



The Sandbox Era Has Begun. The Trusted Execution Era Is Next.

A new architectural layer is emerging for enterprise AI –
and it isn't another security product.

Contents

01	The Convergence – Every Hyperscaler Just Admitted Something	3
02	AI Isn't Creating a Security Problem. It's Exposing an Infrastructure Problem.	4
03	The Inflection Point – What's Happening, By the Numbers	5
04	The Confidence Gap	6
05	Cloud Providers Solved One Problem	7
06	The Exception Economy	8
07	Precedent – Why Separation Works	9
08	The Rise of Trusted Execution	10
09	From Code Execution to Enterprise Execution	11
10	Beyond Isolation – What Trusted Execution Demands	12
11	Conclusion – One Final Thought	13
12	Supporting Data & Sources	14



EXECUTIVE SUMMARY

The Sandbox Era Has Begun. The Trusted Execution Era Is Next.

Over the past year, AWS, Google Cloud, Microsoft Azure, and Cloudflare have each shipped isolated execution environments for AI-generated code. Four competitors, four different architectures, one shared conclusion: some work no longer belongs on trusted infrastructure.

That convergence is not a security feature. It is the first evidence of a new architectural layer forming beneath enterprise AI.

Code sandboxes answer a narrow question – where should AI-generated code run? Enterprises are asking a much bigger one: where should uncertain work run, when that work involves proprietary source code, customer data, security telemetry, and autonomous agents acting on an organization's behalf?

The cloud providers proved the containment primitive. The next decade belongs to the execution layer built on top of it – Trusted Execution.

The Convergence – Every Hyperscaler Just Admitted Something

Within a single quarter, the four largest cloud providers completed the same architectural bet. AWS shipped Lambda MicroVMs, giving each agent session a dedicated Firecracker virtual machine with up to eight hours of runtime and a suspend-resume cycle that preserves memory, disk, and running processes. Google Cloud followed with Cloud Run sandboxes, using gVisor kernel interception, alongside a separate GKE Agent Sandbox for containerized workloads. Microsoft got there first: Azure Container Apps dynamic sessions have run on Hyper-V isolation boundaries since 2024, and Microsoft now reports that Copilot alone consumes more than 400,000 of those sessions every day. Cloudflare built Sandboxes on top of its Containers platform, with each session isolated in its own VM and orchestrated through Workers and Durable Objects.

AWS

Lambda MicroVMs

Firecracker-based microVMs. Up to eight hours per session, with suspend-resume state preservation.

GOOGLE CLOUD

Cloud Run Sandboxes / GKE Agent Sandbox

gVisor kernel interception; sandboxes share the parent instance's CPU and memory at no added cost.

MICROSOFT AZURE

Container Apps Dynamic Sessions

Hyper-V isolation since 2024. Over 400,000 Copilot sessions run daily.

CLOUDFLARE

Sandboxes on Containers

Per-session VM isolation, orchestrated through Workers and Durable Objects.

Four vendors, four genuinely different isolation stacks, one identical product thesis: send in untrusted code, and the platform runs it somewhere the calling application can't reach, with credential exposure governed separately. As The New Stack put it, every cloud now agrees that agents should not run code on the host – which simply moves the argument to who governs what that code does once it's safely contained.

REPLICA'S HEAD START

Replica was built on this thesis in 2017 – nearly a decade before any hyperscaler shipped an isolated execution primitive. The difference is scope: Replica was never limited to isolating a single code execution. It governs the selection of applications, compute, storage, and network access required to complete human and AI work at scale, across every hyperscaler – not locked inside just one.

SECTION 02

AI Isn't Creating a Security Problem. It's Exposing an Infrastructure Problem.

For thirty years, cybersecurity operated on a remarkably consistent assumption: risky work would happen on trusted infrastructure, and controls – firewalls, identity, endpoint security, DLP, SIEM, governance – would make that risk manageable.

Every major category in the industry attempts to answer the same underlying question: how do we make risky work safer?

That model worked, until AI changed the nature of the work itself. Organizations now ask people – and increasingly autonomous agents to:

- Execute unfamiliar code and evaluate AI models
- Connect proprietary source code to third-party services
- Browse unknown or hostile infrastructure
- Investigate criminal networks and detonate malware
- Use sensitive enterprise data to test new technology

None of this is an edge case anymore. It is becoming routine business, and the infrastructure was never built for it.

