ENSURING THE SECURITY AND COMPLIANCE OF A LARGE-SCALE BIG DATA INFRASTRUCTURE USED FOR PROCESSING SENSITIVE DATA IN A HEALTHCARE ORGANIZATION.

Author:

M.Harish
2nd year
Saveetha School of Engineering
SIMATS

Guide:

Dr. Antony Joseph Rajan
Assistant professor (SG)
Saveetha School of Engineering
SIMATS

AGENDA

- ABSTRACT
- LITERATURE SURVEY
- PROPOSED METHODS
- RESULTS AND DISCUSSIONS
- CONCLUSION
- REFERENCES

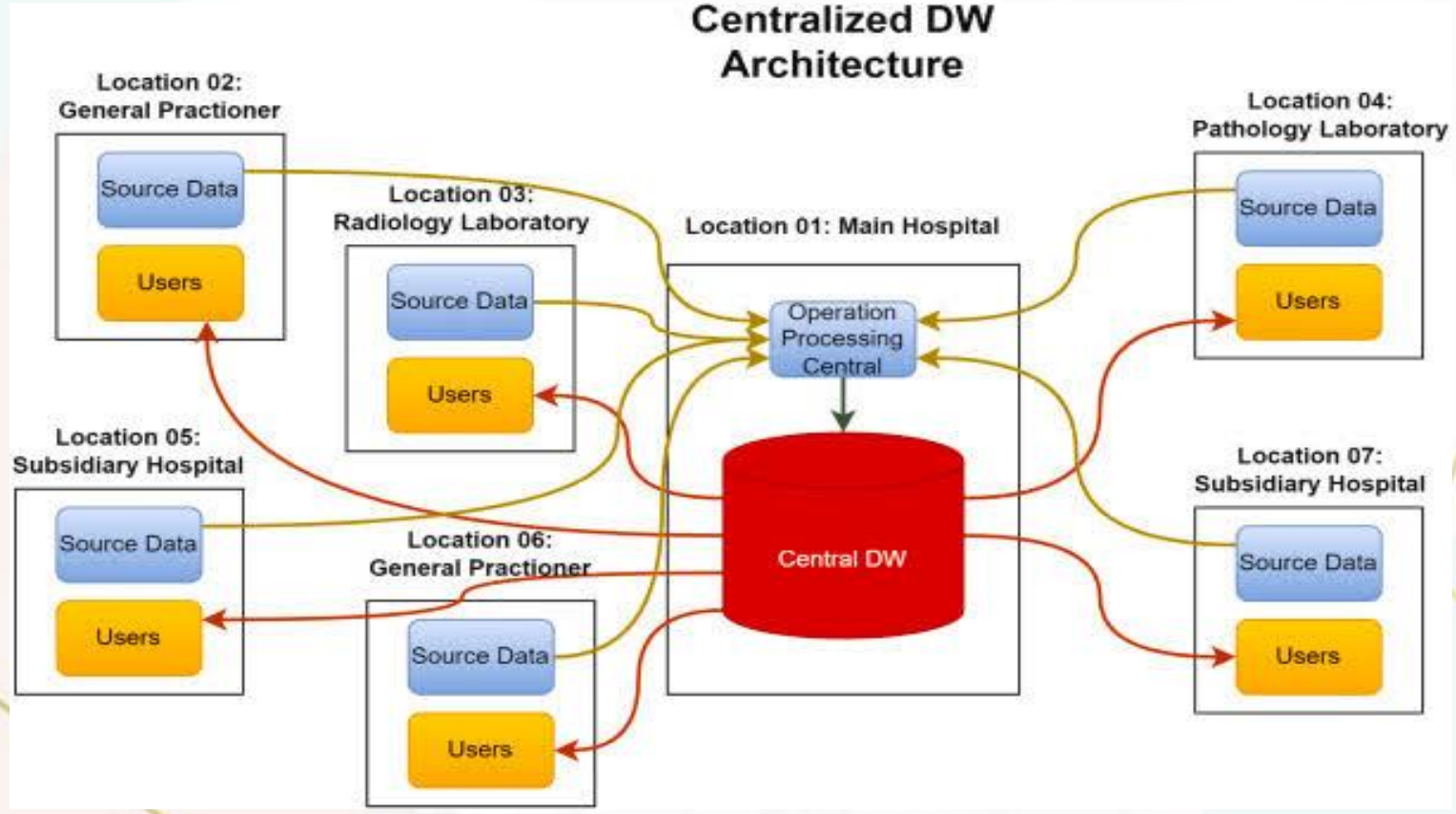
ABSTRACT:

- **Objective:** Secure and comply with regulations for a large-scale big data infrastructure.
- **Issue:** Vulnerabilities threatening sensitive healthcare data.
- **Importance:** Essential for maintaining patient privacy and data integrity.
- **Data Collection:** From system logs, user activities, and network traffic to detect threats.
- **Technology Stack:** Python, Java, Hadoop, Spark, Splunk, ELK Stack, compliance tools.
- **Development Phases:** Design, implementation, testing, and deployment of security measures.
- **Conclusion:** Critical for protecting against threats and ensuring regulatory compliance.

