
This is the **published version** of the bachelor thesis:

Vitrià Pujol, Pau. *Cybersecurity Risks in Smart City Decision-Making Bodies : Immutable Backups as a Defensive Mechanism*. Treball de Final de Grau (Universitat Autònoma de Barcelona), 2026 (Gestió de Ciutats Intel·ligents i Sostenibles). 12 pàgines.

This version is available at <https://ddd.uab.cat/record/335244>

under the terms of the  license.

Cybersecurity Risks in Smart City Decision-Making Bodies: Immutable Backups as a Defensive Mechanism

By: Pau Vitrià Pujol

Abstract

Smart cities rely on digital systems to run everything from traffic lights to water treatment, and that dependency is growing faster than most municipalities can secure it. This paper focuses on a specific part of that problem: the cybersecurity risks facing the organizations that actually make decisions in smart cities, not just the infrastructure they manage. We'll analyze seven documented cyberattacks between 2015 and 2025, and supported by hands-on experimentation in a VMware ESXi environment with Veeam Backup & Replication, we'll show that modern ransomware routinely destroys conventional backups as part of the attack, and that WORM-based immutable storage stops this. The practical tests, run against real Veeam backup files on a Synology NAS, produced clear results: every modification attempt on the immutable repository was blocked, while the conventional backup was entirely compromised. The paper closes with governance recommendations based on both the experimental findings and the NIS2 regulatory framework.

Keywords — Smart cities, cybersecurity, ransomware, immutable backups, WORM storage, Veeam, VMware ESXi, NIS2, critical infrastructure.

I. MOTIVATION, OBJECTIVES, AND METHODOLOGY

The starting point for this paper was a fairly simple observation, most writing on smart city cybersecurity treats the problem at a high level of abstraction, cataloguing threat categories and recommending general best practices. What gets less attention is the specific situation of the organizations that are actually running these systems day to day, such as a city council IT department, a municipal transport authority, or a water management agency. We can see that none of these are a tech company. They don't have a dedicated security operations center, they operate under budget and staffing constraints, and when something goes wrong, the consequences are not measured in lost revenue, although it is important, but rather in the effect it had on all of its users and public security.

This became clearer when I started looking at backup infrastructure specifically, which I had already done some work on previously. The usual assumption behind disaster recovery planning is that if your production systems are compromised, you restore from a backup. That assumption doesn't exactly apply against modern ransomware. According to Sophos [2], 94% of ransomware attacks now attempt to compromise or destroy backup repositories before triggering the main encryption event. The groups behind these attacks know that backup destruction is what forces victims to pay, and they have made this part of their attacks.

The Spain context adds another dimension. The national power outage of April 2025 — though not caused by a cyberattack — showed in real time what widespread infrastructure failure looks like for us ordinary people. Hospitals on generator power, trains stopped, fuel stations unable to process payments. A ransomware attack on a city's decision-making systems could produce similar consequences as the ones mentioned before.

This paper addresses that gap. The goal is not to cover every possible threat to smart city infrastructure, but to focus on one specific, documented, and solvable problem: ransomware destroys conventional backups, and immutable WORM-based backups resist this. The experimental work was designed to demonstrate that.

A. Objectives

- Map the cyber risk landscape for smart city decision-making bodies, focusing on the specific vulnerabilities of municipal organizations.
- Analyze seven real cyberattacks on urban and critical infrastructure to identify consistent patterns in attack vectors, backup impact, and recovery outcomes.

- Establish why ransomware is the most operationally disruptive threat to municipal service continuity.
- Test immutable WORM backup technology against simulated ransomware behavior, using production-grade infrastructure.
- Connect the technical findings to the NIS2 regulatory framework and its practical implications for municipal IT governance.
- Give city decision-makers a set of concrete, proportionate recommendations they can implement.

B. Methodology

The research has two parts. The first is a structured analysis of seven cyberattacks, selected because they are well-documented, span a range of target types and attack techniques, and collectively build a case for why backup resilience matters specifically for organizations managing urban infrastructure. The STRIDE framework is used to categorize threat types.

The second part is experimental. The test environment was a VMware ESXi 7.0 cluster on an HPE ProLiant DL380 Gen10 server, with Veeam Backup & Replication managing backup jobs to two destinations: a conventional SMB share on a Synology NAS, and a WORM-enabled folder on the same device (DSM 7.3.2, Btrfs). A PowerShell script was written to simulate the file-modification behavior of ransomware against both targets. The results are measured with whether the modification attempts succeeded or not.

One methodological note: the ransomware simulation only modifies file extensions rather than encrypting whole files. This is intentional. The goal is to demonstrate the storage-layer immutability enforcement mechanism, not to replicate the cryptographic components of real ransomware. Extension modification is sufficient to render Veeam backup files unrestoreable and is directly comparable to what ransomware does to backup files during a real attack.

II. PROBLEM DEFINITION: CYBER RISK IN SMART CITIES

Before getting into specific cases, it is worth establishing what makes smart cities a distinct cybersecurity problem.

A. Scale and Differences of the Attack Surface

A normal IT environment has a reasonably well-defined perimeter: servers, workstations, network equipment, and cloud services, all managed by the same IT team under a consistent set of policies. A smart city deployment shatters that model. A medium-sized city might have tens of thousands of connected sensors, cameras, parking meters, traffic controllers,

environmental monitors, and public Wi-Fi access points, many produced by different manufacturers with different firmware update cycles, often running for years on default credentials because no one owns the responsibility for maintaining them.

