Journal of Computer Science and Technology Studies
ISSN: 2709-104X
DOI: 10.32996/jcsts
Journal Homepage: al-kindipublisher.com/index.php/jcsts

JCSTS
AL-KINDI CENTER FOR RESEARCH AND DEVELOPMENT

| RESEARCH ARTICLE

Dynamic Clean Room Orchestration Using Infrastructure as Code: A Vendor-Agnostic Reference Architecture and Identity-First Reconstitution Model for Ransomware Recovery

Venkata Ravi Sridhar Potunuru
Bright Horizons Family Solutions, Newton, Massachusetts, United States
Corresponding Author: Pending confirmation, **E-mail:** vrspotunuru@gmail.com

| ABSTRACT

Modern ransomware deliberately targets the systems an organization would rely on to recover, which makes a permanently standing recovery site a liability rather than an asset. This article proposes treating the recovery environment as something that is compiled on demand from version-controlled infrastructure definitions, rather than something that is continuously maintained and rebuilt in a deliberate, identity-first order. However, existing clean-room products, such as Commvault Cleanroom Recovery and Rubrik's cyber-recovery products, do not provide this isolation as an emergent property of the approach. Instead, they isolate as part of a proprietary control plane and order the reconstitution as a designed option, rather than an intrinsic part of the approach. The design and methodology synthesize current industry evidence on backup-repository targeting, ransomware dwell-time compression, and recovery-testing outcomes into a reference architecture with two distinguishing contributions: a vendor-agnostic isolation boundary generated from infrastructure-as-code rather than a running configuration, and an explicit identity-first reconstitution sequencing model in which directory and identity services are rebuilt and independently verified before any workload is restored. The main results, drawn from current industry benchmarking, indicate that backup repositories are targeted in 89% of ransomware attacks, that only 54% of organizations successfully use backups to restore encrypted data, the lowest rate in six years, and that organizations with tested, automated recovery procedures recover measurably faster than those without. The conclusion is that a dynamic, identity-first clean room, compiled on demand rather than maintained as standing infrastructure, offers a more defensible recovery posture than persistent recovery sites that share risk with the systems they are meant to protect, and that reconstitution order deserves the same engineering discipline as the automation that performs it.

| KEYWORDS

Infrastructure as Code, Ransomware Recovery, Clean Room, Identity-First Reconstitution, Disaster Recovery, Cyber Resilience

| ARTICLE INFORMATION

ACCEPTED: 15 August 2026 **PUBLISHED:** 26 September 2026 **DOI:** 10.32996/jcsts.2026.8.9.17

1. Introduction

Ransomware has shifted from a destructive but largely mechanical threat into a patient, identity-aware form of intrusion. Incident-response data compiled by Coveware places average operational downtime following an attack in the range of several weeks in recent reporting periods, and research from Halcyon indicates that the median dwell time between an attacker's initial access and the deployment of ransomware has compressed sharply, falling from roughly nine days in 2022 to four to five days in recent measurement periods (Halcyon, 2025). That compression matters: it leaves defenders less time to detect an intrusion before encryption begins, yet still gives attackers enough time to map directory services, escalate privileges, and locate backup infrastructure before triggering the payload.

That last behavior, deliberately targeting the systems an organization would use to recover rather than simply encrypting files, is what makes traditional disaster-recovery planning insufficient on its own. Sophos's State of Ransomware 2025 report found that backups were used to restore encrypted data in only 54% of incidents, the lowest rate in six years (Sophos, 2025), which indicates that having a backup is no longer the same as having a usable recovery path. This article argues that infrastructure as code, traditionally applied to deployment speed and configuration consistency, can be repurposed as a security control in cyber recovery by generating a temporary, isolated recovery environment only at the moment it is needed and by rebuilding that environment in an order that re-establishes trust before it restores workloads. The remainder of this article proceeds as follows: the literature review positions this work against existing commercial offerings and the broader infrastructure-as-code adoption trend; the methodology section details the proposed reference architecture; the results section presents the framework's validation and staged-cutover model along with current recovery-maturity benchmarking; and the conclusion synthesizes the argument and proposes directions for empirical validation.

2. Literature Review

2.1 The Limits of Persistent Recovery Sites

A standing recovery site is rarely as isolated as organizations assume. Veeam's 2025 Ransomware Trends Report found that 89% of organizations had their backup repositories targeted by ransomware actors, with an average of 34% of backup repositories modified or deleted by attackers once reached (Veeam, 2025). Security advisories tied to widely used enterprise backup platforms have repeatedly noted that backup servers joined to the same identity domain as production systems are particularly exposed, since a single compromised domain account can be enough to reach and disable the very systems meant to enable recovery.