Cloud providers just proved that AI-generated code should never execute directly on trusted infrastructure. Code execution, though, is only the first chapter.

SECTION 03

The Inflection Point – What’s Happening, By the Numbers

Several forces are converging at once – adoption, governance readiness, trust, and identity sprawl are all moving in the same uncomfortable direction simultaneously.

150,000+

AI agents the average Fortune 500 enterprise will run by 2028, up from fewer than 15 in 2025

GARTNER

13%

of organizations believe they have adequate AI agent governance in place today

GARTNER

40%+

of agentic AI projects will be canceled by the end of 2027 over cost, risk, or unclear ROI

GARTNER

34%

of enterprises say they trust the actions their AI agents are taking, despite 86% moving past pilots

BOOMI / FUTURECIO

80:1

ratio of non-human identities to human users in the average enterprise

KPMG

1 in 3

organizations report mature governance and controls for agentic AI, versus rapidly advancing deployment

MCKINSEY

Pause on what these numbers say together. The problem isn’t adoption, capability, or model quality. Agents are being deployed faster than any organization can govern, trust, or even count them. This is the Confidence Gap – and it is widening in both directions at once.

SECTION 04

The Confidence Gap

From “Does It Work?” to “Will It Work With Our Data?”

For decades, enterprise technology evaluation answered a straightforward question: does it work? AI changes that question completely. Today, enterprises need to know whether a solution will work with their proprietary source code, their customer records, their security telemetry, their intellectual property.

SYNTHETIC DATA

Can demonstrate that a technology functions. It cannot demonstrate how that technology will perform inside a specific enterprise.

ENTERPRISE DATA

Is the only thing that can demonstrate organizational confidence – and it’s exactly what most evaluation environments can’t safely use.

PwC FINDS

79%

of companies are already adopting AI agents

66%

report measurable value from those agents

1 in 4+

still name lack of trust as a top-three barrier

Organizations can prove an AI solution works. Far fewer can prove it will work for them – and that gap between technical capability and organizational confidence is becoming one of the largest barriers to enterprise AI adoption.

SECTION 05

Cloud Providers Solved One Problem

Agent sandboxes are a genuinely elegant solution to a genuinely important question: where should AI-generated code execute? The answer – inside an isolated sandbox – is good, and necessary. But it isn't sufficient for what enterprises actually need to accomplish.

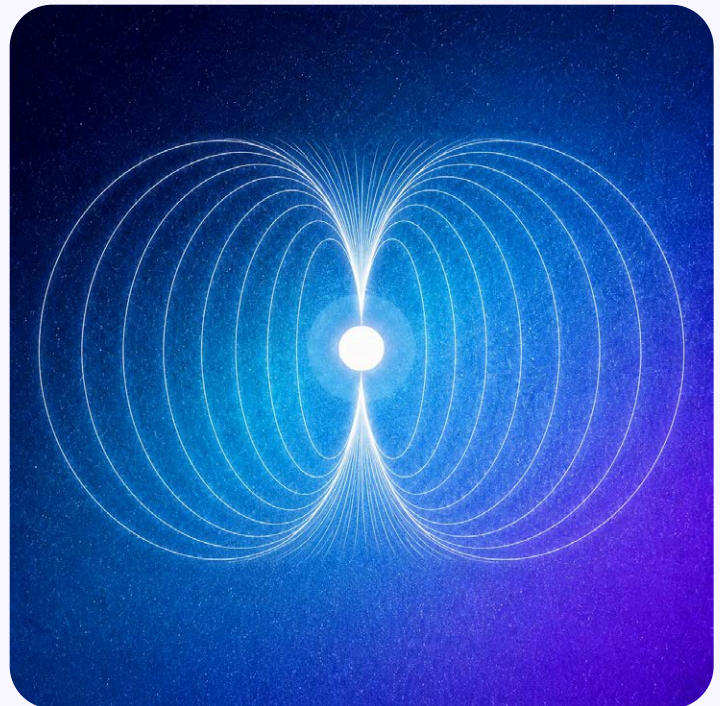
What sandboxes answer

- Where does this code execute?
- Is it isolated from the host?
- Can the session be destroyed cleanly?