LITERATURE SURVEY

S.No	TITLE	YEAR	OBJECTIVE	PROS	CONS
1	Securing Big Data in Healthcare: Challenges and Solutions	Michael A. Brown 2019	To identify key security challenges in healthcare big data and propose a framework for mitigating risks.	Detailed framework for security, focus on regulatory compliance.	Framework not validated through empirical studies.
2	Big Data Analytics for Healthcare: Security and Privacy Challenges	David K. Wilson 2020	To discuss the implications of big data analytics on healthcare security and privacy.	In-depth discussion on privacy-preserving techniques.	Limited coverage of compliance with specific regulations like HIPAA.
3	Enhancing Data Security in Big Data Healthcare Applications	Richard P. Lee 2021	To propose methods for enhancing data security in healthcare big data applications.	Practical methods for data encryption and access control.	Methods may not be scalable for very large datasets.

METHODS


Centralized DW Architecture
Location 02: General Practioner
Source Data
Users
Location 03: Radiology Laboratory
Source Data
Users
Location 01: Main Hospital
Operation Processing Central
Central DW
Location 04: Pathology Laboratory
Source Data
Users
Location 05: Subsidiary Hospital
Source Data
Users
Location 06: General Practioner
Source Data
Users
Location 07: Subsidiary Hospital
Source Data
Users

CODING

HARISH.py - C:/Users/chellapadian/Desktop/HARISH.py (3.11.9)

File Edit Format Run Options Window Help

```
from cryptography.fernet import Fernet
import getpass

# Generate encryption key
key = Fernet.generate_key()
cipher_suite = Fernet(key)

# Example sensitive data (e.g., patient data)
sensitive_data = b"Patient ID: 12345, Name: John Doe, Diagnosis: Diabetes"

# Encrypt sensitive data
encrypted_data = cipher_suite.encrypt(sensitive_data)
print("Encrypted Data:", encrypted_data)

# Decrypt encrypted data (example)
decrypted_data = cipher_suite.decrypt(encrypted_data)
print("Decrypted Data:", decrypted_data.decode())

# Example of secure password input
password = getpass.getpass(prompt="Enter your password securely: ")
print("Entered Password:", password)
```

OUTPUT

```
Encrypted Data: b'gAAAAABmfOZEop6BzBm1nFVsoINtS3qggM7j5PsuiHKgIOeeqJ1a5Nq3EHZy6w
iI4a3GOGNqaIokpfTL0bwV1I8yiyWUf_1NaUqUYtA6nLAjmYHJPa2_UednNoqEFg4YKBwjD84f6FLTAY
19fJTZV-C01mbyrRXpUw=='
Decrypted Data: Patient ID: 12345, Name: John Doe, Diagnosis: Diabetes

Warning (from warnings module):
  File "C:\Program Files\WindowsApps\PythonSoftwareFoundation.Python.3.11_3.11.2
544.0_x64__qbz5n2kfra8p0\Lib\getpass.py", line 100
    return fallback_getpass(prompt, stream)
GetPassWarning: Can not control echo on the terminal.
Warning: Password input may be echoed.
Enter your password securely: Harish
Entered Password: Harish
```

CONCLUSION

- Robust security measures protect sensitive healthcare data from unauthorized access and breaches.
- Automation in security monitoring enhances operational efficiency and reduces response time to threats.
- Implementing and maintaining such systems can be resource-intensive but essential for data integrity and compliance.
- Real-time monitoring and advanced detection algorithms enable proactive prevention of data breaches.
- Ensures regulatory compliance and builds trust among patients and stakeholders.

FUTURE SCOPE

- **Enhanced Automation:** Further automation in security measures to improve efficiency and reduce manual intervention.
- **Advanced Algorithms:** Development of more sophisticated detection algorithms to identify new and evolving threats.
- **AI Integration:** Leveraging artificial intelligence for predictive analytics and proactive threat management.
- **Scalability:** Enhancing the scalability of security systems to accommodate growing data volumes.
- **Regulatory Updates:** Continuous adaptation to evolving healthcare regulations and compliance requirements.
- **User Education:** Increased focus on training and awareness programs for users to minimize human-related security risks.
- **Collaboration:** Fostering collaboration between healthcare organizations and cybersecurity experts for shared insights and best practices.