The 2024 statistics are not encouraging. Over 96% of vulnerabilities discovered in smart city systems traced back to web applications and connected devices [3], with critical vulnerabilities increasing by 83% compared to 2023. A significant portion of that attack surface is invisible to municipal IT teams, because it was procured and deployed by different departments, under service contracts with private vendors, without any central security oversight.

This is the IT/OT convergence problem. Operational technology systems — SCADA controllers for water treatment, building management systems, traffic signal controllers — were historically isolated from internet-connected networks. The whole point of smart city integration is to connect them for easier monitoring and remote management. That connectivity is also what makes them reachable by an attacker who has gained access anywhere in the city's network.

B. Continuity Constraints and the Ransom Incentive

What separates smart city attacks from typical enterprise ransomware incidents is the public safety aspect of continuity failure. When a company's email system goes down, employees switch to phones. When a city's traffic management system goes down, emergency vehicles may not get through. The stakes are much different, and attackers know this.

This creates a real and uncomfortable situation: municipalities under pressure to restore critical services face stronger incentives to pay ransoms than most private organizations. Baltimore chose not to pay \$76,000 and spent \$18 million on recovery. That is a defensible decision in retrospect, but it required accepting weeks of degraded services across 15 departments. For an organization running time-critical infrastructure such as a hospital network, an emergency dispatch system, or a water treatment facility, that kind of downtime may simply not be an option.

The point is not that cities should pay ransoms. The point is that the pressure to pay is real, and the only structural way to remove that pressure is to have recovery capability that does not depend on what the attacker leaves intact.

C. Fragmented Governance

Cybersecurity responsibility in smart cities tends to be spread across multiple departments, contractors, and utility operators, each with their own procurement processes and security standards. This is not unique to

smart cities, but the scale of the problem is. A single smart city deployment might involve dozens of private vendors operating components on behalf of the municipality, none of whom are necessarily held to any consistent security standard.

A relatively minor illustration: in 2022, a Cork City Council parking system breach exploited a vulnerability in a private company's software, but it was the council that faced public accountability. The council did not control the vulnerable code and had no direct visibility into the vendor's security practices. This governance fragmentation is structural, and it means that even a well-resourced municipal IT team cannot fully control its own attack surface.

D. STRIDE Threat Taxonomy

The STRIDE model provides a useful structure for categorizing the specific threats that smart city decision bodies might face:

Threat	Example in Smart City Context
Spoofing	Injecting false data from spoofed sensors into traffic or utility systems
Tampering	Modifying stored emergency management records or SCADA configurations
Repudiation	Denying actions in procurement systems where audit logs have been deleted
Info Disclosure	Exfiltrating citizen PII from municipal databases
DoS	Encrypting systems until ransom is paid, rendering services unavailable
EoP	Using compromised service account to gain admin access to backup infrastructure

Table I. STRIDE threat model applied to smart city decision bodies.

In practice, ransomware attacks combine multiple STRIDE categories simultaneously: they tamper with data (encryption), create denial of service (systems unavailable), and often involve elevation of privilege to reach backup systems. That combination is what makes them so difficult to recover.

III. EVIDENCE: CASE ANALYSIS OF CYBERATTACKS

The seven cases below were chosen to cover different attack types, geographies, and target sectors, while all connecting to the central question of how backup and recovery capability affected outcomes. They are ordered chronologically.

A. Ukraine Power Grid (2015)

The Ukraine power grid attack is worth starting with because it established that cyberattacks could cause real physical infrastructure failure, not just data loss or service disruption. The attack, attributed to the Sandworm group with ties to Russian military intelligence, caused blackouts affecting around 230,000 people across three distribution companies in western Ukraine.

What made the attack significant was its multi-layer structure. The initial compromise came through spear-phishing emails to corporate network users, delivering BlackEnergy malware. After several months of quiet reconnaissance, attackers used their access to take control of SCADA systems and remotely switch substations offline. They then deployed the KillDisk wiper to destroy servers, workstations, and RTUs, preventing operators from using IT tools to restore the grid. A simultaneous denial-of-service attack on the utility's call center cut off information flow to operators and the public. Emergency power at the operations center was also disabled but separately.

From a backup and recovery perspective, the KillDisk wiper is the most important element. Wiping IT infrastructure is not incidental to the attack, it is part of the attack strategy. Extending recovery time and preventing operators from using their own systems are direct objectives, not side effects. This pattern, destroying the tools and data that would enable recovery, is the same logic that drives backup destruction in modern ransomware.

B. San Francisco Muni Ransomware (2016)

The HDDCryptor attack on the San Francisco Municipal Transportation Agency took down all ticketing infrastructure over the Thanksgiving weekend of 2016. The agency offered free rides while recovery was being done, which softened the public impact but did not eliminate it. Over 900 machines were encrypted using MBR-level encryption that prevented systems from loading and working.

The attackers asked for 100 Bitcoin (around \$73,000 at the time). Muni declined and recovered from existing backups over several days. The interesting aspect here is what the attack targeted: ticketing infrastructure, not operational control of the trains themselves. Passengers could still ride, but the attack demonstrated that attacking the revenue and data systems of a transport authority, which are considerably easier to reach than the actual train control systems, can still produce significant disruption and cost.

It also demonstrated something about backup effectiveness: recovery was possible because the

backups were largely intact. That outcome did not repeat in later cases where attackers specifically targeted the backup systems.

