USE CASE 12

GOLDILOCK FIREBREAK™ – AI KILL SWITCH USE CASE

AI Needs A Kill Switch: Human Control Over Autonomous Systems
Hardware-Enforced Isolation for AI-Enabled Infrastructure, Autonomous Systems, and Critical Operational Environments

EXECUTIVE OVERVIEW

Artificial intelligence is rapidly becoming embedded across critical infrastructure, operational technology (OT), datacentres, industrial automation, and security operations. As AI systems become increasingly autonomous and capable of acting at machine speed, organisations face a profound new challenge: maintaining human control when software systems are compromised, manipulated, operating beyond human response times, or experiencing unintended drift.

Governments, regulators, and critical infrastructure operators worldwide are actively debating AI safety, emergency intervention mechanisms, and **AI kill-switch capabilities**. The fundamental security challenge is no longer simply preventing cyber attacks. It is ensuring that organisations can contain AI-driven incidents, limit attack blast radius, preserve operational continuity, and maintain ultimate authority over critical infrastructure when software defenses inevitably fail.

Goldilock FireBreak™ delivers a hardware-enforced Layer-1 isolation controller—functioning as an unassailable **Physical AI Kill Switch**. By creating an independent, physical air-gap triggered via non-IP out-of-band commands, FireBreak provides the ultimate fail-safe: enabling organisations to instantly isolate compromised networks, protect high-value AI assets, and guarantee human intervention at machine speed.

THE CHALLENGE: THE RESILIENCE GAP IN AI-ENABLED ENVIRONMENTS

Modern enterprise and operational environments increasingly depend on software-based security controls to manage and defend AI-enabled systems. However, those very software controls operate within the exact same environments they are intended to protect.

As artificial intelligence accelerates both offensive cyber attacks and autonomous decision-making loops, relying solely on software firewalls, virtual local area networks (VLANs), and logical security guardrails creates a dangerous resilience gap. When software defenses are overwhelmed, exploited by zero-day vulnerabilities, or manipulated by autonomous malware, organisations are left with no independent means of containment.

Organisations require deterministic, infrastructure-level controls that remain fully operational and accessible even when software-based security has been completely compromised. They require a physical **AI Kill Switch.**

PAIN POINTS & CRITICAL THREAT VECTORS

- **Machine-Speed Exploitation:** Advanced threat actors and autonomous malware identify network weaknesses, exploit zero-day vulnerabilities, and execute attack paths at machine speed, far exceeding human decision and response cycles.
- **Accelerated Lateral Movement:** AI-enabled attacks dramatically accelerate lateral movement across interconnected IT, OT, and cloud environments, expanding the attack blast radius before traditional Security Operations Center (SOC) teams can react.
- **Inherent Software Vulnerabilities:** Software security tools share operating environments, kernels, and network stacks with the systems they protect, leaving them susceptible to elevation of privilege, configuration tampering, and adversarial AI bypass.
- **Human Response Lag:** When autonomous systems operate outside intended operational guardrails, decisions occur in milliseconds. Human operators cannot intervene effectively without an independent, deterministic physical kill switch.
- **Rogue AI & Model Drift:** Autonomous systems in industrial automation, grid management, or cooling control can experience unintended drift or command manipulation, leading to kinetic equipment damage if not physically contained.
- **AI Model & Data Compromise:** Multi-billion dollar LLM training weights, proprietary datasets, and AI model supply chains represent primary targets for state-sponsored exfiltration, data poisoning, and persistent backdoor injection.
- **Severe Operational Paralysis:** Once an AI-driven attack spreads unchecked across converged networks, incident response teams are forced into full network shutdowns, causing catastrophic downtime and financial loss.
- **Compliance & Governance Demands:** Regulatory frameworks (including EU NIS2, DORA, and national CNI mandates) increasingly require demonstrable governance, auditable intervention records, and provable operational resilience.

AI RESILIENCE THROUGH DEEP SEGMENTATION

AI resilience reframes the notion that cybersecurity is about building ever-higher walls. Advanced AI can now identify weaknesses, exploit vulnerabilities, and move through networks at a speed that makes prevention alone an increasingly unrealistic strategy.

