IJSES
Volume 9, Issue 1, pp. 30-35, 2025.
International Journal of Scientific Engineering and Science
ISSN (Online): 2456-7361

Developing Highly Resilient Architecture for Critical Systems to Mitigate Operational Risks

Raoul Hira
Principal Security Architect at Vistra, USA

Abstract— *The development of a highly resilient architecture for mission-critical systems is an integrated approach aimed at minimizing operational risks and ensuring the continuity of vital services. In the face of growing threats, including natural disasters and man-made disasters, the sustainability of infrastructure is becoming a key aspect of security and stability. The basic principles of development include reliability, flexibility, and environmental friendliness, which allows the system not only to withstand external influences but also to quickly recover from failures including bad updates in support software like Microsoft Windows Operating Systems and Endpoint Detection and Response (EDR) such as CloudStrike. The implementation of such architectures requires risk analysis, diversifying security and/or management tooling, integration of backup systems, regular monitoring, and staff training. The practical implementation of highly resilient architectures in an organization can significantly reduce the likelihood of failures and minimize the consequences of unforeseen situations such as seen by the recent Microsoft Windows and CloudStrike updates that impacted over 8 million computers which appears to be the latest failure of IT infrastructure. This approach ensures both the protection of information and data and the maintenance of the health of critical business processes*.

Keywords— *Highly resilient architecture, critical systems, operational risks, sustainability, backup systems, monitoring, flexibility, reliability, ecosystem integration, risk management*.

I. INTRODUCTION

In the context of the rapid development of information technologies and globalization, critical systems such as energy grids, transportation infrastructure, and healthcare systems play a key role in ensuring the stable functioning of modern society. The resilience of these systems to both external and internal threats and operational impacts has become a priority for organizations aiming to minimize operational risks and prevent significant disruptions. A highly resilient infrastructure architecture, as the foundation of reliability and security for critical systems, represents a key area of research and practical development in the fields of information technology and engineering.

The relevance of developing a highly resilient infrastructure architecture is driven by the increasing number and complexity of threats faced by critical systems.

The objective of this work is to explore and develop a highly resilient infrastructure architecture for critical systems to minimize operational risks and ensure the continuity of their functioning.

1. Incident Analysis: The Major CrowdStrike Outage of 2024

In July 2024, CrowdStrike, a leading global provider of cybersecurity solutions, experienced a severe service outage caused by a faulty software update. This incident led to immediate operational halts across various sectors of the global economy and sparked widespread discussions about cybersecurity dependence, software update protocols, and corporate responsibility.

The event occurred on July 19, 2024, at 04:09 UTC, when CrowdStrike released an update for its Falcon security platform. The update inadvertently triggered failures in millions of Microsoft Windows systems worldwide, resulting in "blue screens of death" and endless reboot cycles, which required manual intervention to restore functionality. The timing of the incident exacerbated its impact: it struck during peak working hours in Oceania and Asia, early morning in Europe, and late at night in the Americas, causing varying levels of disruption across regions.

The immediate consequences were vast. In the aviation sector, flights were delayed or canceled, with Delta Airlines particularly affected, leaving thousands of passengers stranded. In healthcare, hospitals faced disruptions to critical digital systems, potentially jeopardizing patient care. The financial sector experienced transaction failures, resulting in estimated losses exceeding one billion dollars. Emergency services were also affected: several 911 call centers were offline for days, potentially impacting public safety.

Technical analysis revealed that the root cause of the failure was a logical error in the Falcon sensor update, specifically in Channel File 291. System design flaws included inadequate boundary checks in the content interpreter and insufficient testing protocols, which failed to detect parameter mismatches before the update deployment. These oversights highlighted the need for stricter quality control and validation procedures in software update processes.

The legal and financial repercussions were significant. Although CrowdStrike was found to have minimal liability in most jurisdictions due to service agreement terms, the incident sparked debates about the legal responsibility of cybersecurity providers. The global economy suffered substantial losses: the healthcare and banking sectors reported losses of $1.94 billion and $1.15 billion, respectively. Delta Airlines reported losses of $500 million, leading to public disagreements between the CEOs of Delta and CrowdStrike.