C. Atlanta SamSam Ransomware (2018)

The Atlanta SamSam attack is probably the clearest example of what happens when a city has neither adequate backups nor a credible recovery plan. SamSam was unusual in that it did not rely on phishing. Instead, it used brute-force attacks against exposed RDP endpoints, which means user training would not have helped. Once inside, it spread through the network before triggering encryption across virtually all of the city's departments.

The ransom demand was \$51,000. Atlanta declined. Recovery cost \$2.7 million. Multiple departments ran on paper for weeks. Some records, including police incident reports and years of dashboard camera footage, were permanently lost because they either had not been backed up, or the backups were also encrypted. That last point is important: it was not just that the production systems were compromised. The backup copies were also gone. Without an immutable copy, there was nothing to restore from.

Up to 6 million residents were affected to some degree or another. The city's credit rating outlook was revised. And all of this started from an exposed RDP port and a weak password.

D. Baltimore Ransomware (2019)

Baltimore in 2019 is the case that most clearly illustrates the economics of the backup problem. RobbinHood ransomware, delivered through a combination of stolen NSA tools and lateral movement, encrypted systems across 15 municipal departments simultaneously: email, phone lines, payment systems, property databases, and more. The city was effectively offline.

The ransom demand was \$76,000. Baltimore chose not to pay. The recovery cost was \$18 million. That ratio, 237:1, is not an argument for paying ransoms. It is an argument for investing in infrastructure that means you never face that choice. Even at \$500,000 for a comprehensive immutable backup and network segmentation overhaul, the economics of prevention against recovery are not exactly up for debate.

The reach of the attack, 15 departments across the whole city, is also important. It reflects a flat network where lateral movement is relatively unconstrained. Once an attacker has domain credentials, there is very little stopping them from reaching everything, including backup servers.

E. Colonial Pipeline (2021)

Colonial Pipeline is included not because it is a municipal system, but because it illustrates failure modes

that translate directly to smart city governance. The entry point was a single inactive VPN account with a reused password and no multi-factor authentication. DarkSide ransomware used that access to encrypt operational data and, crucially, billing and scheduling systems. Colonial shut down the pipeline voluntarily because it could not track product flow, not because the operational technology was directly compromised.

The company paid \$4.4 million within hours. US law enforcement later recovered about half of it, but that is the exception. What followed was fuel shortages across the eastern US, panic buying, flight cancellations, and a declared state of emergency in several states, all starting with one inactive account that nobody had disabled.

For smart city operators, the lesson about inactive accounts is direct. Municipal IT environments tend to accumulate service accounts, contractor credentials, and legacy system logins that persist long after they should have been decommissioned. Any one of them is a potential entry point.

F. UAB Cyberattack (2021)

The Autonomous University of Barcelona attack is the most locally relevant case in this analysis. In November 2021, ransomware disabled all virtualized systems at UAB, including the email server, virtual campus, official website, library systems, and internal network. 50,000 students, teaching staff, and administrators were affected during an active academic period. The university spent months working with paper-based processes and physical alternatives while systems were gradually restored.

What is technically interesting about UAB is the virtualization angle. When a hypervisor or its management interface is reachable and poorly segmented, a single compromise can cascade to every VM on the cluster simultaneously. The university had virtualized practically everything for efficiency, which is sensible, but had not correctly isolated the management systems from systems that could be reached through the broader network.

The UAB case also received significant media coverage in Catalonia, which makes it a useful anchor for the local context of this paper.

G. Olsztyn Smart City Transport Attack (2023)

Olsztyn in Poland is the most directly relevant case because it involved a purpose-built smart city infrastructure. The city had deployed an integrated transport management system including traffic signal coordination, intersection violation monitoring, and real-time passenger information. In June 2023, an attack paralyzed around 100 intersections in the city center,

took down passenger information displays, and disrupted ticket purchasing.

The transport authority (ZDiT) responded by physically disconnecting servers, which stopped the spread but also took down whatever functionality there was. This tension, between containing the attack and maintaining some level of service, is one of the defining operational challenges of smart city incident response. Conventional IT incident response can often afford to isolate and investigate. Smart city systems might not be able to absorb the downtime that isolation requires.

This case matters for the argument of this paper because it shows that smart city integration is already producing the vulnerabilities that the theory predicts. Olsztyn is not a hypothetical. It happened three years ago.

H. Cross-Case Comparison

Case	Year	Attack Type	Primary Target	Backup Impact	Cost
Ukraine	2015	APT+Wiper	SCADA/OT	Wiped	N/A
SF Muni	2016	Ransomware	Transport IT	Intact	~\$50K
Atlanta	2018	Ransomware	Municipal IT	Compromised	\$2.7M
Baltimore	2019	Ransomware	Municipal IT	Destroyed	\$18M
Colonial	2021	Ransomware	OT/Pipeline	N/A	\$4.4M
UAB	2021	Ransomware	Education IT	Partial	Months offline
Olsztyn	2023	Unknown	Smart City OT	Unknown	City paralyzed

Table II. Comparative analysis of seven urban cyberattack cases.

The pattern across these cases is quite consistent. When backups were intact and tested (SF Muni), recovery was possible at relatively low cost. When backups were compromised or absent (Atlanta, Baltimore), costs multiplied by orders of magnitude. The single most useful intervention across all cases would have been backup infrastructure that ransomware could not reach and destroy.

IV. CASE OF STUDY: RANSOMWARE AS THE KEY THREAT

Ransomware is not the only threat to smart city decision bodies, but it is the one that most consistently produces severe and costly outcomes. Of the seven cases above, five directly involved ransomware. The other two

(Ukraine and Olsztyn) involved different attack types but shared the common feature of destroying or disabling backups.