What enterprises still need

- Can the agent safely use sensitive customer data?
- Can it access proprietary source code or browse hostile infrastructure?
- Can every action be governed and audited?

Those questions have almost nothing to do with running Python. They're about safely executing enterprise work – a fundamentally different problem than code containment, and one the industry's current “agent security” framing is too narrow to capture.



The Exception Economy

Where Risky Work Actually Happens Today

Replica Cyber's research set out to understand what organizations do when risky work doesn't fit safely inside existing infrastructure. The pattern was strikingly consistent: they don't refuse the work. They find a way around the infrastructure – security exceptions, temporary environments, personal devices, ad hoc virtual machines, unmanaged cloud resources. The work still happens. It simply happens outside the systems built to support it.

RESEARCH INSIGHT – THE EXCEPTION ECONOMY

100%

of organizations surveyed granted security or compliance exceptions in the previous 12 months to keep high-risk digital work moving.

When every organization grants exceptions, the issue is no longer an isolated policy decision – it is evidence that business requirements have outgrown the infrastructure built to support them.

SOURCE: REPLICIA CYBER, TRADING SAFETY FOR SPEED: THE RISE OF THE EXCEPTION ECONOMY, APRIL 2026.

The same research found that organizations abandon approximately 20% of potential high-risk digital work outright, because they can't find a safe place to perform it. One-fifth of innovation never happens – not for lack of ideas, but for lack of environments. That's not a governance failure. It's an infrastructure failure, and it's exactly why every major cloud provider is now investing in sandboxes: isolation isn't merely another security feature anymore. It's becoming infrastructure.

SECTION 07

Precedent – Why Separation Works

Every mature industry that handles inherently dangerous processes has already learned this lesson. It is not a new idea – it is a well-tested one, applied to a new category of risk.



NUCLEAR

Reactors are separated from personnel and administrative operations.



MANUFACTURING

Hazardous processes are isolated from the broader facility.



FINANCE

Duty separation and vault security contain insider and external risk.



CLOUD

Tenant isolation ensures one customer's compromise cannot reach another's.



MARITIME

Watertight compartments prevent a single breach from sinking the ship.



CRITICAL INFRASTRUCTURE

Operational systems are separated from business networks by design.

THE PRINCIPLE

When an activity is inherently risky, separate it from the assets you cannot afford to lose. No policy can make unknown code trustworthy, and no governance framework can make uncertainty disappear – but containment can limit what happens if something goes wrong.

SECTION 08

The Rise of Trusted Execution

Historically, cybersecurity evolved through distinct layers – network security, endpoint security, identity, cloud security, data security, and most recently AI governance. Each addressed an emerging problem in its moment. Trusted Execution is something different: not another security control, but a dedicated execution layer where uncertainty can exist without exposing the systems organizations rely on every day.

