УДК 004.056 doi: 10.15827/0236-235X.142.615-631 2023. Т. 36. № 4. С. 615–631

Обеспечение информационной безопасности научного суперкомпьютерного центра

А.В. Баранов
П.М. Корепанов
Е.Е. Кузнецов

Ссылка для цитирования
Баранов А.В., Корепанов П.М., Кузнецов Е.Е. Обеспечение информационной безопасности научного суперкомпьютерного центра // Программные продукты и системы. 2023. Т. 36. № 4. С. 615–631. doi: 10.15827/0236-235X.142.615-631
Информация о статье
Поступила в редакцию: 25.08.2023 После доработки: 28.09.2023 Принята к публикации: 28.09.2023

Аннотация. В большинстве научных суперкомпьютерных центров (СКЦ) коллективного пользования обрабатывается открытая информация. Для ее защиты, как правило, применяются штатные технологии информационной безопасности, встроенные в используемые операционные системы, системы хранения данных, сетевые устройства. Наблюдается рост как числа угроз безопасности информации, так и проводимых в отношении СКЦ компьютерных атак и состоявшихся инцидентов, что несет для центров репутационные и финансовые риски. В статье рассмотрены особенности обработки информации в СКЦ, существенно ограничивающие применение известных мер и средств защиты информации. К таким особенностям отнесены свобода пользователя СКЦ в выборе инструментальных средств и прикладных программных пакетов для решения своих исследовательских задач, необходимость обеспечения максимальной скорости расчетов на предоставленных пользователям суперкомпьютерных ресурсах, ограниченность применения защищенных операционных систем и средств обновления системного программного обеспечения. Обоснована актуальность разработки комплексного системного подхода к защите информации, при котором достаточный уровень информационной безопасности СКЦ обеспечивается без существенных ограничений спектра и снижения качества предоставляемых пользователям услуг по высокопроизводительным вычислениям. Рассмотрены актуальные угрозы безопасности информации СКЦ, приведена классификация обрабатываемых данных, определен перечень актуальных мер защиты информации. С учетом исследованных особенностей защиты информации в СКЦ представлен вариант построения системы информационной безопасности центра, основанный на разделении информационно-вычислительной инфраструктуры центра на зоны безопасности и применении средств контроля сетевого периметра и анализа событий безопасности.

Ключевые слова: информационная безопасность, защита информации, суперкомпьютерный центр, средства контроля периметра сети, NTA, SIEM
Благодарности. Работа выполнена в МСЦ РАН в рамках государственного задания по теме FNEF-2022-0016

Введение. Главной задачей *суперкомпьютерных центров* (СКЦ), функционирующих в сфере науки и образования, является предоставление услуг по высокопроизводительным вычислениям многочисленным пользователям – исследователям и студентам. В основном обрабатываемая в СКЦ коллективного пользования информация является открытой, и для ее защиты, как правило, применяются штатные технологии информационной безопасности, встроенные в используемые операционные системы, *системы хранения данных* (СХД), сетевые устройства. Последствиями компьютерных инцидентов в СКЦ могут быть искажение и потеря информации, что может привести к задержке в проведении научных исследований, но, как правило, не влечет за собой прямые убытки. Другими словами, основные риски являются не финансовыми, а репутационными, связанными с утратой доверия со стороны пользователей. В последние годы наблюдается рост как числа угроз безопасности информации, так и проводимых в отношении СКЦ компьютерных атак и состоявшихся инцидентов, и это позволяет говорить об актуальности комплексного системного подхода к обеспечению информационной безопасности СКЦ. Следует отметить, что тема информационной безопасности СКЦ все чаще поднимается в мировом научном сообществе, по данному направлению проводятся различные конференции, такие как IEEE/ACM S-HPC [1].

Как известно, любые защитные меры в отношении информационной инфраструктуры ограничивают возможности ее применения. Например, повсеместно используемый полностью автоматизированный публичный тест Тьюринга для различения компьютеров и людей (CAPTCHA) [2] увеличивает время входа в информационную систему, и нередко пользователи испытывают трудности при его прохождении. Предоставляемые научным СКЦ услуги по высокопроизводительным вычислениям предполагают свободу пользователя в выборе инструментальных средств и прикладных программных пакетов для решения своих исследо-

вательских задач, однако применение некоторых мер по защите информации может ограничить эту свободу. Кроме того, для пользователя СКЦ важным является достижение максимальной скорости расчетов, то есть работа суперкомпьютерных ресурсов с производительностью, по возможности приближенной к пиковой. Современные программные средства защиты могут потреблять значительную долю процессорного времени, их использование на вычислительных узлах суперкомпьютера может отрицательно сказаться на производительности вычислений. В этих условиях актуальной научно-технической задачей является разработка такого комплекса мер защиты информации в СКЦ, который, обеспечивая достаточный уровень информационной безопасности, не ограничивает (или несущественно ограничивает) спектр предоставляемых пользователям услуг по высокопроизводительным вычислениям и не снижает их качество. Настоящая статья посвящена поиску возможных решений этой задачи.

Обзор актуальных исследований

Основными направлениями исследований в области высокопроизводительных вычислений являются архитектура суперкомпьютерных систем, процессоров и коммуникационных сред, эффективные параллельные алгоритмы, системное и инструментальное ПО суперкомпьютеров, а также методы и средства эффективного использования их вычислительных ресурсов. Информационная безопасность СКЦ пока находится вне их фокуса, и публикации по данной тематике немногочисленны. Тем не менее, в области защиты информации в СКЦ можно выделить несколько направлений исследований.

Часть публикаций посвящена вопросам стандартизации подходов к обеспечению информационной безопасности высокопроизводительных вычислительных систем. Так, в работе [3] рассмотрены наиболее распространенные угрозы безопасности СКЦ, приведены классификации возможных угроз по источникам их возникновения и уязвимостей компонентов СКЦ, обусловливающих эти угрозы. Предложены методы противодействия угрозам и модель защиты информации в виде трехдольного графа. Система безопасности, представляющая собой множество средств защиты информации, выступает в качестве некоего барьера между защищаемыми объектами (компонентами СКЦ) и угрозами. Считается, что система безопасна, если достигается полное перекрытие, предусматривающее не менее одного барьера между угрозами и объектами защищаемой системы. В работе [4] представлена эталонная модель системы высокопроизводительных вычислений, разделенной на четыре функциональные зоны: вычислений, данных, управления и доступа. Для каждой зоны определены их возможные конфигурации и выделены основные виды угроз. Предложенные эталонная модель, способы и рекомендации по обеспечению безопасности информации в СКЦ были использованы в ходе настоящего исследования.

В исследовании [5] автор выделяет принципиальные отличия защиты информации в суперкомпьютерах от традиционных систем обработки данных и определяет две ключевые проблемы, решение которых будет освещаться и в данной работе. Первая проблема заключается в необходимости предоставлять непосредственный доступ к вычислительным ресурсам большому количеству пользователей, которым на выделенных ресурсах разрешено выполнять произвольный программный код, вторая – в том, что многие распространенные решения по защите информации расходуют значительную вычислительную мощность и оказываются по этой причине неэффективными в области высокопроизводительных вычислений. Основное внимание автор уделяет угрозам, связанным с целостностью и доступностью информации, в числе которых изменение кода или данных, неправомерное использование вычислительных ресурсов, отказ в обслуживании суперкомпьютерной системы или сети.

Другая область исследований связана с алгоритмами шифрования данных. Например, в работе [6] на базе суперкомпьютера Sunway TaihuLight реализована модель защиты данных с использованием параллельных алгоритмов AES и SHA3, позволяющая достичь высокой эффективности шифрования/дешифрования данных и гарантировать их целостность. Модель защиты данных состоит из двух частей: шифрования и дешифрования. В работе предложены стратегии оптимизации параллельного алгоритма AES, учитывающие особенности вычислительной архитектуры и иерархии памяти. Авторам в результате оптимизации алгоритма удалось увеличить пропускную способность модели защиты данных почти в два раза.

Организации доступа пользователей к суперкомпьютерным ресурсам посвящена публикация [7]. Авторы рассматривают используе-

мые инструменты и процесс включения портала MIT SuperCloud Portal для федеративной аутентификации с федерацией InCommon и инфраструктурой открытых ключей правительства США. В результате внедрения системы пользователи получили возможность применения надежных существующих систем многофакторной аутентификации, развернутых в их основных учреждениях. Описан также процесс самостоятельной регистрации и проверки ключей ssh.

Важной областью исследований являются методы и средства анализа системных журналов и поиска аномалий в работе суперкомпьютера. С этой точки зрения информационная безопасность суперкомпьютерных систем рассматривается в [8]. Анализ системных журналов увязывается с известными атаками и распространенными методами и средствами защиты, в частности, методами обнаружения вторжений. Авторы делают вывод о недостаточном использовании системных журналов. В работе [9] представлены подробный обзор и оценка шести методов обнаружения аномалий, проведено сравнение их точности и эффективности на двух наборах реальных данных, а также предложены программные инструменты с открытым исходным кодом для реализации рассмотренных методов. В [10] авторы предлагают свою систему обнаружения аномалий/вторжений на основе анализа журналов. Обнаружение аномалий в суперкомпьютерных системах рассматривается как последовательный процесс принятия решений с применением методов обучения с подкреплением. Публикация [11] посвящена программной платформе обнаружения и диагностики аномалий в системных журналах в режиме онлайн с использованием длинной краткосрочной памяти (LSTM) для моделирования системного журнала как последовательности естественного языка.

