Платформа Радар

О Платформе

Версия 3.1.0

ООО Пангео Радар

Оглавление

1. Описание

1.1. Назначение и область применения

Специализированное программное обеспечение «**Платформа Радар**» (далее – **СПО ПР**) представляет собой совокупность взаимосвязанных программ, предназначенных для автоматизации процессов сбора, обработки и корреляции событий информационной безопасности с целью выявления инцидентов и организации реагирования на них.

СПО ПР загружается на технические средства (серверы) автоматизированных информационных систем центров кибербезопасности при создании **программно-аппаратного комплекса** «**Платформа Радар**» (далее – **ПАК ПР**).

СПО ПР после внедрения функционирует в автоматизированном режиме под управлением администратора Заказчика.

1.2. Основные характеристики

СПО ПР в существующей вычислительной сети Заказчика не накладывает ограничений на функционирование серверов и рабочих станций Заказчика, подключаемых к системе в качестве источников событий ИБ.

Доступ к **СПО ПР** должен осуществляться через графический интерфейс (далее Web-интерфейс). Управление и доступ к подсистемам осуществляется в едином окне Web-интерфейса.

СПО ПР, установленное на технические средства в составе **ПАК ПР**, обеспечивает решение следующих задач:

- сбор информации об активах, их учет и управление записями об активах;
- сбор событий ИБ от активов и/или от установленного на активах ПО;
- автоматическая обработка поступивших событий (нормализация, обогащение);
- загрузка и анализ результатов работы сканеров уязвимостей;
- корреляция событий, создание записей об инцидентах и управление ими;
- автоматизированный контроль реагирования на инциденты, контроль устранения;
- получение, обновление и использование информации об угрозах (TI);
- ручной анализ событий ИБ при расследовании инцидентов;
- формирование отчетов, в том числе в графическом виде (дашборды).

СПО ПР состоит из программных модулей. Объединения модулей по функциональному назначению образуют подсистемы (функциональные наборы модулей).

1.3. Форма поставки

СПО ПР поставляется в виде дистрибутива, состоящего из набора установочных и конфигурационных файлов и скриптов.

1.4. Лицензирование программы

Лицензионное соглашение - это юридически обязывающий договор о пользовании Платформой Радар между пользователями и компанией ООО "Пангео Радар" об использовании специализированного программного обеспечения "Платформа Радар".

Перед установкой Платформы Радар необходимо внимательно ознакомиться с лицензионным соглашением. Этот документ включен в комплект поставки.

Лицензия – это переданное неисключительное право на использование Платформы на определенный в тексте лицензии срок, предоставляемое на основании лицензионного соглашения.

Лицензия включает в себя право на:

1. использование Платформы Радар в соответствии с условиями лицензионного соглашения;
2. получение технической поддержки;
3. обновление до актуальной версии.

Если пользователь при работе с Платформой превысит средний поток событий более чем на 1000 EPS в течении 7 дней, то станет недоступным механизм:

- добавления новых компонентов сбора и обработки событий;
- подключения новых источников.

Остальные функции Платформы продолжат работать в штатном режиме.

Ежегодное продление услуги технической поддержки включает в себя право на обновление Платформы Радар.

Срок лицензии, план и максимальное количество событий в секунду можно посмотреть в правом углу внизу рабочего стола Платформы Радар.

Система обеспечивает работоспособность всех своих функций без необходимости продления технической поддержки производителя ПО.

Базовая техническая поддержка платформы оказывается на русском языке с понедельника по пятницу с 9-30 до 18-00. Возможно подключение расширенной технической поддержки в режиме 24/7.

2. Структура

2.1. Основные подсистемы

СПО ПР состоит из программных модулей – отдельных программ, каждая из которых предназначена для выполнения определенных функций. Модули могут объединяться в подсистемы (наборы модулей) в зависимости от их функционального назначения.

Выделяют шесть основных подсистем:

1. Подсистема управления платформой (ядро) — **RADAR-CORE**
2. Подсистема обработки событий — **RADAR-WORKER**
3. Подсистема корреляции событий — **RADAR-CORRELATOR**
4. Подсистема хранения событий — **RADAR-EVENT-STORAGE**
5. Подсистема анализа трафика — **RADAR-NIDS**
6. Подсистема сбора событий — **RADAR-LOG-COLLECTOR**

Помимо основных подсистем в состав **ПАК ПР** входят следующие подсистемы:

1. Подсистема мониторинга работоспособности — **RADAR-MONITORING**
2. Подсистема инфраструктурных модулей — **RADAR-INFRASTRUCTURE**
3. Подсистема балансировщика событий — **RADAR-BALANCER**
4. Подсистема справочной информации об угрозах — **RADAR-TI**

> При установке **ПАК ПР** последние 4 подсистемы могут устанавливаться как самостоятельные модули, так и в составе подсистемы управления платформой **RADAR-CORE**.

> **Важно!** При передаче информации между программными модулями Платформы используется защищенное соединение (TLS).

2.2. Состав подсистем

Ниже в Таблице 1 приведен состав и назначение модулей для каждой подсистемы Платформы.

Таблица 1 - Перечень модулей в составе подсистем Платформы

Подсистема	Наименование модуля	Назначение модуля	Шифр разработки
RADAR-CORE	Радар главный интерфейс	Представление главного веб-интерфейса пользователя	radadr-ui
	Радар менеджмент-центр	Управление внутренними сервисами и предоставление доступа к ним через API на языке Ruby	rmca
	Радар менеджмент-центр АПИ	Управление внутренними сервисами и предоставление доступа к ним через API на языке Go	cruddy
	Радар кластер-агент	Обеспечение управления серверами кластера со стороны кластер-менеджера	cluster-agent
	Радар кластер-менеджер	Управление кластером серверов при конфигурировании ПАК ПР	cluster-manager
	Радар межсервисный шлюз	Организации межсервисного взаимодействия (API Gateway)	cerberus
	Радар госсопка	Автоматизированный обмен информацией с НКЦКИ и ФИНЦЕРТ	event-ant
	Радар сканер активов	Сканирование и инвентаризация активов	sonar