This is why a warm or hot standby that shares identity services, network paths, or administrative credentials with production offers a false sense of security. Public threat intelligence on ransomware families documents a consistent pattern in which attackers map directory services and locate backup servers before deploying their payload, specifically to deny the victim a clean recovery path. A growing share of the industry has responded by separating recovery infrastructure from production trust boundaries entirely and by re-scanning backup data in an isolated staging area before reintroducing it to production (Veeam, 2025). Clean, isolated recovery has moved from a niche idea toward a mainstream expectation, and the CISA #StopRansomware Guide reflects this shift in its own emphasis on maintaining offline, encrypted, and regularly tested backups as a baseline control rather than an advanced practice (Cybersecurity and Infrastructure Security Agency, n.d.).

2.2 Positioning Against Existing Clean-Room Offerings

The isolated recovery environment is not a new idea, and this article does not claim it as one. Commvault Cleanroom Recovery provisions a secure, isolated environment in Microsoft Azure on demand; bills only for the period it is in use; integrates threat scanning; and uses last-clean-point identification to select a recovery point (Commvault, 2024; TechTarget, 2024). Rubrik offers comparable isolated-recovery capabilities: pre-staged recovery plans define the sequence, dependencies, and target environment for full application stacks, and Rubrik orchestrates recovery into an isolated AWS, Azure, or GCP environment with no connection to the compromised on-premises infrastructure (Rubrik, n.d.). Any serious proposal in this space must state plainly what it adds beyond those products, and this section does so directly rather than asserting novelty without comparison.

Two elements distinguish the approach proposed here. First, the isolation boundary is defined and generated from version-controlled infrastructure-as-code rather than provisioned and managed by a proprietary control plane, which keeps the reference architecture vendor-agnostic and, more importantly, means the network boundary does not exist as a running configuration for an adversary to discover ahead of time. A boundary compiled only at the moment of use, from a definition stored in source control, leaves no idle-state footprint to reconnoiter, unlike a standing isolated environment that, however well secured, still exists continuously as a target. Second, and central to this article, is an explicit identity-first reconstitution sequencing model: a defined, auditable order in which trust is re-established and independently verified before any workload is restored. Commercial offerings allow customizable recovery sequences, but the ordering discipline itself, and the argument that identity must be verified clean before anything depends on it, is treated here as a first-class engineering contribution rather than a configuration option left to the administrator's discretion during an active incident, when judgment is least reliable.

2.3 Infrastructure as Code as a Security Control

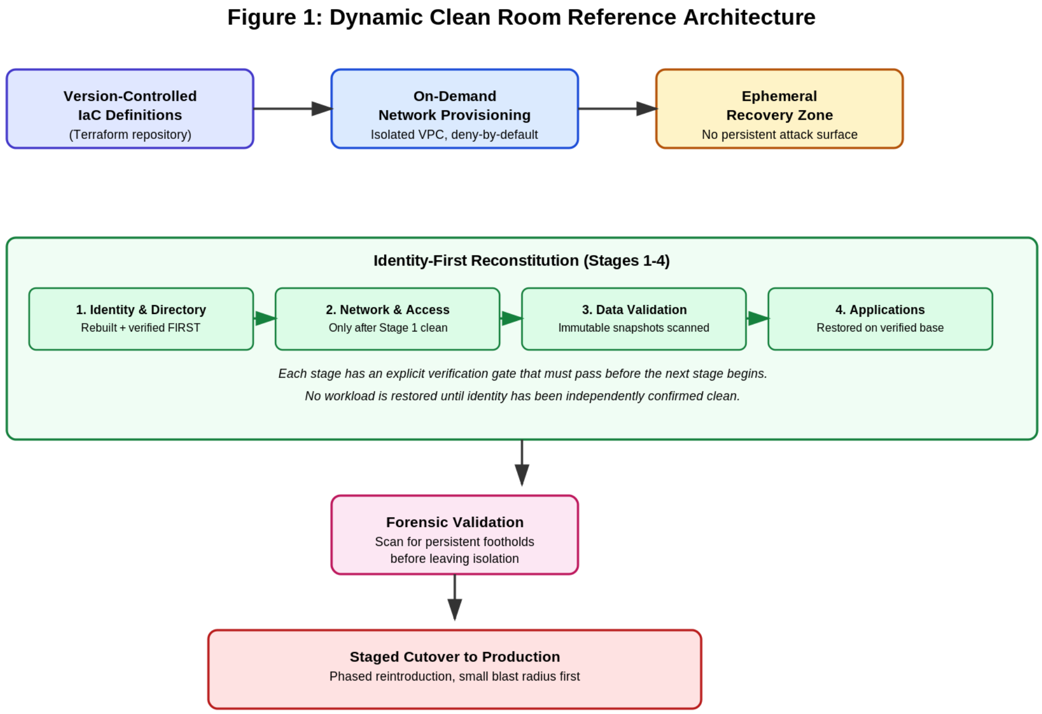
The infrastructure-as-code market has grown substantially, from an estimated $1.06 billion in 2024 to a projected $9.40 billion by 2034, a compound annual growth rate of approximately 24.39% (Precedence Research, n.d.-a). The related disaster-recovery-as-a-service market shows comparable momentum, projected to grow from $22.40 billion in 2025 to $195.71 billion by 2034 (Precedence Research, n.d.-b). This growth reflects infrastructure as code's traditional application to deployment speed and configuration consistency, not its use as a dedicated security control. This article's contribution is to reframe the same underlying capability, the ability to define infrastructure declaratively and regenerate it on demand, as a security property specifically for cyber recovery: an isolation boundary that exists only when needed carries fundamentally different risk characteristics than one that exists continuously, regardless of how well the continuously-running version is hardened.

