info@cloud-spheres.com
cloud-spheres.com

Aegiron: Umsetzungsstand einer mehrschichtigen Sicherheitsarchitektur für mandantenfähiges Gameserver-Hosting

Technisches Whitepaper der Cloud Spheres UG (haftungsbeschränkt).

Zusammenfassung

Aegiron ist die Sicherheitsarchitektur der Cloud Spheres UG (haftungsbeschränkt) für das Hosting von Gameservern auf gemeinsam genutzter Infrastruktur. Sie verteilt die Schutzmaßnahmen auf acht Schichten in drei Zonen und stellt ein Fundament aus Datenträgerverschlüsselung darunter. Das Paper beschreibt den Umsetzungsstand im Quelltext zum 30. Oktober 2026: die Bedrohungen, gegen die die Architektur gebaut ist, zwölf Entwurfsinvarianten, die Bausteine jeder Schicht mit ihren Konstruktionsparametern und die Prüfungen, mit denen sich jede Invariante belegen lässt. Kapitel 11.3 nennt je Invariante, wie weit die Umsetzung reicht, und Kapitel 12.2 die Bausteine, die entworfen, aber noch zu bauen sind.

Im Mittelpunkt stehen vier Teile. Erstens die Isolation fremden Codes auf gemeinsam genutzten Spielknoten mit gehärteten Containern, einer Egress-Sperre je Server, einem eigenen unprivilegierten Konto je Server als Standard und einem Flutschutz auf dem Knoten. Zweitens die Autorisierung zwischen Diensten in der selbst entwickelten Bibliothek serviceauth-go. Drittens die Laufzeiterkennung mit Falco und die automatische Reaktion durch Falco Talon. Viertens ein Entwurfsbeitrag, der die Eindämmung nach der Stärke der Beweise abstuft: Verhalten, blockierte Probe, erfolgreicher Durchbruch und berührter Köder lösen unterschiedliche Reaktionen aus. Die Evidenzklassen sind im Node-Agent umgesetzt, und das Werkzeug für das Katalog-Replay, das belegen soll, welche Signale in keinem legitimen Lauf vorkommen, ist gebaut; die Läufe über den Katalog sind zu erheben. Das Paper ordnet außerdem offen ein, was Standard ist, was eigen entwickelt wurde und was davon einen eigenen Beitrag darstellt.

1 Einleitung

Cloud Spheres vermietet Gameserver, also Prozesse wie einen Minecraft-, Factorio- oder Satisfactory-Server, die mit festem Arbeitsspeicher und zugewiesenen Prozessorkernen auf gemeinsam genutzten Rechnern laufen. Mehrere Kunden teilen sich damit dieselbe physische und virtuelle Infrastruktur. Diese Eigenschaft wird in der Literatur als Mandantenfähigkeit bezeichnet und ist der Grund, warum ein einzelner erfolgreicher Angriff nicht bei einem Kunden enden darf.

Gameserver sind dabei ein besonders unfreundlicher Fall. Kunden installieren Mods, Plugins und Skripte, die auf dem Server beliebigen Code ausführen, und sie tun das mit ausdrücklicher Absicht des Produkts. Die Architektur geht deshalb davon aus, dass im Prozess eines Kundenservers jederzeit fremder, ungeprüfter oder bösartiger Code läuft, und sie ist so gebaut, dass dieser Code weder auf Nachbarn noch auf die Steuerungsebene der Plattform durchgreifen kann. Hinzu kommt, dass Spielverkehr empfindlich auf Latenz reagiert. Eine Schutzmaßnahme, die jedem Paket zusätzliche Verzögerung hinzufügt, ist für dieses Produkt keine.

Die Praxisarbeit vom Januar 2026 hat für diese Aufgabe ein Modell aus acht Schichten vorgeschlagen. Dieses Paper führt das Modell weiter und ersetzt es. Es ergänzt einen Edge-Proxy für Spielverkehr, eine Zulassungskontrolle im Cluster, ein Overlay-Netz mit versionierten Zugriffsregeln, die Anmeldung der Spielknoten über einen Hub, Feldverschlüsselung für Nutzerdaten und die Abstufung der Eindämmung nach Beweisstärke.

Cloud Spheres UG (haftungsbeschränkt)
Registered Office: Paderborn
Address: Biberweg 24, 33102 Paderborn
Registration Court: Amtsgericht Paderborn
Commercial Register Number: HRB 18602
Managing Directors: Silas Timmermann, Alan Khudhur

Andere Bausteine des Entwurfs sind entfallen oder ersetzt worden, etwa ein Ingress-Controller mit Client-Zertifikaten und eine clusterweite Sperrlogik nach dem Muster von Fail2Ban.

Das Paper ist wie folgt aufgebaut. Kapitel 2 legt den Geltungsbereich fest. Kapitel 3 beschreibt System und Bedrohungsmodell, Kapitel 4 die Invarianten und Kapitel 5 die Schichten im Überblick. Die Kapitel 6 bis 8 behandeln die drei Zonen im Detail. Kapitel 9 stellt den Entwurfsbeitrag vor, Kapitel 10 ordnet Standard und Eigenentwicklung ein, Kapitel 11 beschreibt die Prüfungen, und Kapitel 12 nennt die Grenzen der Architektur und den Ausbaupfad.

2 Geltungsbereich und Lesehinweise

Das Paper beschreibt Aegiron im Umsetzungsstand des Quelltexts vom 30. Oktober 2026, also die Software, die Konfigurationen und die Auslieferungsdefinitionen in den Repositories von Cloud Spheres, so wie sie gebaut, getestet und versioniert sind. Aussagen im Präsens beziehen sich auf diesen Stand. Bausteine, die standardmäßig abgeschaltet ausgeliefert werden, nennt der Text als zuschaltbar. Das Paper trifft keine Aussage darüber, welche Bausteine in einer bestimmten Umgebung zu einem bestimmten Zeitpunkt eingeschaltet sind. Bausteine, die Entscheidungen automatisch durchsetzen, sind so gebaut, dass sie zuerst als reine Meldung oder im Schattenmodus laufen können, in dem sie Entscheidungen protokollieren, ohne zu handeln, und erst danach erzwingend. Wer aus diesem Dokument eine Eigenschaft eines laufenden Systems ableiten will, muss den Betriebsstand der jeweiligen Umgebung prüfen. Bausteine, die entworfen, aber im Quelltext nicht vorhanden sind, stehen nicht in den Kapiteln 6 bis 9, sondern in Kapitel 12.2.

Die genannten Parameter, also Größen von Zwischenspeichern, Fristen, Schwellen und Lebensdauern, sind die Konstruktionswerte der Komponenten und stammen aus deren Quelltext und Konfiguration. Sie sind keine Leistungsaussagen. Als Messung sind ausschließlich der Mikro-Benchmark der Bibliothek serviceauth-go in Kapitel 7.3.5 und die Latenzmessung des Edge-Proxys in Kapitel 6.4 gekennzeichnet. Beide stammen aus Testläufen und sagen nichts über den Betrieb aus. Wo eine Zeit oder ein Durchsatz nicht gemessen ist, nennt das Paper keine Zahl, sondern beschreibt, wie sie gemessen wird.

Um die Angriffsfläche nicht zu vergrößern, veröffentlichen wir keine Zugangsdaten, keine Geheimnispfade, keine Adressen oder Namen einzelner Rechner, keine Portnummern interner Dienste und keine Namen von Kunden. Mechanismen werden beschrieben, ihre Adressen nicht.

3 System- und Bedrohungsmodell

3.1 Systemmodell

Die Plattform besteht aus fünf Kubernetes-Clustern auf Basis von k3s, die nach Aufgaben getrennt sind: ein Cluster für Orchestrierung und Betriebswerkzeuge, einer für die Website, einer für die Microservices, einer für die Datenbanken und einer für die Sicherung der Datenbanken. Staging und Produktion werden von getrennten ArgoCD-Instanzen verwaltet. Die Cluster sind über ein Tailscale-Overlay-Netz verbunden, und das Pod-Netz nutzt flannel mit VXLAN über die Tailscale-Schnittstelle.

Die Gameserver der Kunden laufen nicht in diesen Clustern. Sie laufen auf gemieteten virtuellen Servern, die wir Spielknoten nennen und die von einem eigenen Dienst, dem Node-Agent, verwaltet werden. Der Node-Agent startet jeden Gameserver entweder als gehärteten Container oder als Prozess auf dem Knoten, für den standardmäßig ein eigenes unprivilegiertes Konto angelegt wird, und meldet sich bei einem Hub, dem Connection-Gateway, das ihn

Cloud Spheres UG (haftungsbeschränkt)
Registered Office: Paderborn
Address: Biberweg 24, 33102 Paderborn
Registration Court: Amtsgericht Paderborn
Commercial Register Number: HRB 18602
Managing Directors: Silas Timmermann, Alan Khudhur

per gegenseitig authentifiziertem TLS anbindet. Diese Trennung ist für die Sicherheit wesentlich: Die Kubernetes-Mechanismen wie NetworkPolicies und Zulassungsregeln schützen die Steuerungsebene, während die Isolation der Kundenworkloads auf den Spielknoten selbst hergestellt wird.

Die Steuerungsebene besteht aus mehr als zwanzig in Go geschriebenen Microservices, die intern über gRPC kommunizieren (mit wenigen HTTP-Ausnahmen), einer Website auf Basis von Next.js und den Infrastrukturkomponenten, die in den folgenden Kapiteln vorkommen. Ein Dienst namens tenant-service verwaltet Marken und Märkte, ist also ein Register der geschäftlichen Einheiten und keine Sicherheitsgrenze zwischen Kunden. Wenn in diesem Paper von der Isolation zwischen Mandanten die Rede ist, sind die Kunden gemeint, deren Server sich einen Spielknoten teilen.

3.2 Schutzziele und Angreifer

Die Architektur verfolgt vier Schutzziele. Personenbezogene und zahlungsnahe Daten der Kunden dürfen nicht offengelegt werden. Die Steuerungsebene, die Server anlegt, Rechnungen erzeugt und Zugangsdaten verwaltet, muss in ihrer Integrität geschützt bleiben. Laufende Gameserver sollen verfügbar bleiben, auch wenn ein Nachbar angegriffen wird oder sich fehlverhält. Und Sicherheitsrelevantes muss nachvollziehbar sein, damit sich ein Vorfall rekonstruieren lässt.

Es werden sechs Klassen von Angreifern betrachtet. Der erste greift von außen die Website, die Anmeldung und die Schnittstellen an, etwa durch das Durchprobieren von Zugangsdaten oder durch Missbrauch teurer Anfragen. Der zweite adressiert den Spielverkehr, also die öffentlich erreichbaren Spielports, durch Verbindungsfluten oder manipulierte Handshakes. Der dritte ist ein Kunde, oder ein Angreifer im Namen eines Kunden, der Code auf seinem eigenen Gameserver ausführt und aus dessen Umgebung auszubrechen versucht, Nachbarn angreift, das Netz scannt oder Rechenzeit für Kryptominer missbraucht. Der vierte ist ein kompromittierter interner Dienst, der sich seitlich im Cluster bewegen will. Der fünfte betrifft die Lieferkette, also manipulierte Container-Images, Abhängigkeiten oder Zugänge der Build-Pipeline. Die sechste Klasse bilden Innentäter mit legitimem Zugang.

Ausdrücklich nicht betrachtet werden Angriffe des Infrastrukturanbieters auf die virtuellen Maschinen, physische Angriffe auf Rechenzentren und volumetrische Angriffe oberhalb der Kapazität des vorgelagerten Netzes.

4 Invarianten

Die Architektur ist aus zwölf Invarianten entworfen, also aus Aussagen, die für das System gelten sollen. Aus ihnen folgen die Maßnahmen je Schicht. Jede Invariante hat eine Prüfung, mit der sie belegt wird, und eine Abweichung ist als Alarm gedacht und nicht als Befund für die nächste Durchsicht. Wie weit Umsetzung und Automatisierung der Prüfungen je Invariante reichen, steht in Kapitel 11.3.