Подсистема	Наименование модуля	Назначение модуля	Шифр разработки
	Радар база знаний	Предоставление программных интерфейсов (API) для работы с базой знаний	knowledgebase
	Радар мониторинг источников	Наблюдение за источниками событий	pluto
	Радар служба оповещений	Оповещения пользователей	toller
	Радар менеджер прав	Авторизация и идентификация пользователей, разграничения прав доступа	keycloak
	Радар сервис TI	Обновление и наполнение справочников информации об угрозах для termite и logmule	ti-updater
	Радар провайдер мультиарендности	Модуль провайдера мультиарендности	karaken
	Радар отчетность	Модуль предоставления программных интерфейсов для работы с отчетностью	datasapi
	Радар табличные списки	Модуль предоставления программных интерфейсов для работы с табличными списками	rvs_api
	Радар интеграционный слой	Модуль предоставления программных интерфейсов для организации интеграционного слоя	pgr-wal-listener
	Радар генератор трафика	Модуль генерации искусственного трафика для проведения нагруженного тестирования	turntable
RADAR-WORKER	Радар обработчик событий	Разбор, нормализация, обогащение входящих событий	termite
RADAR-CORRELATOR	Радар коррелятор	Корреляция обработанных событии на основе предустановленных правил	logmule
RADAR-EVENT-STORAGE	Радар хранилище событий	Хранение и поиск обработанных событий	elasticsearch

Подсистема	Наименование модуля	Назначение модуля	Шифр разработки
RADAR-EVENT-STORAGE	Радар сервис ротации	Модуль ротации индексов подсистемы хранения событий	auto-rollover
RADAR-LOG-COLLECTOR	Радар лог-коллектор	Сбор событий от источников и их передача в Платформу	logcollector-agent
RADAR-NIDS	Радар анализатор сетевого трафика	Обеспечение функции анализа трафика	suricata
	Радар трафик менеджер	Управление модулем NIDS	nids-agent
RADAR-MONITORING*	Радар визуализатор метрик	Визуализация метрик работы Платформы	grafana
	Радар сборщик метрик платформы	Сбор и хранение метрик работы Платформы	prometheus
	Радар сборщик метрик узлов кластера	Агент для сбора метрик с узлов, входящих в состав кластера Платформы	node_exporter
	Радар сборщик метрик балансировщика	Агент для сбора метрик с RADAR-BALANCER	kafka_exporter
	Радар сборщик метрик хранилища	Агент для сбора метрик с RADAR-EVENT STORAGE	elasticsearch-exporter
	Радар менеджер уведомлений	Сигнализация о некорректной работе Платформы (с использованием сервиса toller)	alert-manager
RADAR-INFRASTRUCTURE*	Радар основное хранилище данных	Хранения данных Платформы, включая базу знаний, настройки компонентов и др.	postgresql
	Радар хранилище коррелятора	Хранение промежуточных значений для сервиса RADAR LOGMULE (документо-ориентированное хранилище)	mongodb
	Радар очередь сообщений	Организации очереди обмена сообщений между сервисами RADAR TERMITE и RADAR LOGMULE	rabbitmq

Подсистема	Наименование модуля	Назначение модуля	Шифр разработки
	Радар хранилище очереди сообщений	Хранение временных значений для сервисов rcma и knowledgebase (хранилище ключ-значение)	redis
RADAR-BALANCER*	Радар балансировщик событийРадар ретранслятор запросов TI	Балансировка входящего потока событий.	kafka
	Радар приемник событий	Прием входящих событий от источников и/или сборщиков событий	rsyslog

* — подсистемы, которые могут устанавливаться как автономно, так и в составе подсистемы **RADAR-CORE**

2.3. Структурная схема решения

Платформа имеет модульную архитектуру, приведенную ниже на Рисунке 1.

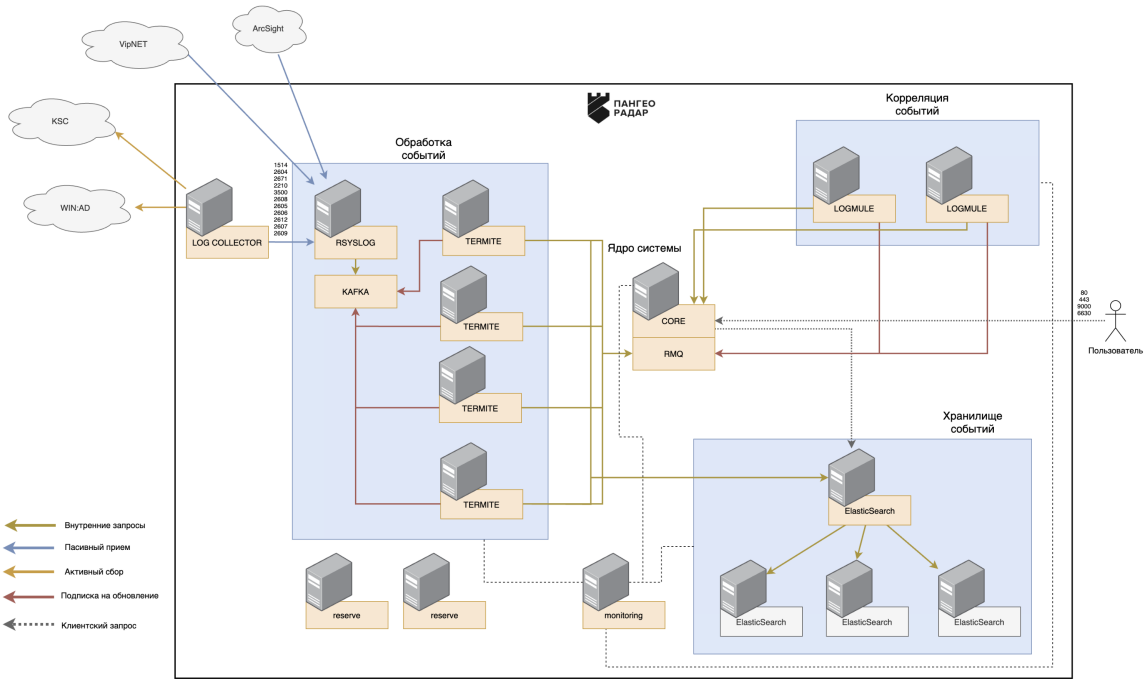

Рисунок 1 - Общая структурная схема решения

Модули сбора событий и модуль управления ими составляют **подсистему сбора событий** (Рисунок 2). Для сбора событий может быть использовано аналогичное ПО сторонних производителей (NXlog и т. п.).

