УДК 658.8:004.9

ПРОЕКТИРОВАНИЕ ЦЕНТРА МОНИТОРИНГА ИНФОРМАЦИОННОЙ БЕЗОПАСНОСТИ (SOC) В ИНФРАСТРУКТУРЕ УНИВЕРСИТЕТА С ПРИМЕНЕНИЕМ BIG DATA И ADVANCED ANALITYCS

Г.К. Стамкулова
Доцент кафедры программного обеспечения компьютерных систем КГТУ
g.stamkulova@kstu.*kg*

С.М. Салиев
Студент 4 курса кафедры программного обеспечения компьютерных систем КГТУ
salievsamat0@gmail.com

Стамкулова Г.К.

Окончила Национальный университет отделение прикладной математики, магистратуру по направлениям "Информатика и вычислительная техника" в КГТУ и "Информационная безопасность" в МИФИ. Область научных интересов связана с вопросами обеспечения информационной безопасности, проектированием распределённых систем мониторинга и организацией учебного процесса в техническом университете.

Салиев С.М.

Оканчивает 4 курс по направлению подготовки "Информационная безопасность" Кыргызского Государственного Технического университета им.И.Раззакова. Область научных интересов связана с мониторингом событий информационной безопасности, проектированием архитектур SOC, применением технологий Big Data в системах кибербезопасности

Аннотация. В статье рассматривается проектирование архитектуры учебно-экспериментального центра мониторинга информационной безопасности (SOC) для инфраструктуры университета с применением технологий Big Data. Предлагается концептуальная модель сбора, хранения и потоковой обработки событий безопасности на основе Apache Kafka и Elastic Stack. Обоснован выбор Java в качестве платформы разработки аналитического модуля корреляции событий. Представлены алгоритмы обнаружения типовых инцидентов и описана модель их формализации. Работа носит проектный характер и направлена на создание основы для дальнейшей практической реализации.

Ключевые слова: SOC, Big Data, информационная безопасность, потоковая аналитика, корреляция событий, Java, Apache Kafka, Elasticsearch.

Введение. Современный университет представляет собой распределенную информационную систему, включающую доменную инфраструктуру, системы дистанционного обучения, почтовые сервисы, сетевую инфраструктуру, базы данных и рабочие станции пользователей. Количество событий информационной безопасности, генерируемых такой инфраструктурой, исчисляется тысячами записей в минуту. В условиях роста кибератак на образовательные учреждения возникает необходимость централизованного мониторинга и анализа событий безопасности. Традиционные средства защиты не обеспечивают целостной картины происходящего и не позволяют выявлять сложные сценарии атак. Одним из решений является создание центра мониторинга информационной безопасности (SOC), основанного на принципах Big Data и потоковой аналитики [1].

Анализ особенностей инфраструктуры университета. Информационная инфраструктура современного университета представляет собой распределённую многоуровневую систему, включающую серверные ресурсы, сетевую инфраструктуру, пользовательские рабочие станции и внешние сервисы. В отличие от корпоративных сетей коммерческих организаций, университетская среда характеризуется высокой динамичностью и большим числом разнородных пользователей.

Ключевыми особенностями являются:

- значительное количество пользователей (студенты, преподаватели, административный персонал)
- регулярная смена контингента обучающихся
- использование модели BYOD (Bring Your Own Device)
- наличие удалённого доступа через VPN
- использование систем дистанционного обучения (LMS)
- активное применение веб-сервисов и облачных платформ
- наличие публично доступных сервисов

Инфраструктура зачастую включает:

- контроллеры домена (Active Directory)
- серверы баз данных
- файловые серверы
- веб-серверы
- почтовые серверы
- сетевые устройства
- точки доступа Wi-Fi
- виртуализированную серверную среду

Большой объём логируемых событий и распределённость инфраструктуры создают сложность централизованного анализа и выявления инцидентов информационной безопасности.