In response to the crisis, CrowdStrike released a preliminary incident report and pledged to provide a detailed root cause analysis. The company implemented enhanced testing and validation processes, developed a phased rollout strategy to prevent simultaneous impact on all users, and offered clients more flexible control over update installations. Industry and

government bodies, including the U.S. Cybersecurity and Infrastructure Security Agency (CISA), worked closely with CrowdStrike to accelerate recovery efforts, while Microsoft provided tools and guidance to restore affected systems.

The incident underscored the risks of over-reliance on a single cybersecurity provider and fueled discussions about the need for diversification in the protection of critical infrastructure. Debates emerged about potentially tightening regulations around software update processes in critical sectors. Additionally, the need for automated and remote recovery solutions was highlighted, which may drive innovation in system resilience.

The 2024 CrowdStrike outage serves as an important lesson about the interconnectedness of modern digital infrastructure and the potential cascading effects of a single point of failure. The incident raises key questions about legal liability, the need for robust testing regimes, and the importance of diversifying critical system dependencies to mitigate similar risks in the future.

2. Fundamentals of Highly Resilient Architecture

The foundational principles for creating resilient infrastructure provide a set of recommendations and actions aimed at significantly enhancing resilience at a national level. This includes improving the continuity of key services such as energy supply, transportation systems, water resources, wastewater treatment systems, waste management, and digital communication, which are essential for the stable functioning of sectors like healthcare and education [1].

The key indicators for assessing system resilience are presented in Table 1.

TABLE 1. Key indicators used in assessing the stability of systems [2].

Key Indicators for System Resilience Assessment	**Description**
Availability	This indicator reflects how long a system remains operational without interruptions. High availability means the system can endure minor disruptions with minimal downtime. In data architecture, this often requires the use of redundant components and failover mechanisms to minimize the impact of failures.
Reliability	This parameter measures the likelihood of system failures. High reliability means failures occur less frequently. To improve reliability, high-quality components, efficient error-checking processes, and regular testing under different conditions are used to identify and address vulnerabilities.
Recovery Time Objective (RTO)	This indicator defines the maximum acceptable downtime after a failure. Designing with a low RTO requires the creation of robust infrastructure and fast failover to backup resources, which may include automated recovery procedures and pre-prepared action plans.
Recovery Point Objective (RPO)	This indicator defines the maximum allowable data loss over time. For example, if the RPO is 30 minutes, the system must ensure that data loss does not exceed this time in the event of a failure. Achieving this requires regular backups and data synchronization to minimize potential losses.

To successfully design with failure considerations in mind, system and security architects must account for various potential scenarios and their consequences. This includes:
- Risk analysis: Identifying potential failure points in the system and assessing their impact.
- Redundancy: Creating backup systems or components that can take over in the event of a failure.
- Effective backup and recovery solutions: Regular backups and recovery plans to minimize data loss and speed up service recovery.
- Constant monitoring and testing: Regularly tracking system performance and conducting stress tests to identify vulnerabilities.
- Team training: Ensuring all team members understand the risks and procedures to follow in different failure scenarios.

The third aspect of resilience requires protecting infrastructure through thoughtful design aimed at risk prevention. This implies:
- Exceeding the minimum requirements for critical infrastructure elements.
- Careful design of the interconnections between different systems to prevent cascading failures.
- Developing and implementing detailed emergency response plans.

Finally, resilience requires ongoing education and knowledge updating. This means that developers and managers must regularly review and test their strategies in practice to optimize the infrastructure's ability to address new challenges [3].

3. Methods for Reducing Operational Risks in Critical Systems

A system that plays a key role in an organization's operations consists of components whose performance directly affects the viability and resilience of the business. Depending on the structure of the hosting environment and the specifics of the tasks performed, such systems may have a wide range of functionalities or be narrowly specialized, focusing on specific operations. Typically, critical systems are integrated with other subsystems and applications, ensuring the execution of core business processes and the achievement of key performance indicators.

If a failure occurs in the operation of a critical system or its functionality is compromised, it leads to significant consequences for the company's operational activities. Examples of such systems include aviation control systems, energy systems, and emergency communication systems. Although the risk to human life may not always be direct, for many companies, critical business systems form the foundation for creating value and accomplishing key tasks.
A system can be considered critical for an organization's tasks if at least one of the following conditions is present, as reflected in Figure 1.

To ensure reliable management of critical systems, it is necessary to call out diversifying critical tooling such as EDRs (avoiding another cloudstrike) and supporting agents that have privileged access on the systems. VM agents, patching agents, etc.

