**Log Intelligence System Using AWS & AI/ML - Requirement Document**

**1. Overview**

This system will collect, store, process, and analyze logs from non-AWS devices and software, filtering logs based on severity (Critical/High/Warnings) and notifying the respective team via Microsoft Teams or Email. The system will leverage AWS and AI/ML services for intelligent log processing and anomaly detection.

**2. Functional Requirements**

**2.1 Multi-Source Log Collection**

* Collect logs from multiple sources:
  + On-premises servers (Windows/Linux)
  + Network devices (firewalls, routers, switches)
  + Applications (databases, ERP, CRM, SaaS tools)
  + IoT devices (sensors, industrial machines)
* Support for different log formats:
  + Syslog, JSON, CSV, XML, Flat files
  + Custom log structures
* Support **real-time ingestion** (AWS Kinesis, Firehose) and **batch ingestion** (S3 uploads, DataSync).

**2.2 Intelligent Log Processing & Classification**

* Standardize log formats (AWS Glue, Lambda functions).
* AI/ML-based log classification:
  + Categorize logs as **Critical, High, Medium, Low, Warning**.
  + Use **Amazon Comprehend or SageMaker ML models**.
* Remove redundant logs (deduplication & noise filtering).
* Log enrichment: Append **geolocation, user identity, metadata**.

**2.3 Advanced Search, Query & Analytics**

* **Fast search & indexing** using **Amazon OpenSearch**.
* Custom **query language** for log analysis.
* **Anomaly detection** (CloudWatch Insights, SageMaker anomaly detection).

**2.4 Real-Time Alerting & Notifications**

* Multi-channel alerting:
  + **Microsoft Teams, Slack, Email (SES), SMS, PagerDuty, Webhooks**.
* Auto-remediation workflow (AWS Systems Manager, Lambda, Step Functions).
* Incident correlation with past logs for historical context.

**2.5 Secure & Scalable Storage Management**

* Store logs in **Amazon S3 (Standard, Glacier, Intelligent-Tiering)**.
* **Retention policies** based on compliance (90 days, 1 year, etc.).
* **Immutable logs** with S3 Object Lock & WORM (Write Once, Read Many).

**2.6 Role-Based Access Control (RBAC) & Auditing**

* **IAM Role-based access** to logs based on teams.
* **CloudTrail logging** for tracking access and modifications.
* Compliance audit logs stored securely.

**2.7 Dashboards & Reporting**

* **Amazon QuickSight / Grafana dashboards** for trends, system health.
* **Scheduled reports** (weekly/monthly security logs in PDF, Excel).
* **Root Cause Analysis (RCA) insights** for incidents.

**3. Non-Functional Requirements (NFRs)**

**3.1 Scalability & High Availability**

* Support for **millions of log entries per second**.
* Auto-scale ingestion using **Kinesis & Lambda**.
* Multi-region deployment for **disaster recovery**.

**3.2 Performance & Latency**

* Real-time processing **(<5 seconds delay)**.
* Query response time **<2 seconds for recent logs, <10 seconds for archived logs**.

**3.3 Security & Compliance**

* **End-to-end encryption** (in transit & at rest using AWS KMS).
* **MFA & SSO authentication** for dashboard access.
* Compliance with **GDPR, HIPAA, ISO 27001, SOC 2**.

**3.4 Reliability & Fault Tolerance**

* **Multi-AZ deployments** for high availability.
* **Auto-retries & failover mechanisms** for processing pipelines.

**3.5 Cost Optimization**

* **Tiered storage** (S3 Standard for active logs, Glacier for archives).
* **Event-driven processing (Lambda, Kinesis Firehose)** to avoid unnecessary costs.
* **Lifecycle policies** to auto-delete old logs based on retention.

**4. Security Enhancements**

**4.1 Data Ingestion Security**

* **TLS encryption for log transfer**.
* IAM-based **access control** for log pushers.
* **Signature validation (AWS SigV4)** for API requests.

**4.2 Storage Security**

* **S3 encryption using KMS**.
* **S3 Object Lock & Versioning** to prevent tampering.
* **S3 access logging & CloudTrail monitoring**.

**4.3 Processing Security**

* **Private VPC for SageMaker & Lambda**.
* **Data integrity checks (Glue DataBrew, Lambda validation)**.

**4.4 Monitoring & Threat Detection**

* **AWS GuardDuty & Macie** for threat detection.
* **Security Lake & SIEM integration** for advanced security monitoring.

**4.5 Notification & Response Security**

* **SNS topic policies** to restrict unauthorized alerts.
* **Teams Webhooks secured via OAuth/JWT authentication**.
* **Automated incident response via AWS Systems Manager & Security Hub**.

**5. Future Enhancements**

* **AI-Powered Insights**: Predict potential failures using ML models.
* **Voice Alerts**: Integrate with Alexa for voice-based log alerts.
* **Custom Chatbot**: Deploy an AI chatbot (Amazon Lex) for log queries.

**Conclusion**

This document outlines the comprehensive **functional, security, and performance requirements** for the Log Intelligence System. The system will leverage **AWS & AI/ML** to enhance log analysis, automate anomaly detection, and enable real-time alerting. By integrating security best practices, the system will ensure **data integrity, compliance, and scalability**.