A. Why Municipalities Are Attractive Targets

There is a main reason why ransomware groups target municipalities. It is not about the value of the data, though that matters. It is about the combination of high recovery pressure, limited defensive capability, and political constraints on the response.

Municipal governments cannot simply take systems offline for weeks without consequence. Essential services depend on those systems, and elected officials face immediate public accountability for any visible disruption. That creates genuine coercive leverage for ransomware operators, independent of the technical sophistication of the attack. The LockBit, Conti, and RobbinHood groups all explicitly factored this pressure asymmetry into their targeting criteria, according to post-operation analysis [1].

The numbers reflect this. Ransomware attacks increased 11% globally in 2024 to 5,414 documented incidents. Attacks on US government entities alone rose 149% year-over-year in early 2025 [1]. Average ransom payments reached \$2.73 million. These are not numbers that suggest the threat is declining.

B. How Backup Targeting Changed Everything

Early ransomware, including the 2017 WannaCry and NotPetya waves, encrypted whatever was accessible on the network and hoped it was enough. The recovery response was often to restore from backup, pay for decryption tools, or rebuild. That response made ransomware a manageable if expensive problem for organizations with decent backups.

The shift came when ransomware operators recognized that backup destruction was more valuable than mass encryption. If the backups are gone, there is no alternative to paying the ransom or losing the data. Modern ransomware-as-a-service platforms include dedicated modules for identifying and destroying Veeam, Acronis, Windows Server Backup, and cloud backup installations. This is not an optional feature; it is built into the standard attack workflow.

The result is that 89% of ransomware attacks now target backup repositories, and attackers successfully destroy 73% of those they reach [4]. Only 32% of organizations that paid ransoms in 2024 successfully recovered their data. The conventional disaster recovery model is broken, and it has been for several years.

C. The Kill Chain in Practice

For a smart city decision body, the realistic attack sequence looks like this:

- Initial access through a phishing email, an exposed RDP endpoint, a reused credential, or a vulnerable public-facing application. The entry point does not need to be sophisticated.
- Quiet lateral movement and privilege escalation over hours or days, using legitimate administrative tools to avoid detection.
- Identification of backup infrastructure, backup server credentials, and any cloud backup accounts.
- Backup deletion or encryption, which may occur silently days before the final event.
- Mass encryption of production data, followed by a ransom note and demand.

The average time from initial access to ransom demand has dropped to 24 hours [5], meaning detection windows are narrow and the backup destruction may already be done before anyone notices that something is wrong.

D. RaaS and the Democratization of the Threat

One aspect of the current threats that is easy to underestimate is the degree to which ransomware has been productized. Ransomware-as-a-service platforms provide ready-to-deploy toolkits, that include backup-targeting modules, negotiation interfaces, and payment infrastructure, to people who take a share of ransom proceeds. The difficulty of doing a municipal attack has dropped dramatically.

This matters for the smart city context because it means the threat is no longer confined to highly specialised criminal groups. Anyone that's motivated and has a few hundred dollars and access to a dark web forum can now deploy a ransomware attack that includes ways to destroying Veeam backups. The backup-targeting component is included by default.

V. IMMUTABLE BACKUPS: ROLE, BENEFITS, AND LIMITATIONS

The core idea behind immutable backups is simple: data written to WORM (write once read many) storage cannot be modified or deleted until a certain defined time has passed, regardless of what credentials are used to attempt the modification. This sounds like quite a simple solution but at the moment it is the one which can provide backup safety.

A. Why Storage-Layer Enforcement Matters

Most backup protection mechanisms operate at the application or operating system layer. A read-only flag on a folder, a backup software setting that restricts deletion, or filesystem ACLs restricting write access: all of these are enforced by the OS or application, which means they can be bypassed by anyone with administrative credentials to that system. A ransomware

attacker who has obtained administrator access, which is a standard step in the current techniques, can simply disable or bypass these security controls.

WORM storage enforces immutability in the storage device firmware, below the OS layer. When a file is written to a WORM folder on a Synology NAS running DSM 7.2+, the device's firmware records a retention lock on that file. Any attempted modification or deletion requests are rejected because the backup is outside the control domain of the compromised production system, unlike the previous example. There is no way for a process running on the host OS to avoid this, because the enforcement is not happening in the OS, but rather on the hardware.

This is why the experimental results in Section VI show modification attempts on the WORM repository returning Access Denied, while the same script successfully modifies files on the conventional SMB share. The difference is not permissions. It is where the enforcement happens.

B. Governance Mode vs Compliance Mode

Synology (and most enterprise WORM implementations) offer two modes with different properties:

- Governance mode retains a bypass mechanism for a designated administrator with multi-factor authentication. This makes it practical for operational environments where occasional manual backup management may be needed. It still provides full protection against ransomware, which cannot supply the MFA bypass, but it does not provide absolute protection against a determined administrator with legitimate access.
- Compliance mode removes the bypass entirely. Once a file is written, no one, including the storage administrator, can delete it before the retention period expires. This is the mode required for regulatory compliance in sectors with strict data retention mandates. For testing and operational environments, it requires careful retention period planning, because there is no undo.

For the experimental work done in this paper, Enterprise (Governance) mode was used with a 7-day retention period, which is perfect for a test environment and still provides full protection from ransomware.