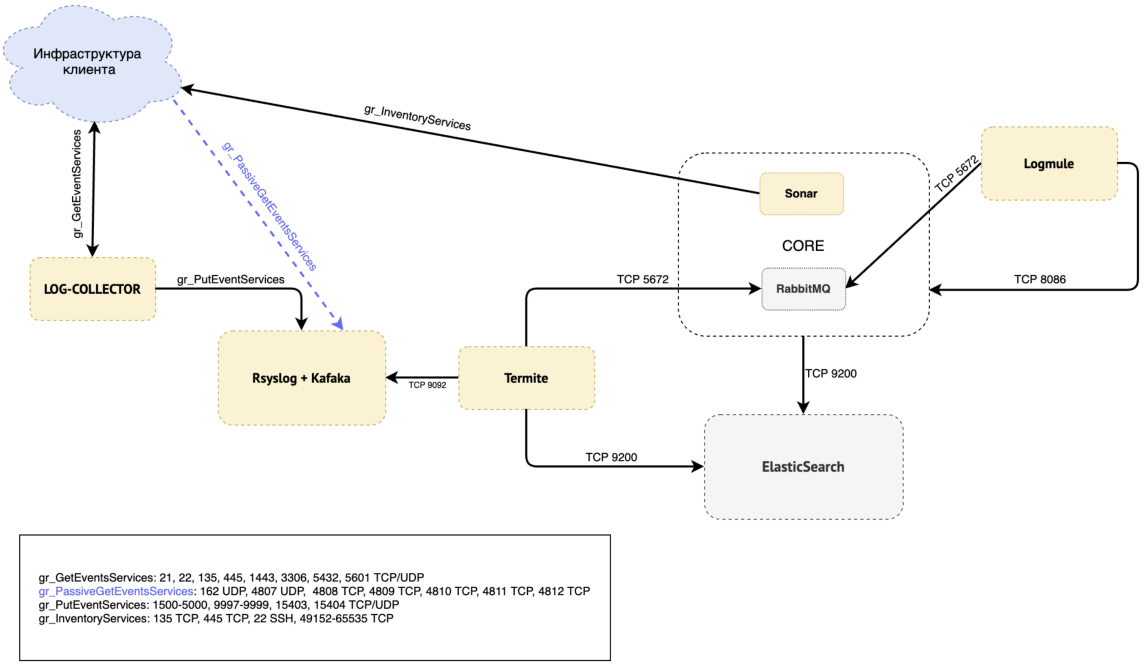

Рисунок 2 - Структура подсистемы сбора событий

> **Важно!** Установка подсистемы сбора событий и стороннего ПО выполняется отдельно от установки основных модулей Платформы и в настоящем документе не описывается.

2.4. Основные функции подсистем

Ниже представлено краткое функциональное описание подсистем Платформы.

2.4.1. Подсистема ядра платформы (RADAR-CORE)

Основные функции подсистемы **RADAR-CORE**:

1. Конфигурирование платформы:
    - первичная настройка платформы для работы в конкретной инфраструктуре
    - управление пользователями и группами
2. Сканирование и инвентаризация активов
3. Контроль ПО и мониторинг состояния модулей платформы
4. Обеспечение резервного копирования и ротации данных
5. Наполнение и обновление справочников информации об угрозах
6. Автоматизированный обмен информацией с НКЦКИ и ФИНЦЕРТ
7. Обеспечение деятельности специалистов по реагированию:
    - управление активами
    - управление правилами корреляции
    - управление инцидентами
    - управление результатами сканирования систем на предмет уязвимостей
    - обеспечения процесса расследования
    - контроль SLA

2.4.2. Подсистема обработки событий (RADAR-WORKER)

Основная функция подсистемы **RADAR-WORKER** — потоковая обработка событий, поступающих в систему, включая:

- разбор событий
- нормализация событий с использованием универсальной схемы
- обогащение событий (данными DNS, Geo-IP, Threat Intelligence)
- запись «сырых» и обработанных событий в модуль хранения данных

Подсистема обработки событий масштабируется до требуемого потока событий.

2.4.3. Подсистема корреляции событий (RADAR-CORRELATOR)

Основные функции подсистемы **RADAR-CORRELATOR**:

1. Корреляция события на основе имеющихся в Платформе правил корреляции
2. Создание инцидентов по результатам корреляции

Подсистема корреляции масштабируется для обработки требуемого потока событий.

2.4.4. Подсистема хранения событий (RADAR-EVENT-STORAGE)

Подсистема **RADAR-EVENT-STORAGE** построена на базе СУБД Elasticsearch и обеспечивает распределенное отказоустойчивое хранение событий.

Основные функции подсистемы **RADAR-EVENT-STORAGE**:

1. Хранение «сырых» и обработанных событий, поступивших в платформу
2. Аналитическая работа с событиями, включающая:
    - поиск, в том числе полнотекстовый
    - агрегации и группировка
    - представление в виде таблиц и графиков

Модули СУБД в составе подсистемы горизонтально масштабируются для обработки требуемого потока событий и обеспечения нужной глубины хранения.

Подсистема хранения событий позволяет выполнять несколько поисковых запросов параллельно. В случае, если подсистема хранения представлена в многосерверной конфигурации, система обрабатывает поисковые запросы с распределением по всем серверам хранения.

2.4.5. Подсистема анализа трафика (RADAR-NIDS)

Основная функция подсистемы **RADAR-NIDS** — мониторинг сетевого трафика в режиме реального времени и выявления в нем аномалий.

Модули подсистемы анализа сетевого трафика интегрируются с сервисом **RADAR-TERMITE** в качестве источника информации.

> **Важно!** Подсистему анализа сетевого трафика рекомендуется устанавливать на отдельно выделенный сервер.

2.4.6. Подсистема сбора событий (RADAR-LOG-COLLECTOR)

Основные функции подсистемы **RADAR-LOG-COLLECTOR**:

1. Управление агентами сбора событий, их централизованное конфигурирование
2. Централизованный мониторинг агентов сбора событий

Для активного сбора событий источников используется модуль **logcollector-agent**

2.4.7. Подсистема мониторинга работоспособности (RADAR-MONITORING)

Основные функции подсистемы **RADAR-MONITORING**:

1. Мониторинг работоспособности Платформы
2. Сбор, хранение и визуализация метрик работы Платформы
3. Сигнализация о некорректной работе Платформы

2.4.8. Подсистема инфраструктурных модулей (RADAR-INFRASTRUCTURE)

Основные функции подсистемы **RADAR-INFRASTRUCTURE**:

1. Хранение данных Платформы в реляционной БД PostgreSQL (включая базу знаний, настройки модулей и т.д.)
2. Хранение промежуточных значений коррелятора (в документо-ориентированном хранилище mongodb)
3. Организация очереди между сервисами **RADAR-TERMITE** и **RADAR-LOGMULE**
4. Хранение временных значений для сервисов **rmca** и **knowledgebase** (в хранилище типа ключ-значение redis)

2.4.9. Подсистема балансировщика событий (RADAR-BALANCER)

Подсистема **RADAR-BALANCER** применяется при масштабировании обработчиков событий, его основными функциями являются:

1. Прием входящего потока событий
2. Распределение потока событий для нескольких обработчиков событий
3. Предоставление временного буфера хранения потока событий

2.4.10. Подсистема справочной информации об угрозах (RADAR-TI)

Основные функции подсистемы **RADAR-TI**:

1. Наполнение справочников информации об угрозах
2. Обновление справочной информации об угрозах для **termite** и **suricata**
3. Трансляция запросов и информации об угрозах с использованием DMZ

2.5. Модуль сбора событий logcollector-agent

Модуль **logcollector-agent** входит в состав подсистемы сбора событий **RADAR-LOG-COLLECTOR**. Предназначен для активного сбора событий с широкого спектра источников (операционные системы, средства защиты, сетевые устройства и др.) и передаче их в подсистему обработки событий **RADAR-WORKER**.

