture by reducing complexity. It also reduces costs, as it eliminates the need for multiple tools.

Detect Cloud Threats

CNAPP can be the single platform used by internal teams, from developers to incident and response teams, to provide context for alerts. The platform can be shared across teams for their own use and help triage incidents for multicloud deployments. Integration with external cloud providers can also be further integrated internally with existing security information and event management (SIEM) deployments.

Sharing Responsibility with a CNAPP

Shared responsibility is a key trait of cloud deployments. Security tools that don't understand or account for shared responsibility simply won't work as well within and across cloud providers in the way that a cloud-native toolset does. But beyond the shared responsibility model, obtaining the extended attention and support for your specific needs is another strength of CNAPP deployment and a true differentiator enabling organizations to achieve next-level results.

See and secure your Google Cloud environment

Google provides significant and extensive architectural guidance for cloud deployments and playing nice with others. That architectural guidance is freely available but is also sometimes difficult to sort out. This book helps navigate some of the technologies available with Google Cloud with a specific focus on security-related technologies and how those offerings can be used to help secure cloud-based applications.

Inside…

- Google Cloud security
- Defense in depth
- What is Shared Fate?
- Keeping the software supply chain safe
- Strategies for CNAPP success

WIZ

Steve Suehring is a media consultant with years of experience in computer security at scale. Steve has written books on several computing-related subjects and is focused on securing crypto- and AI-related tech.

ISBN: 978-1-394-24585-7
Not For Resale

9 781394 245857

for dummies®
A Wiley Brand