C. Technology Comparison

	WORM NAS	Air-Gap Tape	Cloud Object Lock	Linux Hardened Repo
Ransomware resistant	Yes	Yes	Yes	Yes
Network accessible	Yes	No	Yes	Yes
Recovery speed	Fast	Slow	Medium	Fast
Setup complexity	Low	Medium	Low	High
Approx. entry cost	~400 EUR	~800 EUR	Pay-per-GB	Free (labour cost)
Offline protection	Partial	Full	No	Partial

Table III. Comparison of immutability approaches for municipal backup environments.

D. Benefits in the Municipal Context

- Backup integrity under attack: as the experimental results confirm, a fully compromised production environment and conventional backup cannot corrupt a copy protected by WORM.
- Eliminates ransom leverage: if there is always a recoverable backup, the pressure to pay a ransom is almost fully removed.
- Forensic value: WORM-protected backup logs provide tamper-proof records of system conditions before the attack, which matter for both incident investigation and reporting.
- NIS2 alignment: WORM backup satisfies the backup management requirements of NIS2 Article 21, giving municipalities documented compliance evidence.

E. Honest Limitations

WORM backup is not a complete security solution, and it is worth being clear about what it does and does not address. It should be a last line of defense in an organizations systems.

- It does not prevent the initial attack. If an attacker can encrypt production systems, they will, regardless of whether the backup is protected. The organization still faces downtime, potential data exfiltration, and reputational damage.

- It does not protect against exfiltration and double extortion. Modern ransomware often exfiltrates data before encrypting it, threatening publication as additional leverage. Immutable backups do not help with this.
- Storage costs increase because WORM retention prevents deletion until expiry. For large backups with long retention periods, this can become quite significant.
- Credential isolation is not automatic. A WORM repository whose management credentials are accessible from production systems can potentially have its retention policy modified before the lock-in takes effect, depending on the WORM mode configured.

However, WORM protection is not unconditional. Three scenarios could reduce or eliminate its effectiveness.

First, if the attacker obtains the NAS administrator credentials and the WORM folder is configured in Governance mode, they could potentially extend or remove the retention lock using the administrative override, provided they can also supply the required multi-factor authentication. Compliance mode removes this vector entirely.

Second, if the retention lock has not yet activated at the time of the attack — because the backup was written very recently and the lock-in delay has not elapsed — files may still be modifiable. This is mitigated by configuring immediate lock-on-write rather than delayed locking.

Third, misconfigured retention policies, such as a retention period of zero or a folder that was not correctly provisioned as WORM, would obviously provide no protection. These scenarios do not undermine the value of WORM storage, but they do reinforce that the protection is only as strong as the system around it.

VI. PRACTICAL DEMONSTRATION

A. Test Environment

The test was run on a production-grade VMware ESXi environment rather than a desktop virtualization setup. This was done on purpose: the goal was to produce results that are meaningful in the context of a real small-to-medium IT infrastructure, not a controlled lab that does not resemble reality.

The host is an HPE ProLiant DL380 Gen10 running VMware ESXi 7.0.3 (build 23794027), with an Intel Xeon Gold 5118 processor providing 48 logical cores, 6 NICs, and 11 days of uptime at the time of the test. It sits

in the vSphere cluster, which runs production virtual machines alongside the test VM.

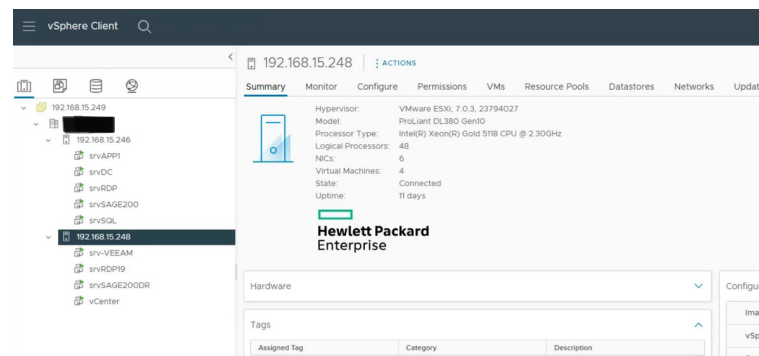


Fig. 1. vSphere Client view of the ESXi host (192.168.15.248) in the company cluster. Production VMs (srvDC, srvRDP, srvSAGE200, srvSQL, srvAPP1) are visible alongside the 'Proves TFG' test VM deployed for this work.

The test VM ('Proves TFG') was created via the vSphere wizard with Windows 10 (64-bit), 4 vCPUs, 8 GB RAM, and a 48 GB virtual disk on the G10 HDD datastore. The VMXNET3 network adapter connects it to the VM Network.

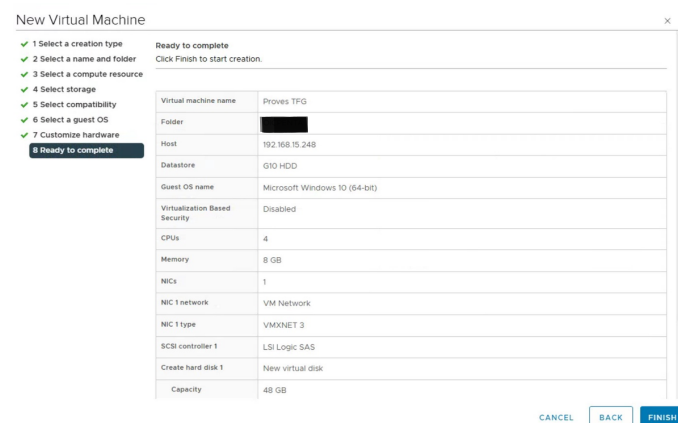


Fig. 2. vSphere VM creation wizard showing the final configuration for 'Proves TFG': Windows 10 64-bit, 4 CPUs, 8 GB RAM, 48 GB disk, VMXNET3 NIC, LSI Logic SAS controller.