Модуль **logcollector-agent** может быть установлен как на выделенный сервер, так и непосредственно на источник событий.

Возможно разворачивание модуля на следующих платформах:

- RHEL
- CentOS
- Debian
- Ubuntu
- SuSE Linux
- Microsoft Windows
- Apple OS X

Модуль **logcollector-agent** поддерживает следующие технологии сбора событий:

- Event Log
- ODBC
- WMI
- ETW
- Opsec Lea
- SSH
- SMB
- FTP
- SFTP
- Netflow
- TCP
- HTTP (пассивный прием)
- File read
- UDP
- HTTP (Удаленный сбор)
- SNMP trap

2.6. Серверные роли

Для удобства управления установкой и масштабированием, наборам модулей или подсистемам (установленным на сервере и выполняющим определенные функции для пользователей или других серверов **ПАК ПР**), присваиваются серверные роли.

Серверная роль определяет основную функцию сервера, при этом одному серверу могут быть назначены несколько ролей, и одна роль может исполняться несколькими серверами. Перечень используемых в **ПАК ПР** ролей приведён в таблице ниже.

Управление серверными ролями происходит через веб-интерфейс Платформы в разделе управления кластером.

Таблица 2 - Серверные роли подсистем и модулей **ПАК ПР**

Наименование роли	**Базовая функция**	**Ролевой состав**

Наименование роли	Базовая функция	Ролевой состав
MASTER	Управление Платформой	Включает подсистемы RADAR-CORE и RADAR-TI
BALANCER	Балансировка входящего потока событий	Включает подсистему RADAR-BALANCER
WORKER	Обработка входящего потока событий	Включает подсистему RADAR-TERMITE
CORRELATOR	Корреляция обработанного потока событий	Включает подсистему RADAR-LOGMULE
INFRASTRCTURE	Обеспечение работы платформы инфраструктурными модулями	Включает подсистему RADAR-INFRASTRUCTURE
MONITORING	Мониторинг работоспособности Платформы	Включает подсистему RADAR-MONITORING. Часто устанавливается вместе с ролью MASTER
DATA	Хранение данных обработанных событий	Включает подсистему RADAR-EVENT-STORAGE
NIDS	Анализ сетевого трафика	Включает подсистему RADAR-NIDS
LOG-COLLECTOR	Сбор событий с агентов	Включает модуль LOGCOLLECTOR-AGENT.

> **Важно!** Модуль LOGCOLLECTOR-MANAGER устанавливается в составе роли MASTER.

2.7. Возможности масштабирования

Платформа построена на базе сервисно-ориентированной архитектуры, что позволяет масштабировать каждый модуль **ПАК ПР** (Рисунок 1).

Каждый модуль **ПАК ПР** может масштабироваться следующим образом:

- «вертикально» — путем наращивания мощности конфигурации серверов
- «горизонтально» — путем увеличения количества параллельно работающих единиц аппаратного обеспечения с копией масштабируемого модуля или подсистемы

В случае необходимости увеличения производительности конкретной подсистемы Платформы можно оперировать как отдельно взятыми модулями и/или подсистемами, так и серверными ролями.

3. Требования к ПО

3.1. Общие требования к ПО

Для нормального функционирования **СПО ПР** требуется установка на выделенные ресурсы следующего программного обеспечения с версией не ниже указанной:

- **ОС Debian Linux 10**.
- **СУБД PostgreSQL**, **версия 11** — система хранения данных.
- **Elasticsearch**, **версия 6.8.13** — поисковая система для работы с системами управления базами данных (СУБД).
- **RabbitMQ**, **версия 3.9.5** — сервер управления очередями сообщений.
- **Apache Kafka**, **версия 2.7.0** — распределённый программный брокер сообщений.
- **Redis Server**, **версия 5.0.3** — резидентная система управления базами данных.
- **MongoDB**, **версия 4.4.8** — СУБД.
- **rsyslog**, **версия 8.1901.0** — серверное специализированное ПО.

Может быть предусмотрен вариант поставки **СПО ПР**, при котором все перечисленные элементы среды функционирования входят в состав комплекта поставки (включая либо исключая операционную систему). Некоторые элементы среды функционирования могут быть описаны в документации в качестве компонентов **СПО ПР** с учетом их функционального назначения.

Для работы с графическим интерфейсом **СПО ПР** на АРМ пользователя должен быть установлен один из следующих браузеров:

- **Microsoft Edge**
- **Google Chrome**
- **Mozilla Firefox**
- **Яндекс.Браузер**

3.2. Требования к СУБД, используемой как хранилище событий

Для организации хранилища событий используется СУБД класса NoSQL **Elasticsearch**.

СУБД для организации хранилища данных должна обеспечивать выполнение следующих задач:

- обработка параллельных запросов СУБД;
- сжатие хранимых данных;
- индексирование данных;
- репликация и распределенное хранение данных.

СУБД не имеет программных ограничений на срок online-хранения событий. Срок online-хранения зависит только от аппаратных ресурсов серверов СУБД.

Платформа Радар поддерживает сжатие данных при хранении журналов событий со средней степенью сжатия до 30-50% для оперативного хранения и до 80% для архивного хранения событий.

Платформа позволяет использовать для хранения событий как локальные хранилища, так и внешние (сетевые). В случае необходимости масштабирования долгосрочного хранилища событий не потребуется глобальных изменений архитектуры решения.

4. Требования к ТО

Технические требования для работы Платформы рассчитываются для обеспечения штатного функционирования в случае одновременной работы всех пользователей Заказчика.

Данный раздел содержит:

- минимальные требования к аппаратному обеспечению Платформы Радар;
- рекомендации по выбору оборудования для серверов, на которых работает Платформа Радар.

4.1. Минимальные аппаратные требования

Минимальные аппаратные требования, предъявляемые модулями Платформы Радар:

Модули	CPU(cores)	RAM	IOPS	HDD	Net(Гбит)
MASTER INFRASTRUCTURE MONITORING (устанавливаются на один сервер)	16	32	1500	500	1Gbs
WORKER	16	4	1000	120	1Gbs
BALANCER	4	4	1500	1000	1Gbs
CORRELATOR	4	4	1000	120	1Gbs
DATA	16	32	1500	2000	1Gbs

4.2. Выбор оптимальной конфигурации оборудования

Выбор оптимальной конфигурации оборудования зависит от многих факторов. При подборе оборудования для развертывания СПО ПР следует учитывать следующие :

1. Установка ПО будет централизованной или распределенной?
2. Будут ли на сервере работать какие-либо приложения, не относящиеся к Платформе Радар?
3. Сколько событий в секунду должен обрабатывать сервер? Сколько событий в день должен обрабатывать сервер? Оба фактора являются переменными, при этом дневная цифра более важна для определения размера сервера.
4. Какой средний размер события?
5. Сколько источников будет подключено к Платформе Радар?
6. Какие требования предъявляются к обеспечению отказоустойчивости?
7. Длительность хранения события?
8. Есть ли необходимость хранить исходное сообщение или достаточно только нормализованного варианта?