Модель угроз для образовательной среды. Университетская среда обладает повышенным уровнем уязвимости по сравнению с классическими корпоративными инфраструктурами. Это обусловлено открытостью информационной среды и разнообразием пользователей.

Угрозы можно классифицировать следующим образом.

1. Внешние угрозы
- перебор паролей (brute force) на сервисах удалённого доступа
- фишинговые атаки на учётные записи сотрудников и студентов
- эксплуатация уязвимостей веб-приложений
- атаки типа «отказ в обслуживании» (DoS).

2. Внутренние угрозы
- компрометация учётной записи студента
- распространение вредоносного программного обеспечения внутри локальной сети
- несанкционированный доступ к информационным ресурсам.

3. Инсайдерские угрозы
- превышение полномочий сотрудниками
- несанкционированное копирование данных
- использование служебных ресурсов в личных целях
- Для университета характерно сочетание внешних и внутренних угроз, что требует комплексного мониторинга событий безопасности.

Обоснование необходимости центра мониторинга. Не связанные друг с другом средства защиты генерируют собственные журналы событий, которые не всегда анализируются в совокупности. Отсутствие централизованного мониторинга затрудняет

выявление сложных сценариев атак, включающих последовательность взаимосвязанных событий.

Создание центра мониторинга информационной безопасности (SOC) позволяет:

- централизовать сбор событий безопасности
- обеспечить корреляцию данных из различных источников
- выявлять аномалии поведения пользователей
- формировать отчёты для руководства университета
- сокращать время обнаружения инцидентов.

Таким образом, особенности университетской инфраструктуры и специфика угроз обосновывают необходимость проектирования SOC, основанного на технологиях Big Data и потоковой аналитики.

Объект и предмет исследования. Объектом исследования является информационная инфраструктура университета, включающая серверные ресурсы, сетевую инфраструктуру, рабочие станции пользователей и системы дистанционного обучения. Предметом исследования являются процессы централизованного сбора, обработки и анализа событий информационной безопасности с применением технологий Big Data и потоковой аналитики.

Цель работы. Целью работы является разработка архитектурной модели учебно-экспериментального центра мониторинга информационной безопасности (SOC) для университета, обеспечивающего централизованный сбор, хранение и анализ событий безопасности в условиях высокой нагрузки и разнородности источников данных.

Задачи исследования. Для достижения поставленной цели необходимо решить следующие задачи:

- Проанализировать особенности инфраструктуры университета и характерные угрозы информационной безопасности.
- Обосновать применение технологий Big Data для обработки потоков событий безопасности.
- Разработать концептуальную архитектуру SOC с выделением уровней сбора, транспортировки, хранения и аналитики.
- Сформировать модель потоков данных системы мониторинга.
- Определить алгоритмы выявления типовых инцидентов информационной безопасности.
- Предложить структуру взаимодействия SOC с подразделениями университета.

Ограничения и допущения. В рамках настоящей работы предполагается, что проектируемый SOC носит учебно-экспериментальный характер и ориентирован на использование open-source технологий. Реализация системы рассматривается на концептуальном уровне без детальной проработки аппаратной части. Также предполагается наличие стандартной университетской инфраструктуры, включающей доменную среду, сетевые устройства и веб-сервисы

Обоснование применения Big Data в SOC университета. Информационная инфраструктура университета характеризуется значительным объёмом и разнообразием событий информационной безопасности, генерируемых серверами, сетевыми устройствами и рабочими станциями пользователей. При численности обучающихся и сотрудников в несколько тысяч совокупный объём журналов может достигать миллионов записей в сутки.

Потоки событий обладают следующими характеристиками:

- Volume – большой объём данных, формируемый распределённой инфраструктурой
- Velocity – высокая скорость поступления событий, требующая обработки в режиме, близком к реальному времени
- Variety – разнообразие форматов журналов (системные логи, сетевые события, веб-доступ, аутентификация и др.).

В условиях таких параметров традиционные централизованные решения мониторинга могут испытывать ограничения, связанные с масштабируемостью и гибкостью настройки аналитических механизмов.