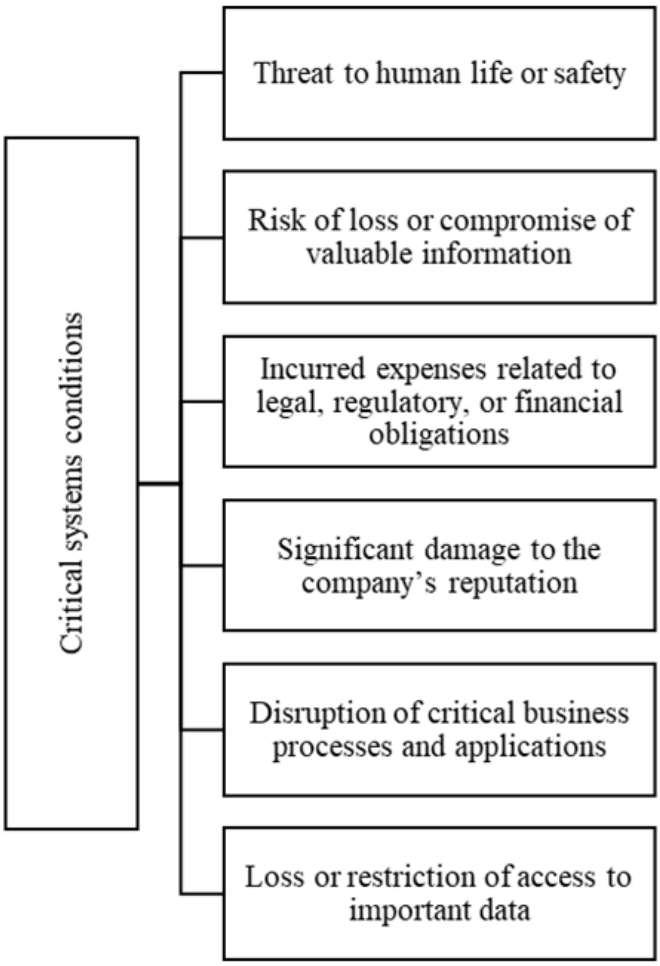

Fig.1. Critical systems conditions [4].

4. Diversification Strategies in Critical System Architecture

The development of highly resilient architecture for critical systems requires the application of diversification strategies, particularly to mitigate operational risks such as those highlighted by the CrowdStrike incident. In this context, diversification refers to creating an ecosystem where the failure of one component does not trigger a cascading effect across the entire system.

One key area is the diversification of endpoint detection and response (EDR). These devices are crucial for immediate action during failures or attacks. To avoid reliance on a single type or supplier, it is recommended to use EDRs from different manufacturers, thereby preventing specific vulnerabilities or a broken update from becoming a single point of failure.

Diversification of virtual machines and hypervisors also plays a critical role in preventing system failures caused by software bugs or vulnerabilities in virtualization platforms. Using different hypervisor technologies (e.g., Azure, AWS, Citrix, VirtualBox, NetApp, VMware, KVM, Hyper-V) across infrastructure segments limits the impact of attacks or errors specific to a particular hypervisor. The adoption of container-based solutions, such as Docker or Kubernetes, alongside traditional virtual machines, allows for lighter, more portable application environments, reducing dependence on a single virtualization strategy. Regular testing and workload migration between different hypervisors ensure compatibility and operational readiness.

Diversification of other critical system components also contributes to increased resilience. In authentication systems, using a variety of identity providers prevents widespread failures or compromises if one provider is disrupted. As noted in [6], 'secondary authentication providers can serve as backup solutions, ensuring access to corporate resources in the event of a failure in the primary system or when critical vulnerabilities are identified'. Employing different network management tools for various parts of the network prevents the entire network from being compromised due to a vulnerability in a single tool [6]. Utilizing mixed data storage solutions, including on-premise, cloud, and hybrid storage, protects against failures in both data centers and cloud service providers.

Effective management of critical systems requires a systematic and iterative approach that involves several stages, as outlined in Table 2.