Перечисленные выше факторы (а также, при необходимости, дополнительные факторы) прорабатываются во время разработки проектного внедрения.

Также вопросы по подбору оптимальной конфигурации оборудования можно адресовать в службу технической поддержки, когда требуется определить размер и параметры оборудования "с нуля" или оценить возможности уже существующей потенциальной конфигурации.

4.3. Подбор параметров серверного оборудования

В данном разделе рассматриваются особенности подбора серверного оборудования для установки Платформы Радар по следующим критериям:

- производительность Процессора;
- объём ОЗУ;
- объем и производительность дисковой подсистемы;
- требования к Сети.

> Дисковая подсистема является наиболее частым узким местом.

> Производительность ЦП - второе по популярности узкое место.

> Производительность сети обычно является узким местом только в случае установки распределенной инсталляции.

4.3.1. Подбор Процессора по производительности

При подборе серверного оборудования следует учитывать следующую информацию:

- Все модули Платформы Радар поддерживают 64-разрядные процессоры.
- Так как большинство модулей Платформы Радар, чувствительных к производительности, являются многопоточными, то ресурсы ЦП сервера можно представить как умножение количества ядер ЦП на скорость каждого ядра.

Два значения, которые часто включаются в спецификации ЦП (сервера),- это количество ядер ЦП и количество потоков ЦП. Например, ЦП может иметь 4 ядра и 8 потоков.

При выборе сервера рекомендуется рассматривать производительность ЦП с точки зрения количества потоков, так как данная метрика более актуальна для производительности Платформы Радар, чем физическое кол-во ядер ЦП.

4.3.2. Подбор объема ОЗУ

Разработчик СПО ПР рекомендует для каждого сервера Платформы Радар минимум **16Гб** оперативной памяти. Дополнительная оперативная память может потребоваться в зависимости от требований к производительности Платформы.

Увеличение объема установленной ОЗУ - эффективный способ снизить накладные расходы на операции дискового ввода-вывода.

Следующие компоненты являются основными пользователями оперативной памяти Платформы Радар:

- **PostgreSQL** - в идеальном случае оперативная память должна обеспечивать буферизацию всей базы данных. В большинстве случаев это невозможно, но чем выше процент базы данных, которая может быть буферизована в ОЗУ, тем лучше с точки зрения производительности. Объем дискового пространства, потребляемого PostgreSQL, будет рассмотрен в разделе Требования к дисковому пространству этого документа.
- **Kafka** - в большинстве случаев Kafka может работать с пространством кучи (heap) 6 ГБ памяти. При таком режиме требуется кэш-память файловой системы размером до 28–30 ГБ на машине с 32 ГБ. Для повышенных производственных нагрузок рекомендуется использовать машины 32 ГБ ОЗУ и выше. В этом случае дополнительная оперативная

память будет использоваться для поддержки кеширования страниц ОС и повышения пропускной способности клиентов. Kafka также может работать и с меньшим объемом оперативной памяти, но при этом его способность справляться с нагрузкой затрудняется. Для нормальной работы Kafka потребуется достаточно памяти для буферизации активных читателей и писателей. Можно сделать предварительную оценку потребностей в памяти, исходя из необходимости иметь возможность буферизования в течение 30 секунд. Тогда потребность в памяти вычисляется как write_throughput * 30. Менее 32 ГБ ОЗУ, как правило, непродуктивно (в конечном итоге понадобится много маленьких машин).

- **RADAR TERMITE** - оптимальный размер оперативной памяти для сервиса 16ГБ.
- **RabbitMQ** - для данного ПО оптимально оборудование с 32 ГБ ОЗУ. Оборудование с 16 ГБ ОЗУ в большинстве случаев
  непродуктивно, так как в конечном итоге требует большого объема быстрого дискового пространства в случае использования персистентых (с гарантированной доставкой) очередей.
- **RADAR LOGMULE** - объем ОЗУ определяется характером правил корреляции. Общая рекомендация не менее 8Гб ОЗУ на инстанс.
- **ElasticSearch** - стандартная рекомендация для производительных кластеров - 32Гб на ноду кластера ElasticSearch.
- **Буферы файловой системы** - ОС обычно выделяет большую часть оставшейся оперативной памяти в этой области. Основная область, в которой буферы файловой системы могут улучшить производительность, - это балансировщик событий, очередь обмена сообщениями и хранилище событий.

Объем дискового пространства, который занимает балансировщик, очередь и хранилище очередь, будет рассмотрен в разделе
Требования к дисковому пространству этого документа.

4.3.3. Подбор дисковой подсистемы

4.3.3.1. Рекомендации по вводу-выводу дисковой подсистемы

Для Платформы Радар в большинстве ситуаций производительность произвольного ввода-вывода дисковой подсистемы более важна, чем производительность последовательного чтения и записи, и может оказаться узким местом до ЦП или ОЗУ. Особенно это характерно для централизованной установки Платформы, при которой происходит много операций записи и чтения разными
модулями Платформы, установленными на один сервер.

До определенного уровня производительности можно использовать накопители на магнитных дисках. Но для максимальной производительности рекомендуется использовать твердотельные накопители (SSD). Хотя твердотельные накопители дороги по сравнению с магнитными хранилищами с точки зрения объема хранимых данных, однако при сравнении производительности произвольного ввода-вывода могут оказаться более рентабельными.

Один из ключевых показателей, на который следует обратить внимание при выборе дисковой подсистемы для использования,- это производительность в IOPS (операций ввода-вывода в секунду) как для случайного чтения, так и для произвольной записи.

4.3.3.2. Рекомендации по подбору файловой системы

Рекомендуется использовать файловые системы XFS и избегать EXT4.

- XFS - это высокопроизводительная масштабируемая файловая система, которая обычно развертывается в самых требовательных приложениях. RHEL 7 является файловой системой по умолчанию и поддерживается на всех архитектурах. XFS имеет свои преимущества, но при настройке JBOD она не дает особых преимуществ.
- Ext4 не масштабируется до того же размера, что и XFS.