3. Methodology: Reference Architecture for the Ephemeral Recovery Zone

3.1 Isolated Network Provisioning

The first building block is a network boundary that does not exist until it is needed. Using declarative definitions in a tool such as Terraform, an organization can describe an isolated virtual network together with strict traffic rules that deny all inbound and outbound connections to production by default. As the definition lives in a repository rather than as a long-running configuration, there is no persistent attack surface to discover and exploit ahead of time. Figure 1 depicts this reference architecture end to end, from the version-controlled definition through provisioning, reconstitution, validation, and staged cutover.

Figure 1. Dynamic Clean Room Reference Architecture



3.2 Presenting Clean Data Without Exposing Backup Control Systems

Once an isolated environment exists, it must be supplied with trustworthy data without exposing the broader backup platform to risk. Enterprise backup vendors increasingly separate the data plane, where backup data physically resides, from the control plane, where jobs are configured, so that compromise of one does not automatically grant access to the other. The clean-room architecture proposed here follows the same separation, presenting immutable snapshots to the isolated environment through a narrow, purpose-built integration rather than granting it broad access to the backup platform's management console. This topic has direct economic stakes: organizations whose backups are compromised during an attack face materially higher recovery costs than those whose backups remain intact (Veeam, 2025), which means protecting the integrity of the backup control plane during recovery is one of the larger variables determining total recovery cost, not a peripheral detail easily deferred to a later hardening phase.

3.3 Automated, Identity-First Reconstitution

Post-deployment configuration automation, using tools such as Ansible and Ansible Tower, rebuilds systems inside the clean room in a deliberate order rather than restoring everything simultaneously. As directory and identity services determine who and what every other system trusts, they are rebuilt and independently verified first. Only after that trust anchor is confirmed clean are network security controls and access policies rebuilt, followed by validation of the underlying data, and only then are business applications restored on top of that verified foundation. This ordering is the identity-first reconstitution sequencing model, summarized in Table 1.

Table 1. Identity-First Reconstitution Sequencing Model

Stage	Action	Verification Gate
1. Identity & Directory Services	Rebuilt and independently verified first, since every other trust decision depends on this layer	Directory service integrity confirmed clean before proceeding
2. Network Security Controls & Access Policy	Rebuilt only after Stage 1 is confirmed clean	Access policies validated against known-good baseline
3. Data Validation	Immutable snapshots scanned and confirmed against known-clean baselines	Forensic scan confirms no persistent footholds

Stage	Action	Verification Gate
4. Business Application Restoration	Restored only on top of the verified foundation from Stages 1-3	Staged cutover with phased monitoring before full production return

Sequencing matters independently of how quickly the automation runs. Halcyon's research on ransomware response indicates that automated, playbook-driven containment substantially outperforms manual, ad hoc response in both speed and consistency (Halcyon, 2025). Separately, CloudZero's analysis of the October 2025 AWS and Azure outages found that organizations that had tested their recovery procedures before those outages recovered significantly faster than those that had not (CloudZero, 2026), a finding about the value of rehearsed, disciplined processes rather than about cloud infrastructure choice per se. Taken together, these findings support the case that the order in which systems are rebuilt, and whether that order has been tested in advance, deserves the same engineering rigor as the automation that rebuilds them. An organization that automates its recovery pipeline but restores systems in an arbitrary or convenience-driven order gains speed without gaining the trust guarantees that ordered reconstitution is specifically designed to provide.