Помимо анализа системных журналов, поиск связанных с безопасностью аномалий в работе суперкомпьютерной системы может осуществляться при помощи мониторинга. Так, в [12] рассмотрен подход к автоматизированному обнаружению аномалий в HPC-системах с использованием машинного обучения. Авторы применили особый тип нейронных сетей, называемый автокодировщиком (autoencoder). Сети обучены нормальному поведению каждого вычислительного узла на основе его исторических телеметрических данных о «хорошем» поведении. Автокодировщики обучаются и тестируются на серийном суперкомпьютере и развертываются как расширение встроенных устройств мониторинга. В работе [13] продемонстрировано, как данные системы мониторинга Nagios могут быть использованы для решения задачи обнаружения и предсказания аномалий в высокопроизводительных вычислительных системах, причем без применения супервизорных подходов.

Некоторые из публикаций посвящены защите инженерной инфраструктуры СКЦ. В работе [14] показано, как злоумышленник может атаковать СКЦ, нарушив работу систем управления состоянием окружающей среды в помещениях, где размещены вычислительные узлы. На реальных данных продемонстрировано использование злоумышленником систем управления, обеспечивающих подачу воды в подсистему охлаждения, в качестве точек входа для косвенного воздействия на вычислительные ресурсы СКЦ. В исследовании [15] затронута тема влияния повышения потребляемой мощности на сбои в работе вычислительных систем, такие как отключение питания при перегрузке электросети, что, в свою очередь, приводит к потере результатов вычислений при выключении узлов и увеличении времени ожидания в очереди из-за их остановки. Также авторы предлагают согласованный подход к сбору, обработке и хранению статистики энергопотребления для пользовательских приложений, которые могут выполняться на разных суперкомпьютерах одного центра.

Важнейшей областью обеспечения информационной безопасности СКЦ является защита его сетевой инфраструктуры. В [16] предложен вариант межсетевого экрана, который позволяет установить набор правил, регулирующих соединения во внутренней сети суперкомпьютера с помощью Linux netfilter. В работе [17] рассмотрены два метода «грубой силы» обнаружения атак, применяемых в Корейском институте науки и технологий. Один метод заключается в анализе системных журналов на предмет выявления событий неудачной аутентификации. Считается, что атака обнаружена, если количество неудачных попыток входа превышает заданное пороговое значение. Другой метод состоит в анализе событий межсетевого экрана, заключающемся в отсеивании определенных подключений правилами экрана. Такие события группируются по адресам источника и назначения отсеянного подключения, и, если количество событий в какой-либо группе превышает пороговое значение, считается, что атака обнаружена.

Анализ публикаций последних лет дает основания утверждать, что обеспечение информационной безопасности СКЦ коллективного пользования является актуальной научно-технической задачей, решение которой находится на начальной стадии. Необходимы исследования и разработки в области безопасных архитектур суперкомпьютерных систем и центров, обеспечивающих в то же время эффективное и удобное использование вычислительных ресурсов для проведения высокопроизводительных расчетов.

Угрозы безопасности информации в научном СКЦ

Угрозы несанкционированного доступа к информации возникают практически везде, где применяются средства вычислительной техники. Научные СКЦ не являются исключением – инциденты безопасности возникают в таких средах, по крайней мере, с 1980-х годов [18, 19] и регулярно происходят в настоящее время (см., например, https://csirt.egi.eu/attacks-on-multiple-hpc-sites/). Так же, как и любой другой компьютер, суперкомпьютеры подключаются к сети и работают под управлением стандартных операционных систем, как правило, на базе Linux, и, соответственно, подвергаются многим традиционным атакам, таким как подбор паролей и сканирование портов. В работе [4] дана следующая классификация угроз безопасности информации в СКЦ.

- Угрозы, которые могут привести к утечке данных или нарушению их целостности. Опасность этих угроз связана с большим числом пользователей СКЦ и, соответственно, с относительно высокой вероятностью нарушения (злонамеренного или нет) кем-то из пользователей правил информационной безопасности.
- Злоупотребление вычислительными ресурсами: использование суперкомпьютерных мощностей вычислительных комплексов для выполнения действий, нарушающих законодательство, например, для майнинга криптовалюты.
- Вредоносное ПО. Суперкомпьютеры могут стать целью атак, направленных на установку вредоносного ПО, которое обусловливает компрометацию учетных данных пользователей, нарушение работоспособности системы или утечку конфиденциальной информации.
- Сетевые угрозы [17, 20]. Суперкомпьютеры, как правило, подключены к сети Интернет, что делает их уязвимыми для различных сетевых атак, таких как отказ в обслуживании (DoS), перехват сетевого трафика (сниффинг), атаки на протоколы прикладного уровня, такие как Secure Shell, эксплуатация уязвимостей сетевого оборудования, сканирование периметра сети и другие.
- Угрозы инженерной инфраструктуре СКЦ: несанкционированный физический доступ к оборудованию, повреждение компонентов [9], пожары, наводнения и другие чрезвычайные ситуации.
- Уязвимости приложений и операционных систем, например, в параллельных файловых системах [21]. Наличие таких уязвимостей может позволить злоумышленникам получить несанкционированный доступ к суперкомпьютерным ресурсам и данным пользователей, а также осуществить иные типы атак. Возможность таких атак обусловлена вероятными уязвимостями в коде ПО, ошибками конфигурации или недостаточными мерами безопасности.
- Социальная инженерия. Угрозы, основанные на социальной инженерии, направлены на манипулирование сотрудниками или пользователями СКЦ с целью получения несанкционированного доступа к системе или конфиденциальной информации. Такие угрозы могут включать фишинг [22], а также вовлечение сотрудников в устный или письменный диалог с целью получения от них информации, необходимой для осуществления той или иной атаки.
- Внутренние угрозы, которые могут представлять серьезную угрозу для безопасности суперкомпьютера. В роли злонамеренных или невольных нарушителей могут выступать авторизованные пользователи суперкомпьютерных систем, системные администраторы и администраторы информационной безопасности, лица, привлекаемые для установки, настройки, испытаний, пусконаладочных и иных видов работ, лица, обеспечивающие функционирование СКЦ: охрана, уборщики и др. Перечисленные потенциальные внутренние нарушители могут осуществить в отношении суперкомпьютерных систем несанкционированный доступ или иные злонамеренные либо ошибочные действия [23].

Типовая схема обработки данных в научном СКЦ

Прежде чем переходить к вопросам обеспечения информационной безопасности в научном СКЦ, необходимо рассмотреть типовую схему организации такого центра (рис. 1), а также категории обрабатываемой в нем инфор-

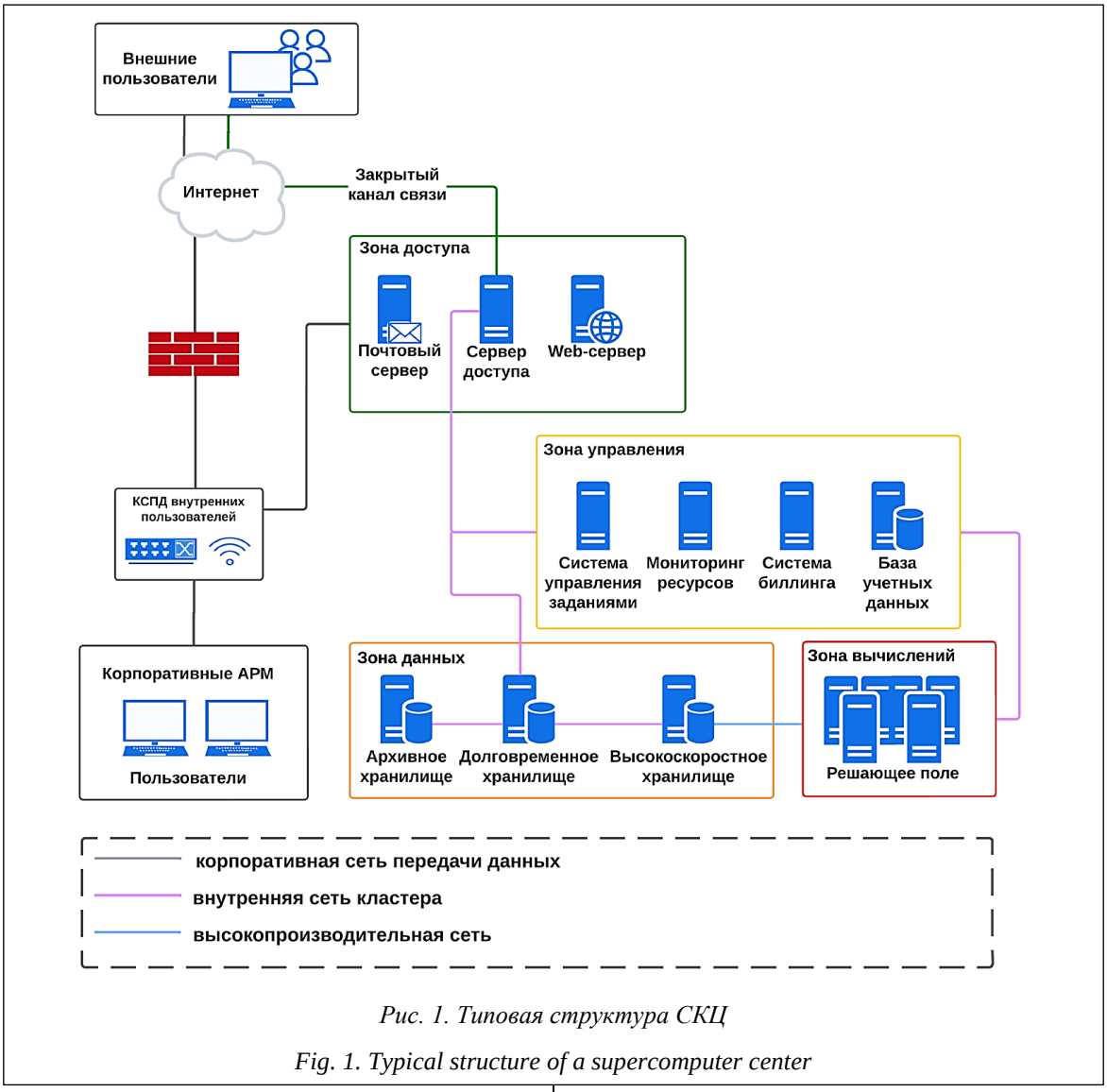

Рис. 1. Типовая структура СКЦ

Fig. 1. Typical structure of a supercomputer center

мации. Воспользовавшись результатами работы [4], разделим компоненты СКЦ на зоны безопасности.