Nr.	Invariante	Prüfung
I1	Jeder Verkehr zwischen Workloads ist authentifiziert und für die konkrete Methode autorisiert. Netzwerkposition und Dienstname sind kein Ausweis.	Mesh mit Default-Deny und namentlichen Aufrufern, Autorisierung in der Anwendungsschicht, täglicher Test mit einem nicht berechtigten Aufrufer

Cloud Spheres UG (haftungsbeschränkt)
Registered Office: Paderborn
Address: Biberweg 24, 33102 Paderborn
Registration Court: Amtsgericht Paderborn
Commercial Register Number: HRB 18602
Managing Directors: Silas Timmermann, Alan Khudhur

I2	Nichts kommuniziert ohne ausdrückliche Freigabe. Das gilt für Namespaces, Knoten und Container, für eingehenden und ausgehenden Verkehr.	Erzeugte Default-Deny-Richtlinie je Namespace, CI-Test, der einen Namespace ohne Richtlinie ablehnt, Flussprotokolle mit Abgleich gegen die Freigaben
I3	Es läuft nur Code, dessen Herkunft belegt ist. Jedes Image ist an einen Commit und an die Identität der Pipeline gebunden, die es gebaut hat.	Signaturprüfung bei der Zulassung, feste Digests, Herkunftsnachweis je Image, Ablehnung von Images ohne Nachweis
I4	Kundencode läuft nie mit administrativen Rechten, nie mit Zugriff auf Schlüssel der Steuerungsebene und immer hinter zwei voneinander unabhängigen Isolationsgrenzen.	Automatische Knotenprüfung nach der Bereitstellung und laufend, Angriffssimulation gegen die Grenzen in Staging
I5	Kein Datum ist im Ruhezustand unverschlüsselt, auch nicht in Sicherungen und auf Spielknoten.	Prüfung des Verschlüsselungsstatus aller Datenträger und Sicherungen, Aufnahme in den Pool nur bei bestandener Prüfung
I6	Alle Zugangsdaten sind kurzlebig, werden automatisch erneuert und lassen sich sofort widerrufen. Es gibt keine statischen Langzeitgeheimnisse.	Inventar aller Anmeldeverfahren mit Ablaufzeit, Alarm bei jedem Token ohne Ablauf
I7	Jede sicherheitsrelevante Entscheidung erzeugt einen Eintrag in einem Protokoll, dessen Manipulation nachweisbar ist.	Kryptografische Verkettung mit täglicher Verifikation, externer Anker, Schreibrecht nur zum Anhängen
I8	Jede Erkennung löst eine festgelegte Reaktion aus, und die Zeit von der Erkennung bis zur Wirkung ist gemessen und begrenzt.	Ende-zu-Ende-Messung als Histogramm, vierteljährliche Übungen mit echten Auslösern
I9	Jede Kontrolle ist wirksam und das lässt sich beweisen. Der Zustand im Repository und der Zustand im Cluster weichen nicht ab.	Kontrolltests je Regel, täglicher Abgleich der ausgelieferten Objekte mit dem Repository
I10	Der Ausfall einer Kontrolle hat ein bekanntes Verhalten. Kontrollen, die Geld, Schlüssel oder Rechte betreffen, schließen bei einem Ausfall geschlossen. Wo aus Verfügbarkeitsgründen offen gefallen wird, ist die Dauer begrenzt und ein Alarm ausgelöst.	Fehlerinjektion je Kontrolle mit erwartetem Ergebnis, Alarm bei jedem Rückfall
I11	Der Ausfall oder die Kompromittierung einer Schicht führt nicht zum Ausfall einer zweiten. Schichten teilen weder Schlüssel noch Konfigurationsquelle noch Administratorzugang.	Abhängigkeitsanalyse der Kontrollen, Test mit gezielt ausgeschalteter Schicht

Cloud Spheres UG (haftungsbeschränkt)
Registered Office: Paderborn
Address: Biberweg 24, 33102 Paderborn
Registration Court: Amtsgericht Paderborn
Commercial Register Number: HRB 18602
Managing Directors: Silas Timmermann, Alan Khudhur

I12	Keine einzelne Person kann Schlüsselmaterial oder Produktionsdaten freigeben oder eine Sicherheitsrichtlinie abschalten.	Aufgeteilte Schlüsselanteile, Vier-Augen-Freigabe, Alarm bei jeder Änderung von Richtlinien

5 Architekturübersicht

Das Modell gliedert sich in drei Zonen. Zone A regelt, wer überhaupt in das System gelangt und wohin sich Verkehr bewegen darf. Zone B regelt, was innerhalb des Systems erlaubt ist, und umfasst die Identität und Autorisierung von Diensten sowie die Verwaltung von Geheimnissen. Zone C liegt quer über den anderen und sorgt dafür, dass Verhalten erkannt, sichtbar gemacht und automatisch begrenzt wird. Das Fundament L0 bildet die Verschlüsselung der Datenträger der Spielknoten. Zusätzlich gibt es eine Zulassungskontrolle im Cluster, die keiner einzelnen Schicht zugeordnet ist und im Kapitel zu Zone B beschrieben wird.

Schicht	Aufgabe	Bausteine
L0	Daten im Ruhezustand	LUKS für den Datenträger der Spielknoten, Schlüsselhinterlegung über das Gateway in Vault
L1	Netzwerksegmentierung und Isolation	NetworkPolicies, Tailscale-Zugriffsregeln als Code, Host-Firewall mit Flutschutz, gehärtete Container mit Egress-Sperre, unprivilegiertes Konto je Server als Standard
L2	Kante und Transport	Cloudflare und Tunnel für die Website, eigener Edge-Proxy für Spielverkehr, cert-manager, gegenseitiges TLS für Knoten
L3	Service Mesh	Linkerd mit gegenseitigem TLS und kurzlebigen Zertifikaten
L4	Autorisierung zwischen Diensten	Eigenentwicklung serviceauth-go mit Policies aus Vault
L5	Geheimnisse	HashiCorp Vault, Vault Secrets Operator, Transit-Verschlüsselung für Nutzerdaten
L6	Laufzeitsicherheit	Falco auf Clustern und Spielknoten, Falco Talon, evidenzgestufte Eindämmung auf dem Knoten
L7	Beobachtbarkeit	Prometheus, Loki, Tempo und Grafana
L8	Alarmierung und Gegenmaßnahmen	Alertmanager mit Chat-Routen und einem zuschaltbaren Pager-Empfänger, Kontosperren, Ratenbegrenzung, gestufte Sperren
Querschnitt	Zulassungskontrolle	Kyverno mit durchsetzenden und prüfenden Richtlinien

Cloud Spheres UG (haftungsbeschränkt)
Registered Office: Paderborn
Address: Biberweg 24, 33102 Paderborn
Registration Court: Amtsgericht Paderborn
Commercial Register Number: HRB 18602
Managing Directors: Silas Timmermann, Alan Khudhur

6 Zone A: Zugang und Segmentierung

6.1 Segmentierung im Cluster

Als Netzwerk-Plugin dient flannel, die Durchsetzung von NetworkPolicies übernimmt der in k3s eingebaute kube-router. Jeder Dienst bringt eine eigene NetworkPolicy mit, die die eigenen Pods auswählt, Ein- und Ausgang auf das Notwendige beschränkt und in der Regel den Linkerd-Port für Mesh-Verkehr freigibt. Ergänzend kann Kyverno für Namespaces, die ein entsprechendes Label tragen, automatisch eine Richtlinie erzeugen, die Ein- und Ausgang sperrt und nur die Namensauflösung und die Steuerungsebene des Mesh freigibt, sodass ein Pod ohne eigene Freigabe isoliert ist.

Eine technische Besonderheit wirkt sich auf die Formulierung der Regeln aus. kube-router bewertet ausgehende Regeln nach der Adressumsetzung, sodass Quellen, die über einen NodePort eintreffen, in Pod-Richtlinien nicht immer unter ihrer ursprünglichen Adresse erscheinen. Zugriffe von außerhalb des Clusters werden deshalb zusätzlich über die Zugriffsregeln des Overlay-Netzes und über gezielte Regeln der Host-Firewall für einzelne Ports gesteuert, nicht über die Richtlinien der Pods.

6.2 Overlay-Netz und Zugriffsregeln

Die Kommunikation zwischen den Clustern und der Zugang für Administratoren laufen über ein Tailscale-Netz. Die Zugriffsregeln sind als Datei beschrieben. Ausgangszustand ist Default-Deny. Rollen werden über Markierungen der Knoten abgebildet, und die Freigaben zwischen den Clustern nennen überwiegend konkrete Ziele und Ports, etwa von Website-Knoten zur Registry oder von den Clustern zum Speicher für Traces. Die Datei enthält außerdem Tests, die erwartete Zugriffe und erwartete Sperren als Aussagen festhalten, sodass eine unbeabsichtigte Öffnung beim Anwenden auffällt. Die Datei enthält auch die Regeln für SSH. Für ausgewählte Ziele und Benutzer verlangen sie eine periodische erneute Bestätigung.

6.3 Der Weg der Website

Der vorgesehene öffentliche Weg der Website führt von Cloudflare über einen Tunnel zum Dienst im Website-Cluster. Einen Ingress-Controller gibt es nicht, und die Ingress-Ressource des Website-Charts ist in den Umgebungswerten ausgeschaltet. TLS wird bei Cloudflare beendet. Regeln der Web Application Firewall, Ratenregeln und die Bot-Prüfung sind in einem Terraform-Modul im Repository beschrieben, zusammen mit einem Skript, das den Stand bei Cloudflare gegen den Code abgleicht. Cloudflare Turnstile ist zuschaltbar und schützt das Anfordern eines Passwort-Resets und das Supportformular. Die Prüfung schlägt bei Ablehnung oder Nichterreichbarkeit fehl, und eine Einstellung macht die Schlüssel zur Pflicht: Ohne sie startet die Website nicht, statt die Prüfung zu überspringen.

Innerhalb der Website wirkt eine zweistufige Ratenbegrenzung. Eine Schicht in der Middleware begrenzt pro IP-Adresse im Speicher, mit strengeren Grenzen für Zahlungswege und Anmeldung als für allgemeine API-Anfragen. Fällt die Begrenzung selbst aus, antwortet sie mit einem Fehler, anstatt Anfragen ungeprüft durchzulassen. Eine zweite Schicht in den Routen zählt mit einem Fenster in Redis, mit eigenen Grenzen für Anmeldeversuche und Zahlungsanfragen, und fällt bei einem Ausfall von Redis auf einen lokalen Zähler zurück. Für kontobezogene Grenzen dient ein Ausschnitt des Sitzungstokens als Schlüssel. Die Client-Adresse wird vor der Verwendung als Schlüssel auf eine gültige Form und eine Höchstlänge geprüft, und IPv6-Adressen zählen je Präfix von 64 Bit.

Cloud Spheres UG (haftungsbeschränkt)
Registered Office: Paderborn
Address: Biberweg 24, 33102 Paderborn
Registration Court: Amtsgericht Paderborn
Commercial Register Number: HRB 18602
Managing Directors: Silas Timmermann, Alan Khudhur

Strukturierte Daten im Format JSON-LD, die die Seiten einbetten, werden so serialisiert, dass Zeichen, die ein Skriptelement beenden oder einen Kommentar beginnen könnten, als Unicode-Escapes geschrieben werden. Ein Test deckt diesen Fall ab.

6.4 Spielverkehr: der Edge-Proxy

Spielverkehr nimmt einen anderen Weg als Webverkehr. Er läuft nicht über Cloudflare. Ein eigener Edge-Proxy, geschrieben in Go, läuft als DaemonSet mit Host-Netzwerk auf ausgewählten Knoten des Website-Clusters und nimmt Spielverkehr entgegen. Der Name der Spieladresse ist ein Wildcard-Eintrag, der auf die öffentlichen Adressen der Edge-Knoten zeigt. Ein eigener Wächter, dns-guard, gleicht diese Einträge in einem konfigurierbaren Takt mit den tatsächlich dienstbereiten Knoten ab und entfernt tote Knoten aus der Namensauflösung, ohne je eine leere Menge zu schreiben. Das dient der Verfügbarkeit.

Der Edge-Proxy kennt zwei Weiterleitungsarten. Im Portmodus lauscht er pro Server auf einem eigenen Port. Im Hostnamenmodus lauscht er auf dem Standardport von Minecraft Java und liest den Handshake des Protokolls, um den Zielserver anhand des angefragten Hostnamens zu bestimmen. Der Parser ist eng begrenzt: Er erwartet den Handshake innerhalb einer konfigurierbaren Frist, begrenzt die Größe des Pakets, die Länge des Hostnamens und die Länge der variabel codierten Ganzzahlen und bricht bei jeder Abweichung ab. Der Parser und weitere Dekodierer für fremde Bytes haben je ein Fuzz-Ziel von Go, insgesamt dreizehn, darunter die Weiterleitung von QUIC-Verbindungen anhand des Servernamens, das Sperrlistendokument und die Schlüssel von Ratenbegrenzung und Zulassungstoren. Die CI führt sie bei jeder Pipeline mit einem kleinen Zeitbudget aus. Die Zuordnung von Servern zu Zielen erhält der Proxy laufend vom Orchestrator über eine gRPC-Verbindung, die sich mit gegenseitigem TLS betreiben lässt. Auf TCP-Verbindungen kann der Proxy den Client mit dem PROXY-Protokoll in der Version 2 an den Zielserver melden. Für Datenbank-Ziele ist das abgeschaltet, weil ein zusätzlicher Vorspann Protokolle stört, bei denen der Server zuerst spricht.

Der Proxy begrenzt Verbindungen an mehreren Stellen. Für neue Verbindungen und neue UDP-Sitzungen gilt pro Quelladresse ein Token-Bucket, für Spielziele und für Datenbankziele mit getrennten, konfigurierbaren Werten. Zweitens gilt pro Ziel eine Obergrenze gleichzeitiger Verbindungen, die für Datenbanken deutlich niedriger liegt als für Spielziele. Die Zähler laufen für beide Zielarten getrennt, sodass eine Beitrittswelle bei Spielen den Zugriff auf eine Datenbank nicht verdrängt. Überschreitungen führen zum Verwerfen der Verbindung und werden protokolliert. Die Metrikschnittstelle ist auf interne Adressbereiche beschränkt, und keine Metrik trägt eine Quelladresse als Label.

Ein Satz zuschaltbarer Schutzmaßnahmen ergänzt diese Grenzen gegen Verbindungsfluten und Missbrauch. Alle Schalter sind standardmäßig abgeschaltet, und mit allen Schaltern aus bleibt der Datenpfad der ohne die Erweiterung. Der Konfigurationslader bricht den Start bei widersprüchlichen Werten ab. Die folgende Tabelle ordnet jeder Bedrohung ihre Maßnahme zu.