4. Results and Discussion

4.1 Forensic Validation and Staged Return to Production

Before any system leaves the isolated environment, it should pass automated scanning to confirm that no persistent footholds, such as hidden administrative accounts or scheduled tasks left behind by an attacker, have survived into the restored copy. Once validation is complete, workloads move through a staged cutover into the rebuilt production environment rather than all at once. A phased return reintroduces a small number of systems first and monitors them closely, so that any unexpected issue is caught while the blast radius of a mistake is still small, before the remaining workloads follow the same validated path.

Table 2 compares this framework against Commvault Cleanroom Recovery and Rubrik's cyber recovery offerings on three dimensions: isolation model, sequencing discipline, and vendor dependency.

Table 2. Comparison Against Existing Commercial Clean-Room Offerings

Dimension	Commvault Cleanroom Recovery	Rubrik Cyber Recovery	This Framework
Isolation model	Proprietary managed Azure environment	Proprietary orchestrated multi-cloud environment	IaC-generated, vendor-agnostic boundary
Sequencing discipline	Configurable; not identity-first by default	Configurable; not identity-first by default	Identity-first, engineered and auditable
Vendor dependency	Commvault platform	Rubrik platform	None — portable across cloud providers

4.2 Benchmarking Recovery Maturity

Industry-wide recovery data shows meaningful improvement alongside continued risk. Sophos's State of Ransomware 2025 report found that 53% of organizations recovered fully within one week of an attack, up sharply from 35% the previous year, and that organizations performing frequent backup and recovery testing recovered substantially faster than those testing annually or not at all (Sophos, 2025). Healthcare organizations continue to face the longest recovery timelines of any sector, extending well beyond the cross-industry average because of patient-safety validation and regulatory obligations that demand more extensive review before systems return to use.

Figure 2. Illustrative Conceptual Comparison of Recovery Timelines

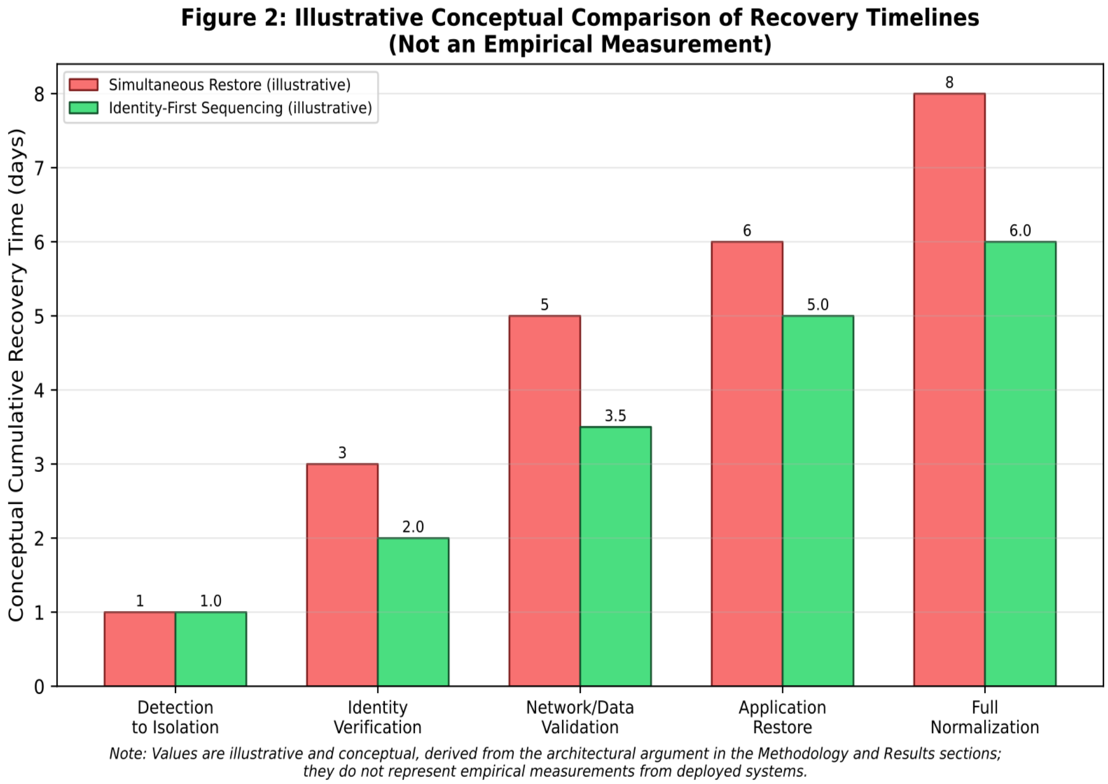


Figure 2 presents an illustrative, conceptual comparison of recovery-time distribution under an identity-first sequencing approach versus a simultaneous-restore approach, derived from the architectural argument developed in this section rather than from an empirical dataset.