Применение технологий Big Data и потоковой архитектуры обработки данных позволяет обеспечить:

- устойчивость к пиковым нагрузкам
- горизонтальное масштабирование
- централизованный анализ разнородных источников
- возможность реализации механизмов корреляции и выявления аномалий.

Таким образом, особенности университетской инфраструктуры объективно обосновывают выбор распределённой архитектуры мониторинга, основанной на принципах Big Data [5].

Общая архитектурная модель SOC. Проектируемый центр мониторинга информационной безопасности представляет собой многоуровневую систему, обеспечивающую централизованный сбор, транспортировку, хранение и анализ событий безопасности. Архитектура строится по принципу разделения функциональных уровней, что позволяет обеспечить масштабируемость, отказоустойчивость и гибкость внедрения аналитических механизмов. Общая схема архитектуры представлена на рисунке 1.

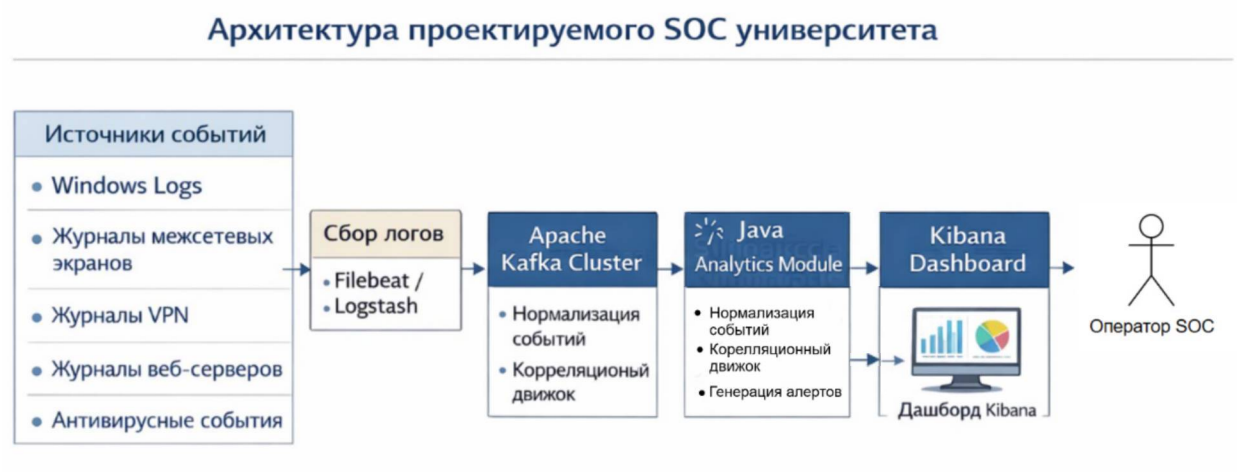

Рисунок 1. Архитектура проектируемого SOC

Уровень сбора данных. На уровне сбора осуществляется получение журналов событий от различных источников инфраструктуры университета. В качестве источников выступают:

- серверы доменной инфраструктуры
- серверы приложений
- сетевые устройства
- VPN-шлюзы
- рабочие станции пользователей.

Для централизованной передачи данных предполагается использование специализированных агентов сбора логов, обеспечивающих:

- нормализацию форматов
- фильтрацию несущественных событий
- передачу данных в потоковую систему.

Данный уровень играет критическую роль, поскольку именно здесь формируется первичный поток событий безопасности.

Уровень транспортировки. Для передачи и буферизации потоков событий предлагается использование распределённого брокера сообщений [2].

Данный компонент обеспечивает:

- устойчивость к пиковым нагрузкам;
- асинхронную обработку данных;
- горизонтальное масштабирование;
- разделение потоков по категориям событий.
- Использование потоковой архитектуры позволяет избежать потери данных при временных перегрузках системы аналитики.

Уровень хранения. На уровне хранения осуществляется индексирование и долговременное сохранение событий безопасности.