Bedrohung	Maßnahme
Viele Quelladressen, damit der Limiter einen Verursacher vergisst, auch durch rotierende Adressen innerhalb eines IPv6-Präfixes von 64 Bit	Verdrängungsfeste Limiter-Tabelle fester Größe statt eines LRU-Speichers, IPv6 je Präfix von 64 Bit
Viele gleichzeitige Verbindungen einer Quelle oder insgesamt	Zulassungstore mit Obergrenze insgesamt, Grenze je Quelle und einer weichen Schwelle, oberhalb der nur bereits bekannte Quellen zugelassen werden

Cloud Spheres UG (haftungsbeschränkt)
Registered Office: Paderborn
Address: Biberweg 24, 33102 Paderborn
Registration Court: Amtsgericht Paderborn
Commercial Register Number: HRB 18602
Managing Directors: Silas Timmermann, Alan Khudhur

Langsame und halboffene Handshakes auf dem gemeinsamen Port	Absolute Frist für den gesamten Handshake, harte Größengrenze, begrenzte Zahl wartender Handshakes insgesamt und je Quelle
Unlesbare oder fremde Protokolle, Handshakes für Hostnamen, die niemand besitzt	Sperre der Quelle nach wiederholten unlesbaren Handshakes, Zeitüberschreitungen oder unbekannten Hostnamen
Verursacher, die nach Ablauf einer Sperre wiederkehren	Eskalationsleiter: jede Wiederholung dauert länger, bis zu einem Höchstwert, bei begrenztem Speicher
Fluten neuer UDP-Sitzungen	Grenzen für Sitzungen insgesamt und je Quelle, bereits etablierte Quellen behalten ihre Sitzungen
Ein Server verbraucht die Bandbreite des Randes	Bandbreitengrenze je Server und Richtung
Fehlerhafte Eingaben an Parsern	Begrenzte Parser und Fuzz-Ziele in der CI, immer aktiv

Die Limiter-Tabelle nutzt aus, dass ein wieder aufgefüllter Token-Bucket keine Information trägt: Ihn zu vergessen kostet nichts, einen noch begrenzten zu vergessen schon. Ein Neuling erhält einen freien oder einen aufgefüllten Platz, nie den Platz einer begrenzten Quelle. Findet er keinen, teilt er sich mit anderen unbekannten Quellen ein gemeinsames Überlaufbudget fester Größe. Neulinge, die dieses Budget abweist, sammeln keine Verstöße, damit eine Flut keine unbeteiligten Spieler in die Sperre treibt. Die Zulassungstore entscheiden in der Annahmeschleife, bevor eine Verbindung eine Goroutine, eine Verbindung zum Ziel oder eine Protokollzeile kostet. Die Reihenfolge, in der Last abgeworfen wird, ist festgelegt: zuerst eine Quelle über ihrer eigenen Grenze, dann Quellen, die der Proxy zuvor nicht bedient hat, zuletzt alle. Bestehende Verbindungen werden nicht angetastet.

Die Adressen der eigenen Infrastruktur sind von den Toren für TCP ausgenommen, weil sich eine TCP-Adresse nicht fälschen lässt. Für UDP gilt keine Ausnahme, weil sich eine UDP-Absenderadresse fälschen lässt. Aus demselben Grund zählen Verstöße über UDP nur auf ausdrückliche Einstellung für Sperren, und solche Sperren sind in ihrer Dauer nach oben begrenzt, damit ein Dritter mit gefälschter Absenderadresse niemanden lange aussperren kann. Sperren gelten je Edge-Knoten, knotenübergreifend wirkt nur die Sperrliste. Zuschaltbare Alarmregeln melden eine hohe Sperrrate, Wiederholungstäter, abgeworfene Verbindungen, die Annäherung an die Obergrenzen, gehäufte Handshake-Fehler und die Sättigung des Limiters; sie werden mit promtool geprüft.

Gegen Angriffe, die die Leitung des Edge-Knotens oder das vorgelagerte Netz füllen, bevor ein Paket den Prozess erreicht, hilft keine Software auf dem Edge-Knoten (Kapitel 12.1). Eine Verbindung, die den Proxy umgeht und einen Spielport unmittelbar erreicht, unterliegt den Grenzen des Proxys nicht. Dafür dient die optionale Beschränkung der Spielports auf die Edge-Adressen in Kapitel 6.5.

Der Aufschlag des Proxys ist gemessen, und zwar mit einem Go-Benchmark in der CI auf einer virtuellen Linux-Maschine mit Intel Xeon: drei Läufe zu je 2000 Durchgängen, ein Echo-Ziel auf derselben Maschine, Loopback, angegeben sind Spannen. Gegenüber einer direkten Verbindung zum Ziel kommen etwa 105 bis 115 Mikrosekunden je neuer Verbindung und etwa 12 bis 14 Mikrosekunden je Round Trip einer bestehenden Verbindung hinzu (64 Byte). Der Aufschlag je neuer Verbindung entsteht vor allem durch die zusätzliche Verbindung, die der Proxy zum Ziel aufbaut. Eine Entscheidung des Limiters kostet 117 bis 132 Nanosekunden, der Durchgang durch ein Zulassungstor 147 bis 155 Nanosekunden. Beides fällt einmal je neuer Verbindung an und nicht je Paket. Mit allen Schutzmaßnahmen zusammen ist der Verbindungsaufbau im Rahmen der Streuung nicht messbar langsamer. Beim Weiterleiten liegt der Unterschied in der Streuung der gemeinsam genutzten CI-Maschine, ein kleiner Aufwand je Paket ist dadurch nicht ausgeschlossen. Die Messung sagt nichts über die Latenz

Cloud Spheres UG (haftungsbeschränkt)
Registered Office: Paderborn
Address: Biberweg 24, 33102 Paderborn
Registration Court: Amtsgericht Paderborn
Commercial Register Number: HRB 18602
Managing Directors: Silas Timmermann, Alan Khudhur

zwischen Spieler und Server und nichts über den Betrieb aus.

6.5 Isolation der Kundenworkloads auf den Spielknoten

Die Trennung zwischen Kunden, deren Server auf demselben Knoten laufen, ist die Sicherheitsaufgabe mit den größten Folgen und liegt im Node-Agent und in der Bereitstellung des Knotens. Der Node-Agent kennt zwei Startarten für Gameserver. Im Containermodus läuft jeder Gameserver in einem Container mit einem festen Satz von Einschränkungen. Alle Linux-Capabilities werden entfernt, und die Ausweitung von Rechten durch neue Prozesse ist untersagt. Ein eigenes Seccomp-Profil arbeitet als Positivliste, die Standardaktion ist also die Ablehnung, und es lehnt zusätzlich 53 ausdrücklich benannte Systemaufrufe ab, darunter mount, ptrace, bpf, unshare, setns, io_uring und die Ladefunktionen für Kernelmodule. Das Wurzeldateisystem ist schreibgeschützt, und die beschreibbaren temporären Verzeichnisse sind als nicht ausführbar eingehängt. Der Prozess läuft unter einer unprivilegierten Benutzerkennung, Arbeitsspeicher und Swap sind gleich gesetzt, die Zahl der Prozesse ist begrenzt, der IPC-Namensraum ist privat, und die Rechenzeit wird über Anteile in einer eigenen cgroup-Gruppe verteilt. Der Docker-Daemon läuft mit der Neuzuordnung von Benutzerkennungen, und der Node-Agent startet im Containermodus nicht, wenn weder diese Einstellung noch ein Daemon ohne Administratorrechte (rootless) vorliegt.

In der zweiten Startart läuft der Spielprozess auf dem Knoten. Im Hostmodus legt der Node-Agent standardmäßig ein eigenes unprivilegiertes Konto je Server an und legt die Rechte vor der Ausführung des Spielprogramms ab. Eine Ausnahme setzt eine ausdrückliche Eintragung in einer Liste voraus; ein unbekannter, leerer oder fehlerhafter Spieltyp bleibt unprivilegiert. Der Start eines Servers wird abgelehnt, wenn für ein Spiel, das unprivilegiert laufen muss, kein Konto zugewiesen werden kann, und ein Test stellt sicher, dass jede Installationsvorschrift eines Spiels eine ausdrückliche Einordnung trägt. Eine zuschaltbare Knotenprüfung kann die Spielprozesse eines Knotens auf administrative Rechte untersuchen und die Zahl als Metrik ausweisen. Kann sie die Prozesse nicht sehen, meldet sie diesen Zustand und nicht null. Die Erkennungsregeln aus Kapitel 8.1 erfassen die Container und die Prozesse dieser Konten. Eine zweite, vom Kernel des Knotens unabhängige Isolationsgrenze durch eine Sandbox-Laufzeit ist als abschaltbare Option je Knotengruppe angelegt, aber nicht der Standard, weil sie in einer Referenzmessung den Durchsatz deutlich senkte; dedizierte Knoten gehören zum Ausbaupfad (Kapitel 12.2).

Ausgehender Verkehr ist je Container standardmäßig gesperrt. Über Regeln der FORWARD-Kette, die an die Adresse des Containers gebunden sind, wird zunächst alles verworfen und danach nur freigegeben, was der Server laut seiner Liste erlaubter Ziele braucht, dazu Namensauflösung und bereits aufgebaute Verbindungen. Unabhängig von der Liste wird ausgehender Verkehr auf dem SMTP-Port 25 verworfen, ebenso der Zugriff auf die Adresse des Metadatendienstes der Cloud-Anbieter. Lässt sich die Egress-Regel nicht setzen, wird der Container wieder entfernt, damit nie ein ungebundener Container an den Aufrufer zurückgegeben wird. Die Host-Firewall lehnt eingehenden Verkehr standardmäßig ab und öffnet SSH, SFTP und die Ports der Server des Knotens in TCP und UDP, und im Containermodus wird die Verwaltungsschnittstelle der Spiele ausschließlich an die lokale Schnittstelle gebunden.

Für Mods und Modpacks erlaubt der Mod-Manager nur bekannte Herkunftsadressen für Modpack-Dateien, und zwar nur über HTTPS, verlangt für jeden Download eine Prüfsumme (SHA-256 oder SHA-1), bricht ohne sie ab und entpackt Archive gegen Pfadangriffe und Archivbomben abgesichert.

Gegen Verbindungsfluten auf die Spielports kann der Node-Agent einen Flutschutz auf dem Knoten einrichten. Er legt eine eigene Regelkette an, die je geschütztem Spielport ungültige Pakete und unsinnige TCP-Flag-Kombinationen verwirft, neue Verbindungen und neue UDP-Flüsse je Quelladresse und je Port begrenzt, die Länge

Cloud Spheres UG (haftungsbeschränkt)
Registered Office: Paderborn
Address: Biberweg 24, 33102 Paderborn
Registration Court: Amtsgericht Paderborn
Commercial Register Number: HRB 18602
Managing Directors: Silas Timmermann, Alan Khudhur

des ersten UDP-Pakets nach oben begrenzt und auf Wunsch Quellen für begrenzte Zeit sperrt. Ein Kernel-Profil ergänzt die Kette (SYN-Cookies, Rückwärtspfadprüfung, keine Quellrouten und Umleitungen, begrenzte ICMP-Raten, Größe der Verbindungsverfolgung). Es wird vor der ersten Änderung gesichert und beim Entfernen wiederhergestellt. Der Node-Agent prüft den Regelsatz vorab mit dem Parser von iptables, spielt ihn atomar ein, liest ihn zurück, kontrolliert ihn laufend auf Abweichungen und repariert diese. Ein Trockenlauf protokolliert den Regelsatz, ohne ihn anzuwenden, und ein Befehl entfernt alles auch ohne den Dienst. Von den Grenzen je Quelle sind die Adressen der Edge-Proxys ausgenommen, weil viele Spieler hinter wenigen Adressen stehen. Das Overlay-Netz wird über seine Schnittstelle erkannt und nie über einen Adressbereich, den ein Angreifer fälschen könnte.

Optional beschränkt der Node-Agent die Spielports auf die Edge-Proxys. Neue direkte Verbindungen zu einem Spielport werden dann verworfen, bestehende bleiben erhalten. Diese Betriebsart verlangt eine ausdrückliche Bestätigung in der Konfiguration, verweigert sich, wenn die Schnittstelle des Overlay-Netzes fehlt, und fällt auf den einfachen Flutschutz zurück, wenn der Regelsatz abgelehnt wird. Sie verengt den direkten Weg und ist keine Authentifizierung: Bei UDP kann eine gefälschte Absenderadresse die Prüfung bestehen. Flutschutz und Beschränkung sind standardmäßig abgeschaltet. Bei der Bereitstellung eines Knotens lässt sich der Flutschutz auf Wunsch mit einrichten, ebenso das Kernel-Profil und eine ratenbegrenzende Firewall-Regel für SSH und SFTP.

Der Flutschutz schützt nicht gegen volumetrische Angriffe, die die Anbindung des Knotens füllen. Er schützt auch nicht gegen Fluten mit gefälschten Absendern, die sich pro Quelladresse nicht trennen lassen; ein Ventil je Port hält dann nur die Rechenlast in Grenzen und verwirft dabei auch legitime neue Verbindungen. Angriffe innerhalb einer bestehenden Spielsitzung deckt er ebenfalls nicht ab. Die Regeln werden in Tests gegen eine nachgebildete Umgebung geprüft; vor jeder Zuschaltung folgt die Abnahme gegen einen echten Kernel.

Der Node-Agent besitzt eine zuschaltbare Selbstprüfung, die im Containermodus das Seccomp-Profil gegen die eingebettete Fassung und die Neuzuordnung der Benutzerkennungen kontrolliert und in beiden Startarten den laufenden Falco, den verschlüsselten Datenträger und eine Firewall mit Standardablehnung. Eine Prüfung, die sich nicht auswerten lässt, gilt als nicht bestanden. Das Ergebnis steht als Metriken und als maschinenlesbarer Bericht zur Verfügung, und der Node-Agent kann den Bericht an das Connection-Gateway senden. Das Gateway speichert den letzten Bericht je Knoten und leitet daraus ein Kennzeichen ab. Ein Netztest für den Containermodus startet zwei gehärtete Wegwerf-Container und prüft, dass der eine den anderen nicht erreicht; er läuft nur auf Anforderung, etwa aus einem vom Betreiber eingerichteten Zeitgeber, und die Selbstprüfung liest sein Ergebnis.

6.6 Anmeldung der Knoten und Datenträgerverschlüsselung

Ein neuer Spielknoten meldet sich mit einem einmaligen Bootstrap-Token beim Gateway an. Das Token wird geprüft und verbraucht, bevor die Zertifikatsanfrage signiert wird, sodass ein wiederholter Einsatz scheitert. Das Zertifikat des Knotens stammt aus der PKI von Vault, ist 48 Stunden gültig und wird vom Node-Agent im letzten Drittel der Laufzeit erneuert. Der eigentliche Verkehr läuft über einen Listener, der ein gültiges Client-Zertifikat verlangt und TLS 1.3 als Mindestversion durchsetzt. Das Gateway begrenzt die Rate geöffneter Streams pro Quelladresse und insgesamt, und die Bootstrap- und Schlüssel-Endpunkte begrenzen pro Knoten und lehnen bei vollem Zähler eher ab, als Anfragen durchzulassen.