4.3.3.3. Рекомендации по использованию твердотельного накопителя

Платформа Радар в настоящее время использует два разных уровня SSD-накопителей для собственных стендов:

- На серверах, для которых требуется скорость обработки до 10К событий в секунду или меньше, используются твердотельные накопители Intel 320 Series и X-25M. На текущий момент времени используются приводы серии 320 для новых установок. Приводы X-25M были установлены на стендах до выпуска приводов серии 320.
- На серверах, которым требуется скорость обработки более 10К событий в секунду, были использованы SSD-диски корпоративного класса, такие как Intel S3700 Series или Kingston E100. Данные твердотельные накопители стоят значительно дороже, чем диски Intel 320 Series и X-25M, но при этом обладают значительно большей производительностью.

Все модели SSD-накопителей, используемые на стендах Платформы Радар, зарекомендовали себя как производительные и надежные (для своих уровней) и могут быть рекомендованы для использования в составе серверного оборудования СПО ПР. Ниже приведены ссылки на показатели производительности для некоторых из вышеупомянутых дисков:

- Intel 320 серии.
- Intel серии S3700.
- Kingston E100.

Также могут подойти другие модели SSD-накопителей.

Рекомендуется включать TRIM на SSD-дисках (если это возможно в конкретной модели) и использовать правильное выравнивание разделов.

В некоторых случаях уязвимым местом дисковой подсистемы становится диск или RAID-контроллер, поэтому при использовании твердотельных накопителей также рекомендуется проверить производительность диска или RAID-контроллера, к которому будут подключены твердотельные накопители.

4.3.3.4. Рекомендации по использованию магнитного накопителя

Если SSD-накопители не подходят, то рекомендуется использовать самую быструю доступную конфигурацию магнитного накопителя. Например:

- Использовать накопитель с 7200 оборотов в минуту, а еще лучше 10к оборотов или 15к оборотов в минуту вместо дисков 5400 RPM.
- Использовать диски SAS, так как они обычно быстрее, чем диски SATA.
- Использовать объединение несколько магнитных жестких дисков в один массив RAID 10. Даже если в конкретном случае не нужна совокупная емкость хранилища, то это один из способов увеличения производительности доступного дискового ввода-вывода.

4.3.3.5. Рекомендации по комбинированию твердотельных накопителей и магнитных дисков

Можно использовать комбинацию твердотельных накопителей и магнитных накопителей, чтобы воспользоваться преимуществами каждого из них:

- Компоненты, интенсивно использующие дисковый ввод-вывод, такие как база данных PostgreSQL, MongoDB, RabbitMQ, Kafka, ElasticSeach, могут храниться на SSD.
- Компоненты с низким объемом операций ввода-вывода, такие как операционная система, журналы и резервные копии, могут храниться на магнитных дисках.
- Также рекомендуется переносить неиспользуемые индексы ElasticSearch при длительном хранении на магнитные диски.

4.3.3.6. Рекомендации по использованию RAID-массивов

Чтобы оптимизировать производительность и обеспечить избыточность в случае отказа жесткого диска, рекомендуется
использовать массивы RAID 1 и/или RAID 10.

Не рекомендуется использовать массивы RAID 5 и RAID 0.

При прочих равных условиях надежность Платформы Радар с использованием массива RAID 10 обычно превосходит надежность с
использованием массивов RAID 5 и RAID 0.

4.3.3.7. Рекомендации по кэшированию чтения и записи на контроллерах RAID

Некоторые контроллеры RAID имеют возможность включить кэш чтения и/или записи.

Для **кэша записи** рекомендуется выполнить следующие действия:

- Отключить кэш записи RAID-контроллера, если нет заведомо исправного BBU (блока резервного питания от батареи). Это связано с тем, что кэш записи создает риск потери данных, когда нет работающего BBU.
- Если установлен заведомо исправный BBU, часто имеет смысл включить кэш записи RAID-контроллера. Выполнение этого
  на массиве RAID, который использует магнитные диски, почти всегда повысит производительность. Выполнение
  этого на RAID-массиве, в котором используются SSD-диски, часто, но не всегда, улучшает производительность. Это связано
  с тем, что SSD-диски достаточно быстры и используют собственное кэширование, поэтому иногда дополнительные накладные расходы на выполнение кэширования
  записи оказывают влияние на производительность.
- Некоторые контроллеры RAID также имеют параметр конфигурации для отключения кэширования записи в случае отказа BBU. При наличии данного параметра конфигурации рекомендуется его включить.

Статья «Диски с точки зрения файловой системы» на ACM.org содержит более подробную информацию о том, как работает кэширование записи.

Для **кэша чтения** рекомендуется выполнить следующие действия:

- Перед включением кэша записи RAID-контроллера рекомендуется отключить его кэш чтения, чтобы для записи можно было

выделить больше ресурсов кэша.
- Если кэш записи RAID-контроллера отключается, то имеет смысл включить кэш чтения RAID-контроллера.

Выполнение вышеприведенных действий по включению/отключению кэш на массиве RAID, который использует магнитные диски, в большинстве случаев улучшит производительность.

Выполнение вышеприведенных действий на массиве RAID, в котором используются SSD-диски, часто, но не всегда, улучшит производительность. Это связано с тем, что SSD-диски достаточно быстрые, поэтому иногда дополнительные накладные расходы на выполнение кэширования чтения влияют на производительность.

> Мы рекомендуем отключить кэш чтения RAID-контроллера.

Оперативная память операционной системы может служить кэшем чтения и более доступна для ЦП, чем кэш RAID-контроллера.

4.3.4. Требования к дисковому пространству {#dr}

Требования к дисковому пространству определяются следующими характеристиками:

- длительность хранения событий;
- количество событий в секунду, которые необходимо обработать Платформе Радар.

Вопросы по подбору оптимального дискового пространства или запросы для оценки существующего дискового пространства можно адресовать в службу технической поддержки.

4.4. Требования к параметрам сети

Быстрая и надежная сеть - важный компонент производительности в распределенной системе. Низкая задержка гарантирует, что узлы могут легко обмениваться данными, а высокая пропускная способность помогает перемещению и восстановлению сегментов.
Современные сети центров обработки данных (1 GbE, 10 GbE) достаточны для подавляющего большинства кластеров.

Высокой пропускной способности при работе Платформы Радар не получится достичь при использовании сетевой подсистемы ниже чем 1GbE.

4.5. Совместимость Платформы Радар с технологиями виртуализации

Платформа Радар совместима с некоторыми технологиями виртуализации, которые обеспечивают 64-разрядный процессор и установка ОС Debian 9/10.

Ниже приведены технологии виртуализации, совместимые с Платформой Радар:

- VMware ESX (i) и vSphere.
- Сервер VMware, использующий Linux или Windows в качестве ОС хоста.
- KVM.
- Xen .
- Сервер Microsoft Hyper-V.