TABLE 2. Stages of Effective Critical System Management [4]

Stage Name	Stage Description
Identification	Conducting thorough assessments and creating an organizational map to identify all key systems that play a significant role in the company's operations. Critical processes should be isolated if necessary, and data should be regularly backed up.
Interrelationship Analysis	Defining connections between systems to assess potential risk impacts. This helps establish the importance of each system and its role in organizational functionality.
Implementation	Critical system management is an ongoing process that requires integrating new data and business processes into existing security standards. It is important to consider the impact of changes on critical systems.
Monitoring and Analysis	Continuous real-time monitoring and risk analysis allow for the timely identification of potential threats and mitigation measures. Regular reports and audits from external IT providers offer additional data for decision-making [4].

Thus, diversification in critical system architecture goes beyond merely having backups; it involves creating a system where the failure of one component does not lead to a domino effect. By integrating these diversification strategies, organizations can achieve greater resilience to operational risks, ensuring that incidents like the CrowdStrike outage are less likely to result in system-wide failures. Such an approach requires careful planning, investment in diverse technologies, and fostering a culture of resilience that values redundancy through diversity.

It is important to note that, despite the importance of risk management in operational activities, only about 46% of companies participating in a recent ERM Initiative study have an official risk management policy. For instance, to prevent data loss and disruptions to key business networks, data centers and servers must be protected from various risks, including cyberattacks, power outages, and hardware failures. Proper management of these aspects is essential for ensuring organizational operational stability.

International Journal of Scientific Engineering and Science
Volume 9, Issue 1, pp. 30-35, 2025. ISSN (Online): 2456-7361

The constant increase in connected devices introduces new security threats affecting critical systems. According to the latest Edgescan report, 19% of vulnerabilities in 2018 were related to web applications, while 81% were tied to network vulnerabilities [4]. This highlights the need for a comprehensive approach and clear organization of processes in managing critical systems, ensuring the company's stable operation and protection from various risks.

5. Practical Implementation of Highly Resilient Architecture in Organizations

To demonstrate the practical application of highly resilient architecture principles, particularly through the diversification of security tooling, a case study of the Metropolitan Health Center (MHC) is presented. MHC is an anonymized large-scale hospital with over 800 inpatient beds, serving a major metropolitan area in the United States.

The MHC infrastructure architecture employs a comprehensive diversification strategy that extends beyond a simple assortment of security tools. The primary production environment is hosted in AWS, secured by an Endpoint Detection and Response (EDR) solution from Vendor 1, while the disaster recovery (DR) environment is hosted on-premises using a different EDR solution from Vendor 2. This dual-environment approach ensures multi-layered resilience by incorporating the following:

- Infrastructure Diversity: Primary systems hosted in the AWS cloud, with local on-premises backup systems for disaster recovery.
- Security Tool Diversity: Different EDR solutions deployed in each environment.
- Network Diversity: Separate network paths and controls for each environment.
- Geographical Diversity: Physical separation between the primary and backup sites.

Figure 2 illustrates the current and future state of the architecture.