Beim Bereitstellen eines Knotens kann der Datenträger mit LUKS verschlüsselt werden, und der Schlüssel wird über einen eigenen mTLS-Endpunkt des Gateways in Vault hinterlegt. Die Hinterlegung ist nur einmal möglich und kann einen vorhandenen Schlüssel nicht überschreiben. Beim Abruf prüft das Gateway zuerst eine Sperrmarkierung des Knotens und verweigert die Herausgabe, wenn die Markierung gesetzt ist oder sich nicht lesen lässt. Der Node-

Cloud Spheres UG (haftungsbeschränkt)
Registered Office: Paderborn
Address: Biberweg 24, 33102 Paderborn
Registration Court: Amtsgericht Paderborn
Commercial Register Number: HRB 18602
Managing Directors: Silas Timmermann, Alan Khudhur

Agent prüft in seiner Selbstprüfung, ob der Datenträger auf dm-crypt liegt.

6.7 Zertifikate und Lebensdauern

Zertifikate im System haben sehr unterschiedliche Lebensdauern, die in der folgenden Tabelle zusammengestellt sind. Kurze Lebensdauern verringern das Zeitfenster eines gestohlenen Schlüssels und setzen eine Erneuerung voraus, die je nach Zertifikat automatisch oder von Hand erfolgt. Die Rotation des Vertrauensankers von Linkerd ist als Ablauf mit einer Überlappungszeit beschrieben, in der der alte und der neue Anker gemeinsam gelten. Ein Probejob spielt die ersten Phasen ohne Änderung durch; die vollständige Durchführung gehört zum Ausbaupfad.

Zertifikat	Aussteller	Lebensdauer (Konstruktionswert)
Linkerd Vertrauensanker	cert-manager, selbstsigniert, ECDSA P-256	zehn Jahre
Linkerd Identitätsaussteller	Vertrauensanker	30 Tage, Erneuerung zehn Tage vor Ablauf
Linkerd Workload-Zertifikat	Identitätsaussteller	24 Stunden
Interne CA für Infrastruktur	cert-manager, ECDSA P-256	zehn Jahre
Zertifikat eines Spielknotens	Vault PKI	48 Stunden, Erneuerung im letzten Drittel der Laufzeit
Client-Zertifikat Edge-Proxy zum Orchestrator	Vault PKI	90 Tage bei manueller Ausstellung, 30 Tage bei Erneuerung über den Operator, Neuladen im Proxy ohne Neustart zuschaltbar
Wildcard-Zertifikat der Spieladresse	Let's Encrypt über DNS-01	90 Tage, Erneuerung 30 Tage vor Ablauf
Redis- und Token-Handler-Zertifikate	cert-manager	ein Jahr in Produktion

7 Zone B: Interne Sicherheit

7.1 Zulassungskontrolle mit Kyverno

Kyverno ist für alle Cluster definiert und prüft Objekte beim Anlegen. Die Registry-Richtlinie erlaubt Images grundsätzlich nur aus der eigenen Harbor-Registry und aus einer Liste weiterer Quellen, mit Ausnahmen für das Mesh, den Bau im Cluster und ausgewählte Werkzeuge. Eine Signaturrichtlinie, die ein Image gegen die Identität der Build-Pipeline und einen Image-Pfad der eigenen Registry prüft, und eine Richtlinie, die die Referenz über den Digest verlangt, liegen als Entwurf im Audit-Modus im Quelltext vor und sind standardmäßig abgeschaltet. Die Richtlinie gegen Privilegien verbietet privilegierte Container, das ausdrückliche Setzen der Rechteausweitung, Einhängungen von Host-Pfaden und besonders gefährliche Capabilities. Eine Richtlinie im Audit-Modus, standardmäßig abgeschaltet, prüft eine Auswahl aus den Pflichten der Pod-Security-Stufe restricted, etwa den Betrieb ohne Administratorrechte, ein schreibgeschütztes Wurzeldateisystem und gesetzte Ressourcengrenzen.

Eine weitere Richtlinie sorgt dafür, dass nur berechtigte Administratoren das Label setzen dürfen, das einen Pod von der automatischen Beendigung bei Shell-Aufrufen ausnimmt, und schließt damit einen Umgehungsweg an Falco Talon. Eine andere prüft das Regions-Label der ArgoCD-Cluster-Objekte, damit keine Vorlagen in die

Cloud Spheres UG (haftungsbeschränkt)
Registered Office: Paderborn
Address: Biberweg 24, 33102 Paderborn
Registration Court: Amtsgericht Paderborn
Commercial Register Number: HRB 18602
Managing Directors: Silas Timmermann, Alan Khudhur

Generatoren eingeschleust werden können. Eine Richtlinie im Audit-Modus prüft Betriebsregeln wie feste Service-Accounts und gesetzte CPU-Anforderungen, und eine Mutation setzt für Images mit Digest oder Commit-Tag die Pull-Richtlinie IfNotPresent, was der Verfügbarkeit bei einem Ausfall der Registry dient. Im Produktionsprofil ist der Zulassungs-Webhook mit drei Replikas und der Ausfallrichtlinie Fail definiert, sodass ein Ausfall von Kyverno das Anlegen von Pods verhindert. Ausnahmen für Plattform-Namespaces, deren Komponenten die verbotenen Rechte benötigen, sind namentlich geführt, und ein Lint in der CI prüft, dass jede Ausnahme ein Ablaufdatum trägt.

Die Build-Pipeline enthält einen Zusatzjob, der Images ohne festen Schlüssel mit Cosign signiert, mit Trivy auf behebbare Schwachstellen der Stufen HIGH und CRITICAL prüft, wobei ein Fund den Job abbricht, und mit Syft eine Software-Stückliste sowie einen Herkunftsnachweis erzeugt, der Commit, Pipeline und Bauumgebung nennt. Der Job ist in einer ersten Gruppe von Diensten eingeführt; seine Ausweitung und seine Verknüpfung mit der Auslieferung gehören zum Ausbaupfad. Außerdem prüfen die Pipelines mehrerer Dienste die heruntergeladenen Werkzeuge (Helm, kubectl und das Vault-Werkzeug) vor der Installation gegen feste SHA-256-Werte, die aus den Veröffentlichungen der Hersteller übernommen sind, und ein Fehlschlag bricht den Job ab.

7.2 Service Mesh

Linkerd verschlüsselt den Verkehr zwischen Pods per gegenseitigem TLS und gibt jedem Workload eine kryptografische Identität. Die Edge-Version ist für den Modus mit Init-Container, also ohne CNI-Plugin, konfiguriert. Der Vertrauensanker und die Aussteller sind in Kapitel 6.7 aufgeführt. Die Steuerungsebene ist mit drei Replikas definiert, der Injektor-Webhook mit der Ausfallrichtlinie Fail, damit kein Pod versehentlich ohne Proxy startet. Für einen Teil der Dienste liegen Server- und Autorisierungsobjekte mit namentlich genannten Aufrufern vor. Einzelne Infrastrukturkomponenten laufen bewusst ohne Proxy.

Das Mesh liefert damit Verschlüsselung, Identität und eine erste Zugriffsentscheidung auf der Ebene von Verbindungen. Welcher Dienst welche Methode aufrufen darf, entscheidet zusätzlich die Anwendungsschicht L4. Beide Entscheidungen sind unabhängig voneinander, sodass der Fehler in einer von ihnen nicht zu einem offenen Zugang führt.

7.3 Autorisierung mit serviceauth-go

serviceauth-go ist die eigene Go-Bibliothek für die Authentifizierung und Autorisierung zwischen Diensten. Sie wird als gRPC-Interceptor eingehängt und für REST als HTTP-Middleware, sodass die Fachlogik der Dienste keine Berechtigungsprüfung enthält. Neben den drei Paketen für Linkerd, gegenseitiges TLS und die hybride Variante enthält sie Pakete für das Laden der Policies, das Erneuern des Vault-Tokens, das Sichtbarmachen abgeschalteter Funktionen und die Prüfung von Konfigurationswerten.

7.3.1 Identität

Bei gRPC liest die Bibliothek die Identität aus dem Metadatenfeld l5d-client-id, das der Linkerd-Proxy des Empfängers aus dem geprüften Zertifikat des Aufrufers setzt. Die Bibliothek prüft das Format der Identität und, in der strengen Betriebsart, das vollständige Identitätsformat einschließlich der Vertrauensdomäne. Die strenge Betriebsart lässt sich je Dienst einschalten und nimmt die Vertrauensdomäne als Parameter. Die Bibliothek verifiziert das Zertifikat nicht selbst und verlässt sich darauf, dass der Proxy das Feld setzt und von außen gelieferte Werte entfernt. Die NetworkPolicies beschränken, wer einen Dienst erreichen kann. Für Verbindungen über ein Multicluster-Gateway kann die Bibliothek bestimmte Weiterleiter als vertrauenswürdig kennzeichnen und übernimmt

Cloud Spheres UG (haftungsbeschränkt)
Registered Office: Paderborn
Address: Biberweg 24, 33102 Paderborn
Registration Court: Amtsgericht Paderborn
Commercial Register Number: HRB 18602
Managing Directors: Silas Timmermann, Alan Khudhur

dann den ursprünglichen Aufrufer aus dem letzten Eintrag des Forwarded-Headers nach RFC 7239. Fehlt der Header, wird abgelehnt.

Bei REST-Schnittstellen mit Client-Zertifikat stellt die Bibliothek eine TLS-Konfiguration mit TLS 1.3 als Mindestversion bereit, verlangt ein gültiges Client-Zertifikat mit der Verwendung für Clients und leitet den Dienstnamen aus dem ersten Label des Common Name ab. Eine Prüfung von Sperrlisten ist zuschaltbar und schließt dann geschlossen.

7.3.2 Policy-Modell

Die Policy eines Dienstes ist ein Dokument im KV-Speicher von Vault, benannt nach der Dienstkennung. Es enthält die Rollen des Dienstes, die erlaubten Endpunkte, eine Liste erlaubter Dienste als beratende Angabe, Werte für die Ratenbegrenzung und ein Ablaufdatum. Ein Endpunkt wird als vollständiger gRPC-Methodenname angegeben, als Stern oder mit einem Stern am Ende als Präfix. Ein Stern in der Mitte trifft nie. Eine leere Liste erlaubter Endpunkte lehnt alles ab, und solche Policies werden als Metrik ausgewiesen. Der Namensraum gehört nicht zum Schlüssel der Policy, er trennt lediglich Einträge im Zwischenspeicher. Bereitgestellt werden sie von der CI-Pipeline des jeweiligen Dienstes.

7.3.3 Zwischenspeicher, Singleflight und Circuit Breaker

Damit nicht jeder Aufruf Vault belastet, arbeitet die Bibliothek mit zwei Zwischenspeichern. Der erste hält die aufgelösten Claims eines Aufrufers in einem LRU-Speicher mit 1000 Einträgen und ist nur begrenzt gültig. Der zweite hält die geladenen Policies in einem LRU-Speicher mit 500 Einträgen. Er hat keine Gültigkeitsprüfung beim Lesen, sondern wird durch einen Hintergrundprozess aktualisiert, der in einem festen Takt die Liste der Policies in Vault abfragt und jede einzelne liest, mit einem zufälligen Versatz beim Start und mehreren parallelen Arbeitern. Einträge, die in Vault nicht mehr vorhanden sind, werden verworfen, ein gelöschter Dienst verschwindet also spätestens beim nächsten Durchlauf aus dem Policy-Speicher; bereits aufgelöste Claims wirken für lesende Methoden noch über mehrere Gültigkeitsdauern nach, für kritische Methoden bis zum Ablauf ihres Zwischenspeichers. Für den Entzug einer Berechtigung bietet die Bibliothek einen Hook, der den betroffenen Eintrag sofort aus den Zwischenspeichern entfernt.

Gleichzeitige Auflösungen desselben Dienstes werden im Authenticator mit Singleflight zusammengefasst, sodass mehrere Goroutinen nur eine Anfrage auslösen. Drei getrennte Circuit Breaker auf Basis von gobreaker schützen den Authenticator, den Vault-Anbieter und die Zertifikatsverwaltung der PKI. Sie sind gleich eingestellt: Im geschlossenen Zustand öffnen sie bei einer hohen Fehlerquote nach einer Mindestzahl von Anfragen in einem kurzen Fenster, bleiben eine feste Zeit offen und lassen im halboffenen Zustand wenige Testanfragen durch. Unbekannte Dienste lassen sich kurz negativ zwischenspeichern (zuschaltbar), damit ein wiederholter Aufruf nicht jedes Mal eine Auflösung auslöst.

7.3.4 Verhalten bei Ausfall

Fällt Vault aus, kennt die Bibliothek zwei Verhaltensweisen. Die erste greift auf den veralteten Eintrag im Claims-Speicher zurück, solange dieser ein festgelegtes Vielfaches seiner Gültigkeit nicht überschreitet (Dienste setzen eigene Grenzen). Die zweite lehnt Anfragen ab. Ein Dienst ordnet seine Methoden über die Option für kritische Methoden zu: Für diese Methoden prüft die Bibliothek, wie alt die zuletzt geladene Policy-Quelle ist, und lehnt ab, sobald die Aktualisierung über die erwartete Frist hinaus ausbleibt. Lesende Methoden dürfen auf

Cloud Spheres UG (haftungsbeschränkt)
Registered Office: Paderborn
Address: Biberweg 24, 33102 Paderborn
Registration Court: Amtsgericht Paderborn
Commercial Register Number: HRB 18602
Managing Directors: Silas Timmermann, Alan Khudhur

zwischengespeicherte Berechtigungen zurückgreifen, wobei sich die Dauer je Dienst begrenzen lässt, und die Bibliothek zählt jeden Rückgriff als Metrik. Ist der Speicher leer und Vault nicht erreichbar, wird abgelehnt. Eine Bereitschaftsprüfung, die ein Dienst einbinden kann, meldet einen Pod erst als bereit, wenn die Policies geladen sind. Die Entscheidung ist in einem Architekturdokument festgehalten. Mehr als zehn Dienste, darunter Abrechnung, Token-Handler, der DSGVO-Dienst und mehrere Dienste der Server-Bereitstellung, haben die Einstufung übernommen; die Ausweitung auf weitere Dienste gehört zum Ausbaupfad.