Система хранения должна обеспечивать:

- быстрый поиск по временным меткам
- агрегацию данных
- построение аналитических выборок
- хранение исторических данных для ретроспективного анализа.

Использование распределённого хранилища позволяет обеспечить масштабируемость и устойчивость к отказам [4, 6].

Уровень аналитики. Уровень аналитики является ключевым элементом проектируемого SOC. Он выполняет:

- корреляцию событий из различных источников
- выявление типовых атак
- анализ поведенческих отклонений
- формирование алертов.

В рамках концепции предполагается реализация модуля потоковой обработки событий [3], функционирующего в режиме, близком к реальному времени.

Диаграмма вариантов использования. Функциональная модель проектируемого SOC представлена на рисунке 2 в виде диаграммы вариантов использования. Диаграмма отражает основные категории пользователей системы и базовые сценарии их взаимодействия с платформой мониторинга.

В рамках концепции SOC предполагается разграничение уровней анализа инцидентов, однако на представленной диаграмме показаны обобщённые роли пользователей, без детального разделения на уровни L1–L3. Такое представление обусловлено концептуальным характером работы.

Акторы системы. На диаграмме представлены следующие обобщённые категории пользователей: Оператор SOC – Осуществляет мониторинг алертов, первичный анализ событий и обработку уведомлений. Администратор системы - Выполняет конфигурацию и техническое сопровождение платформы. В рамках практической реализации предполагается возможное выделение аналитика второго уровня (L2), выполняющего углублённый анализ инцидентов. Данный функционал логически вытекает из сценария «Расследование инцидента, представленного на диаграмме.

Связь сценариев использования. Диаграмма отражает базовые сценарии, такие как:

- просмотр алертов
- анализ событий
- настройка правил корреляции
- формирование отчётов.

Следует отметить, что отдельные действия пользователей являются логическим продолжением других сценариев. Так, расследование инцидента развивается из процедуры анализа алерта, а формирование отчёта является результатом завершённого расследования. Таким образом, даже если отдельные сценарии не выделены как самостоятельные элементы диаграммы, они функционально включены в общий процесс обработки инцидента.

Значение функциональной модели. Представленная диаграмма позволяет:

- определить границы взаимодействия пользователей с системой;
- формализовать ключевые функции SOC;
- показать связь между архитектурной моделью и эксплуатационными процессами.

Функциональная модель носит укрупнённый характер и может быть детализирована на этапе практической реализации проекта.

Рисунок 2. Диаграмма вариантов использования

Интеграция проектируемого SOC. Интеграция центра мониторинга информационной безопасности в инфраструктуру университета предполагает включение его в существующую сетевую и серверную архитектуру без нарушения текущих бизнес-процессов. Схема интеграции представлена на рисунке 3.

Структура университетской сети. Типовая инфраструктура университета условно разделяется на несколько логических зон:

Периметр сети

В периметре размещаются:

- межсетевой экран
- VPN-шлюз
- публичные веб-сервисы
- почтовый сервер.

Данная зона обеспечивает взаимодействие университета с внешней сетью Интернет и является наиболее уязвимой с точки зрения внешних атак.

DMZ (демилитаризованная зона)

В DMZ размещаются сервисы, доступные извне:

- веб-портал университета
- система дистанционного обучения
- сервер электронной почты.

Эти ресурсы генерируют значительный объём журналов доступа и сетевых событий.

Внутренняя сеть

Внутренняя сеть включает:

- контроллеры домена
- файловые серверы

- серверы баз данных
- рабочие станции сотрудников
- Wi-Fi инфраструктуру

Именно в этой зоне сосредоточено большинство событий аутентификации и доступа к ресурсам.

Механизм интеграции SOC. Проектируемый SOC интегрируется в инфраструктуру путём подключения к источникам журналов событий. Передача логов осуществляется:

- через агенты сбора данных
- через экспорт журналов сетевых устройств
- посредством передачи событий в потоковую систему.