B. Backup Infrastructure Setup

Two backup destinations were configured in Veeam Backup & Replication:

The conventional repository ('Backup Repository TFG_NS') was set up as an SMB share pointing to \\192.168.15.241\TFG_NS on the Synology NAS, with the TFG account for authentication. The repository has no additional protection: anyone with SMB credentials can write on it or delete from it.

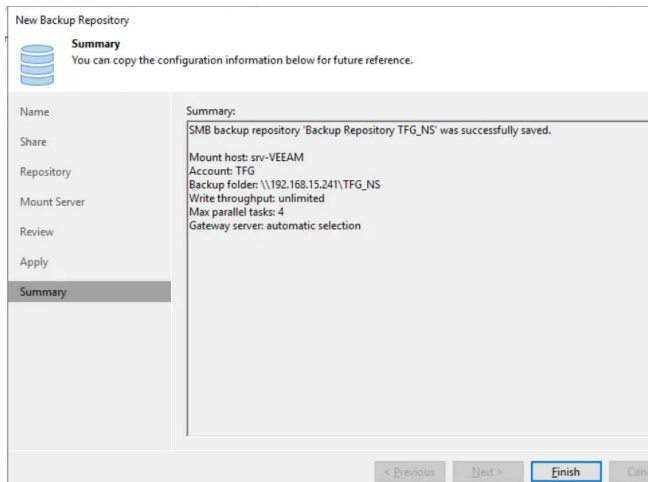


Fig. 3. Veeam repository creation summary for 'Backup Repository TFG_NS'. SMB share at \\192.168.15.241\TFG_NS, mount host srv-VEEAM, account TFG, 4 parallel tasks.

The immutable repository was configured on the same NAS using a separately provisioned WORM-enabled shared folder. The WORM folder was created in DSM Control Panel with Enterprise mode and a 7-day minimum retention period. It was mapped to the Windows test VM as the Y: drive for the attack simulation.

C. Running the Backup Job

A Veeam VMware backup job ('Backup Job TFG') was run targeting the 'Proves TFG' VM. The job completed cleanly: 60 GB processed, 4.7 GB read at 236 MB/s, 2.3 GB transferred (2.1x compression), 0 errors, duration 2:58. Change Block Tracking was used for efficient disk reads. This time reflects the small size of the backup, real backups on municipal servers would naturally take longer.

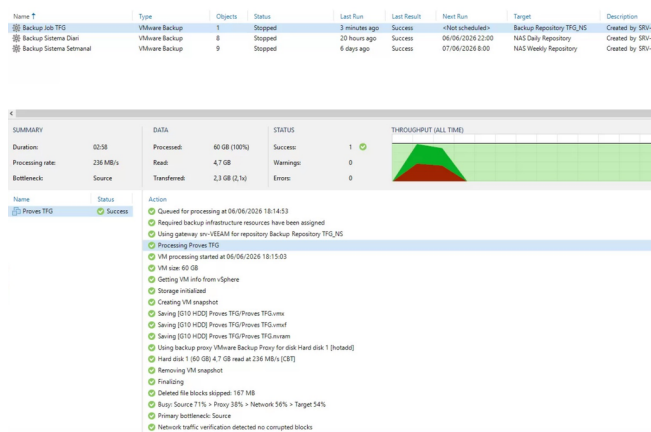


Fig. 4. Veeam backup job completion summary: 60 GB processed, 4.7 GB read at 236 MB/s, 2.3 GB transferred, 0 errors, 02:58 duration. CBT used for disk read efficiency.

The job produced the standard Veeam backup file set: a .vbk full backup, two .vib incremental files, and a .vmb metadata file. These were stored in the SMB repository. A Veeam Backup Copy Job ('Backup Copy Job INMS')

then replicated the same files to the WORM-enabled folder, producing 6 files there.

D. The Attack Simulation

The PowerShell script was written to do one thing clearly: attempt to rename every targeted file by changing its extension to .xifrat, which is what ransomware does when it encrypts files, and record whether each attempt was successful or was blocked. The script runs against three locations simultaneously: the production server folder, the conventional backup repository, and the immutable repository.

The script asks for explicit confirmation of isolated environment before running, logs every operation with a timestamp, and produces a summary at the end. It targets Veeam file extensions (.vbk, .vib, .vmb) as well as standard document formats, because the point of the test is to show what happens to backup files specifically.

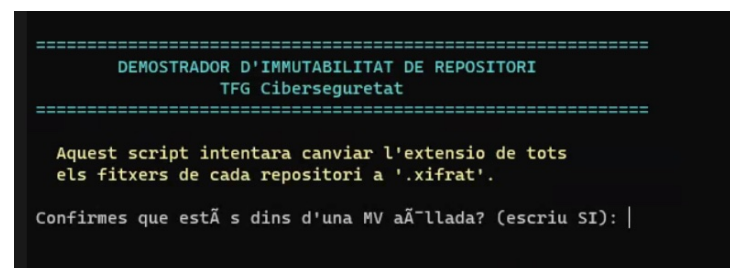


Fig. 5. Script startup screen. Before doing anything, it lists the target locations and asks for confirmation that it is running in an isolated VM.

E. Results

The script ran at 19:43:07 and completed within one second, which is a real approximation of the speed at which ransomware processes files.