С точки зрения производительности рекомендуется с осторожностью использовать мощности облачного провайдера для
установки Платформы Радар.

5. Архитектура решения

Платформа имеет модульную архитектуру, при этом каждый модуль может масштабироваться как «вертикально» (наращиванием мощности конфигурации серверов), так и «горизонтально» (увеличением количества единиц аппаратного обеспечения с копией каждого модуля для параллельной работы).

Описание и состав Платформы приведён в документе «Описание».

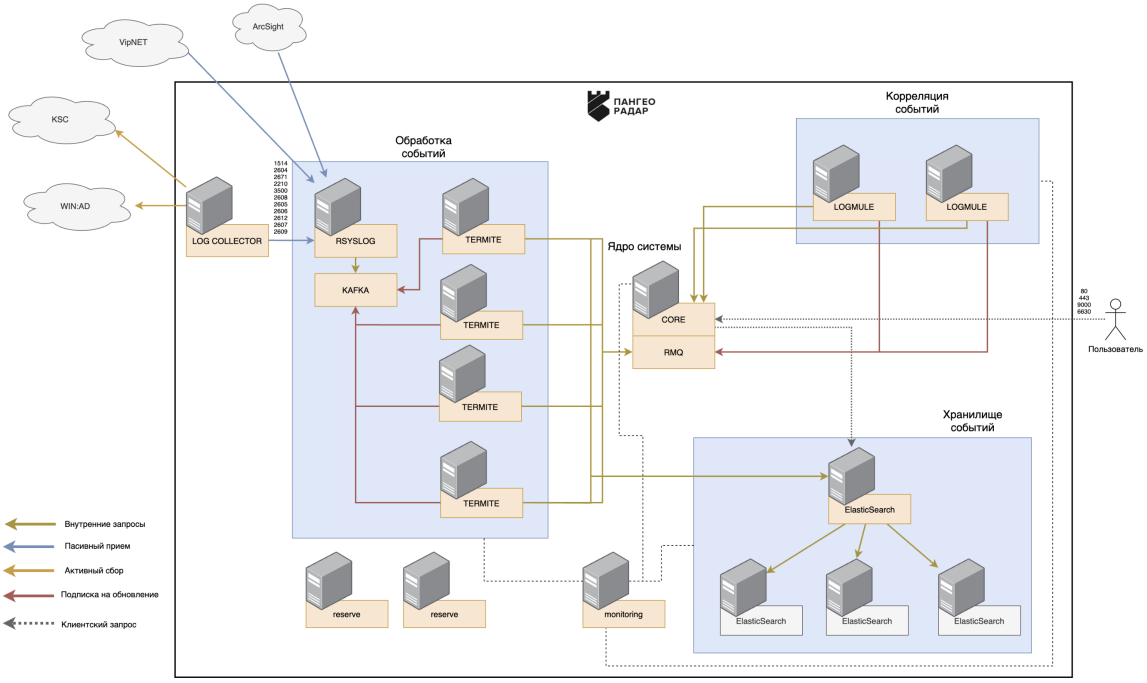

Рисунок 3

Модули сбора событий и модуль управления этими модулями составляют подсистему сбора событий. Для сбора событий также может быть использовано аналогичное ПО сторонних производителей (NXlog и т. п.).

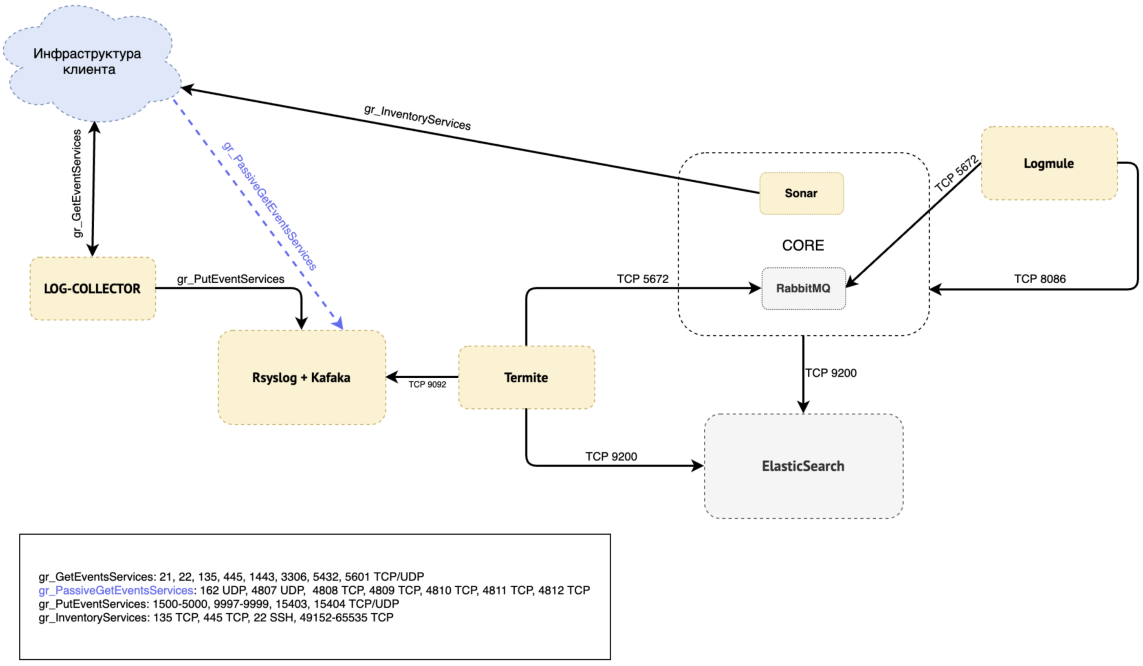

Рисунок 4

> Установка подсистемы сбора событий и стороннего ПО выполняется отдельно от установки основных модулей Платформы и в настоящем документе не описывается

6. Примеры конфигураций

Данные конфигурации представлены для вариантов распределенной установки.
Централизованную установку Платформы (все в
одном - All In One) рекомендуется использовать в случае потока количества событий менее 5K в секунду (EPS), а также когда
нет необходимости выстраивать отказоустойчивое решение с долговременным хранением.