Важно отметить, что SOC не участвует в маршрутизации сетевого трафика и не влияет напрямую на работу сервисов. Он функционирует как система мониторинга и аналитики.

Значение интеграционной модели. Представленная модель интеграции демонстрирует:

- отсутствие необходимости кардинальной перестройки инфраструктуры
- возможность поэтапного внедрения системы
- масштабируемость решения
- адаптацию SOC к существующим условиям эксплуатации.

Таким образом, проектируемый центр мониторинга может быть встроен в инфраструктуру университета с минимальным воздействием на её функционирование.

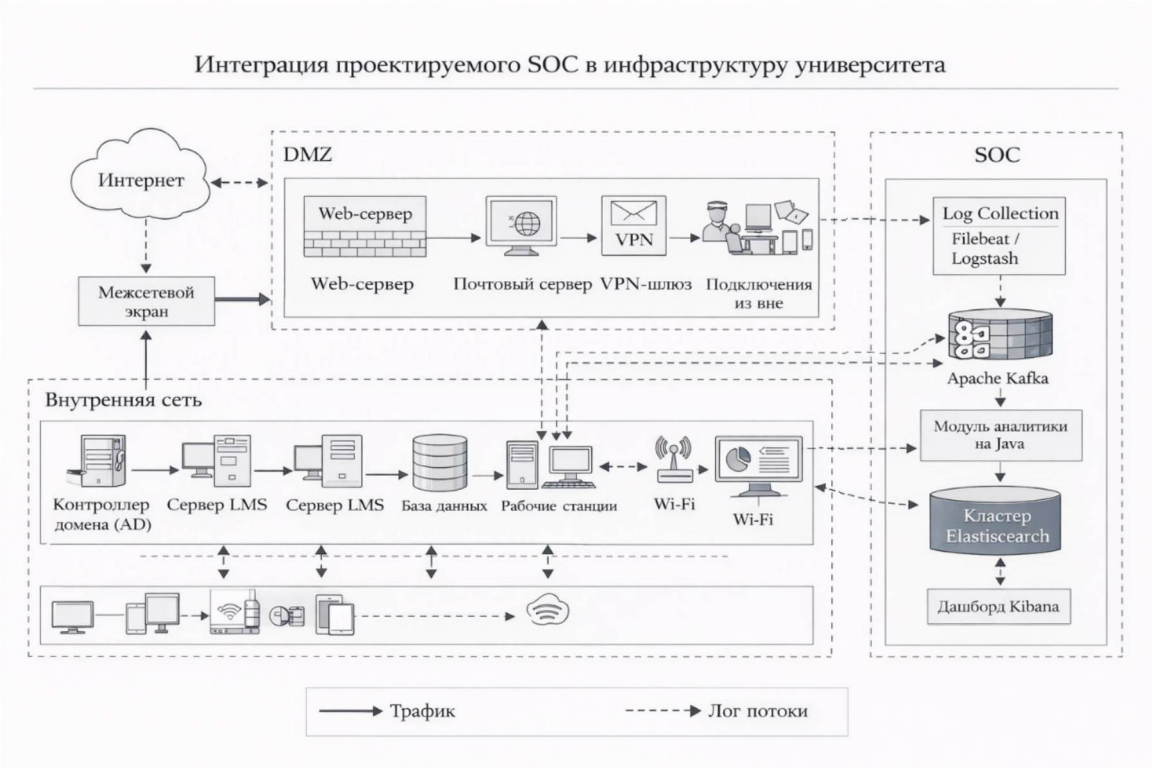

Рисунок 3. Инфраструктура университета с интегрированным SOC

Модель потоков данных. Для формализации взаимодействия проектируемого центра мониторинга информационной безопасности с внешней средой используется контекстная диаграмма потоков данных (DFD Level 0), представленная на рисунке 4. На данном уровне система рассматривается как единый процесс, без детализации внутренней структуры.