Zur Beobachtung veröffentlicht die Bibliothek Metriken für Anfragen nach Ergebnis, Dauer der Prüfung, Fehler, begrenzte Anfragen, Ladevorgänge, geöffnete Breaker, Rückfälle auf veraltete Einträge und zusammengefasste Anfragen. Alarmregeln melden Ablehnungen wegen fehlender Berechtigung und wegen fehlender Identität. Der Rate-Limit-Interceptor im jeweiligen Dienst begrenzt die Last, je nach Dienst insgesamt, je Aufrufer oder je Methode.

7.3.5 Leistung

Die Vorarbeit nannte weniger als eine Millisekunde bei einem Treffer im Zwischenspeicher. Diese Größenordnung ist gemessen: Ein Benchmark der Bibliothek, ausgeführt auf einem Rechner mit AMD Ryzen 7 3700X unter Windows mit einem sequentiellen Aufruf in einer Schleife, stummem Logger und einem Anbieter im Speicher, ergab (bestes Ergebnis aus fünf Durchläufen mit einem Prozessorkern, als Referenz im Repository abgelegt) für einen Treffer im Zwischenspeicher 1623 Nanosekunden pro Aufruf bei 8 Allokationen, für ein Nachladen der Policy über Singleflight und Breaker 2796 Nanosekunden bei 16 Allokationen und für den gesamten unären Interceptor um einen Treffer 5127 Nanosekunden bei 28 Allokationen. Sie belegen, dass der Overhead der Bibliothek im Bereich weniger Mikrosekunden liegt.

Die Messung sagt nichts über die Dauer eines Zugriffs auf Vault aus, weil der Anbieter im Speicher lag, und nichts über den Durchsatz eines Servers unter Last. Beides ist mit einem Lasttest gegen ein echtes Vault zu messen; der Benchmark der Bibliothek läuft unterdessen mit einer Regressionsgrenze in der CI. Für den Zugriff auf Vault und den Durchsatz nennt das Paper deshalb keine Zahl.

7.4 Die Interceptor-Kette

Die gRPC-Server der Dienste folgen im Regelfall einer gemeinsamen Reihenfolge: Wiederherstellung nach Panics, Anfragekennung, Logging, Metriken, Authentifizierung mit serviceauth-go, Ratenbegrenzung und Validierung, jeweils für unäre Aufrufe und Streams. Die Tracing-Instrumentierung mit OpenTelemetry ist bei den meisten Diensten als StatsHandler des Servers eingehängt und läuft damit vor allen Interceptoren, bei einigen als Interceptor. Einzelne Dienste ergänzen die Kette, etwa um die Sprache der Anfrage und eine Ratenbegrenzung pro Methode, und Server für besondere Aufgaben haben verkürzte Ketten. Listener für besondere Aufgaben, etwa die Anbindung eines Dienstes außerhalb des Mesh, ersetzen die Authentifizierung durch die Prüfung des Common Name eines gegenseitig authentifizierten TLS-Zertifikats. Im Community-Dienst begrenzen die Handler des Creator-Programms die Größe des gelesenen Anfragekörpers, und ein Test prüft, dass das Lesen oberhalb der Grenze abbricht.

7.5 Geheimnisse mit Vault

Vault ist je Cluster-Gruppe als Raft-Verbund mit drei Replikas, mit Anti-Affinität und einem Disruption Budget definiert. Der Zugriff auf die API ist mit TLS geschützt. Die Entsiegelung ist als Transit-Siegel angelegt: Beim Start entschlüsselt ein eigener, dedizierter Transit-Vault den Hauptschlüssel. Für den Notfall gibt es eine dokumentierte

Cloud Spheres UG (haftungsbeschränkt)
Registered Office: Paderborn
Address: Biberweg 24, 33102 Paderborn
Registration Court: Amtsgericht Paderborn
Commercial Register Number: HRB 18602
Managing Directors: Silas Timmermann, Alan Khudhur

Wiederherstellung, die auf Sicherungen mit AES-256 und einer geteilten Hinterlegung des Sicherungsschlüssels beruht, ergänzt um regelmäßige Snapshots und Alarme bei fehlgeschlagenen oder veralteten Sicherungen.

Genutzt werden vier Funktionen. Der KV-Speicher in der Version 2 hält Geheimnisse und Policy-Dokumente, mit höchstens zehn Versionen und einer Löschfrist von einem Jahr. Die PKI mit einer Root-CA und einer Zwischen-CA stellt Zertifikate für Knoten, Gateways und den Edge-Proxy aus. Die Transit-Engine verschlüsselt Nutzerdaten der Anwendungen (Kapitel 7.6). Als Authentifizierung dienen die Kubernetes-Methode für Dienste und der Operator sowie AppRole und JWT mit dem OIDC-Token der Pipeline für die Automatisierung der CI. Die Secret-IDs der CI-Rollen sind befristet, und die Lebensdauer der Tokens ist je Rolle festgelegt.

Dienste erhalten ihr Vault-Token je nach Dienst über ein Kubernetes-Secret als periodisches Token, das der Dienst erneuert, oder über die Kubernetes-Anmeldung im Pod. Der Vault Secrets Operator übernimmt Geheimnisse aus Vault in Kubernetes-Secrets und Zertifikate in die Pods, und die zugehörigen Rollen vergeben Tokens mit begrenzter Lebensdauer. Wo Tokens erneuert werden, prüft das Paket zur Erneuerung jede Erneuerung mit einer Abfrage des eigenen Tokens, erneuert in einem festen Takt und stellt einen Zustand für die Beobachtung bereit, der nach mehreren aufeinanderfolgenden Fehlern ungesund meldet. Das Audit-Gerät von Vault protokolliert jeden Zugriff, und ein Notzugang ist dokumentiert.

7.6 Identität und Datenschutz auf Anwendungsebene

Die Anmeldung wird vom Dienst IAM abgewickelt, der Token an den Token-Handler delegiert. Der Token-Handler stellt Token wahlweise als JWT (signiert mit RS256) oder als JWE (verschlüsselt mit RSA-OAEP-256 und A256GCM) aus, und IAM verwendet das JWT als Zugriffstoken und das JWE für das Profil-Token. Die Schlüsselkennung wird deterministisch aus dem öffentlichen Schlüssel abgeleitet, sodass alle Replikas dieselbe Kennung erzeugen. Signatur und Verschlüsselung können getrennte Schlüssel verwenden, und die Trennung lässt sich erzwingen. Bei einer Rotation bleibt der alte öffentliche Schlüssel für eine Übergangszeit gültig. Zugriffstoken sind kurzlebig, Refresh-Token gelten 720 Stunden. Die Refresh-Token verwaltet IAM und speichert sie als SHA-256-Wert in Postgres. Wird ein bereits widerrufenes Refresh-Token außerhalb einer kurzen Toleranzfrist für gleichzeitige Erneuerungen erneut vorgelegt, gilt das als Diebstahlsindiz, und IAM widerruft die gesamte Token-Familie. Der Widerruf von Zugriffstoken liegt in Redis und schließt für privilegierte Token bei einem Ausfall von Redis geschlossen. Über einen NATS-JetStream-Strom werden Widerrufe beim Abmelden zwischen Regionen verteilt. Das kurzlebige Token, mit dem IAM eine Registrierung an den Profildienst übergibt, wird dort per HMAC geprüft, und Aussteller, Zielgruppe und Ablaufzeit sind Pflichtangaben. Ein Token ohne sie oder mit fremden Werten wird abgelehnt.

Passwörter werden mit bcrypt gespeichert. Nach mehreren fehlgeschlagenen Versuchen wird das Konto für eine begrenzte Zeit gesperrt, wobei Fehlversuche während einer laufenden Sperre diese nicht verlängern, und der Anmeldepfad auch bei unbekannten Konten dieselbe Rechenarbeit leistet, um keine Rückschlüsse auf die Existenz eines Kontos zuzulassen. Die zweite Stufe nutzt TOTP nach RFC 6238 mit HMAC-SHA1, sechs Stellen, 30 Sekunden Schrittweite und einer Toleranz von einem Schritt. Ein Zähler in Redis verhindert die Wiederverwendung eines Codes. Das TOTP-Geheimnis wird auf Anwendungsebene mit AES-GCM verschlüsselt, zehn Backup-Codes werden als bcrypt-Hash abgelegt, und eine Herausforderung erlaubt nur wenige Fehlversuche. Als externe Anmeldeverfahren stehen Google, Discord, Steam und Microsoft zur Verfügung.

Personenbezogene Felder werden über die Transit-Engine verschlüsselt. Die Bibliothek dbcrypto unterstützt dafür je Mandant einen eigenen benannten Schlüssel vom Typ aes256-gcm96 und legt ihn beim ersten Gebrauch an. Für

Cloud Spheres UG (haftungsbeschränkt)
Registered Office: Paderborn
Address: Biberweg 24, 33102 Paderborn
Registration Court: Amtsgericht Paderborn
Commercial Register Number: HRB 18602
Managing Directors: Silas Timmermann, Alan Khudhur

die Suche nach verschlüsselten E-Mail-Adressen dient in IAM ein HMAC-SHA256 als blinder Index.

Der Audit-Log-Dienst nimmt Ereignisse über gRPC entgegen und überschreibt das Feld für den erzeugenden Dienst mit der authentifizierten Identität des Aufrufers, sodass ein Dienst nicht im Namen eines anderen schreiben kann. Der Dienst kann die Einträge verketten; die Verkettung ist zuschaltbar und standardmäßig abgeschaltet. Mit eingeschalteter Verkettung wird jeder Eintrag mit dem vorherigen durch einen Hash verbunden, der Kopf der Kette regelmäßig über die Transit-Engine von Vault signiert und zusätzlich in einem Objektspeicher mit Schreibschutz verankert. Die Datenbankrollen sind dann auf das Anhängen beschränkt, eine tägliche und eine wöchentliche Prüfroutine verifizieren die Kette, und Alarme melden Brüche. Die Integrationstests gegen Postgres gehören nicht zum Standardlauf der CI. Ereignisse liegen 90 Tage in Postgres und werden danach nach geprüftem Hochladen archiviert.

8 Zone C: Erkennung und Reaktion

8.1 Laufzeiterkennung mit Falco

Falco ist für die verwalteten Cluster mit dem Treiber modern eBPF im CO-RE-Verfahren als DaemonSet mit hoher Priorität definiert, das alle Taints toleriert. Die Regeln bestehen dort aus dem Upstream-Regelsatz und eigenen Regeln. Auf den Spielknoten wird eine eigene Instanz als Systemdienst mit einem eigenen Regelsatz eingerichtet. Die Ausgabe erfolgt als JSON. Die Tabelle nennt die eigenen Regeln für die Cluster.

Regel	Auslöser	Priorität
Shell in Container gestartet	Start einer Shell in einem Container	Critical
Rechteausweitung	Start von sudo, su, newgrp oder sg	Critical
Zugriff auf sensible Dateien	Lesen ausgewählter sensibler Dateien wie Schattenpasswörter und Zugangsdaten von Administratorkonten	Critical
Paketmanager gestartet	apt, apk, yum, dnf, pip oder npm in einem Container	Critical
Kryptomining	bekannte Miner-Prozesse oder Stratum-Adressen und Donate-Optionen in der Kommandozeile	Critical
Setuid oder Setgid	Aufruf durch nicht privilegierte, nicht freigegebene Prozesse	Warning
Download-Werkzeug	curl oder wget in einem Container	Warning
Unerwartete ausgehende Verbindung	Verbindung zu einem unüblichen Ziel	Warning
Lesen des Service-Account-Tokens	Zugriff auf das eingehängte Token	Info

Falcosidekick verteilt die Ereignisse. Es schreibt, wo ein Loki-Speicher vorhanden ist, alle Ereignisse ab der Stufe Informational dorthin, sendet ab Critical an den Alertmanager und leitet ab Warning an Talon weiter. Der direkte Weg zu Chatdiensten ist abgeschaltet, Chat-Meldungen laufen nur über den Alertmanager.

Cloud Spheres UG (haftungsbeschränkt)
Registered Office: Paderborn
Address: Biberweg 24, 33102 Paderborn
Registration Court: Amtsgericht Paderborn
Commercial Register Number: HRB 18602
Managing Directors: Silas Timmermann, Alan Khudhur

Die Regeln der Spielknoten sind an die Spielcontainer und an die Prozesse der Konten gebunden, die der Node-Agent je Server anlegt, und decken beide Ergebnisse eines Aufrufs ab. Zwölf Regeln haben die Priorität Critical: Ausbruch über Namensräume oder Mounts, Manipulation von Prozessspeicher, Manipulation des Kernels, Nutzung von BPF, perf und userfaultfd, Ausbruch über File Handles, Schreiben unter Binärverzeichnissen, Kontakt zu den Metadatendiensten der Cloud, das Starten von Angriffswerkzeugen wie Netcat, Compilern oder Debuggern, Kryptominer, Fileless-Ausführung aus dem Speicher, das Lesen sensibler Dateien und der Zugriff auf einen Köder. Sechs Regeln haben die Priorität Warning, darunter unerwartete ausgehende Verbindungen, eine Shell mit dem Spielprozess als Elternprozess, io_uring, Kernel-Keyring, Systemsteuerung und verdächtige Werkzeuge. Eine weitere Regel meldet eine unerwartete Shell mit Notice.