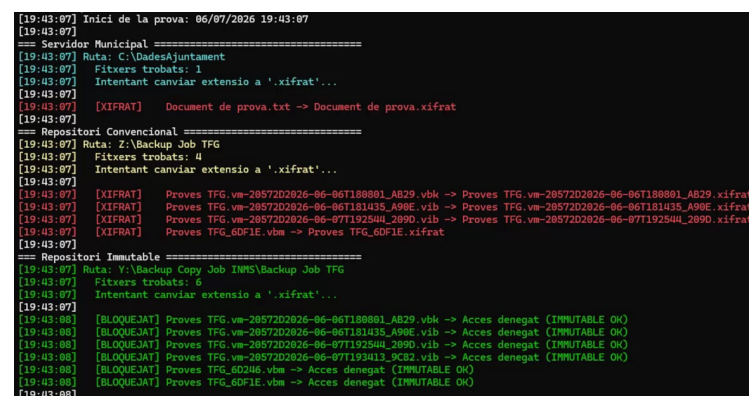


Fig. 6. Attack execution log. Top section: production server file successfully renamed. Middle: 4 Veeam backup files (.vbk, two .vib, .vmb) renamed in the conventional repository. Bottom: all 6 files in the immutable repository blocked with 'Acces denegat (IMMUTABLE OK)'.

```

=== RESUM FINAL =====
[19:43:08]
[19:43:08] Servidor Municipal
[19:43:08] Intents de modificacio: 1
[19:43:08] Modificacions fetes: 1
[19:43:08] Modificacions bloquejades: 0
[19:43:08] Estat: VULNERABLE - modificacions permeses
[19:43:08]
[19:43:08] Repositori Convencional
[19:43:08] Intents de modificacio: 4
[19:43:08] Modificacions fetes: 4
[19:43:08] Modificacions bloquejades: 0
[19:43:08] Estat: VULNERABLE - modificacions permeses
[19:43:08]
[19:43:08] Repositori Immutable
[19:43:08] Intents de modificacio: 6
[19:43:08] Modificacions fetes: 0
[19:43:08] Modificacions bloquejades: 6
[19:43:08] Estat: PROTEGIT - immutabilitat confirmada
[19:43:08]
=====
[19:43:08] Log guardat a: C:\TFG\log_immutabilitat_20260607_194300.txt
PS C:\>

```

Fig. 7. Final summary: Servidor Municipal (1/1 modified, VULNERABLE), Repositori Convencional (4/4 modified, VULNERABLE), Repositori Immutable (0/6 modified, PROTEGIT). Log saved to C:\TFG\log_immutabilitat_20260607_194300.txt.

Location	Files	Modified	Result
Production server (C:)	1	1	VULNERABLE
Conventional repo (Z:)	4	4	VULNERABLE
Immutable repo (Y:)	6	0	PROTEGIT

Table IV. Summary of modification attempts and outcomes.

F. What the Results Mean

Four Veeam backup files were successfully modified in the conventional repository, including the .vbk full backup and both incrementals. In a real attack, these files would be unrestoreable. Recovery without the immutable copy would require rebuilding the VM from scratch or paying the ransom.

Six files in the WORM repository were fully protected. None were modified. The access denied responses came from the Synology firmware. The same credentials that successfully modified the Z: share could not touch the Y: share once the immutable lock was in effect.

The full recovery path from the immutable repository was not tested in this experiment, but Veeam's standard restore process for NAS-hosted repositories would recover the complete VM from the .vbk and .vib files in the WORM folder. Recovery time for a 60 GB VM at network transfer speeds would be about under one hour.

G. Recovery Metrics

This data is all theoretical and based off tests done by others. These tests, however, could be done on the system that was set up as part of a continuation of this project in the future.

Metric	Conventional Backup	Immutable Backup
--------	---------------------	------------------

Backup files after attack	Destroyed (.xifrat)	Intact
RTO	Days to weeks	Under 1 hour
RPO	Total data loss	Last backup point
Ransom payment	Likely required	Not required
Forensic record	Potentially compromised	WORM-protected log

Table V. Recovery metric comparison after simulated ransomware attack.

VII. RECOMMENDATIONS FOR CITY DECISION-MAKERS

The recommendations below are based on the case analysis and findings. They are intended to be proportionate to the real context of municipal organisations, which have limited IT budgets, mixed technical expertise, and pressing service continuity obligations.

A. The 3-2-1-1 Rule as a Starting Point

The conventional 3-2-1 backup rule (3 copies, 2 media types, 1 offsite) is well-established but no longer enough against modern ransomware. Adding a fourth requirement, 1 copy stored as WORM-immutable with credentials isolated from production systems, directly addresses the backup destruction problem.

The credential isolation is as important as the WORM storage itself. If the account that can write to the WORM repository is the same account that the ransomware compromises, the attacker may be able to modify the retention policy or disable WORM protection before the lock-in takes effect. Governance mode requires MFA to override. Compliance mode removes the bypass entirely.

B. Credential Isolation in Practice

In the experimental environment, the conventional repository was accessible from a standard domain account (TFG). In a real attack where domain credentials are compromised, that account would give attackers access to the backup. Separating backup credentials from domain credentials means:

- A dedicated backup administrator account, not a repurposed domain administrator account.
- Multi-factor authentication on all backup management interfaces.
- The backup network VLAN is not reachable from production workstations.
- Regular reviews of which accounts have access to backup storage.

C. Test Recovery, Not Just Backup

A backup that has never been tested for recovery is not a backup; it is a file that might be a backup. Routine recovery tests that simulate a complete VM loss, restore to an isolated network, and measure the actual time of restoration are the only way to know whether a backup is recoverable in the time that is needed. For Veeam environments, SureBackup automates recovery verification without requiring manual restoration.