Ядром любого СКЦ является зона высокопроизводительных вычислений, в которой размещаются одна или нескольких высокопроизводительных вычислительных установок, собственно, и называемых суперкомпьютерами. Каждая установка состоит из пула вычислительных (суперкомпьютерных) узлов, представляющих собой серверы с одним или несколькими многоядерными процессорами. Достаточно часто узлы комплектуются ускорителями вычислений, как правило, на базе графических процессоров. Для возможности информационного взаимодействия узлы суперкомпьютера объединяются так называемым интерконнектом – высокоскоростной низколатентной сетью, такой как Infiniband или Intel OmniPath. Дополнительно к интерконнекту суперкомпьютерные узлы оснащаются транспортной сетью для подключения к СХД и сетью управления. Обратим внимание на важный с точки зрения безопасности факт: каждый узел управляется собственной операционной системой и имеет уникальные сетевые адрес и имя.

Главной задачей суперкомпьютерной системы является обеспечение выполнения параллельных программ. Под параллельной программой подразумевается программа, содержащая отдельные части, обычно называемые ветвями, которые могут выполняться одновременно, причем каждая ветвь на отдельном вычислительном устройстве (процессоре, ядре, вычислительном узле). На практике ветви такой программы представляют собой процессы операционной системы, выполняющиеся на одном или нескольких узлах суперкомпьютера.

Взаимодействие процессов параллельной программы осуществляется через интерконнект при помощи специализированных коммуникационных библиотек, таких как MPI, SHMEM и др. Драйверы интерконнекта, коммуникационные библиотеки, подсистемы управления заданиями и ресурсами, мониторинга и управления учетными записями пользователей формируют специализированный программный стек, полный состав которого на примере Межведомственного суперкомпьютерного центра РАН рассмотрен в [24, 25]. Являясь важнейшим стержневым компонентом, программный стек формируется индивидуально для каждого суперкомпьютера, причем процесс его формирования, конфигурирования и настройки, как показывает практика, требует значительных времени и трудозатрат высококвалифицированного персонала СКЦ. На программный стек суперкомпьютера опираются все запускаемые в зоне вычислений параллельные программы, которые могут быть как разработаны самими пользователями, так и являться компонентами специализированных прикладных программных пакетов для высокопроизводительных вычислений.

Прямой доступ к зоне вычислений и, соответственно, к суперкомпьютерным узлам для пользователей закрыт. Терминальные устройства пользователей (рабочие станции, ноутбуки, планшеты, телефоны и т.п.) расположены вне границ СКЦ, и пользователи с этих устройств соединяются со специально выделенным сервером доступа, размещенным в одноименной зоне. Зона доступа включает узлы, подключенные к внешним сетям, таким как корпоративная сеть СКЦ или Интернет. Эта зона предоставляет средства для авторизации пользователей, разграничения их доступа к информационным и вычислительным ресурсам, обеспечения подключений пользователей и системных администраторов к суперкомпьютерам. Сервер доступа служит для подготовки пользователями параллельных программ, исходных данных, анализа результатов расчетов. Чтобы подготовленная параллельная программа могла быть выполнена на решающем поле, пользователь должен оформить так называемое задание – информационный объект, включающий саму программу, требования к ресурсам (какие узлы суперкомпьютера, в каком количестве и на какое время должны быть предоставлены для выполнения программы) и исходные данные. Подготовленное задание направляется в специальную программную систему управления заданиями [26], отвечающую за ведение очереди заданий и распределение узлов суперкомпьютера между различными параллельными программами разных пользователей.

Система управления заданиями функционирует на выделенном сервере управления, размещенном в одноименной зоне. Зона управления включает в себя множество серверов, через которые системные администраторы могут конфигурировать, настраивать, тестировать и контролировать компоненты зоны вычислений, в том числе формировать или модифицировать программный стек суперкомпьютера. Здесь же размещаются серверы и подсистемы мониторинга суперкомпьютерных ресурсов и учета их потребления пользователями. Непосредственный доступ в зону управления разрешен только администраторам, причем администратор входит сначала в зону доступа и только затем в зону управления. Для пользователей суперкомпьютеров зона управления доступна только через интерфейс системы управления заданиями. Зона управления является наиболее критичной с точки зрения информационной безопасности, поскольку содержит множество точек отказа.

Пользовательские данные размещаются в одной или нескольких СХД, которые в соответствии с [4] образуют зону данных. Особенностью СХД СКЦ является то, что каждый раздел СХД монтируется на всех серверах, включая вычислительные узлы, под одним и тем же именем. Это делает данные одинаково доступными пользователю или администратору из зон доступа, управления и вычислений. Для обеспечения безопасности информации важно понимать иерархию СХД современных СКЦ, которую составляют следующие виды хранилищ данных:

– высокоскоростное хранилище так называемых горячих данных;

– долговременное надежное хранилище оперативных данных;

– долговременное архивное (медленное) хранилище так называемых холодных данных.

Под горячими данными понимаются данные, непосредственно используемые параллельными программами при их выполнении в зоне вычислений, в том числе промежуточные результаты расчетов. Главное назначение хранилища таких данных – обеспечение как можно более быстрого доступа к большим объемам информации, сохраненным в результате выполнения параллельной программы. Пользователи активно используют эти хранилища для

скоростной аккумуляции результатов нескольких подряд выполнившихся заданий, а также для передачи результатов одного задания в качестве входных данных следующего за ним задания. С точки зрения обеспечения безопасности важно, что такие хранилища, как правило, являются гиперконвергентными, то есть используют в качестве запоминающих устройств скоростные локальные диски вычислительных узлов суперкомпьютера. В единое хранилище эти диски объединяются при помощи параллельных файловых систем, таких как GPFS или Lustre. Время хранения горячих данных ограничивается системным администратором, а их длительная сохранность не гарантируется. Пользователям рекомендуется регулярно сохранять критичные для них данные в долговременное оперативное хранилище, представляющее собой, как правило, классическую сетевую СХД. Именно в оперативном хранилище выполняются действия по подготовке параллельных программ, исходных данных для них, здесь же сохраняются результаты расчетов, используемые пользователем для анализа. Архивные хранилища не подразумевают оперативного доступа и служат для сохранения резервных копий оперативных данных, а также для длительного хранения редко используемых так называемых холодных данных.

Кроме перечисленных зон доступа, управления, вычислений и данных, в СКЦ может быть выделена корпоративная сеть внутренних пользователей, включающая рабочие места сотрудников СКЦ и серверы обеспечения их деятельности, такие как сервер веб-сайта СКЦ, почтовый сервер, серверы обеспечивающих подразделений (бухгалтерия, охрана, тендерная группа и пр.).

Рассмотрим подлежащие защите данные, обрабатываемые в СКЦ. С точки зрения возрастания требований безопасности эти данные можно разделить на следующие категории:

– программные коды, исходные данные и результаты расчетов;

– стек системного и инструментального ПО суперкомпьютера (программный стек);

– учетные данные пользователей;

– учетные данные системных администраторов суперкомпьютера.

Наименее критичной к безопасности, хотя и наиболее объемной категорией, являются пользовательские программные коды, исходные данные и результаты высокопроизводительных расчетов, размещаемые в зоне данных. Обычно пользовательские данные не содержат информацию ограниченного доступа (например, персональные данные, служебные документы и т.п.), а интерес со стороны потенциальных нарушителей к необработанным результатам расчетов и пользовательскому программному коду сравнительно низкий из-за высокой сложности, неструктурированности данных и отсутствия документации. Обычно для защиты этой категории данных, если нет дополнительных соглашений с отдельными пользователями, применяются штатные средства разграничения доступа, имеющиеся в составе операционных систем суперкомпьютерных узлов, серверов СХД, зон доступа и управления. Хранилище оперативных данных обеспечивает надежное долговременное хранение данных этой категории, но СКЦ, как правило, при этом не дает никаких юридических гарантий обеспечения их сохранности. От аккуратности, дисциплины и бдительности пользователя во многом зависит защита его данных от несанкционированного доступа: пользователь должен обеспечивать конфиденциальность своей аутентификационной информации (значения паролей, закрытых ключей) и регулярное резервное копирование своих данных в независимое от СКЦ пространство хранения.

Упоминавшийся стек системного и инструментального ПО (программный стек) суперкомпьютера чувствителен к угрозам модификации и уничтожения данных. Злонамеренное или ошибочное внесение некорректных изменений в программный стек может привести к неработоспособности суперкомпьютера и простою его ресурсов, поскольку восстановление разрушенного или уничтоженного стека, как показывает практика, занимает длительное время и требует значительных трудозатрат со стороны персонала СКЦ. Рабочая копия программного стека размещается в зоне данных, резервные копии могут создаваться на специально выделенном сервере зоны управления.

Следующей по критичности категорией являются учетные данные пользователей СКЦ. Компрометация (утечка паролей, закрытых ключей и т.п.) пользовательской учетной записи может повлечь несанкционированный доступ к суперкомпьютеру от имени и с правами пользователя. Под угрозой оказываются данные скомпрометированного пользователя, а также высокопроизводительные вычислительные ресурсы, которые от имени пользователя могут быть применены для несанкционированных расчетов, а также для противозаконных действий.