TRUSTED EXECUTION SUPPORTS

**A purpose-built execution environment
for uncertain enterprise work**

AI Evaluation

Agent Development

Source Code Analysis

Malware Detonation

Threat Intelligence

Fraud Investigation

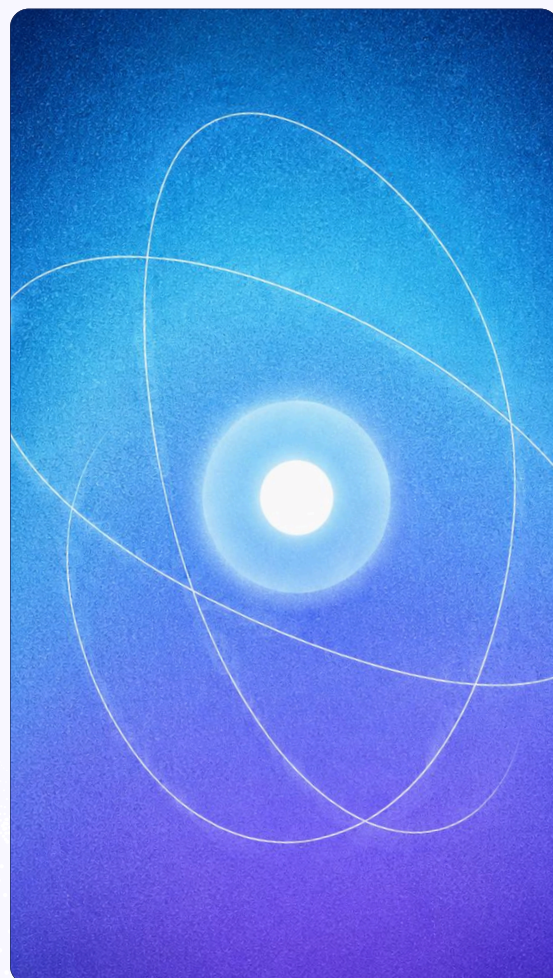
Vendor Evaluation

Regulated Collaboration

Internet Research

Sensitive Workflows

Not because these activities are malicious –
because they are uncertain, and uncertainty
deserves a dedicated place to live.



SECTION 09

From Code Execution to Enterprise Execution

Every foundational infrastructure shift looks narrow at first and becomes the substrate for everything built afterward. Containerization looked like a packaging convenience; it went on to transform software delivery by packaging applications and dependencies into a portable, consistent unit that runs anywhere. Cloud didn't relocate workloads – the servers still sit on racks in real datacenters – it separated the workload from infrastructure ownership, and that separation is what enabled elastic scale. Zero trust didn't eliminate trust; it changed how access decisions get made – continuously and contextually, rather than once at a perimeter.

Agent sandboxes are following the same arc. What looks today like a narrow containment feature is becoming the substrate underneath enterprise AI. And once a layer becomes substrate, competing on it stops making sense.

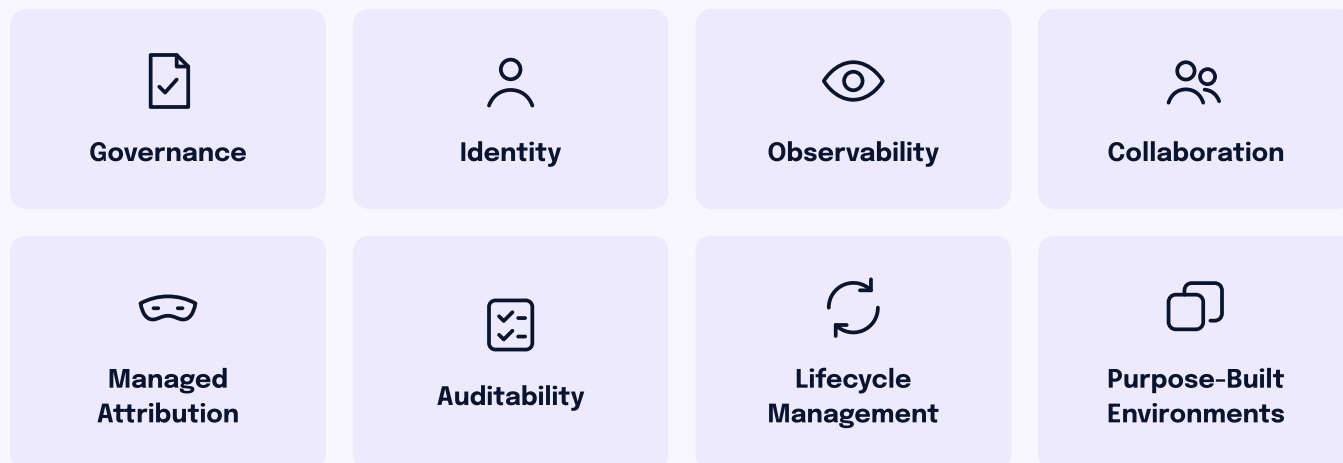
The differentiation moves up the stack: from where does code execute, to where does enterprise work execute.

Governance asks how to reduce the probability something bad happens. Containment asks what happens if it does. Those aren't competing philosophies – they're complementary, and most organizations have built real infrastructure for only one of them. Identity still matters. Detection still matters. Policy still matters. None of that answers where uncertain work should actually happen.