D. NIS2 and the ENS

The NIS2 Directive, transposed into Spanish law through the updated Esquema Nacional de Seguridad (ENS), requires that essential service entities including local governments implement backup management and business continuity measures under Article 21. Immutable backup implementation accomplishes these requirements and generates documented evidence of compliance. For Catalan municipalities, CESICAT provides security assessments and can assist with both technical implementation guidance and compliance documentation.

E. Scaling by Organisation Size

Area	Small (<50 staff)	Large (>200 staff)
Backup	WORM NAS, entry model (~400 EUR)	WORM NAS + offsite replication
Network	Backup on separate VLAN	Full segmentation + DMZ
Access	MFA on remote access	PAM solution + MFA everywhere
Testing	Annual recovery test	Quarterly SureBackup automation
Monitoring	Basic log review	SIEM with backup alert rules

Table VI. Minimum viable security baseline by organisation size.

F. Veeam-Specific Notes

For organisations already running Veeam, adding immutable backup protection requires relatively little additional work:

- Add a Backup Copy Job targeting the WORM repository as a secondary destination.
- Configure at least 14 days of immutable restore points to account for attacks that are detected late.
- Enable Veeam's built-in malware detection scan on restore points to verify backups are clean before recovery.
- Consider the Veeam hardened repository (Linux-based) as an alternative if a NAS WORM upgrade is not feasible.

VIII. CONCLUSIONS

The argument of this paper is quite simple: smart city decision-making bodies are high-value, relatively vulnerable targets for ransomware attacks, and these specifically destroy backup systems to prevent recovery and enforce ransom payment. The case analysis shows this across a few documented incidents. The experimental work demonstrates that WORM-based immutable storage stops the destruction of backups from the ransomware attack.

The results from the VMware ESXi and Veeam environment are also pretty clear. The same script that renamed four Veeam backup files (.vbk, .vib, .vmb) in the conventional repository to unrestoreable .xifrat files was blocked on the WORM-enabled repository. Six modification attempts, six access denied responses, all at the storage firmware level, independent of any OS or application permission. The conventional backup was gone. The immutable backup was intact.

There are things this paper does not address. Immutable backup does not stop the initial compromise. It does not prevent data exfiltration or double extortion. It does not eliminate downtime during the recovery process. A complete cybersecurity programme for a smart city decision body also needs network segmentation, credential management, endpoint protection, incident response planning, and staff training.

But the specific, solvable, well-documented problem of backup destruction enabling ransomware attacks to succeed has a clear, affordable, and technically proven solution. WORM storage is available in professional-grade hardware, supported by enterprise backup platforms like Veeam, and ordered by the NIS2 framework that applies to many of the organisations this paper is addressing. technical.

There is still a lot of work to do regarding public understanding of digital safety and security. And there are many small steps that can be taken to increase the safety of all the systems found in a smart city, this paper however focused on a very clear one. And I think it was demonstrated that the use of immutable backups in our smart cities should not be a question of if, but of how soon they can be implemented.

Acknowledgements

I'd like to give special thanks to Pere Rocamora Quílez from Serma Seguretat Informàtica i Sistemes for the help and guidance provided in the practical experiment aspect of this paper.

I'd also like to give thanks to my tutor during this project Angeles Vazquez Castro, for proposing this topic and for all the help also provided during these few months of work.

REFERENCES

- [1] CyberProof, "Digital Frontlines: The Escalating Cyber War on Municipalities Worldwide," CyberProof Blog, July 2025. Available: <https://www.cyberproof.com>
- [2] NIST SP 800-209, Security Guidelines for Storage Infrastructure, NIST, Oct. 2020
- [3] ENISA, Smart Cities Cybersecurity Recommendations
- [4] Storware, "Understanding Immutable Backups and Their Role in Cyber Resilience," storware.eu, October 2025.
- [5] Veeam Software, "Immutable Backups and Their Role in Cyber Resilience," veeam.com, August 2025.
- [6] Stormshield, "Overview of the Attack Surface of Smart Cities," stormshield.com, May 2024.
- [7] CLTC Berkeley, "The Cybersecurity Risks of Smart City Technologies," cltc.berkeley.edu, May 2023.
- [8] ITSS Verona, "Assessing the Cybersecurity Challenges of Smart Cities," itssverona.it, January 2025.
- [9] ENISA, ENISA Threat Landscape (2024 or 2025 ed.), public administration sector analysis
- [10] IBM Security, "Cost of a Data Breach Report 2024," IBM, 2024.
- [11] Wikipedia, "Atlanta government ransomware attack."
- [12] Wikipedia, "Colonial Pipeline ransomware attack."
- [13] Xataka, "El ciberataque a la UAB afectara hasta finales de ano," xataka.com, 2021.
- [14] European Parliament, "Directive (EU) 2022/2555 (NIS2)," Official Journal of the EU, December 2022.
- [15] Synology Inc., "DSM 7.3 Administrator's Guide: WriteOnce Shared Folders," Synology Documentation, 2024.
- [16] Veeam Software, "Veeam Backup & Replication User Guide for VMware vSphere," Veeam Help Center, 2024.
- [17] Heimdal Security, "The Curious Case of the Baltimore Ransomware Attack," heimdalsecurity.com.
- [18] CISA/FBI/NSA/MS-ISAC, #StopRansomware Guide, Oct. 2023
- [19] NIST SP 800-209
- [20] CISA #StopRansomware Guide