AI resilience is therefore defined by an organisation's ability to maintain control and continue operating even when its defences have been breached by autonomous, fast-moving threats.
The speed and scale of AI-driven attacks fundamentally changes incident response. Organisations must shift from **Fix First to Contain First**, recognising that software-based security alone is no longer sufficient. Governance, identity, access management, backup, and recovery remain essential, but resilience must also be embedded directly into the network architecture.

Deep segmentation is a critical component of this approach. By creating deterministic physical boundaries around critical systems, operational domains, and sensitive data, organisations limit an attack's blast radius, prevent AI-driven lateral movement, and create valuable time for both technical and executive teams to respond—while enabling as much of the organisation as possible to continue operating.

AI resilience is not a replacement for cyber resilience—it is its evolution. As AI-driven threats continue to raise the bar, organisations need a layered approach that combines software-based protection with independent, hardware-enforced infrastructure controls. By embedding deep segmentation and an independent **AI Kill Switch** into the architecture, organisations establish deterministic security boundaries that preserve operational continuity and protect critical systems, even when software-based defences have been compromised.

THE FIREBREAK SOLUTION & CORE BENEFITS

Goldilock FireBreak™ provides an appliance-based, patented physical connection controller that allows authorised personnel to remotely and physically connect or disconnect network segments using non-IP, out-of-band commands.

Key Solution Capabilities

- **Deep segmentation:** Create deterministic, hardware-enforced physical boundaries around critical systems, operational domains, high-density compute clusters, and sensitive data.
- **Infrastructure AI Kill Switch:** Authorised personnel can physically disconnect infrastructure and network connections instantaneously whenever anomalous behavior, rogue AI execution, or systemic compromise is detected.
- **Independent Out-of-Band Control:** Control channels remain completely separate and independent from the production IP network being protected, ensuring management accessibility even during a total software breach.
- **Human Authority:** Guarantees that human operators retain ultimate authority over network connectivity, isolation, recovery, and operational decision-making.
- **Threat-to-Action Automation:** Integrates seamlessly with enterprise SIEM, SOAR, and Intrusion Detection Systems (IDS) to trigger automated physical containment actions at machine speed.
- **Operational Continuity:** Rapidly contains security incidents, minimises the attack blast radius, and enables unaffected business sectors and operational units to continue operating without disruption.

REAL-WORLD FIREBREAK DEPLOYMENT SCENARIOS

Scenario 1: Automated Machine-Speed Threat Containment *(Circuit Breaker)*

- **Challenge:** Containing autonomous malware or adversarial AI tools that traverse IT/OT boundaries within milliseconds, outpacing human SOC intervention.
- **FireBreak Solution:** Deploy FireBreak as an automated **Infrastructure Kill Switch** between core enterprise routing and sensitive operational or compute segments. Integrated via automated Threat-to-Action APIs with SIEM/SOAR platforms, FireBreak physically severs Layer-1 fiber or copper connections the moment anomalous lateral movement is detected. Threat propagation is stopped dead at the physical layer, isolating the affected segment while rest-of-site operations proceed normally.

Scenario 2: Protecting High-Value AI Models & HPC Clusters *(Invisibility Cloak)*

- **Challenge:** Safeguarding multi-million-dollar proprietary AI training weights, LLM codebases, and research data from online exfiltration and data poisoning attacks.
- **FireBreak Solution:** Enforce a "Disconnected-by-Default" architecture around GPU clusters and AI storage arrays. Using FireBreak as an on-demand kill switch mechanism, assets remain physically air-gapped from all internal and external networks. Layer-1 connectivity is opened strictly during authorized, time-bound training or data ingestion sessions and immediately severed upon completion—rendering sensitive IP completely invisible and inaccessible to remote threat actors.

Scenario 3: Autonomous Industrial & OT Infrastructure Control *(Human Authority & Time Lock)*

- **Challenge:** Protecting SCADA systems, industrial robotics, and power distribution grids from rogue autonomous AI commands or remote firmware manipulation.
- **FireBreak Solution:** Position FireBreak controllers between centralized AI analytics platforms and local physical programmable logic controllers (PLCs). Operating networks run in physical isolation. Authorized maintenance crews utilize encrypted out-of-band triggers (such as secure SMS or physical key switches) to activate a physical connection window for telemetry updates, guaranteeing human kill-switch authority over kinetic systems.