Die Meldungen der Knoten gelangen über einen Weiterleiter des Knotens per gegenseitigem TLS an eine eigene Falcosidekick-Instanz, die nach Loki und in den Alertmanager schreibt. Fehlt dem Weiterleiter das Zertifikat, verwirft er das Ereignis, statt es unverschlüsselt zu senden. Ein zweiter, unabhängiger Verbraucher schreibt Meldungen als Zeile in die Konsole des betroffenen Gameservers, sodass der Kunde sie sieht.

8.2 Automatische Reaktion mit Falco Talon

Falco Talon ist in den Clustern mit zwei Replikas und Anti-Affinität definiert und führt bei Ereignissen Aktionen aus. Vor seinem Endpunkt kann ein Proxy ein Token prüfen (zuschaltbar), und eine NetworkPolicy begrenzt den Zugriff auf den Endpunkt. Talon liest genau eine Regeldatei. Sie wird aus einer geprüften Quelldatei im Repository erzeugt, und ein Prüfsummenwert im Pod-Template sorgt dafür, dass eine Regeländerung die Pods neu startet. Eine Prüfung in der CI stellt sicher, dass Staging und Produktion dieselben Regeln tragen, dass die ausgelieferte Datei aus der Quelle erzeugt ist, dass Ausnahmeregeln vor allen anderen stehen und dass die Rechte des Dienstes dem Prinzip der geringsten Rechte entsprechen.

Die Regeln sind nach Schweregrad gestuft. Zuerst stehen Ausnahmeregeln, die nur protokollieren und alle nachfolgenden Regeln überspringen, etwa für Wartungs-Shells in Plattform-Namensräumen und für die Namensräume von Falco und Talon selbst. Ereignisse der niedrigeren Stufen führen zu einer Markierung des Pods, etwa bei Paketverwaltung, Setuid, Download-Werkzeugen oder einem unerwarteten ausgehenden Aufruf. Auf der Stufe Error ist die Isolation des Pods durch eine NetworkPolicy vorgesehen, die nur Ziele in privaten Adressbereichen erlaubt, ohne den Pod zu beenden; das gilt für auffällige ausgehende Verbindungen wie DNS-Tunnel. Auf der Stufe Critical ist das Beenden des Pods ohne Karenzzeit vorgesehen, etwa bei Shells in Containern, Kryptomining, Paketmanagern, Rechteausweitung (mit vorheriger Isolation), Zugriff auf Zugangsdaten, Schreiben unter Binärverzeichnisse, Kopierwerkzeugen für Dateien und Reverse Shells. Für den bestätigten Ausbruch aus einem Container ist als höchste Stufe vorgesehen, den Knoten zu sperren und den Pod zu beenden. Eine letzte Auffangregel erfasst jedes Critical-Ereignis, das keine frühere Regel beansprucht hat.

Jede Regel trägt einen eigenen Schalter für den Trockenlauf: Talon protokolliert dann, was es tun würde, und ruft keine Schnittstelle von Kubernetes auf. Die Wirkstufen Markierung, Isolation, Beenden und Sperre des Knotens werden einzeln freigeschaltet. Eine zerstörende Regel wird erst scharf geschaltet, nachdem ein Abspiel echter Falco-Ereignisse gegen den Regelabgleich gezeigt hat, dass sie nur Ereignisse trifft, für die ein Pod verloren gehen darf, und nachdem die Rechte des Dienstes dafür angepasst sind.

Cloud Spheres UG (haftungsbeschränkt)
Registered Office: Paderborn
Address: Biberweg 24, 33102 Paderborn
Registration Court: Amtsgericht Paderborn
Commercial Register Number: HRB 18602
Managing Directors: Silas Timmermann, Alan Khudhur

8.3 Eindämmung auf dem Spielknoten

Auf den Spielknoten übernimmt der Node-Agent die Eindämmung. Er kennt drei Betriebsarten: nur Meldung, Schattenmodus und Erzwingung. Im Schattenmodus protokolliert er die Entscheidung, die er in der Erzwingung getroffen hätte, ohne zu handeln, sodass eine Regel vor ihrer Scharfschaltung an echtem Verkehr geprüft werden kann. Ein unbekannter Wert lässt den Node-Agent beim Start abbrechen. Die Erzwingung kann den Server bei einem Durchbruch oder einem berührten Köder sofort hart beenden, wobei die Daten erhalten bleiben, und der Gameserver-Manager führt den Server dann als eingefroren mit ausstehender Untersuchung; bei wiederholten Proben sperrt sie nur den ausgehenden Verkehr des Servers. Sie ist je Startart und Evidenzklasse einzeln freizuschalten. Sie löst nach Kapitel 9 nach Beweisstärke aus, und Ratenlimits begrenzen sie, damit ein Fehler in einer Regel nicht einen ganzen Knoten leeren kann: je Server und je Knoten gilt eine feste Obergrenze, und die Aktionen der Klassen E1 und E3 haben eigene, getrennte Grenzen, damit eine Klasse das Budget einer anderen nicht aufbraucht. Ein Bericht wertet die Entscheidungen des Schattenmodus je Regel aus.

8.4 Beobachtbarkeit

Prometheus und Alertmanager sind für die Cluster der Orchestrierung, der Microservices und der Website definiert, Loki als Speicher für die Orchestrierung und die Microservices, dr-heartbeat als gemeinsamer Chart für die Cluster. Prometheus fragt in einem festen Takt ab und wertet Regeln in diesem Takt aus. Jeder Dienst bringt einen ServiceMonitor mit. Loki speichert Protokolle in S3-kompatiblem Objektspeicher und trennt sie nach Namensraum in Mandanten, für ausgewählte Namensräume mit personenbezogenen Daten gilt eine kürzere Aufbewahrung. Der Ruler von Loki ist für die Cluster der Orchestrierung und der Microservices konfiguriert, und seine Regeln erkennen unter anderem gehäufte Ablehnungen und Anmeldefehler in den Protokollen. Tempo speichert Traces. Grafana ist für den Zugriff über das Overlay-Netz vorgesehen.

Alarmregeln, Dashboards und Datenquellen sind Code. Ein Skript in der CI-Validierung stellt bei jeder Änderung sicher, dass jede Regeldatei des Repositorys in mindestens einer Umgebungsdefinition als ausgelieferte Ressource ankommt, sodass Regeln nicht stillschweigend unwirksam bleiben. Für die Steuerungsebene von Linkerd sind Monitore und Alarme definiert.

8.5 Alarmierung

Der Alertmanager gruppiert nach Alarmname, Cluster und Namensraum, wartet kurz auf weitere Alarme derselben Gruppe, fasst Änderungen zusammen und wiederholt Meldungen in festen Abständen. In den Alertmanagern der Cluster für Orchestrierung, Microservices und Website gehen Meldungen der Schwere Critical und Page über eine eigene Route mit sofortiger Zustellung und kürzeren Wiederholungsabständen an den Chat. Ein Pager-Empfänger ist als weitere Route angelegt und lässt sich zuschalten. Warnungen gehen an den Chat, Meldungen der Stufe Info werden gesammelt zugestellt oder verworfen. Inhibit-Regeln unterdrücken Folgealarme, etwa Pod-Meldungen bei einem nicht bereiten Knoten.

Der Aufbau sieht einen täglichen synthetischen Alarm vor, der durch die Kette läuft, und einen Dead-Man-Alarm, der meldet, wenn er ausbleibt. In den Alertmanagern der Orchestrierung und der Microservices melden zwei weitere Alarme, wenn kein Empfänger konfiguriert ist und wenn fast alle Zustellungen fehlschlagen, und ein Wächter im Render-Prozess verhindert bei Alertmanagern, deren Routing im Repository liegt, einen leeren Standardempfänger. Das Modell der Schweregrade ist verbindlich: Page bedeutet, dass innerhalb einer festgelegten Zeit ein Mensch reagieren soll. Als unabhängiger Lebensbeweis ist dr-heartbeat vorgesehen, ein CronJob, der in regelmäßigen

Cloud Spheres UG (haftungsbeschränkt)
Registered Office: Paderborn
Address: Biberweg 24, 33102 Paderborn
Registration Court: Amtsgericht Paderborn
Commercial Register Number: HRB 18602
Managing Directors: Silas Timmermann, Alan Khudhur

Abständen ein Lebenszeichen an einen externen Prüfdienst sendet, sobald ein Endpunkt hinterlegt ist, sodass ein toter Cluster durch das Ausbleiben des Signals auffallen kann. Eigene Alarme melden, wenn dieser Job nicht erfolgreich läuft oder fehlschlägt.

Für Ausfälle, die den Betrieb der Plattform selbst treffen, liefert das Repository Alarmbündel, die auf die Ursache und nicht auf ein Folgesymptom zielen und die Stufe Page tragen. Sie betreffen die Erreichbarkeit des Siegels des Geheimnisspeichers, fehlgeschlagene oder ausbleibende Sicherungen des Geheimnisspeichers, die Zahl der Mitglieder der Steuerungsebene, Speicher- und Arbeitsspeicherdruck auf Knoten, gehäufte Fehler beim Laden von Images, ablaufende oder nicht bereite Zertifikate und die Erreichbarkeit der Anbindungspunkte der Spielknoten, die Blackbox-Sonden prüfen. Jeder Alarm trägt eine einzeilige Handlungsanweisung mit dem ersten Befehl. Ergänzend melden Regeln des Loki-Rulers Zeilen im Protokoll des Geheimnisspeichers, die ein nicht erreichbares Siegel oder unvollständige Snapshots anzeigen. Ein Testskript prüft die Alarmregeln mit positiven und negativen Fällen und ist Teil der CI.

8.6 Automatische Gegenmaßnahmen im Überblick

Die folgende Tabelle fasst zusammen, bei welchen Signalen das System selbstständig handelt. Sie enthält bewusst nur Maßnahmen mit Sicherheitsbezug.

Signal	Automatische Maßnahme (soweit zugeschaltet)	Ort
Mehrere fehlgeschlagene Passwortprüfungen	Konto zeitlich begrenzt gesperrt	IAM
Mehrere falsche Codes bei einer MFA-Herausforderung	Herausforderung verbraucht, bei Ausfall von Redis abgelehnt	IAM
Vorlage eines bereits widerrufenen Refresh-Tokens außerhalb der Toleranzfrist	Widerruf der gesamten Token-Familie	IAM
Überschreiten eines Anfragebudgets	Ablehnung mit 429 oder dem gRPC-Fehler ResourceExhausted	IAM, Website, weitere Dienste
Wiederholter Missbrauch von einer Quelle bei Anmeldung und Konto	Gestufte Sperre mit Ablauf, Begründung und Audit-Eintrag (wenn zugeschaltet, zunächst im Schattenmodus)	IAM
Wiederholte Verstöße an einem Spielport (Verbindungsrate, unlesbarer oder abgelaufener Handshake, unbekannter Hostname, Verbindungsgrenze)	Sperre der Quelle mit Ablaufzeit, bei Wiederholung länger	Edge-Proxy
Viele gleichzeitige oder halboffene Verbindungen	Abweisen in der Annahmeschleife, unbekannte Quellen zuerst	Edge-Proxy
Verbindungsflut auf einen Spielport eines Knotens	Verwerfen oberhalb der Grenzen je Quelle und Port	Node-Agent
Ablehnung durch Turnstile	Anfrage abgewiesen (wenn Turnstile zugeschaltet ist)	Website

Cloud Spheres UG (haftungsbeschränkt)
Registered Office: Paderborn
Address: Biberweg 24, 33102 Paderborn
Registration Court: Amtsgericht Paderborn
Commercial Register Number: HRB 18602
Managing Directors: Silas Timmermann, Alan Khudhur

Pod verletzt Zulassungsregel	Anlegen abgelehnt	Kyverno
Falco Critical im Cluster	Markierung des Pods; Beenden, je nach Regel nach Isolation, sobald die Regel freigeschaltet ist	Falco Talon
Falco Error im Cluster	Isolation des Pods durch NetworkPolicy und Markierung, sobald die Regel freigeschaltet ist	Falco Talon
Bestätigter Container-Ausbruch	Sperre des Knotens und Beenden des Pods, sobald die Regel freigeschaltet ist	Falco Talon
Erfolgreicher Durchbruch oder berührter Köder auf dem Spielknoten	Sofortiges hartes Beenden (bei aktivierter Erzwingung), Server eingefroren, Audit-Eintrag	Node-Agent
Wiederholte blockierte Proben eines Servers	Sperre des ausgehenden Verkehrs dieses Servers (bei aktivierter Erzwingung), Marker und Audit-Eintrag	Node-Agent
Ausfall eines Edge-Knotens	Entfernen aus der Namensauflösung	dns-guard
Schlüsselabruf eines gesperrten Knotens	Herausgabe verweigert	Gateway
Wiederverwendung eines Bootstrap-Tokens	Anmeldung abgelehnt	Gateway
Ausfall der Policy-Quelle in Vault	Kritische Methoden lehnen ab, sobald die Policy-Quelle überfällig ist, lesende Methoden greifen auf Zwischengespeichertes zurück	serviceauth-go
Fehlgeschlagene SSH-Anmeldungen	Zeitlich begrenzte Sperre der Adresse	fail2ban auf den vom Orchestrator eingerichteten Infrastrukturservern

8.7 Sicherung und Wiederherstellung

Der Sicherungsdienst bietet in seiner Cluster-Betriebsart an, Sicherungen der Spielwelten vor dem Ablegen mit einer Hüllverschlüsselung zu schützen (AES-256-GCM, Schlüssel über Transit). Eine Verifikation stellt eine Sicherung testweise wieder her, und ein Job für Wiederherstellungsübungen ist definiert. Beides ist standardmäßig abgeschaltet. Postgres wird mit pgBackRest gesichert, wobei eine Kopie außerhalb des Clusters in einem Objektspeicher liegt, Vault durch verschlüsselte Snapshots. Für die Datenbank stehen abgeleitete Zielwerte für Wiederherstellungszeit und Datenverlust in einem Dokument; eine vollständige Wiederherstellungsübung hat sie zu bestätigen.