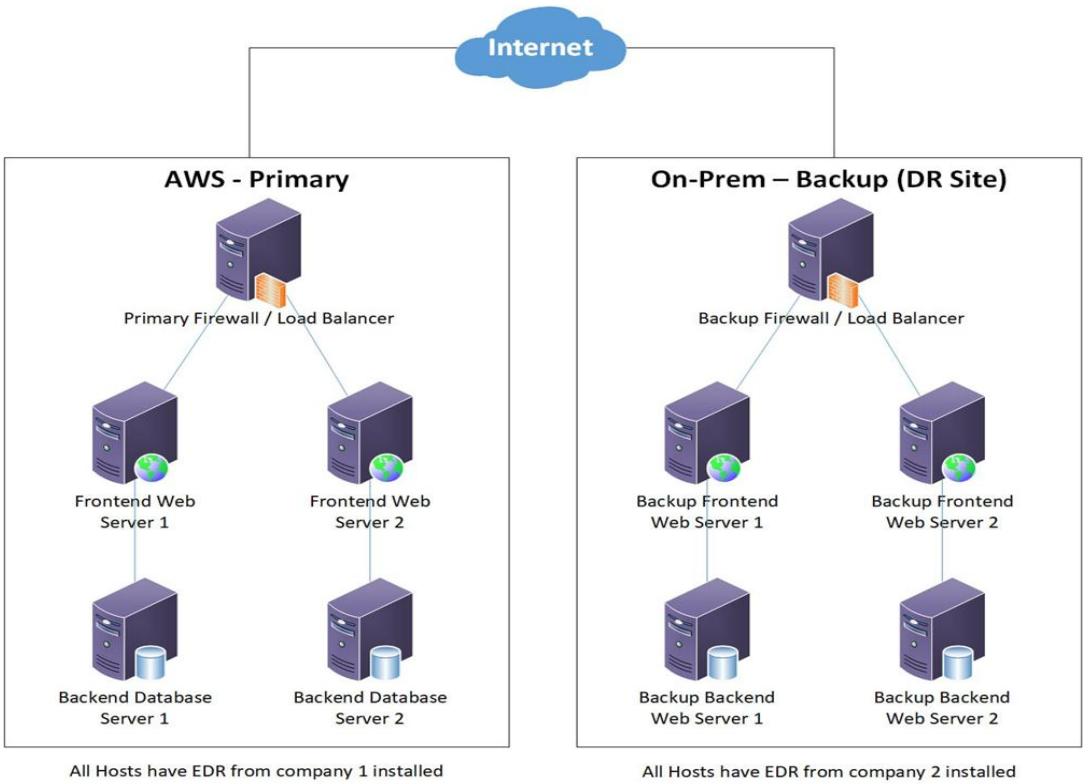

Fig.2. The future and current state of architecture

From Figure 2, the primary environment (AWS) includes:

- Interface web servers with load balancing;
- Backup database servers;
- Primary firewall and load balancer;
- EDR solution from Vendor 1.

The disaster recovery environment (local DR site) includes:

- Corresponding configurations of interface and server systems;
- Backup firewall and load balancer;
- EDR solution from Vendor 2;
- Independent network and security management tools.

Key systems include an Electronic Health Records (EHR) system, medical device integration and monitoring platforms, laboratory and radiology information systems, a pharmacy management system, and hospital resource management and scheduling tools. Given the life-critical nature of its services and the sensitive patient data it handles, MHC recognized the imperative need for a highly resilient architecture to mitigate risks associated with potential security incidents or system failures.

In the aftermath of the major CrowdStrike outage of 2024, MHC conducted a comprehensive risk analysis of its IT

infrastructure. The analysis revealed a critical vulnerability: the organization's heavy reliance on a single Endpoint Detection and Response (EDR) solution across all its systems. This dependency posed several significant risks. Firstly, it created a single point of failure in the cybersecurity infrastructure, increasing the potential for widespread system downtime that could adversely affect patient care. Secondly, it heightened the organization's vulnerability to specific types of cyber attacks, particularly those targeting healthcare institutions. Thirdly, in the event of an EDR system failure, extended recovery times could compromise patient safety. Lastly, potential non-compliance with healthcare data protection regulations due to security gaps could have legal and financial repercussions.

Recognizing these risks, MHC prioritized addressing them by focusing on diversifying its EDR solutions as a starting point for enhancing overall resilience.

MHC developed a phased approach to implement a dual-EDR strategy, aiming to enhance its cybersecurity resilience while minimizing disruption to its critical operations.

Phase 1: Planning and Preparation (3 months)

During this initial phase, MHC conducted a thorough evaluation of alternative EDR solutions, taking into account healthcare-specific requirements. A secondary EDR provider was selected based on factors such as compatibility with existing systems, complementary features, and experience within the healthcare industry. A new security architecture incorporating both EDR solutions was designed, and new Security Operations Center (SOC) procedures were developed for managing the dual EDR systems. A detailed implementation plan and timeline were created, carefully considering the 24/7 nature of hospital operations.

Phase 2: Pilot Implementation (4 months)

In this phase, the new EDR solution was deployed on 5% of endpoints, focusing on non-critical administrative systems to minimize risk. The SOC team received training on the new procedures and tools, with an emphasis on healthcare-specific threat scenarios. Baseline metrics for selected Key Performance Indicators (KPIs) were replicated from the primary EDR solution. Initial testing and optimization were conducted to ensure there was no disruption to patient care.

Phase 3: Deployment to a Single Critical System (3 months, planned)

The next phase involves deploying the new EDR solution to the secondary environment of a single critical production system, such as the backup Electronic Health Records system. Automated failover mechanisms between the EDR systems will be implemented in this environment. New backup and recovery procedures specific to this critical system will be developed and tested. Targeted training will be conducted for staff managing this critical system.