Наиболее защищаемая категория данных – это учетные данные системных администраторов, осуществляющих непосредственное управление суперкомпьютером. Компрометация любой учетной записи этой категории может нанести максимальный ущерб, вплоть до приведения суперкомпьютера в нерабочее состояние (например, путем разрушения программного стека), а также хищения или повреждения всего объема пользовательских данных. Следует отметить, что учетные данные пользователей и системных администраторов могут быть сравнительно просто отделены от остальных категорий и размещены в защищенном сегменте зоны управления.

Особенности обеспечения информационной безопасности СКЦ

Высокопроизводительные вычислительные системы, размещаемые в научных СКЦ, представляют собой отдельный класс научного оборудования, применяемого для исследований преимущественно в режиме коллективного пользования. Предъявляемые требования информационной безопасности не должны препятствовать решению основной задачи СКЦ – организации совместного использования суперкомпьютерных ресурсов множеством пользователей в целях проведения актуальных научных исследований. Кроме этого, развитие методов и средств обеспечения информационной безопасности высокопроизводительных научных вычислений следует рассматривать как отдельную специализированную область науки. Перед тем как выявить проблемные места и отличительные особенности в обеспечении информационной безопасности СКЦ, рассмотрим традиционно применяемые в СКЦ методы и средства защиты информации.

- Аутентификация и авторизация пользователей и сотрудников СКЦ, а также системных администраторов суперкомпьютерных систем. Это базовый метод обеспечения информационной безопасности, повсеместно применяемый в СКЦ. Широко распространены парольная защита, а также применение ключей и сертификатов для подключения к суперкомпьютерным системам. Использование многофакторной аутентификации, а также таких методов, как применение биометрических данных или аппаратных ключей, хотя и повышает уровень безопасности, пока не имеет широкого распространения по причинам, которые будут рассмотрены далее.
- Контроль доступа к информационным и вычислительным ресурсам СКЦ, а также к данным пользователей. Реализация механизмов контроля доступа позволяет определять для каждого пользователя или групп пользователей набор разрешенных операций доступа (чтение, изменение, создание, удаление, выполнение и пр.) к программам и данным. Обеспечивающие контроль методы могут включать списки контроля доступа, ролевую модель доступа или политики безопасности на уровне операционной системы или приложений.
- Шифрование данных, являющееся эффективным способом защиты конфиденциальности информации. В подавляющем большинстве СКЦ, как минимум, шифруется трафик от терминала пользователя до сервера доступа суперкомпьютера. Кроме этого, могут шифроваться данные пользователей в СХД центра, а также содержимое баз данных и передаваемая по сети информация, в том числе передаваемая внутри коммуникационной подсистемы, объединяющей суперкомпьютерные узлы. Отмечается, что применение алгоритмов шифрования и правильное управление ключами являются главными аспектами эффективной защиты данных [6].
- Мониторинг и обнаружение событий и инцидентов безопасности. Они позволяют своевременно выявлять аномальное поведение, вторжения или другие нарушения безопасности. Механизмы мониторинга настраиваются для сбора и анализа данных из различных источников, включая по возможности журналы событий, данные о сетевых соединениях и иные регистрируемые источники события. Обнаружение атак позволяет оперативно реагировать на угрозы безопасности и предпринимать соответствующие действия для предотвращения инцидентов или минимизации их последствий [27].
- Физическая безопасность СКЦ. Обеспечение физической безопасности играет важную роль в защите от несанкционированного доступа к суперкомпьютерам путем проникновения посторонних лиц в машинный зал, а также от физических повреждений оборудования такими лицами. Среди защитных мер применяются контролируемый доступ в помещения, видеонаблюдение, биометрическая идентификация, а также защита от пожара, протечек и других стихийных бедствий.

Следует отметить, что большинство перечисленных методов и средств защиты информации в СКЦ имеют ограниченное применение. Рассмотрим объективные причины, пре-

пятствующие полноценной реализации мер обеспечения информационной безопасности в научно-исследовательских СКЦ.

- Противоречие между требованиями безопасности информации и удобством работы пользователей СКЦ. Проблема заключается в том, что меры по обеспечению безопасности почти всегда приводят к ограничению возможностей пользователей. Перефразируя известную максиму, можно сказать, что самый защищенный суперкомпьютер – это выключенный суперкомпьютер. Защита информации в СКЦ – это всегда баланс между безопасностью и удобством пользователей, которое определяется прежде всего широтой спектра предоставляемых ему услуг и возможностей по производству научных расчетов. Допустим, пользователям предоставляется возможность передачи прав доступа к своим программам и данным, и в какой-то момент некоторый пользователь временно передает определенные права другому пользователю. Если по истечении заданного времени права не будут отозваны (например, по причине забывчивости пользователя), в суперкомпьютерной системе образуется уязвимость, о которой не знают системные администраторы. Допустим, в качестве меры защиты в этом случае будет принят механизм запросов к системному администратору на передачу прав доступа третьим лицам. У системного администратора будет зафиксировано, кто, кому, когда и какие права передал, но, очевидно, в подобной системе совместная работа пользователей будет осуществляться со значительными трудностями и временными задержками. Аналогичные соображения можно привести по поводу двухфакторной аутентификации пользователей, которая заметно затрудняет вход последних в суперкомпьютерную систему, особенно в случае ошибок пользователя при наборе паролей и кодов. Таким образом, фактор удобства работы пользователя может осложнить защиту высокопроизводительной системы коллективного пользования, система может стать более уязвимой.

- Невозможность выполнения доверенного программного кода. Одной из распространенных мер защиты информации является обеспечение выполнения доверенного программного кода на доверенном вычислительном устройстве. Применение такой меры в научном СКЦ невозможно, поскольку проводимые пользователями исследования прямо предполагают применение новых программных пакетов, которые не могут быть доверенными по определению. Пользователями СКЦ широко применяются ПО с открытым исходным кодом и самостоятельно разработанные программы. Как известно, ПО с открытым исходным кодом уязвимо к угрозам цепочки поставок [27]. Самостоятельно разработанное ПО может содержать ошибки, которые в некоторых случаях могут привести к появлению уязвимостей и, соответственно, к актуализации угроз безопасности информации (конфиденциальности, целостности или доступности данных). В научных вычислительных экспериментах широко применяется специализированное ПО, которое может либо не иметь должных доверия и поддержки, либо содержать в себе новые классы уязвимостей. Другими словами, пользователю СКЦ предоставлена базовая возможность выполнения на вычислительных узлах суперкомпьютера произвольного программного кода, качественную проверку которого невозможно осуществить. Особо следует отметить невозможность изоляции недоверенного кода в рамках некоторой виртуальной машины, поскольку накладные расходы на виртуализацию существенны и значительно снижают производительность расчетов [28]. Одним из актуальных направлений в этой области является исследование безопасности новых технологий, предлагаемых для включения в программный стек суперкомпьютерной системы.

- Ограниченность вычислительных ресурсов для средств защиты информации. Как известно, применение современных средств защиты информации, таких как средства шифрования или глубокого анализа сетевого трафика, способно заметно замедлить работу вычислительной системы. В то же время главной функцией суперкомпьютерных систем является обеспечение максимальной производительности расчетов. Можно сказать, что безопасность информации в СКЦ ценна только в той степени, в какой она не замедляет работу суперкомпьютеров и не препятствует высокопроизводительным вычислениям, и в этой связи необходимо минимизировать негативное влияние на производительность мер защиты информации. Кроме этого, некоторые особенности построения суперкомпьютерных систем делают некоторые средства защиты неприменимыми в принципе. Например, информационные обмены между процессами параллельной программы, выполняющимися на разных вычислительных узлах, осуществляются через интерконнект с применением механизмов прямого доступа одного узла к оперативной памяти другого. Сетевой трафик интерконнекта даже теоретически

невозможно обработать при помощи средств анализа трафика, рассчитанных на работу с сетевыми протоколами стека TCP/IP.

- Ограниченность применимости защищенных операционных систем, таких как Astra Linux SE, которые на текущий момент не поддерживают программный стек суперкомпьютера. Как правило, для этих систем отсутствуют драйверы интерконнекта и эффективные реализации коммуникационных библиотек. Отсутствие поддержки программного стека значительно ограничивает применение защищенных операционных систем в качестве операционных систем вычислительных узлов суперкомпьютера. При этом подобные системы могут использоваться в качестве операционных систем серверов и рабочих станций в зонах доступа, управления или данных, за исключением сервера доступа. Поскольку на сервере доступа пользователи осуществляют трансляцию и сборку своих параллельных программ, на этом сервере в полном объеме должен быть предоставлен программный стек суперкомпьютера. Обычной практикой является идентичность системного и инструментального ПО сервера доступа и вычислительных узлов суперкомпьютера.
- Ограниченность возможности регулярных обновлений системного ПО. Как известно, такие обновления являются неотъемлемыми элементами защиты информации и позволяют устранять известные уязвимости и исправлять ошибки, которые могут быть использованы злоумышленниками для несанкционированного доступа к ресурсам или проведения компьютерных атак. Однако, как уже упоминалось, важнейшим элементом любого СКЦ является программный стек суперкомпьютера. Любое обновление системного ПО влечет за собой неизбежное внесение изменений в программный стек и может привести к его неработоспособности или снижению производительности. Каждое изменение в программном стеке тщательно тестируется специалистами СКЦ, прежде чем внедряется в обслуживающие пользователей суперкомпьютерные системы. Далеко не все обновления системного ПО, даже критические с точки зрения безопасности информации, могут быть применены без ущерба для качества функционирования СКЦ.