В центре модели располагается процесс: «Система мониторинга информационной безопасности (SOC) университета». Он выполняет функции:

- сбора событий безопасности;
- анализа и корреляции данных;
- формирования уведомлений и отчётов.

Входные и выходные потоки данных указаны на рисунке 4.

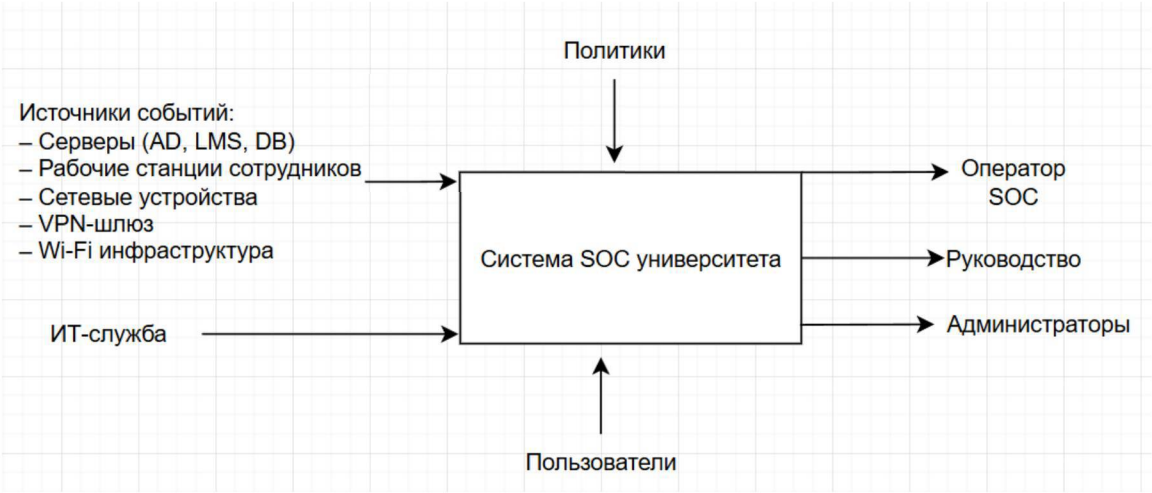

Рисунок 4. Модель потоков данных

Предполагаемый алгоритм функционирования SOC. Проектируемый центр мониторинга информационной безопасности функционирует в соответствии с последовательной логикой обработки событий, обеспечивающей их сбор, анализ и передачу результатов ответственным подразделениям. Алгоритм работы системы включает следующие этапы:

Формирование события. В результате действий пользователя или сетевой активности в инфраструктуре университета формируется событие информационной безопасности. Примерами таких событий являются:

- попытка аутентификации
- обращение к веб-ресурсу
- изменение учётной записи
- сетевое соединение с внешним узлом.

Событие фиксируется соответствующим компонентом инфраструктуры и передаётся в систему мониторинга.

Передача и буферизация данных. Полученные журналы событий направляются в потоковую подсистему, обеспечивающую их буферизацию и асинхронную обработку. Данный этап позволяет:

- избежать потери данных при пиковых нагрузках
- обеспечить устойчивость системы к временным задержкам аналитического модуля
- разделить события по категориям.

Аналитическая обработка. На этапе аналитической обработки осуществляется:

- нормализация форматов событий
- сопоставление событий из различных источников
- проверка условий выявления типовых инцидентов
- анализ отклонений от стандартных сценариев поведения.

В рамках концепции предполагается использование правил корреляции, основанных на пороговых значениях и временных интервалах.

Формирование инцидента. При выполнении условий корреляции событие или совокупность событий классифицируются как потенциальный инцидент информационной безопасности. На данном этапе формируется:

- уведомление (алерт)
- запись в журнале инцидентов
- информация для дальнейшего анализа.

Передача результатов. Сформированные алерты передаются оператору SOC для первичной обработки. В случае подтверждения инцидента информация направляется: системным администраторам - для технического реагирования, руководству - в составе аналитических отчётов.