9 Entwurfsbeitrag: evidenzgestufte Eindämmung und deterministische Tripwires

9.1 Ausgangspunkt

Verhaltensbasierte Erkennung, wie Falco sie mit Regeln über Prozesse, Dateien und Verbindungen leistet, hat eine unvermeidliche Fehlerrate. Für Kundenworkloads wiegt das schwer, denn eine falsche automatische Eindämmung beendet einen zahlenden Server und im Ernstfall die Spielwelt einer Gemeinschaft. Die übliche Antwort ist,

Cloud Spheres UG (haftungsbeschränkt)
Registered Office: Paderborn
Address: Biberweg 24, 33102 Paderborn
Registration Court: Amtsgericht Paderborn
Commercial Register Number: HRB 18602
Managing Directors: Silas Timmermann, Alan Khudhur

automatische Reaktionen zu vermeiden oder sehr konservativ zu drosseln. Damit bleibt die Reaktionszeit an einem Menschen hängen, und ein Angriff, der in Sekunden Nachbarn erreicht, ist dann nicht mehr einzuholen.

Der Entwurfsbeitrag setzt an einer Beobachtung an, die für Gameserver besonders günstig ist. Die Menge der Systemaufrufe, die ein Spielserver legitim ausführt, ist klein, gut bekannt und für jede Version eines Spiels reproduzierbar messbar. Das Seccomp-Profil der Container ist eine Positivliste der Systemaufrufe, die Spielserver legitim benötigen. Es soll damit nicht nur ein Werkzeug der Durchsetzung sein, sondern zugleich ein Messinstrument: Was das Profil ablehnt, soll nicht zum legitimen Verhalten gehören, und die Goldläufe aus Kapitel 9.3 sollen das belegen.

9.2 Evidenzklassen

Die Eindämmung richtet sich nicht nach der Priorität einer Regel, sondern nach der Stärke des Beweises, den ein Ereignis liefert. Es gibt vier Klassen.

Klasse	Beweis	Beispiel	Reaktion
E0 Verhalten	Auffälliges, aber möglicherweise legitimes Verhalten	Unerwartete ausgehende Verbindung, verdächtiges Werkzeug	Meldung an Loki und Alertmanager, Markierung, keine automatische Eindämmung
E1 Probe	Ein blockierter Aufruf eines Systemaufrufs oder Zugriffs, der in keinem legitimen Lauf vorkommen soll	mount oder ptrace mit dem Ergebnis EPERM	Meldung, Zähler je Server, bei Wiederholung innerhalb eines Zeitfensters Sperre des ausgehenden Verkehrs des Servers, Beweispaket, Marker und Audit-Eintrag
E2 Durchbruch	Ein Ereignis aus einer geschlossenen Liste von Regeln für Ausbruchsversuche wurde mit Erfolg des Aufrufs quittiert, das Verbot hat also nicht gegriffen	mount oder setns mit dem Ergebnis SUCCESS	Beweispaket, sofortiges hartes Beenden, Server eingefroren, Audit-Eintrag
E3 Köder	Berührung eines Objekts, das nur ein Angreifer anfassen würde	Berühren eines Köderobjekts	Beweispaket, Beenden und Einfrieren des auslösenden Servers, Marker und Audit-Eintrag, Zuordnung zum Server

Das Ergebnis eines Systemaufrufs trennt dabei zwei Fälle, die eine reine Verhaltensregel nicht unterscheidet. Ein abgelehnter Aufruf zeigt, dass jemand etwas versucht, was das System verbietet. Ein erfolgreicher Aufruf aus derselben Menge zeigt, dass das Verbot nicht gegriffen hat, und ist damit der seltene Fall, in dem die Verteidigung selbst versagt hat. Der zweite Fall rechtfertigt eine sofortige automatische Reaktion, weil an seiner Fehlalarmrate kein legitimes Verhalten hängt.

9.3 Deterministische Tripwires durch Katalog-Replay

Ob ein Signal wirklich in keinem legitimen Lauf vorkommt, soll bewiesen und nicht angenommen werden. Dafür gibt es im Repository der Spiel-Images ein Werkzeug, das für ein Spiel und eine Version einen Goldlauf startet, während

Cloud Spheres UG (haftungsbeschränkt)
Registered Office: Paderborn
Address: Biberweg 24, 33102 Paderborn
Registration Court: Amtsgericht Paderborn
Commercial Register Number: HRB 18602
Managing Directors: Silas Timmermann, Alan Khudhur

ein Beobachter jeden Aufruf mitschreibt, die Läufe zusammenführt und in der CI ein Gate bereitstellt, das auf Versionen ohne Goldlauf hinweist und bei einem neuen Signal in der Menge bricht. Ein Goldlauf durchläuft den Start des Servers, eine Last aus Statusabfragen, Verbindungen und Konsolenbefehlen, das Speichern und das Beenden und liefert die Menge der beobachteten Systemaufrufe und Dateizugriffe. Die Menge der nie legitimen Signale ist die Menge, die in keinem Goldlauf über alle Spiele und Versionen vorkommt und die zusätzlich fachlich für einen Spielserver nicht in Betracht kommt, etwa Einhängen von Dateisystemen, Debugging fremder Prozesse oder das Laden von Kernelmodulen. Ein Signal, das in einem Goldlauf auftaucht, wird aus der Menge entfernt. Das Gate ist bisher nur hinweisend, und Goldläufe liegen für den Katalog bisher nicht vor; die Bindung der Katalogaufnahme an das Gate gehört zum Ausbaupfad.

Sobald die Läufe vorliegen, entsteht für die getesteten Spiele und Versionen eine Menge von Tripwires, deren Fehlalarmrate durch Konstruktion null ist. Die Aussage gilt bewusst nur für die getesteten Konfigurationen. Bis die Goldläufe über den Katalog erhoben sind, arbeitet der Durchbruch-Fall E2 mit einer festen, geschlossenen Liste von Regeln für Ausbruchsversuche, die Menge aus den Läufen wird als unbewiesen geführt, und ein Lauf über echte Spiele liegt bisher nicht vor. Mods können weitere Aufrufe auslösen, weshalb der Katalog auch häufig genutzte Modpacks als Goldläufe enthalten soll und die Menge im Zweifel nach der Seite der Vorsicht schrumpft. Signale, die nur in einigen Läufen vorkommen, gehören zur Klasse E0 und nicht zu E1 oder E2.

9.4 Köder

Ergänzend kann der Node-Agent je Server Köderobjekte auslegen, die kein legitimer Prozess anfasst. Name, Kennung und Inhalt jedes Köders werden aus der Kennung des Servers und einem geheimen Wert des Knotens abgeleitet, sodass sie je Server verschieden sind und nichts über die Kennung verraten. Ein Zugriff belegt damit einen Angriff und ordnet ihn dem Server zu, dessen Konto den Köder berührt hat. Weitere Arten von Ködern gehören zum Ausbaupfad.

9.5 Erprobung der Eindämmung

Die Erzwingung ist so angelegt, dass sie je Betriebsart zuerst im Schattenmodus betrieben wird, bis die Zahl der Ereignisse und die ausbleibenden Fehlalarme ihre Scharfschaltung tragen, und der Bericht des Node-Agents zeigt dafür je Regel die Ereignisse nach Klasse und Ergebnis, die Zahl der betroffenen Server und die Aktionen, die im Schattenmodus nur berechnet oder tatsächlich ausgeführt wurden. Für jede Startart soll die Erzwingung erst nach einem eigenen Schattenlauf freigegeben werden. Die Liste der Aufrufe für die Klasse E1 ist bis zum Ergebnis des Katalog-Replays vorläufig. Jede Aktion der Klassen E1 und E3 schreibt eine Marker-Datei und einen Audit-Eintrag. Vor einer Eindämmung kann der Node-Agent, wenn die Beweissicherung zugeschaltet ist, ein Beweispaket sichern: die Prozesse und Verbindungen des Servers, den berührten Köder und ausgewählte Protokollauszüge, jeweils mit dem Vorgänger über eine Kette von Prüfsummen verbunden und über eine Senke in ein Dateisystem oder in einen Objektspeicher mit Schreibschutz übertragen. Die Eindämmung wartet dafür höchstens eine begrenzte Zeit, das Übertragen läuft außerhalb dieses Pfades, und ein Werkzeug prüft die Kette. Die Zeit vom Ereignis bis zur Aktion kann als Histogramm je Klasse gemessen werden, und ein Übungswerkzeug spielt synthetische Ereignisse gegen Attrappen ein, die im Betrieb verworfen werden. Ein Test je Tripwire mit echtem Auslöser in Staging und die Messung an echten Ereignissen gehören zum Ausbaupfad (Kapitel 12.2).

Cloud Spheres UG (haftungsbeschränkt)
Registered Office: Paderborn
Address: Biberweg 24, 33102 Paderborn
Registration Court: Amtsgericht Paderborn
Commercial Register Number: HRB 18602
Managing Directors: Silas Timmermann, Alan Khudhur

9.6 Einordnung des Beitrags

Einzelne Elemente sind bekannt: Seccomp-Profile als Positivliste, Köderdateien in der Sicherheitsforschung und Falco-Regeln für Systemaufrufe. Neu ist nach unserer Kenntnis die Verbindung. Das Durchsetzungsprofil ist zugleich Messinstrument, ein Katalog-Replay soll für die getesteten Konfigurationen belegen, welche Signale in keinem legitimen Lauf vorkommen, und die Eindämmung richtet sich nach der Beweisklasse statt nach einer Priorität. Das Ergebnis ist eine automatische Reaktion, die sich für die getesteten Konfigurationen auf einen Beweis stützt und nicht auf eine Wahrscheinlichkeit. Ob diese Verbindung in der Praxis die erhoffte Fehlalarmfreiheit erreicht, zeigen die Goldläufe und die Schattenphase, und die Ergebnisse gehören in den Änderungsverlauf dieses Papers.

10 Einordnung: Standard und Eigenentwicklung

Aegiron ist zum größeren Teil eine sorgfältige Zusammenstellung bewährter Komponenten. Der Nutzen liegt in der Auswahl, in der Konfiguration, in den Nahtstellen zwischen den Bausteinen und in den Prüfungen, die jede Kontrolle belegen. Die folgende Tabelle trennt Standardkomponenten von Eigenentwicklungen.

Baustein	Art	Anmerkung
k3s, flannel, kube-router	Standard	Kubernetes-Distribution und Netzwerk
Linkerd, cert-manager	Standard	Service Mesh und Zertifikatsverwaltung
Vault, Vault Secrets Operator	Standard	Geheimnisse, PKI, Transit
Kyverno, Falco, Falco Talon	Standard	Zulassung, Erkennung, Reaktionsmaschine. Regeln und Stufen sind eigen.
Prometheus, Loki, Tempo, Grafana, Alertmanager	Standard	Beobachtbarkeit
Cloudflare Tunnel, Tailscale	Standard	Zugang und Overlay-Netz. Die Zugriffsregeln sind eigen und als Datei beschrieben.
serviceauth-go	Eigenentwicklung	Autorisierung pro Methode aus Vault-Policies mit begrenzter Veraltung und Breaker. Verwendet bekannte Muster (LRU, Singleflight, Circuit Breaker) in einer für Vault-Policies angepassten Kombination.
Node-Agent mit Isolation und Eindämmung	Eigenentwicklung	Gehärtete Container, Egress-Sperre, Flutschutz auf dem Knoten, Selbstprüfung, evidenzgestufte Eindämmung
Connection-Gateway	Eigenentwicklung	Anmeldung der Knoten mit verbrauchtem Bootstrap-Token, PKI-Ausstellung, Schlüsselhinterlegung

Cloud Spheres UG (haftungsbeschränkt)
Registered Office: Paderborn
Address: Biberweg 24, 33102 Paderborn
Registration Court: Amtsgericht Paderborn
Commercial Register Number: HRB 18602
Managing Directors: Silas Timmermann, Alan Khudhur

Edge-Proxy und dns-guard	Eigenentwicklung	Hostnamen-Routing über den Minecraft-Handshake, getrennte Grenzen je Zielart, verdrängungsfeste Limiter-Tabelle, Zulassungstore und Eskalationsleiter für Sperren, DNS-Abgleich ohne leere Menge. Verwendet bekannte Muster (Token-Bucket, mengenassoziative Tabelle, Semaphor) in einer für Verbindungsfluten angepassten Kombination.
Token-Handler, IAM, dbcrypto, Audit-Log	Eigenentwicklung	Anwendungsschicht für Identität und Datenschutz auf Standardverfahren
Evidenzstufen, Katalog-Replay, Köder	Entwurfsbeitrag	Kapitel 9

Ein Alleinstellungsmerkmal im Sinne einer neuen kryptografischen oder theoretischen Methode beansprucht das Paper nicht. Der eigene Beitrag liegt im Zusammenspiel: Das Durchsetzungsprofil wird zum Messinstrument, die Reaktion richtet sich nach der Beweisstärke, und jede Aussage über die Wirksamkeit soll durch einen automatischen Test belegt werden.

11 Nachweis

11.1 Abdeckung der Bedrohungen

Eine Architektur ist nur dann schlüssig, wenn jeder Angreifer aus Kapitel 3.2 durch Invarianten abgedeckt ist und jede Invariante eine Prüfung hat. Die folgende Tabelle stellt diese Zuordnung her.