Таким образом, защита высокопроизводительных вычислительных систем требует применения комплексного системного подхода, включающего различные методы и механизмы. Решения для обеспечения информационной безопасности должны быть не только направлены на борьбу с потенциальными нарушителями, в том числе и в контексте цепочки поставок оборудования, цепочки поставок ПО и инсайдерских угроз, но и включать в себя сбои, связанные с недостатками и ошибками в аппаратном обеспечении, базовом ПО, операционной системе, библиотеках, компиляторах, а также ошибки проектирования и реализации инфраструктуры высокопроизводительных вычислений. Сюда же следует отнести ошибки пользователей, ошибки в исследовательских программных кодах и рабочих процессах, а также естественные сбои, такие как отказ аппаратных компонентов. Наконец, решения должны быть просто удобными для применения пользователями-исследователями и, в конечном счете, обеспечивать высокие производительность и качество вычислений. В некоторых случаях решением может быть простое отключение критически важных систем от Интернета или, возможно, от любой другой сети [29]. Однако для большинства СКЦ такой подход неприменим.

Средства защиты сетевого периметра и анализа событий безопасности

Рассмотрим концепцию SOC Visibility Triad, предложенную агентством Gartner в 2019 г. [30]. Суть ее заключается в сборе данных из нескольких источников (рис. 2) для обнаружения угроз безопасности, а именно:

– сбор и анализ системных журналов, генерируемых компонентами вычислительной инфраструктуры: терминальными устройствами пользователей, почтовым сервисом, службами аутентификации и авторизации и т.д.;

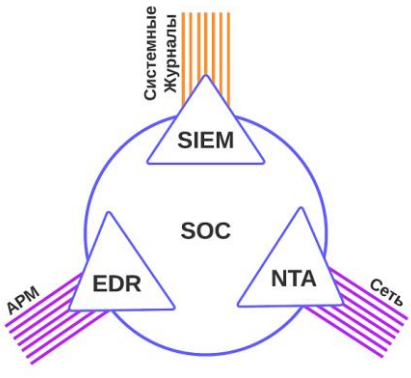

Рис. 2. Источники информации для обнаружения угроз безопасности

Fig. 2. Information sources for detecting security threats

– обнаружение аномалий на терминальных устройствах (автоматизированных рабочих местах) пользователей в корпоративной сети;

– сетецентричное обнаружение потенциальных атак и реагирование на инциденты.

Сетецентричный подход предполагает анализ трафика сетевых устройств с помощью ПО и оборудования для контроля сетевого периметра и трафика внутри сети.

Показанная на рисунке 2 комбинация источников позволяет не только увидеть больше различных событий безопасности, но и сократить время их обнаружения и, следовательно, время реагирования на потенциальные инциденты. Кроме этого, за счет корреляции сетевых аномалий и событий безопасности становится возможным выявление неизвестных атак, для которых у систем контроля сетевого периметра нет сигнатур. При этом, говоря о сетевом периметре, стоит помнить, что стык между корпоративной сетью передачи данных и зонами доступа, управления, вычислений и данных – это тоже граница, часть сетевого периметра, равно как и стыки между самими этими зонами и внутри них, например, стык между сетью управления и интерконнектом внутри зоны вычислений.

Рассматривая возможные решения по обеспечению информационной безопасности, следует учитывать описанные выше особенности научного СКЦ. Например, можно осуществлять контроль границы между транспортной сетью и сетью управления, но не трафика внутри интерконнекта. Применяя концепцию SOC Visibility Triad, выделим из множества средств защиты информации две большие группы: средства контроля сетевого периметра и средства сбора и анализа событий безопасности.

Контроль периметра сети [31, 32] может рассматриваться в качестве одного из основных методов защиты СКЦ. Контроль сетевого периметра включает в себя мониторинг и ограничение доступа к сети центра, а также обнаружение и предотвращение попыток несанкционированного доступа или атак на информационно-вычислительную инфраструктуру СКЦ. Рассмотрим существующие виды средств контроля периметра [32].

- Системы NGFW и UTM. NGFW (Next-Generation Firewall) – это межсетевой экран нового поколения, который комбинирует функции традиционного межсетевого экрана с возможностями контроля приложений и защиты от вторжений. UTM (Unified Threat Management) – интегрированная система управления угрозами, выполняет те же функции, что и системы NGFW, но в отличие от них работает в один поток вместо нескольких, что часто рассматривается как недостаток. При этом стоимость систем UTM ниже по сравнению с NGFW.
- Системы NTA и NDR. NTA (Network Traffic Analysis) – система анализа сетевого трафика, которая позволяет обнаруживать аномальное поведение или подозрительную активность в сети. NDR (Network Detection and Response) представляет собой систему обнаружения и реагирования на сетевые угрозы, которая позволяет выявлять вторжения и другие аномалии в сети, а также принимать меры по их предотвращению.
- Системы IDS и IPS. IDS (Intrusion Detection System) и IPS (Intrusion Prevention System) – соответственно, системы обнаружения и предотвращения вторжений, которые анализируют сетевой трафик и реагируют на подозрительную активность или попытки вторжения. Ключевое отличие систем заключается в том, что IPS может разрывать подозрительное соединение, в то время как IDS только анализирует трафик и сообщает о подозрительной активности. Следует отметить, что эти системы нередко являются частью более сложного и комплексного ПО, но при ограниченном бюджете могут применяться отдельно.

Несомненным преимуществом средств контроля сетевого периметра для СКЦ является то, что они размещаются вне зон суперкомпьютерных систем и таким образом не оказывают негативного влияния ни на производительность вычислений, ни на удобство и порядок работы пользователей СКЦ. В статье [32] приведена таблица рассмотренных средств контроля периметра, позволяющая сравнить функциональные возможности разных видов и классов этих средств. Авторы рекомендуют обратить внимание на NTA-системы, ярким примером которых является PT NAD (Positive Technologies Network Attack Discovery, см. https://www.ptsecurity.com/ru-ru/research/analytics/network-traffic-analysis-2022/). Система PT NAD предоставляет функциональность обнаружения и анализа сетевой активности без вмешательства в работу сети. Она способна выявлять подозрительные пакеты данных, несанкционированные подключения и другие аномалии в сетевом трафике. Система основана на использовании алгоритмов машинного обучения и повышенной гибкости в настройке, что позволяет ей адаптироваться к изменяющимся угрозам и обеспечивать безопасность

СКЦ. Она анализирует не только внешний, но и внутренний трафик и детектирует перемещения злоумышленников внутри сети, попытки эксплуатации уязвимостей, атаки на конечных пользователей в домене и на внутренние сервисы. Обзор других российских NTA-решений можно найти в https://securitymedia.org/analytics/obzor-rossijskih-nta-ndr-sistem.html. Следует отметить, что одной из проблем использования NTA-систем является большое количество генерируемой информации о различных процессах в сети.

Сетевой периметр – это первое, с чем сталкивается внешний нарушитель, поэтому необходимо контролировать входящий и исходящий трафики. Однако только этого недостаточно – необходима организация нескольких уровней защиты сети. Так, Агентство национальной безопасности США рекомендует следующие меры [33]:

– применение пограничного маршрутизатора для подключения к внешней сети, например, к интернет-провайдеру;

– организация нескольких уровней NGFW по всей сети для ограничения входящего трафика, ограничения исходящего трафика и проверки всей внутренней активности между разрозненными регионами сети, причем рекомендуется на каждом уровне применять системы NGFW разных производителей для защиты от атак, использующих одну и ту же уязвимость;

– создание демилитаризованной зоны (DMZ) – изолированной зоны между корпоративной и внешней сетями, в которой размещаются общедоступные системы и сервисы (вебсайт, почта, LDAP и т.п.);

– мониторинг сети при помощи систем обнаружения вторжений (IDS) или NTA-систем;

– выделение нескольких серверов журналирования для выявления корреляции событий безопасности на разных средствах защиты сети;

– наличие резервных средств защиты сети в зонах с наибольшим объемом трафика для обеспечения оперативной балансировки нагрузки с целью увеличения пропускной способности сети и уменьшения задержек.

Таким образом, защищая периметр сети СКЦ и анализируя трафик внутри корпоративной сети, а также защищая публичные сервисы, расположенные в зоне доступа (почтовый сервер, веб-сервер), и выделяя их в отдельную демилитаризованную зону, возможно обеспечить безопасность, а также автоматизировать реагирование на инциденты, что очень важно при большом объеме трафика в сети.

Реализация рассмотренного комплекса мер приводит к тому, что в сети появляется большое количество оборудования и различных систем безопасности, которые постоянно генерируют поток событий. В связи с этим возникает проблема анализа этого потока для выявления атак или признаков компрометации системы. Для этой цели применяются системы управления событиями информационной безопасности (системы SIEM – Security Information and Event Management [34]), способные обнаруживать и реагировать на угрозы, а также выявлять и анализировать события безопасности в режиме реального времени.

Система SIEM предназначена для анализа информации, поступающей из различных источников, таких как системы контроля периметра, антивирусные системы, сетевое оборудование, журналы событий и другие источники. Как только SIEM обнаруживает отклонение, он сразу генерирует событие для администратора безопасности. Следует отметить, что SIEM-системы российских производителей являются одними из лидеров рынка и представлены такими продуктами, как KUMA (Kaspersky Unified Monitoring and Analysis Platform), MaxPatrol SIEM, RuSIEM, KOMRAD Enterprise SIEM. Все эти системы сертифицированы ФСТЭК России. Вместе с коммерческими решениями на рынке также присутствуют свободно распространяемые решения SIEM: ELK Stack (Elasticsearch, Logstash, Kibana), OSSIM (Open Source Security Information Management), Graylog. Сравнительные обзоры представленных на российском рынке SIEM-систем можно найти в работах [35–37].

Для обеспечения информационной безопасности научного СКЦ важно, что SIEM-системы, как и средства контроля сетевого периметра, не потребляют суперкомпьютерные ресурсы и не нарушают порядок работы пользователей СКЦ. Применение их в СКЦ в качестве средств защиты информации более чем целесообразно и оправданно.

Вариант построения системы информационной безопасности научного СКЦ

Проведенный анализ позволяет сформулировать основные принципы построения архитектуры системы информационной безопасности научного СКЦ.