Назначение алгоритма. Представленный алгоритм отражает логическую последовательность функционирования системы мониторинга и связывает архитектурную модель с практическими процессами обработки инцидентов. Алгоритм носит концептуальный характер и может быть детализирован на этапе практической реализации проекта.

Заключение. В работе рассмотрены особенности информационной инфраструктуры университета и проведён анализ характерных угроз информационной безопасности, обусловленных открытостью образовательной среды и высокой динамичностью пользовательского состава. Установлено, что разрозненные средства защиты не обеспечивают комплексного анализа событий и требуют централизованного подхода к мониторингу. В рамках исследования сформулирована задача проектирования учебно-экспериментального центра мониторинга информационной безопасности (SOC) и разработана его концептуальная архитектура, основанная на принципах Big Data и потоковой обработки данных. Предложена многоуровневая модель системы, включающая сбор, транспортировку, хранение и аналитическую обработку событий безопасности. Построены функциональная модель (диаграмма вариантов использования), схема интеграции SOC в инфраструктуру университета и модель потоков данных, формализующая взаимодействие системы с внешней средой. Представлен предполагаемый алгоритм функционирования центра мониторинга, отражающий последовательность обработки событий и формирования уведомлений. Практическая значимость работы заключается в возможности поэтапного внедрения предложенной архитектуры в университетскую среду с использованием open-source технологий и без кардинальной перестройки существующей инфраструктуры.

Дальнейшее развитие проекта может быть связано с:

- реализацией прототипа аналитического модуля;
- расширением механизмов корреляции событий;
- внедрением методов поведенческой аналитики;
- оценкой эффективности обнаружения инцидентов.

Таким образом, предложенная концепция может служить основой для создания масштабируемого центра мониторинга информационной безопасности в образовательной организации.

Список литературы

[1] Kim D., Solomon M. Fundamentals of Information Systems Security. 3rd ed. Burlington: Jones & Bartlett Learning; 2018.

[2] Kreps J., Narkhede N., Rao J. Kafka: The Definitive Guide. Sebastopol: O'Reilly Media; 2017.

[3] Newman S. Building Microservices. Sebastopol: O'Reilly Media; 2016.

[4] Grolinger K., Higashino W., Tiwari A., Capretz M. Data management in cloud environments: NoSQL and Big Data. Journal of Cloud Computing. 2013;2(22).

[5] Романюк М.В., Лещенко Е.А., Марковский С.С. Применение Big Data для защиты компьютерных сетей // BIG DATA and Advanced Analytics. 2024. №2. С. 486–489.

[6] Elastic NV. Elastic Stack Documentation.

Авторский вклад

Стамкулова Гулдана Кубанычбековна – научное руководство исследованием, принимала участие в постановке задачи, формировании концепции работы, корректировке архитектурной модели и редактировании текста статьи.

Салиев Самат Мелисович – анализ особенностей инфраструктуры университета, разработал архитектурную модель проектируемого SOC, построил функциональные и структурные схемы, сформировал модель потоков данных и подготовил основной текст статьи.

DESIGNING A SECURITY OPERATIONS CENTER (SOC) IN THE UNIVERSITY INFRASTRUCTURE USING BIG DATA AND ADVANCED ANALYTICS

G.K. Stamkulova
Associate Professor, Department of Computer Systems Software, KSTU

S.M. Saliev
Fourth-year student in the Department of Computer Systems Software at KSTU

Abstract. The article discusses the design of the architecture of a training and experimental center for monitoring information security (SOC) for university infrastructure using Big Data technologies. A conceptual model for collecting, storing, and streaming security events based on Apache Kafka and Elastic Stack is proposed. The choice of Java as the development platform for the event correlation analysis module is justified. Algorithms for detecting typical incidents are presented and a model for their formalization is described. The work is project-based and aims to create a foundation for further practical implementation.

Keywords: SOC, Big Data, information security, streaming analytics, event correlation, Java, Apache Kafka, Elasticsearch