Phase 4: Testing, Monitoring, and Optimization (6 months, planned)

This phase will involve comprehensive testing of the secondary-EDR setup in the critical system's secondary environment. Continuous monitoring of performance and security metrics will be carried out. EDR configurations will be optimized based on observed performance and any detected issues. Various failure scenarios will be simulated to test resilience and failover capabilities. SOC procedures will be refined based on learnings from this phase.

Phase 5: Expanded Rollout (12–18 months, planned)

The final phase entails a gradual rollout of the optimized secondary-EDR solution to other critical production systems, prioritizing based on criticality and insights gained from previous phases. The refined automated failover mechanisms will be implemented across these systems. New backup and recovery procedures will be extended to each system as it is incorporated into the secondary-EDR architecture. Comprehensive staff training on new security protocols will be conducted for each affected system, including clinicians and support staff. Monitoring, testing, and optimization will continue as the rollout progresses.

The implementation of the dual-EDR strategy is expected to face several challenges. Increased operational complexity is anticipated due to the introduction of additional systems and procedures. This will be mitigated through a phased implementation, extensive staff training, and the development of clear, documented procedures tailored to each critical system.

Higher costs associated with deploying and maintaining an additional EDR solution are another concern. These costs will be justified through a cost-benefit analysis demonstrating long-term savings from reduced downtime, faster incident response, and improved regulatory compliance. The phased approach allows for better cost management and justification.

An initial performance impact on critical systems may occur during the implementation. This will be addressed through careful tuning and optimization during Phase 4, ensuring no degradation in system performance before wider rollout.

Integrating the new EDR solution with existing medical systems presents technical challenges. These will be resolved through close collaboration with both EDR vendors, in-house IT teams, and medical device manufacturers. Learnings from each phase will inform subsequent integrations.

In turn, synchronization across environments: must be addressed through robust data replication mechanisms and regular synchronization checks between AWS and on-premises systems.

Compatibility between different environments: This should be ensured through extensive testing of all applications and systems in both AWS and on-premises environments.

Network latency considerations: This should be reduced by optimizing network paths and monitoring inter-network communications.

Backup site readiness: Maintained through regular testing and verification of the local environment's ability to perform critical operations.

Complex failover procedures: Simplified through automation and clear documentation, as well as regular drills to ensure staff preparedness.

Finally, the requirement for minimal disruption to hospital operations necessitates meticulous planning. This will be achieved by utilizing secondary environments for initial deployment and carefully planning production rollouts to avoid impacting patient care.

II. Conclusion

In conclusion, the development of a highly resilient infrastructure architecture for critical systems is an essential part of ensuring the resilience and security of organizations in the modern world. The methods and approaches discussed demonstrate how the integration of flexible and reliable architectural solutions contributes to reducing operational risks and protecting key business processes. An important aspect is the continuous monitoring, testing, and adaptation of systems to new conditions and challenges, ensuring their long-term functionality. It is recommended to continue research in this field, with particular attention to new technologies and approaches that enhance the resilience of infrastructures against future threats.

References

1. Sherman J. A. et al. A resilient architecture for the realization and distribution of Coordinated Universal Time to critical infrastructure systems in the United States: Methodologies and recommendations from the National Institute of Standards and Technology (NIST). – 2021.
2. Shahbazi A. et al. Hybrid stochastic/robust optimization model for the resilient architecture of distribution networks against extreme weather conditions //International Journal of Electrical Power & Energy Systems. – 2021. – T. 126. – C. 106576.
3. Gardoni P., Murphy C. Society-based design: promoting societal well-being by designing sustainable and resilient infrastructure //Sustainable and Resilient Infrastructure. – 2020. – T. 5. – №. 1-2. – C. 4-19.
4. Abdelgawad M., Ray I., Vasquez T. Workflow Resilience for Mission Critical Systems //International Symposium on Stabilizing, Safety, and Security of Distributed Systems. – Cham: Springer Nature Switzerland, 2023. – C. 498-512.
5. Garagad V. G., Iyer N. C., Wali H. G. Data integrity: a security threat for Internet of things and cyber-physical systems //2020 International Conference on Computational Performance Evaluation (ComPE). – IEEE, 2020. – C. 244-249.
6. Raoul Hira. (2024). Implementing Comprehensive Identity Continuity Plans To Counteract Cyber Threats. The American Journal of Engineering and Technology, 6(12), 59–67. https://doi.org/10.37547/tajet/Volume06Issue12-07