Angreifer	**Invarianten**	**Wesentliche Kontrollen**
Extern gegen Website und Anmeldung	I2, I6, I7, I8, I10	Regeln der Firewall als Code, Bot-Prüfung an ausgewählten Formularen, gestufte Sperren, kurzlebige gebundene Token
Extern gegen Spielverkehr	I2, I8, I10, I11	Edge-Proxy mit Ratengrenzen je Quelle und Ziel, verdrängungsfestem Limiter, Zulassungstoren, Sperrliste und Sperren mit Eskalation, Bandbreitenlimit, Fuzz-Tests der Parser, Abgleich der Namensauflösung, Flutschutz auf dem Knoten
Kundencode auf dem Gameserver	I4, I5, I2, I8, I11	Gehärtete Container oder unprivilegiertes Konto je Server als Standard, Egress-Sperre, Verschlüsselung des Datenträgers, Selbstprüfung des Knotens, evidenzgestufte Eindämmung

Cloud Spheres UG (haftungsbeschränkt)
Registered Office: Paderborn
Address: Biberweg 24, 33102 Paderborn
Registration Court: Amtsgericht Paderborn
Commercial Register Number: HRB 18602
Managing Directors: Silas Timmermann, Alan Khudhur

Kompromittierter interner Dienst	I1, I2, I6, I10, I11	Netzwerkrichtlinien je Dienst, Mesh-Identität, Methodenautorisierung mit Fail-Closed für kritische Methoden, Entzug per Abfrage
Lieferkette	I3, I9, I12	Signatur, Schwachstellenprüfung, Stückliste und Herkunftsnachweis im Zusatzjob der Pipeline, geprüfte Prüfsummen heruntergeladener Werkzeuge, Signaturrichtlinie im Entwurf
Innentäter	I7, I12, I6	Zuschaltbar verkettetes Protokoll, Escrow des Sicherungsschlüssels mit Schwellenaufteilung, versionierte Zugriffsregeln des Overlay-Netzes

11.2 Abnahmekriterien

Die Wirksamkeit der Architektur soll an messbaren Kriterien belegt werden, die möglichst automatisch geprüft werden. Verbindliche Zielwerte für Zeiten werden erst nach der ersten Messkampagne festgelegt, weil ein Wert ohne Messung nur eine Behauptung wäre. Bis dahin gelten als Richtwerte, die die Messung bestätigen oder korrigieren soll, eine Erkennung innerhalb einer Sekunde und eine Eindämmung innerhalb von zehn Sekunden im 95. Perzentil.

Kriterium	Prüfung	Messgröße
Jeder Namespace besitzt Default-Deny	CI-Prüfung und laufender Abgleich	Abweichungen gleich null
Kein Pod ohne geprüfte Signatur	Zulassungsprotokoll und Stichprobe im Betrieb	Ablehnungen bei ungültiger Signatur gleich 100 Prozent
Kein Spielprozess mit administrativen Rechten	Automatische Knotenprüfung nach Bereitstellung und laufend	Verstöße gleich null
Alle Datenträger verschlüsselt	Statusprüfung aller Knoten	Anteil gleich 100 Prozent
Kein Container erreicht einen zweiten Container	Täglicher Netztest je Knoten	Erfolgreiche Verbindungen gleich null
Der Edge-Proxy hält seinen Latenzaufschlag	Benchmark in der CI bei jeder Pipeline	Aufschlag je neuer Verbindung und je Round Trip innerhalb der dokumentierten Spannen
Entzug einer Berechtigung wirkt zügig	Messung über die Benachrichtigung	95. Perzentil im Bereich weniger Sekunden, nach Messung festzulegen
Zeit bis zur Eindämmung	Ende-zu-Ende-Messung und Übungen	95. Perzentil unter dem festgelegten Zielwert
Fehlalarmfreiheit der Tripwires	Katalog-Replay bei jeder neuen Spielversion, Schattenphase	Fehlalarme gleich null in den Goldläufen
Alarme erreichen einen Menschen	Täglicher synthetischer Alarm bis zum Pager	Zustellung innerhalb der festgelegten Zeit, Ausbleiben ist ein Alarm

Cloud Spheres UG (haftungsbeschränkt)
Registered Office: Paderborn
Address: Biberweg 24, 33102 Paderborn
Registration Court: Amtsgericht Paderborn
Commercial Register Number: HRB 18602
Managing Directors: Silas Timmermann, Alan Khudhur

Regeln im Repository gleich Regeln im Cluster	Täglicher Abgleich	Abweichungen gleich null
Audit-Log unverändert	Tägliche Verifikation der Kette	Brüche gleich null
Wiederherstellung funktioniert	Vierteljährliche vollständige Wiederherstellung	Gemessene Zeit und gemessener Datenverlust innerhalb der Ziele
Kein Langzeitgeheimnis	Inventar aller Zugangsdaten	Zugangsdaten ohne Ablauf gleich null

11.3 Umsetzungsstand je Invariante

Die folgende Tabelle ordnet jeder Invariante zu, was im Quelltext umgesetzt ist. Umgesetzt heißt: gebaut, mit Tests versioniert und für die Auslieferung definiert. Einzelne Bausteine sind standardmäßig abgeschaltet. Die Tabelle sagt nichts darüber, welche Bausteine in einer bestimmten Umgebung eingeschaltet sind (Kapitel 2).

Nr.	Im Quelltext umgesetzt	Stand
I1	Mesh-Identität, Methodenautorisierung in serviceauth-go mit Methodenklassen in mehr als zehn Diensten, Server- und Autorisierungsobjekte für einen Teil der Dienste	teilweise
I2	NetworkPolicy je Dienst, Default-Deny durch Kyverno auf Anforderung, Zugriffsregeln des Overlay-Netzes als Datei mit Tests, zuschaltbarer Flutschutz und Beschränkung der Spielports auf dem Knoten, zuschaltbare Zulassungstore im Edge-Proxy	teilweise
I3	Signatur- und Digest-Richtlinie als abgeschalteter Entwurf im Audit-Modus, Signatur, Scan, Stückliste und Herkunftsnachweis im Zusatzjob einer ersten Dienstgruppe, Prüfsummen für heruntergeladene Werkzeuge in mehreren Pipelines	teilweise
I4	Gehärteter Container, unprivilegiertes Konto je Server als Standard, Egress-Sperre, Selbstprüfung. Die zweite Isolationsgrenze ist als Option angelegt, dedizierte Knoten gehören zum Ausbaupfad	teilweise
I5	LUKS-Einrichtung mit Schlüsselhinterlegung im Gateway (zuschaltbar), Hüllverschlüsselung im Sicherungsdienst (zuschaltbar), Transit-Verschlüsselung von Nutzerdaten	teilweise

Cloud Spheres UG (haftungsbeschränkt)
Registered Office: Paderborn
Address: Biberweg 24, 33102 Paderborn
Registration Court: Amtsgericht Paderborn
Commercial Register Number: HRB 18602
Managing Directors: Silas Timmermann, Alan Khudhur

I6	Kurzlebige Zertifikate im Mesh und in der PKI, Erneuerung der Vault-Tokens der Dienste	teilweise
I7	Hash-Kette mit Signatur, Verankerung, eingeschränkten Datenbankrechten und Verifikation im Audit-Log-Dienst (zuschaltbar)	teilweise
I8	Gestufte Reaktion durch Talon (Trockenlauf je Regel) und Node-Agent, Schattenmodus mit Auswertung, Beweispaket und Latenzmessung auf dem Knoten (zuschaltbar). Die Messung an echten Ereignissen gehört zum Ausbaupfad	teilweise
I9	Abgleich der Regeldateien mit den ausgelieferten Ressourcen in der CI, Kontrolltests der Alarmkonfiguration und der Alarmregeln, CI-Prüfung der Talon-Regeln auf Abweichung	teilweise
I10	Fail-Closed für kritische Methoden und für den Widerruf privilegierter Token, Turnstile-Prüfung mit konfigurierten Schlüsseln	teilweise
I11	Getrennte Kontrollen je Schicht mit eigenen Schlüsseln und Konfigurationsquellen. Eine Abhängigkeitsanalyse gehört zum Ausbaupfad	teilweise
I12	Escrow des Sicherungsschlüssels mit Schwellenaufteilung. Weitere Härtung der Schlüsselhierarchie und der Änderungsüberwachung gehört zum Ausbaupfad	teilweise

12 Grenzen der Architektur und Ausbaupfad

12.1 Grenzen

Auch diese Architektur ist nicht unverwundbar, und ihre Grenzen gehören zur Beschreibung. Eine bisher unbekannte Schwachstelle im Kernel, im Hypervisor oder in einer Sandbox-Laufzeit kann die Isolation durchbrechen, und mehrere voneinander unabhängige Grenzen machen das schwieriger, aber nicht unmöglich. Der Hypervisor unseres Infrastrukturanbieters ist eine Vertrauensannahme, weil wir ihn nicht prüfen können. Ein Innentäter, der mehrere Schlüsselanteile oder mehrere Freigaben gleichzeitig kontrolliert, umgeht die Aufteilung. Eine Kompromittierung der vorgelagerten Software, aus der die Images gebaut werden, wird durch Herkunftsnachweise erschwert und nicht verhindert. Volumetrische Angriffe oberhalb der Kapazität des vorgelagerten Netzes wirken sich auf die Verfügbarkeit aus, und Fluten mit gefälschten Absendern lassen sich durch Grenzen je Quelladresse nicht von legitimem Verkehr trennen. Die Fehlalarmfreiheit der Tripwires gilt für getestete Konfigurationen und nicht für jeden denkbaren Mod.

Cloud Spheres UG (haftungsbeschränkt)
Registered Office: Paderborn
Address: Biberweg 24, 33102 Paderborn
Registration Court: Amtsgericht Paderborn
Commercial Register Number: HRB 18602
Managing Directors: Silas Timmermann, Alan Khudhur

Das Ziel ist deshalb nicht Unverwundbarkeit, sondern ein System, in dem jeder erfolgreiche Angriff mehrere unabhängige Grenzen überwinden muss, in dem er in bekannter Zeit erkannt und begrenzt wird und in dem sich hinterher belegen lässt, was geschehen ist. Dieses Paper wird bei Änderungen der Architektur fortgeschrieben. Jede Änderung einer Aussage erhöht die Hauptversion, und der Änderungsverlauf am Ende zeigt, was sich wann geändert hat.

12.2 Ausbaupfad

Die folgenden Bausteine sind entworfen, aber im Quelltext nicht oder nur teilweise vorhanden. Sie bilden den Ausbaupfad der Architektur, und jede Fassung dieses Papers nennt, welche davon hinzugekommen sind. Zur Isolation gehören dedizierte Knoten für einzelne Kunden als zweite Grenze sowie weitere Isolationsgrenzen und ihre regelmäßige Überprüfung. Dazu kommt die weitere Härtung der Netz- und Alarmierungsebene, der Authentifizierung, der Schlüsselhierarchie, der Änderungsüberwachung und der Beweissicherung im Cluster.

In der Erkennung und Reaktion gehören dazu die Messung der Reaktionszeit an echten Ereignissen, Tests mit echten Auslösern und der weitere Ausbau der Beobachtung. Zu den Kontrollen gehören außerdem die flächendeckende Anwendung der Kontrollen aus den Kapiteln 6 und 7, die Ausweitung der Prüfungen in der Build-Pipeline und der Einstufung kritischer Methoden auf alle Dienste und die vollständige Rotation des Vertrauensankers im Mesh. Für den Entwurfsbeitrag gehören dazu die Erhebung der Goldläufe über den gesamten Katalog mit Bindung der Katalogaufnahme an das Gate und weitere Arten von Ködern. Wird ein Baustein umgesetzt, erhöht das die Hauptversion, und der Änderungsverlauf nennt ihn.

13 Literatur

Gilman, E., Barth, D. (2017): Zero Trust Networks. O'Reilly Media.

National Institute of Standards and Technology (2020): Zero Trust Architecture. NIST Special Publication 800-207. https://nvlpubs.nist.gov/nistpubs/SpecialPublications/NIST.SP.800-207.pdf

National Institute of Standards and Technology (2017): Application Container Security Guide. NIST Special Publication 800-190. https://nvlpubs.nist.gov/nistpubs/SpecialPublications/NIST.SP.800-190.pdf

National Security Agency, Cybersecurity and Infrastructure Security Agency (2022): Kubernetes Hardening Guidance, Version 1.2. https://media.defense.gov/2022/Aug/29/2003066362/-1/-1/0/CTR_KUBERNETES_HARDENING_GUIDANCE_1.2_20220829.PDF

Kubernetes Authors: Network Policies; Multi-tenancy. https://kubernetes.io/docs/concepts/services-networking/network-policies/

Linkerd Authors: Automatic mTLS; Architecture. https://linkerd.io/2/features/automatic-mtls/

HashiCorp: Vault documentation, Kubernetes auth method, Transit secrets engine, Seal wrap and auto-unseal. https://developer.hashicorp.com/vault/docs

Falco Authors: Falco documentation. https://falco.org/docs/

Falco Talon Authors: Falco Talon documentation. https://falco-talon.github.io/docs

Kyverno Authors: Kyverno documentation. https://kyverno.io/docs/

Cloud Spheres UG (haftungsbeschränkt)
Registered Office: Paderborn
Address: Biberweg 24, 33102 Paderborn
Registration Court: Amtsgericht Paderborn
Commercial Register Number: HRB 18602
Managing Directors: Silas Timmermann, Alan Khudhur

Internet Engineering Task Force: RFC 6238, TOTP: Time-Based One-Time Password Algorithm (2011). RFC 7239, Forwarded HTTP Extension (2014). RFC 7519, JSON Web Token (2015). RFC 7516, JSON Web Encryption (2015). RFC 8555, Automatic Certificate Management Environment (2019). RFC 4987, TCP SYN Flooding Attacks and Common Mitigations (2007). RFC 3704 (BCP 84), Ingress Filtering for Multihomed Networks (2004). https://www.rfc-editor.org

Tarreau, W.: The PROXY protocol, Versions 1 and 2. https://www.haproxy.org/download/1.8/doc/proxy-protocol.txt

The Go Authors: Go Fuzzing. https://go.dev/doc/security/fuzz/

Cloud Spheres UG (haftungsbeschränkt)
Registered Office: Paderborn
Address: Biberweg 24, 33102 Paderborn
Registration Court: Amtsgericht Paderborn
Commercial Register Number: HRB 18602
Managing Directors: Silas Timmermann, Alan Khudhur