- Для зоны вычислений (вычислительного кластера) не могут быть применимы многие современные средства защиты информации. Тре-

буются эффективное использование штатных механизмов операционных систем вычислительных узлов, проверка безопасности исходного кода программного стека суперкомпьютера, строгое сегментирование сетевого взаимодействия и разделяемых пользовательских вычислительных ресурсов, максимально возможные меры физической защиты.

- Для зоны хранения данных, помимо указанных принципов, допускаются средства защиты при условии их незначительного влияния на скорость передачи данных между СХД и вычислительным кластером. Однако применение гиперконвергентных хранилищ горячих данных, построенных на основе вычислительных узлов суперкомпьютера, осложняет разграничение зон вычислений и данных. В этом случае возможны либо объединение этих зон в единую зону вычислений и данных, либо перенос хранилища горячих данных в зону вычислений.
- Для всех остальных зон (доступа, управления, корпоративной сети передачи данных) возможно применение современных средств защиты, традиционно используемых для защиты типовой корпоративной сетевой инфраструктуры, в том числе с географически распределенными подразделениями. При этом следует помнить о том, что на сервере доступа пользователям должен быть предоставлен программный стек суперкомпьютера для возможности подготовки параллельных программ.

С учетом особенностей обработки информации в СКЦ рассмотрим вариант построения системы информационной безопасности с применением средств контроля сетевого периметра и SIEM. Предлагаемый вариант представлен на рисунке 3.

На уровне сети защита периметра строится с использованием следующих мер:

– публичные сервисы, расположенные в зоне доступа, выделяются в отдельную демилитаризованную зону (сегмент публичных сервисов);

– осуществляется межсетевое экранирование с контролем состояния соединений между общедоступным, корпоративным сегментами и сегментом публичных сервисов;

– сенсоры NTA получают информацию о трафике на периметре сети и в сегментах *корпоративной сети передачи данных* (КСПД) со SPAN-портов коммутаторов; полученный трафик проверяется на наличие признаков нарушителя в сети с помощью сигнатурного и поведенческого анализа;

– осуществляется фильтрация трафика по категориям и отдельным ресурсам с помощью протокола ICAP (см. https://datatracker.ietf.org/doc/html/rfc3507);

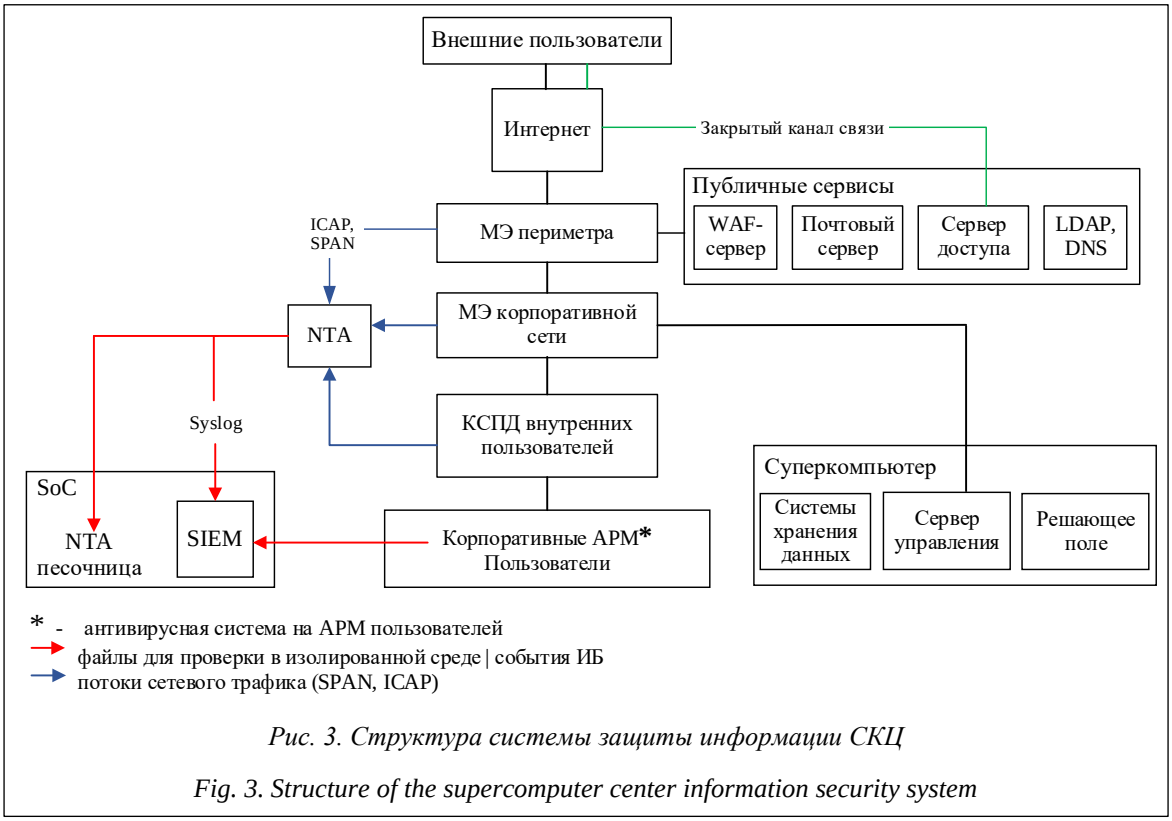

Рис. 3. Структура системы защиты информации СКЦ

Fig. 3. Structure of the supercomputer center information security system

– копируется и затем расшифровывается SSL/TLS-трафик (SSL mirroring);

– обнаруженные в трафике файлы и ссылки отправляются на проверку в локальную песочницу, а также производителю NTA для проверки файлов и ссылок;

– в песочнице файлы и ссылки проверяются с помощью сигнатурного и поведенческого анализа, а также выполняется их запуск в изолированной среде с детальной инспекцией и журналированием действий;

– полученную информацию NTA отправляет в SIEM в виде событий.

На уровне компонентов информационно-вычислительной инфраструктуры СКЦ (серверы, АРМ пользователей и сотрудников и т.п.) устанавливаются антивирусные системы, которые собирают события безопасности, информацию о процессах, работе с файлами, сетевом трафике. При выявлении подозрительной активности принимаются необходимые меры для изоляции соответствующего сервера или АРМ и сбора информации для последующего анализа. Собранная информация отправляется в SIEM, которая объединяет все события и предоставляет интерфейс для мониторинга событий безопасности и реагирования на инциденты. Для повышения общего уровня защищенности инфраструктуры к решениям NTA и SIEM добавляется сетевой экран уровня приложений (Web Application Firewall – WAF). Он обеспечивают защиту веб-серверов при их публикации в сети общего пользования. Кроме этого, реализуются функционал обратного прокси-сервера, расширенные варианты аутентификации, расшифровка SSL, специализированная защита от атак на веб-приложения.

Заключение

Проведенный анализ актуальных исследований в области обеспечения информационной безопасности научного СКЦ показал, что научная разработка вопросов защиты информации в СКЦ находится на начальной стадии, публикации по этой тематике немногочисленны. Обработка информации в СКЦ имеет ряд характерных особенностей, ограничивающих применение известных методов и средств защиты информации. К основным особенностям относятся отсутствие ограничений пользователей на выполняемый ими программный код, невозможность его проверки, недопустимость снижения производительности суперкомпьютерных ресурсов из-за применения средств защиты, ограниченные возможности по применению защищенных операционных систем и обновлению системного ПО вычислительных узлов суперкомпьютера. Для обеспечения защиты информации в СКЦ предлагается разделение информационно-вычислительной инфраструктуры СКЦ на зоны с выделением зон вычислений и данных, для защиты которых предлагается применять средства контроля сетевого периметра и анализа событий безопасности. Предложенный вариант системы информационной безопасности СКЦ позволяет реализовать технологию защиты информации в СКЦ, учитывающую выявленные особенности.

Список литературы

1. Message from the S-HPC 22 workshop chairs. Proc. IEEE/ACM First Int. Workshop on Cyber S-HPC, 2022, pp. iv–iv. doi: 10.1109/S-HPC56715.2022.00004.

2. Xu X., Liu L., Li B. A survey of CAPTCHA technologies to distinguish between human and computer. Neurocomputing, 2020, vol. 408, pp. 292–307. doi: 10.1016/j.neucom.2019.08.109.

3. Абрамов Н.С., Фраленко В.П. Угрозы безопасности вычислительных комплексов: классификация, источники возникновения и методы противодействия // Программные системы: теория и приложения. 2015. Т. 6. № 2. С. 63–83. doi: 10.25209/2079-3316-2015-6-2-63-83.

4. Guo Y., Chandramouli R., Wofford L., Gregg R. et al. High-Performance Computing (HPC) security: Architecture, threat analysis, and security posture. NIST SP, 2023, no. 800-223 ipd. doi: 10.6028/NIST.SP.800-223.ipd.

5. Peisert S. Security in high-performance computing environments. Communications of the ACM, 2017, vol. 60, no. 9, pp. 72–80. doi: 10.1145/3096742.

6. Chen Y., Li K., Fei X., Quan Z., Li K. Implementation and optimization of a data protecting model on the Sunway TaihuLight supercomputer with heterogeneous many-core processors. Concurrency and Computation: Pract. and Experience, 2019, vol. 31, no. 21, art. e4758. doi: 10.1002/cpe.4758.

7. Prout A., Klein A., Michaleas P. et al. Securing HPC using Federated authentication. Proc. IEEE HPEC, 2019, pp. 1–7. doi: 10.1109/HPEC.2019.8916255.

8. Luo Z., Qu Z., Nguyen T.T., Zeng H., Lu Z. Security of HPC systems: From a log-analyzing perspective. ICST Transactions on Security and Safety, 2019, vol. 6, no. 21, art. 163134. doi: 10.4108/eai.19-8-2019.163134.