Scenario 4: Secure AI Supply Chain Quarantine & Inspection *(Security Airlock)*

- **Challenge:** Safely ingesting open-source AI libraries, vendor model updates, and third-party software without risking immediate cross-contamination across production environments.
- **FireBreak Solution:** Establish a hardware-enforced inspection airlock. Incoming updates are downloaded into a physically isolated staging environment. FireBreak physically disconnects the external ingress pathway *before* opening the internal pathway to the sandbox network, ensuring that latent backdoors or malicious code cannot traverse into core enterprise systems during verification.

Scenario 5: Immutable Air-Gapped AI Model Checkpoints & Backup Recovery

- **Challenge:** Preventing AI-driven ransomware from simultaneously infecting live production databases and secondary backup servers required for disaster recovery.
- **FireBreak Solution:** Place FireBreak as a physical air-gap barrier shielding immutable backup servers and gold-master AI checkpoints. Systems connect exclusively during automated backup windows and immediately return to a zero-connectivity state. In the event of a catastrophic production compromise, clean data restoration is guaranteed from physically pristine environments.

WHY ENTERPRISE & INFRASTRUCTURE LEADERS CHOOSE GOLDILOCK FIREBREAK

Absolute Invisibility = Zero Attack Surface
Software firewalls block network traffic; Goldilock FireBreak eliminates the physical pathway entirely. When critical AI systems, datacentre compute racks, SCADA nodes, or backup arrays are physically isolated by FireBreak, they carry no IP address, no open ports, and zero digital footprint. An autonomous threat cannot scan, exploit, or target a system that physically does not exist on the network.

True Zero Trust, Enforced by Hardware
FireBreak elevates Zero Trust from a logical policy to a physical reality. By implementing deep physical segmentation, organisations enforce just-in-time, role-based access to critical infrastructure. Systems connect only when explicitly required and strictly for the authorized duration, backed by immutable hardware-level audit logs for full compliance verification.

Rapid, Non-Disruptive Deployment
Engineered as an appliance-based Layer-1 physical controller, FireBreak drops seamlessly into existing optical fiber and copper network architectures. It operates completely independent of protocol, operating system, or software stack—requiring no complex endpoint agents, no changes to existing IP addressing schemes, and no costly forklift redesigns.

Proven Protection for Critical National Infrastructure
Trusted across defense, government, financial services, and energy sectors, FireBreak is battle-tested in environments where operational downtime carries severe societal and economic consequences. Its out-of-band architecture guarantees deterministic control when all software defenses have been bypassed.

Peace of Mind in a Volatile Threat Landscape
Knowing that mission-critical assets can be instantly rendered unreachable provides unmatched executive confidence. FireBreak simplifies regulatory compliance under NIS2, DORA, ISO 27001, and national CNI mandates, demonstrating a tangible commitment to physical resilience and human authority over autonomous systems.

SUMMARY

Goldilock FireBreak offers a pioneering, hardware-enforced cybersecurity solution purpose-built for the AI era. By providing instant physical-layer disconnection of critical compute infrastructure, AI model repositories, OT networks, and backup environments, FireBreak delivers absolute protection against software bypass, autonomous threat escalation, and machine-speed lateral movement.

This deterministic approach establishes true offline security when systems are disconnected—a capability that no software-based firewall, virtual VLAN, or logical air-gap can match. By embedding deep segmentation and an independent physical kill switch into network architecture, organisations achieve true AI resilience: preserving operational continuity, protecting critical assets, and retaining ultimate human authority over autonomous systems.

AI Needs A Kill Switch. FireBreak Provides One.

Discover how Goldilock FireBreak combines hardware-enforced isolation, independent out-of-band control, and deep segmentation to help organisations build AI resilience, maintain human authority over autonomous systems, and preserve operational continuity—even when software-based defences have been compromised.