Graylog SIEM Capacity Planning

(Catnip Games International)

Introduction

This document outlines the capacity planning for the Security Monitoring Competency Experience project. It defines the resources required to meet the system's performance, storage, and availability requirements for both the production environment and the prototype implementation.

Infrastructure Overview

Catnip Games International's gaming infrastructure requires comprehensive security monitoring across several key components:

Our gaming environment consists of:

- 300 Linux game servers distributed across two data centers
- Player authentication systems managing login attempts and sessions
- Development environments for game creation and testing
- Network infrastructure connecting all components

Each component generates specific security-relevant events that our SIEM system must monitor:

- Game server logs capture potential DDoS attacks and performance issues
- Authentication systems track unauthorised access attempts
- Development environment logs identify suspicious activities
- Network traffic logs enable security threat analysis

Production Requirements

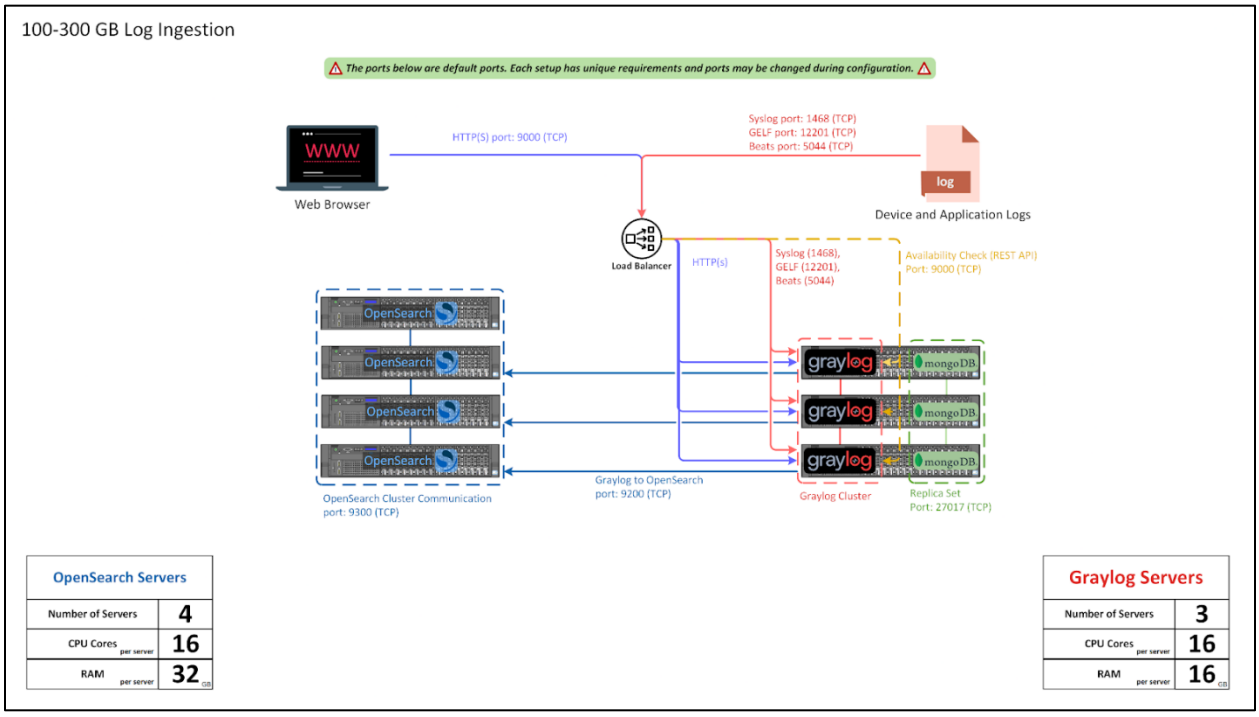
Our production system needs to meet the following performance metrics:

- Process 10,000 Events Per Second (EPS)
- Maintain 30 days of security data in hot storage
- Provide query responses within 5 seconds
- Ensure 99.9% system availability for continuous monitoring

Resource Analysis

Reference Architecture

The following architecture shows Graylog's recommendation for deployment sizes between 100–300GB of log data per day. We referred to it and Graylog's official documentation regarding planning the deployment.



Based on Graylog's recommendations for environments processing 100-300GB (Range that Catnip Games falls into) of log data daily, the reference architecture suggests:

Core Components:

- Three Graylog servers: 16 CPU cores, 16GB RAM each
- Four OpenSearch servers: 16 CPU cores, 32GB RAM each
- Load balancer for traffic distribution
- High availability configuration for system resilience

Storage Requirements

Our calculations for security event storage indicate:

- Average individual event size: 200 bytes, but we'll overestimate slightly to 300 bytes.
- Daily data volume: 259 GB (10,000 EPS × 300 bytes × 86,400 seconds (per day))

- Monthly storage need: 7.8 TB raw data
- Compressed storage requirement: 2.5-3.9 TB
- Recommended allocation: 4-5 TB fast storage for operational buffer

Prototype Implementation

Our prototype is designed to implement core functionality at approximately 5-10% of the original production load. This scaling allows us to demonstrate the system's capabilities while working within the constraints of our available resources.

VM Specifications

Specification	Ubuntu Server VM	Ubuntu Desktop VM
Purpose	Hosts core SIEM components; Graylog, Elasticsearch, MongoDB.	Provides access to Graylog's web interface for Dashboard Creation, Log Search, Analysis.
Role	Data Processing and Storage	User Interface and Interaction
CPU	4 vCPUs	2 vCPUs
RAM	8000 MB	4096 MB
Storage	65 GB	30 GB
Operating System	Ubuntu Server	Ubuntu Desktop

Scaled Requirements

Processing Capacity:

- 500-1,000 EPS (scaled down from 10,000 EPS based on our resources and working environment)
- 3-5 days retention (scaled from 30 days)
- Approximately 13 GB daily raw security data at 500 EPS
- 5 second query response time

Planned Implementation Timeline

Week 1: Architecture Setup

- VM configuration
- Basic SIEM deployment

Weeks 2-3: Data Collection

- Security log ingestion setup
- Parser optimisation
- Initial dashboard development

Weeks 4-5: Security Features

- Security correlation rules
- Alert configuration
- Performance optimisation
- Documentation preparation

Week 6: Assessment Preparation

- Final testing
- Demonstration preparation
- Documentation completion

Technical Considerations

Assumptions and Constraints

Assumptions:

- Average log size: 200 bytes (However, based on research)
- Compression ratio: 2-3×
- Peak traffic: 1,000 EPS

Constraints:

Timeline: 6-week implementation

Limitations

Our prototype has several constraints:

- Limited simultaneous event processing
- Reduced data retention period
- Restricted storage capacity

Risk Mitigation

To address these limitations:

- We will focus on core security monitoring capabilities
- Scaling considerations have been documented above for production deployment
- Maintain clear performance expectations

This capacity plan ensures that the prototype, while operating at reduced capacity, will demonstrate essential security monitoring capabilities while documenting considerations for full production deployment. The proposed resource allocation is based on best practices and documentation.

References

Graylog. (n.d.). *Planning your deployment.* Graylog Documentation. Retrieved February 22, 2025, from https://go2docs.graylog.org/6-0/planning_your_deployment/planning_your_deployment.html

IBM. (n.d.). Estimating the size of message logs. IBM Documentation. Retrieved February 27, 2025, from https://www.ibm.com/docs/en/sva/11.0.0?topic=logging-estimating-size-message-logs