9. He S., Zhu J., He P., Lyu M.R. Experience report: System log analysis for anomaly detection. Proc. IEEE 27th ISSRE, 2016, pp. 207–218. doi: 10.1109/ISSRE.2016.21.

10. Luo Z., Hou T., Nguyen T.T., Zeng H., Lu Z. Log analytics in HPC: A data-driven reinforcement learning framework. Proc. IEEE INFOCOM, 2020, pp. 550–555. doi: 10.1109/INFOCOMWKSHPS50562.2020.9162664.

11. Du M., Li F., Zheng G., Srikumar V. DeepLog: Anomaly detection and diagnosis from system logs through deep learning. Proc. CCS'17, 2017, pp. 1285–1298. doi: 10.1145/3133956.3134015.

12. Borghesi A., Libri A., Benini L., Bartolini A. Online anomaly detection in HPC systems. Proc. IEEE AICAS, 2019, pp. 229–233. doi: 10.1109/AICAS.2019.8771527.

13. Borghesi A., Molan M., Milano M., Bartolini A. Anomaly detection and anticipation in high performance computing systems. Proc. IEEE Transactions on Parallel and Distributed Systems, 2022, vol. 33, no. 4, pp. 739–750. doi: 10.1109/TPDS.2021.3082802.

14. Chung K., Formicola V., Kalbarczyk Z.T., Iyer R.K., Withers A., Slagell A.J. Attacking supercomputers through targeted alteration of environmental control: A data driven case study. Proc. IEEE Conf. on CNS, 2016, pp. 406–410. doi: 10.1109/CNS.2016.7860528.

15. Kiselev E., Baranov A., Telegin P., Kuznetsov E. System for collecting statistics on power consumption of supercomputer applications. In: LNCS. Proc. RuSCDays, 2022, vol. 13708, pp. 548–561. doi: 10.1007/978-3-031-22941-1_40.

16. Prout A., Arcand W., Bestor D., Bergeron B. et al. Enhancing HPC security with a user-based firewall. Proc. IEEE HPEC, 2016, pp. 1–4. doi: 10.1109/HPEC.2016.7761641.

17. Lee J.-K., Kim S.-J., Hong T. Brute-force attacks analysis against SSH in HPC multi-user service environment. INDJST, 2016, vol. 9, no. 24, pp. 1–4. doi: 10.17485/ijst/2016/v9i24/96070.

18. Stoll C. Stalking the wily hacker. Communications of the ACM, 1988, vol. 31, no. 5, pp. 484–497. doi: 10.1145/42411.42412.

19. Bishop M. UNIX security in a supercomputing environment. Proc. Supercomputing'89, 1989, pp. 693–698. doi: 10.1145/76263.76341.

20. Addai P., Freas R., Tesfa E.M., Sellers M., Mohd T.K. Prevention and detection of network attacks: A comprehensive study. In: LNBIP. Proc. ICDSST, 2023, vol. 474, pp. 56–66. doi: 10.1007/978-3-031-32534-2_5.

21. Wilhelm F. General Pr0ken Filesystem – Hacking IBM's GPFS. ERNW INSINUATOR, 2015. URL: https://insinuator.net/2015/04/general-pr0ken-filesystem-hacking-ibms-gpfs/ (дата обращения: 04.09.2023).

22. Alkhalil Z., Hewage Ch., Nawaf L., Khan I. Phishing attacks: A recent comprehensive study and a new anatomy. Front. Comput. Sci., 2021, vol. 3. doi: 10.3389/fcomp.2021.563060.

23. Liu L., De Vel O., Han Q.-L., Zhang J., Xiang Y. Detecting and preventing cyber insider threats: A survey. IEEE Communications Surveys & Tutorials, 2018, vol. 20, no. 2, pp. 1397–1417. doi: 10.1109/COMST.2018.2800740.

24. Savin G.I., Shabanov B.M., Telegin P.N., Baranov A.V. Joint supercomputer center of the Russian Academy of Sciences: Present and future. Lobachevskii J. Math., 2019, vol. 40, no. 11, pp. 1853–1862. doi: 10.1134/S1995080219110271.

25. Baranov A.V., Savin G.I., Shabanov B.M., Shitik A.S., Svadkovskiy I.A., Telegin P.N. Methods of jobs containerization for supercomputer workload managers. Lobachevskii J. Math., 2019, vol. 40, no. 5, pp. 525–534. doi: 10.1134/S1995080219050020.

26. Reuther A., Byun Ch., Arcand W. et al. Scalable system scheduling for HPC and big data. JPDC, 2018, vol. 111, pp. 76–92. doi: 10.1016/j.jpdc.2017.06.009.

27. Boyens J., Smith A., Bartol N., Winkler K., Hplbrook A., Fallon M. Cybersecurity supply chain risk management practices for systems and organizations. NIST SP, 2022, no. 800-161r1. doi: 10.6028/NIST.SP.800-161r1.

28. Aladyshev O., Baranov A., Ionin R., Kiselev E., Shabanov B. Variants of deployment the high performance computing in clouds. Proc. IEEE EIConRus, 2018, pp. 1453–1457. doi: 10.1109/EIConRus.2018.8317371.

29. Martinez J., Connor C., Hollander B., Paschke K. HPC infrastructure security – ensuring storage and network safety, integrity, efficiency, and performance. Proc. Virtual Conf. Tapia, 2020. doi: 10.2172/1657098.

30. Barros A., Chuvakin A., Belak A. Applying network-centric approaches for threat detection and response. Gartner Research, 2019. URL: https://www.gartner.com/en/documents/3904768 (дата обращения: 04.09.2023).

31. Uctu G., Alkan M., Dogru I.A., Dorterler M. Perimeter network security solutions: A survey. Proc. ISMSIT, 2019, pp. 1–6. doi: 10.1109/ismsit.2019.8932821.

32. Ким Д., Рыжков В. NTA, IDS, UTM, NGFW – в чем разница? // Positive Technologies. 2021. URL: https://www.securitylab.ru/analytics/517592.php?ysclid=lia4byjqfe792657504 (дата обращения: 04.09.2023).

33. Network Infrastructure Security Guide. National Security Agency, Cybersecurity Tech. Report, 2022. URL: https://media.defense.gov/2022/Jun/15/2003018261/-1/-1/0/CTR_NSA_NETWORK_INFRASTRUCTURE_SECURITY_GUIDE_20220615.PDF (дата обращения: 04.09.2023).

34. Vielberth M. Security Information and Event Management (SIEM). In: Encyclopedia of Cryptography, Security and Privacy, 2021. doi: 10.1007/978-3-642-27739-9_1681-1.

35. Лазарева Н.Б., Кармадонов М.С. Анализ программных продуктов SIEM-систем с целью выбора решения для защиты ИБ предприятия // Far East Math: мат. науч.-практич. конф. 2022. С. 127–136.

36. Стукалин А.А., Назарова О.Б. Системы мониторинга информационных инцидентов (SIEM-системы): обзор и сравнительный анализ // Управленческие механизмы противодействия идеологии экстремизма и терроризма: мат. науч.-практич. конф. 2018. С. 116–121.

37. Вишнякова А.Н., Рябцев А.А., Кондрашова Е.В., Шинаков К.Е. Обзор основных SIEM-систем, представленных на российском рынке // Информационная безопасность и защита персональных данных. Проблемы и пути их решения: сб. науч.-практич. конф. 2023. С. 62–65.

Software & Systems doi: 10.15827/0236-235X.142.615-631 2023, vol. 36, no. 4, pp. 615–631

Information security of a supercomputer center

Anton V. Baranov
Pavel M. Korepanov
Egor E. Kuznetsov

For citation
Baranov, A.V., Korepanov, P.M., Kuznetsov, E.E. (2023) 'Information security of a supercomputer center', *Software & Systems*, 36(4), pp. 615–631 (in Russ.). doi: 10.15827/0236-235X.142.615-631
Article info
Received: 25.08.2023 After revision: 28.09.2023 Accepted: 28.09.2023

Abstract. Most scientific supercomputer centers (SC) process open access information. Typically, they use standard information security technologies built into operating systems, data storage, and network devices. In recent years we can see an increase in both the number of information security threats and computer attacks and incidents against SCs, which causes reputational and financial risks for SCs. The paper discusses information processing features in SC, which significantly limit known information security measures and tools. Such features include SC user's freedom to choose tools and application software packages to solve their research problems; a need to ensure maximum computing performance of supercomputer resources provided to users; a limited use of secure operating systems and software update tools. The paper justifies the relevance of developing a complex system approach to SC information security. An adequate level of SC information security must be ensured without significant restrictions of user abilities and reducing the quality of high-performance computing services provided to users. The paper discusses current threats to SC information security, provides a classification of data processed in SC, and defines a list of current information security measures. The proposed optional SC information security system structure takes into account the reviewed features of SC information security. The proposed approach is based on dividing SC computing infrastructure into security zones and on using network perimeter control tools, as well as access and information security event analysis.
Keywords: supercomputer center, information security, network perimeter and access control, security information and event management
Acknowledgements. The work was within the state task of JSCC RAS on topic FNEF-2022-0016