6.0.1. Конфигурация 1

Параметр	Значение
Период хранения в днях	7
Размер одного события в KB	4
EPS	5000
Кол-во правил корреляции	15
Кол-во потоков данных	2
Нужна отказоустойчивость?	Нет

Серверные роли	Кол-во	CPU ядра	RAM GB	Хранилище GB	TB	Сеть
DATA (hot)	1	8	32	200	0,2	1Gbps
DATA (cold)	1	8	16	3600	3,6	1Gbps
CORE	1	24	64	1000	1,0	1Gbps
LOG-COLLECTOR	1	4	4	200	0,2	1Gbps
Суммарно	4	44	108	HDD 4800 Gb		
				SSD 1200 Gb		

6.0.2. Конфигурация 2

Параметр	Значение
Период хранения в днях	7

Параметр	Значение
Размер одного события в KB	4
EPS	10000
Кол-во правил корреляции	15
Кол-во потоков данных	2
Нужна отказоустойчивость?	Нет

Серверные роли	Кол-во	CPU ядра	RAM GB	Хранилище GB	TB	Сеть
DATA (hot)	2	8	48	1200	1,2	1Gbps
DATA (cold)	1	8	16	7200	7,2	1Gbps
CORE	1	10	16	1000	1,0	1Gbps
WORKER	2	16	4	200	0,2	1Gbps
CORRELATOR	2	8	4	200	0,2	1Gbps
LOG-COLLECTOR	2	4	4	200	0,2	1Gbps
Суммарно	11	94	156	HDD 10 Tb		
				SSD 3400 Gb		

6.0.3. Конфигурация 3

Параметр	Значение
Период хранения в днях	7
Размер одного события в KB	8
EPS	30000
Кол-во правил корреляции	15
Кол-во потоков данных	2
Нужна отказоустойчивость?	Да

Серверные роли	Кол-во	CPU ядра	RAM GB	Хранилище GB	TB	Сеть	OS
DATA (hot)	3	8	64	1800	1,8	1Gbps	Debian 10
DATA (cold)	1	8	48	22000	22,0	1Gbps	Debian 10
DATA (coordinator)	1	8	16	1000	1,0	1Gbps	Debian 10

Серверные роли	Кол-во	CPU ядра	RAM GB	Хранилище GB	TB	Сеть	OS
CORE	1	16	32	1000	1,0	1Gbps	Debian 10
INFRA	1	16	16	1000	1,0	1Gbps	Debian 10
WORKER	6	16	4	200	0,2	1Gbps	Debian 10
CORRELATOR	6	8	4	200	0,2	1Gbps	Debian 10
BALANCER	1	4	4	1000	1,0	1Gbps	Debian 10
LOG-COLLECTOR	6	4	4	200	0,2	1Gbps	WIN10-2016/Debian 10
Суммарно	26	244	380		HDD 28 TB		
					SSD 6,4 TB		

6.1. Примеры производительности

6.1.1. Общие данные по тестированию и параметры тестового стенда

В данном раздел приведены примеры работы Платформы Радар. В частности, приведены результаты тестов скорости, которые были
проведены на аппаратном кластере разработчиков Платформы Радар.

Данные результаты следует рассматривать как информацию, которая поможет потенциальному заказчику в принятии решений
относительно его оборудования.

Реальные результаты сильно зависят от многих факторов, не в последнюю очередь из которых: размер события , время
хранения событий, сложность парсинга, нормализации и правил корреляции.

Технические характеристики гипервизора, на котором проводились тесты:

- ЦП: AMD EPYC 7601 2.2 Ghz
- ОЗУ: 256 GB;
- Диск : 1280 GB SSD;
- Сеть: 40 Gbps;
- VMware ESX.

6.1.2. Тестирование обработчика событий

Параметры сообщений для формирования нагрузки на WORKER:

Сценарий	Размер	Формат
1	3 Kb;	JSON
2	107 bytes	raw

Размеры буфера:

Сценарий	Master	Worker	Количество worker модулей
1	100000	1000	-
2	2000	2000	15

Измерение количества фактически обработанных EPS.

Проведение измерений для следующих конфигураций:

- JSON-сообщение, без обогащения;
- Raw-сообщение, без обогащения;
- Raw-сообщение, с обогащением.

Результат WORKER:

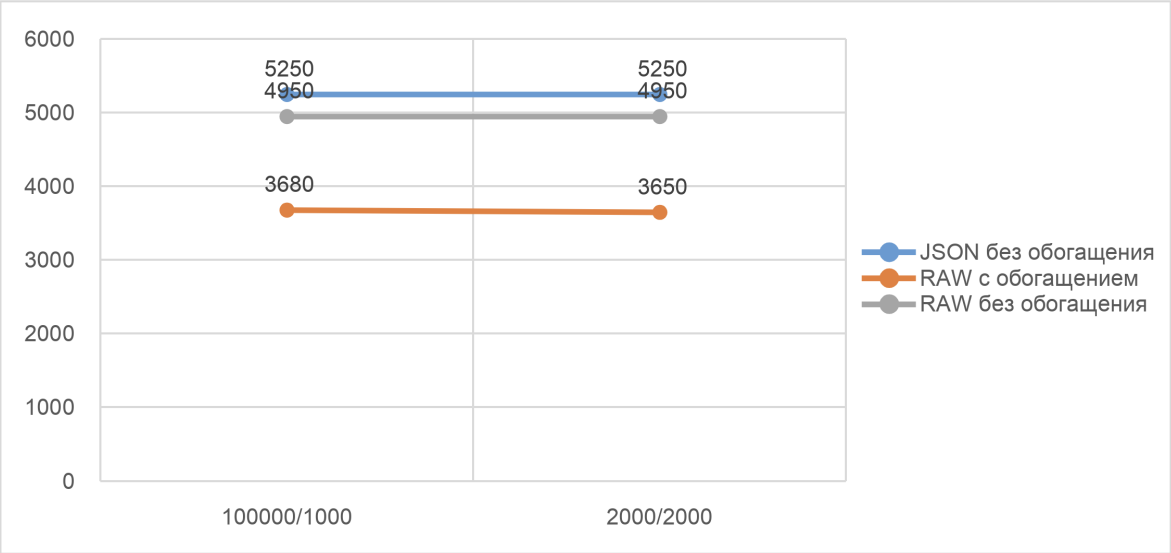

Рисунок 5

Тип	100000/1000	2000/2000
JSON без обогащения	5250	5250
RAW с обогащением	3680	3650
RAW без обогащения	4950	4950

6.1.3. Тестирование распределенного стенда

Роль	CPU	RAM
BALANCER	4C	8Gb
CORRELATOR	4	8Gb
CORRELATOR	4	8Gb
WORKER	16	32GB

Роль	CPU	RAM
WORKER	16	32GB
WORKER	16	32GB
CORE + DATA + MONITORING + INFRA	32	64GB

Параметры сообщений для формирования нагрузки:

- Размер: 3 Kb;
- Формат: JSON.
- Поток: 10k EPS

Рисунок 6

Publish - кол-во событий в секунду выходящих с WORKER.

Deliver - кол-во событий в секунду обрабатываемых CORRELATOR.

Подсистемы	шт	Одновременная запись/чтение	EPS Запись	EPS Чтение
WORKER	3	10 000	14 000	-
CORRELATOR	2	10 000	-	14 000