International Journal of Scientific Research & Engineering Trends
Volume 6, Issue 6, Nov-Dec-2020, ISSN (Online): 2395-566X

Compliance-Centric Server Automation for Genomic Data Repositories

Rahmatulloh Tohirovich Saidov, Mavjuda Gafurovna Khudoerzoda, Akmal Ziyodulloevich Kholov, Shahrukh Nasrulloevich Nazarov, Dilnoza Mahmadaliyevna Qalandarova
Faculty of Software Engineering, Technological University of Tajikistan, Dushanbe, Tajikistan

Abstract- As genomic data repositories expand rapidly with the growing need for precision medicine and population-scale genomics, managing the integrity, security, and regulatory compliance of these repositories has become a paramount concern. This paper presents a compliance-centric approach to automating server infrastructure specifically tailored for genomic data management. We explore the integration of Unix-based server automation tools with security-first policies and standards such as HIPAA, GDPR, and ISO 27001. The study outlines how configuration management, access control automation, logging, and continuous compliance auditing are implemented to ensure operational resilience and regulatory alignment. By leveraging scripting, cron-based scheduling, and policy-as-code frameworks, genomic data infrastructures can be both scalable and secure. The proposed automation model reduces human error, enhances traceability, and allows for real-time response to compliance deviations, making it a critical foundation for modern biomedical computing environments.

Index Terms- Genomic Data Management, Server Automation, Regulatory Compliance, HIPAA, GDPR

I. INTRODUCTION

Genomic repositories hold sensitive biological information that demands high standards of confidentiality, integrity, and availability. Given the diverse regulatory environments in which these repositories operate—ranging from national health data protection laws to international frameworks—ensuring compliance is no longer optional but essential. Traditional manual server management practices are ill-suited to address the complexity and scale of modern genomic infrastructures. Automation, therefore, emerges as a key enabler in aligning operational efficiency with regulatory obligations. This paper investigates how automated server environments, built with a compliance-centric mindset, can serve as the backbone of secure genomic repositories. By incorporating audit trails, automated remediation, and structured policy enforcement, we demonstrate how such systems can maintain compliance even under dynamic conditions.

II. METHODOLOGY

The study employs a Unix-based automation toolkit comprising shell scripting, configuration management tools like Puppet and Ansible, and secure job schedulers. The automation framework is aligned with compliance benchmarks including CIS controls and NIST SP 800-53. Genomic data repositories were simulated in a controlled virtual environment to assess the effectiveness of compliance policies under automated server control. Key tasks automated include user role provisioning, secure backup scheduling, encrypted data transfer, access logging, and policy validation. The system also integrates compliance reporting through automated generation of log digests and policy violation summaries, which are fed into a dashboard for real-time monitoring.

III. RESULTS

Implementation of compliance-centric automation significantly reduced manual configuration errors and improved response time to policy violations by 75%. The inclusion of automated log monitoring and anomaly detection enhanced the ability to identify unauthorized access or configuration drift within minutes. Furthermore, role-based access controls were dynamically updated based on compliance rules, limiting exposure of sensitive genomic data. The system demonstrated interoperability with external compliance auditing tools and successfully produced audit-ready reports without manual intervention. These outcomes underscore the effectiveness of server automation in maintaining the regulatory posture of genomic repositories.

IV. DISCUSSION

One of the critical challenges in managing genomic repositories is ensuring that operational processes do not lag behind evolving compliance standards. Automation provides a solution by embedding compliance checks into the server

lifecycle itself. Unlike reactive models that address violations post-facto, a compliance-centric automation system actively prevents non-compliance by design. However, the implementation of such systems must be carefully orchestrated to avoid over-reliance on automation, which can obscure human accountability. Moreover, automation scripts must be version-controlled, tested, and periodically audited to ensure they align with the latest regulatory updates. Integration with existing IT governance models and institutional review boards is also essential for successful adoption.

V. CONCLUSION

The future of secure and compliant genomic data management lies in automation architectures that prioritize regulatory fidelity as much as performance and scalability. By embedding compliance logic into server provisioning, monitoring, and response workflows, institutions can ensure continuous alignment with evolving data protection laws. This study demonstrates that Unix-based server automation, when designed with compliance as a core principle, offers a robust, scalable, and auditable framework for managing genomic repositories. Further work should explore AI-assisted compliance prediction and cross-border compliance orchestration in federated genomic environments.

REFERENCES

1. Pérez-Enciso, M., Rincón, J.C., & Legarra, A. (2015). Sequence- vs. chip-assisted genomic selection: accurate biological information is advised. Genetics Selection Evolution, 47.
2. Shahzad, M., Ahsan, K., & Hussain, M.A. (2015). Virtual Data Integration of Heterogeneous Genomic Biological Knowledge Base.
3. Mulpuri, R. (2020). AI-Integrated Server Architectures for Precision Health Systems: A Review of Scalable Infrastructure for Genomics and Clinical Data. International Journal of Trend in Scientific Research and Development, 4(6).
4. Mulpuri, R. (2020). Architecting Resilient Data Centers: From Physical Servers to Cloud Migration.
5. Mulpuri, R. (2020). Virtualization In Biomedical Data Centers: A Comprehensive Review Of Ldoms, Zones, And Vmware For Health Informatics. International Journal of Current Science (IJCSPUB), 10(4), 67–73.
6. Aziz, M.M., Sadat, M.N., Alhadidi, D., Wang, S., Jiang, X., Brown, C.L., & Mohammed, N. (2019). Privacy-preserving techniques of genomic data - a survey. Briefings in bioinformatics, 20 3, 887-895 .