These benchmarks suggest a research direction: codifying the sequencing decisions that experienced incident responders already make intuitively into a repeatable, version-controlled framework, rather than leaving recovery quality dependent on which responders happen to be available during a given incident. An identity-first reconstitution model, evaluated against emerging benchmarks such as mean time to clean recovery, offers a way to test whether disciplined sequencing produces measurable improvement over less structured approaches, an empirical question current vendor literature largely leaves unanswered.

5. Conclusion

Persistent recovery infrastructure was designed for a threat model that no longer matches how modern ransomware behaves. Generating an isolated recovery environment from code at the moment it is needed, and rebuilding it in a deliberate, identity-first order rather than restoring everything at once, offers a more defensible approach than maintaining a standing site that shares risk with the systems it is meant to protect. The framework is intentionally vendor-neutral because its load-bearing principles, isolation by default, immutable data presentation, ordered reconstitution beginning with verified identity, and staged return to production are not specific to any one platform. As the disaster-recovery-as-a-service and infrastructure-as-code markets continue to grow, and as emerging benchmarks push the industry to measure trustworthiness rather than raw speed, treating recovery infrastructure as something compiled on demand from version-controlled definitions and treating reconstitution order as an engineered property represents a measurable shift in how resilience is built. Future empirical work should test whether disciplined, identity-first sequencing produces measurable improvement over less structured approaches, an empirical question current vendor literature largely leaves unanswered, and should do so across multiple organizations and incident types rather than relying on a single case.

Statements and Declarations

Funding: This research received no external funding.

Conflicts of Interest: The author declares no conflict of interest.

Acknowledgments: None.

AI Tool Disclosure: Generative AI-assisted drafting and editing tools were used during manuscript preparation for literature synthesis support, structural drafting, and language editing; all core ideas, architectural contributions, source verification, and final content decisions originated from and were reviewed and approved by the human author.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers.

References

[1] Commvault. (2024). Cleanroom recovery. Commvault Documentation. https://documentation.commvault.com/2024e/software/cleanroom_recovery.html

[2] Cybersecurity and Infrastructure Security Agency. (n.d.). #StopRansomware guide. U.S. Department of Homeland Security. https://www.cisa.gov/stopransomware

[3] CloudZero. (2026). How to build cloud resilience before the next AWS or Azure outage. https://www.cloudzero.com/blog/aws-and-azure-outages/

[4] Halcyon. (2025). Ransomware dwell-time and containment research. https://www.halcyon.ai/ransomware-statistics

[5] Precedence Research. (n.d.-a). Infrastructure as code market size to reach USD 9.40 Bn by 2034. https://www.precedenceresearch.com/infrastructure-as-code-market

[6] Precedence Research. (n.d.-b). Disaster recovery as a service market size to surge USD 195.71 Bn by 2034. https://www.precedenceresearch.com/disaster-recovery-as-a-service-market

[7] Rubrik. (n.d.). Cyber recovery: Rapid, reliable restore. https://www.rubrik.com/products/cyber-recovery

[8] Sophos. (2025). The state of ransomware 2025. https://www.sophos.com/en-us/content/state-of-ransomware

[9] TechTarget. (2024). Commvault adds Cleanroom Recovery for ransomware attacks. https://www.techtarget.com/searchdatabackup/news/366583037/Commvault-adds-Cleanroom-Recovery-for-ransomware-attacks

[10] Veeam. (2025). From risk to resilience: 2025 ransomware trends and proactive strategies. https://www.veeam.com/blog/ransomware-trends.html

Author Biography

Venkata Ravi Sridhar Potunuru is a Principal Integration Engineer at Bright Horizons Family Solutions, where he leads hybrid-cloud infrastructure resilience and security automation initiatives spanning identity, disaster recovery, and post-incident recovery architecture. He holds an M.S. in Electrical Engineering from the University of Texas at El Paso and a B.E. from Anna University, and has over 15 years of experience architecting enterprise infrastructure across VMware, Active Directory, and multi-cloud identity platforms. He holds multiple industry certifications in cloud infrastructure and security automation.