SECTION 10

Beyond Isolation – What Trusted Execution Demands

A Firecracker VM or a gVisor sandbox isolates code from the host. It does nothing about what that code does with the credentials a developer deliberately handed it, and it says nothing about who's accountable when an autonomous agent acts. Trusted Execution requires considerably more than a containment boundary:



This is also where the non-human identity problem becomes unavoidable – with machine and agent credentials already outnumbering human users 80 to 1 in the average enterprise, identity governance built for people alone is no longer sufficient.

Trusted Execution treats agent identity, access, and lifecycle as first-class problems – not an afterthought bolted onto a code sandbox.

CONCLUSION

One Final Thought

“If the last twenty years of cybersecurity were about protecting trusted infrastructure, the next twenty may be about building infrastructure for untrusted work.”

The industry is already moving in that direction – the hyperscalers just proved it, four different ways, in the same quarter. The open question isn't whether isolated execution becomes standard infrastructure. It's whether enterprises keep treating sandboxes as a developer tool, or start treating them as the foundation of an entirely new execution layer.

Because once you see that shift, something becomes clear: the sandbox isn't the destination. It's the beginning.

**TRUSTED EXECUTION IS THE NEXT
ARCHITECTURAL LAYER OF ENTERPRISE AI**

Replica Cyber · Trusted Execution for enterprise AI · July 2026

SUPPORTING DATA & SOURCES

Sources

Cited in order of first reference. Figures reflect the most recently published data available as of July 2026.

- 1 AWS News Blog**
AWS Lambda introduces MicroVMs for isolated sandboxes with full lifecycle control.
- 2 Google Cloud Blog**
Google Cloud Run sandboxes enter public preview, built on gVisor isolation.
- 3 Microsoft Azure Documentation**
Azure Container Apps dynamic sessions: Hyper-V isolation and Copilot session volume.
- 4 The New Stack**
Janakiram MSV, on AWS, Google Cloud, Azure and Cloudflare all offering agent sandboxes, July 24, 2026.
- 5 Gartner Newsroom**
Gartner Identifies Six Steps to Manage AI Agent Sprawl, April 28, 2026.
- 6 Gartner Newsroom**
Gartner Predicts Over 40% of Agentic AI Projects Will Be Canceled by End of 2027, June 25, 2025.
- 7 FutureCIO**
Survey commissioned by Boomi: 86% of enterprises have deployed AI agents; 34% trust them.
- 8 KPMG International**
KPMG, Cybersecurity Considerations 2026: non-human identities outnumber human users 80 to 1.
- 9 McKinsey & Company**
State of AI Trust in 2026: Shifting to the Agentic Era – AI Trust Maturity Survey.
- 10 PwC**
PwC AI Agent Survey: enterprise adoption, realized value, and trust as a top-three barrier.
- 11 Cybersecurity Ventures**
Cybercrime projected to exceed \$15 trillion annually by the end of the decade.
- 12 Nasdaq Global Financial Crime Report**
Global financial crime exceeds \$4 trillion annually.
- 13 FBI IC3 2024 Annual Report**
Reported U.S. cybercrime losses increased 33% in 2024.
- 14 FTC 2024 Consumer Fraud Data**
Reported U.S. consumer fraud losses increased 25% in 2024.
- 15 Celent**
Financial institutions spend approximately \$190 billion annually on financial crime compliance.
- 16 LexisNexis True Cost of Fraud**
Every \$1 of fraud loss creates approximately \$5.75 in total downstream costs.
- 17 UNODC**
Money laundering accounts for approximately 2–5% of global GDP annually.
- 18 IBM Cost of a Data Breach Report**
Average global cost of a data breach.
- 19 Replica Cyber**
Trading Safety for Speed: The Rise of the Exception Economy, April 2026. Proprietary research.
- 20 Okta Newsroom**
AI Agents at Work 2026: Securing the Agentic Enterprise – shadow AI and digital labor governance.
- 21 Gartner**
Hype Cycle for Agentic AI, 2026: 17% of organizations have deployed AI agents; 60%+ expect to within two years.
- 22 Microsoft WorkLab**
2026 Work Trend Index: Agents, Human Agency, and Opportunity.