References

1. 'Message from the S-HPC 22 workshop chairs', (2022) *Proc. IEEE/ACM First Int. Workshop on Cyber S-HPC*, pp. iv-iv. doi: 10.1109/S-HPC56715.2022.00004.
2. Xu, X., Liu, L., Li, B. (2020) 'A survey of CAPTCHA technologies to distinguish between human and computer', *Neurocomputing*, 408, pp. 292–307. doi: 10.1016/j.neucom.2019.08.109.
3. Abramov, N., Fralenko, V. (2015) 'Computer systems security threats: Classification, sources of origin and counteraction methods', *Program Systems: Theory and Applications*, 6(2), pp. 63–83 (in Russ.). doi: 10.25209/2079-3316-2015-6-2-63-83.
4. Guo, Y., Chandramouli, R., Wofford, L., Gregg, R. et al. (2023) 'High-Performance Computing (HPC) security: Architecture, threat analysis, and security posture', *NIST SP*, (800-223 ipd). doi: 10.6028/NIST.SP.800-223.ipd.
5. Peisert, S. (2017) 'Security in high-performance computing environments', *Communications of the ACM*, 60(9), pp. 72–80. doi: 10.1145/3096742.
6. Chen, Y., Li, K., Fei, X., Quan, Z., Li, K. (2019) 'Implementation and optimization of a data protecting model on the Sunway TaihuLight supercomputer with heterogeneous many-core processors', *Concurrency and Computation: Pract. and Experience*, 31(21), art. e4758. doi: 10.1002/cpe.4758.
7. Prout, A., Klein, A., Michaleas, P. et al. (2019) 'Securing HPC using Federated authentication', *Proc. IEEE HPEC*, pp. 1–7. doi: 10.1109/HPEC.2019.8916255.
8. Luo, Z., Qu, Z., Nguyen, T.T., Zeng, H., Lu, Z. (2019) 'Security of HPC systems: From a log-analyzing perspective', *ICST Transactions on Security and Safety*, 6(21), art. 163134. doi: 10.4108/eai.19-8-2019.163134.
9. He, S., Zhu, J., He, P., Lyu, M.R. (2016) 'Experience report: System log analysis for anomaly detection', *Proc. IEEE 27th ISSRE*, pp. 207–218. doi: 10.1109/ISSRE.2016.21.
10. Luo, Z., Hou, T., Nguyen, T.T., Zeng, H., Lu, Z. (2020) 'Log analytics in HPC: A data-driven reinforcement learning framework', *Proc. IEEE INFOCOM*, pp. 550–555. doi: 10.1109/INFOCOMWKSHPS50562.2020.9162664.
11. Du, M., Li, F., Zheng, G., Srikumar, V. (2017) 'Deeplog: Anomaly detection and diagnosis from system logs through deep learning', *Proc. CCS'17*, pp. 1285–1298. doi: 10.1145/3133956.3134015.
12. Borghesi, A., Libri, A., Benini, L., Bartolini, A. (2019) 'Online anomaly detection in HPC systems', *Proc. IEEE AICAS*, pp. 229–233. doi: 10.1109/AICAS.2019.8771527.
13. Borghesi, A., Molan, M., Milano, M., Bartolini, A. (2022) 'Anomaly detection and anticipation in high performance computing systems', *Proc. IEEE Transactions on Parallel and Distributed Systems*, 33(4), pp. 739–750. doi: 10.1109/TPDS.2021.3082802.
14. Chung, K., Formicola, V., Kalbarczyk, Z.T., Iyer, R.K., Withers, A., Slagell, A.J. (2016) 'Attacking supercomputers through targeted alteration of environmental control: A data driven case study', *Proc. IEEE Conf. on CNS*, pp. 406–410. doi: 10.1109/CNS.2016.7860528.
15. Kiselev, E., Baranov, A., Telegin, P., Kuznetsov, E. (2022) 'System for collecting statistics on power consumption of supercomputer applications', in *LNCS. Proc. RuSCDays*, 13708, pp. 548–561. doi: 10.1007/978-3-031-22941-1_40.

16. Prout, A., Arcand, W., Bestor, D., Bergeron, B. et al. (2016) 'Enhancing HPC security with a user-based firewall', *Proc. IEEE HPEC*, pp. 1–4. doi: 10.1109/HPEC.2016.7761641.

17. Lee, J., Kim, S., Hong, T. (2016) 'Brute-force attacks analysis against SSH in HPC multi-user service environment', *INDJST*, 9(24), pp. 1-4. doi: 10.17485/ijst/2016/v9i24/96070.

18. Stoll, C. (1988) 'Stalking the wily hacker', *Communications of the ACM*, 31(5), pp. 484–497. doi: 10.1145/42411.42412.

19. Bishop, M. (1989) 'UNIX security in a supercomputing environment', *Proc. Supercomputing'89*, pp. 693–698. doi: 10.1145/76263.76341.

20. Addai, P., Freas, R., Tesfa, E.M., Sellers, M., Mohd, T.K. (2023) 'Prevention and detection of network attacks: A comprehensive study', in *LNBIP. Proc. ICDSST*, 474, pp. 56–66. doi: 10.1007/978-3-031-32534-2_5.

21. Wilhelm, F. (2015) 'General Pr0ken Filesystem – Hacking IBM's GPFS', *ERNW INSINUATOR*, available at: https://insinuator.net/2015/04/general-pr0ken-filesystem-hacking-ibms-gpfs/ (accessed September 04, 2023).

22. Alkhalil, Z., Hewage, Ch., Nawaf, L., Khan, I. (2021) 'Phishing attacks: A recent comprehensive study and a new anatomy', *Front. Comput. Sci*., 3. doi: 10.3389/fcomp.2021.563060.

23. Liu, L., De Vel, O., Han, Q.-L., Zhang, J., Xiang, Y. (2018) 'Detecting and preventing cyber insider threats: A survey', *IEEE Communications Surveys & Tutorials*, 20(2), pp. 1397–1417. doi: 10.1109/COMST.2018.2800740.

24. Savin, G.I., Shabanov, B.M., Telegin, P.N., Baranov, A.V. (2019) 'Joint supercomputer center of the Russian Academy of Sciences: Present and future', *Lobachevskii J. Math*., 40(11), pp. 1853–1862. doi: 10.1134/S1995080219110271.

25. Baranov, A.V., Savin, G.I., Shabanov, B.M., Shitik, A.S., Svadkovskiy, I.A., Telegin, P.N. (2019) 'Methods of jobs containerization for supercomputer workload managers', *Lobachevskii J. Math*., 40(5), pp. 525–534. doi: 10.1134/S1995080219050020.

26. Reuther, A., Byun, Ch., Arcand, W. (2018) 'Scalable system scheduling for HPC and big data', *JPDC*, 111, pp. 76–92. doi: 10.1016/j.jpdc.2017.06.009.

27. Boyens, J., Smith, A., Bartol, N., Winkler, K., Hplbrook, A., Fallon, M. (2022) 'Cybersecurity supply chain risk management practices for systems and organizations', *NIST SP*, (800-161r1). doi: 10.6028/NIST.SP.800-161r1.

28. Aladyshev, O., Baranov, A., Ionin, R., Kiselev, E., Shabanov, B. (2018) 'Variants of deployment the high performance computing in clouds', *Proc. IEEE EIConRus*, pp. 1453–1457. doi: 10.1109/EIConRus.2018.8317371.

29. Martinez, J., Connor, C., Hollander, B., Paschke, K. (2020) 'HPC infrastructure security – ensuring storage and network safety, integrity, efficiency, and performance', *Proc. Virtual Conf. Tapia*, 2020. doi: 10.2172/1657098.

30. Barros, A., Chuvakin, A., Belak, A. (2019) 'Applying network-centric approaches for threat detection and response', *Gartner Research*, available at: https://www.gartner.com/en/documents/3904768 (accessed September 04, 2023).

31. Uctu, G., Alkan, M., Dogru, I.A., Dorterler, M. (2019) 'Perimeter network security solutions: A survey', *Proc. ISMSIT*, pp. 1–6. doi: 10.1109/ismsit.2019.8932821.

32. Kim D., Ryzhkov V. (2021) 'Difference between NTA, IDS, UTM, NGFW', available at: https://www.securitylab.ru/analytics/517592.php?ysclid=lia4byjqfe792657504 (accessed September 04, 2023) (in Russ.).

33. (2022) 'Network Infrastructure Security Guide. National Security Agency', *Cybersecurity Tech. Report*, available at: https://media.defense.gov/2022/Jun/15/2003018261/-1/-1/0/CTR_NSA_NETWORK_INFRASTRUCTURE_SECURITY_GUIDE_20220615.PDF (accessed September 04, 2023).

34. Vielberth, M. (2021) 'Security Information and Event Management (SIEM)', in *Encyclopedia of Cryptography, Security and Privacy*. doi: 10.1007/978-3-642-27739-9_1681-1.

35. Lazareva, N.B., Karmadonov, M.S. (2022) 'Analysis of software products of SIEM systems in order to choose a solution for the protection of enterprise information security', *Proc. Far East Math*, pp. 127–136 (in Russ.).

36. Stukalin, A.A., Nazarova, O.B. (2018) 'SIEM systems: review and comparative analysis', *Management Mechanisms Against Extremism and Terrorism Ideology: Proc. Sci. and Pract. Conf.*, pp. 116–121 (in Russ.).

37. Vishnyakova, A.N., Ryabtsev, A.A., Kondrashova, E.V., Shinakov, K.E. (2023) Review of the main SIEM systems in Russian market', *Information Security and Personal Data Protection. Problems and Solutions: Proc. Sci. and Pract. Conf.*, pp. 62–65 (in Russ.).

Авторы

Баранов Антон Викторович [1], к.т.н., доцент, ведущий научный сотрудник, antbar@mail.ru, abaranov@jscc.ru

Корепанов Павел Михайлович [1], начальник сектора информационной безопасности, kpm@jscc.ru

Кузнецов Егор Евгеньевич [1], младший научный сотрудник, egor57k@ro.ru

[1] Межведомственный суперкомпьютерный центр РАН, г. Москва, 119334, Россия

Authors

Anton V. Baranov [1], Cand. of Sci. (Engineering), Associate Professor, Leading Researcher, antbar@mail.ru, abaranov@jscc.ru

Pavel M. Korepanov [1], Head of Sector Information security, kpm@jscc.ru

Egor E. Kuznetsov [1], Junior Researcher, egor57k@ro.ru

[1] Joint Supercomputer Center of RAS, Moscow, 119334, Russian Federation