**Вариант 1. Крупный онлайн-ритейлер**

**Шаг 1: Определение требований**

* 1. **Объем данных**
     1. Ожидаемый объем 500 ТБ в год
     2. Ожидаемый рост: 50% ежегодно
  2. **Скорость получения данных:** 
     1. Мобильный приложения и веб-сайты: до 5000 транзакция в секунду
     2. Социальные сети: обновление каждые 15 минут.
     3. 1C: управление торговлей: в режиме реального времени.
  3. **Типы данных:**
     1. Структурированные: транзакционные данные (информация о продажах, покупках, возвратах, включая идентификаторы товаров, количество, цена, дата, время транзакции), данные о клиентах, складские данные, данные от 1С – 60%
     2. Полуструктурированные: логи событий (данные о взаимодействиях пользователей с сайтом и мобильным приложением, включая время на странице, клики, навигацию), данные о продуктах (описание категорий атрибутов товаров в формате JSON/XML), обратная связь от клиентов – 30%
     3. Неструктурированные: о товарах (изображения и видео товаров), массовые рассылки (текстовые и мультимедийные данные с предложениями и акциями), посты в социальных сетях – 10%
  4. **Требования к обработке:** 
     1. Прогнозирование спроса: в режиме реального времени.
     2. Анализ поведения пользователей: в режиме реального времени.
  5. **Доступность данных:** 
     1. Время отклика для аналитических запросов: <5 секунд.
     2. Доступность системы: 99.99% (допустимое время простоя ~ 8.5 часов в год)
  6. **Безопасность данных:** 
     1. Шифрование
     2. Соответствие 152-ФЗ и PCI DSS (законы о персональных данных, стандарт безопасности данных платёжных карт)

**Шаг 2: Выбор модели хранилища данных: компоненты архитектуры**

Выбор схемы «Снежинка», т. к. она экономит место, что важно для данной компании, у которой по условию объем обрабатываемых и хранимых данных 500 ТБ в год, да еще и ожидается ежегодный рост этого значения на 50%. Также анализ может включать сложные иерархии и отношения.

**2.1. Источники данных**

⎯ Мобильное приложение и веб-сайт.

⎯ Социальные сети, медиа.

⎯ 1C: управление торговлей.

⎯ Внутренние документы.

**2.2. Слой сбора данных**

⎯ Apache Kafka – для сбора потоковых данных (высокая пропускная способность (десятки тысяч сообщений в секунду), горизонтальное масштабирование для роста объема данных, интеграция с ML моделями для прогнозирования на лету, репликация, шифрование, открытое решение).

⎯ Fluentd – для сбора логов (обработка десятков тысяч событий в секунду, имеет плагин kafka2 для прямой отправки в Kafka, шифрование, открытое решение).

**2.3. Слой хранения данных**

⎯ Cloud.ru Object Storage для неструктурированных данных (встроенное S3 Versioning для версионности файлов, решение от Сбера, шифрование, CDN → Интеграция с G-Core Labs для раздачи медиа (сокращение загрузки картинок в 5 раз)).

⎯ PostgreSQL для структурированных и полуструктурированных данных. — надежная реляционная СУБД, подходит для аналитических запросов к нормализованным таблицам (схема «Снежинка»), есть формат хранения данных в бинарном виде JSONB для полуструктурированных, hstore хранит в формате ключ-значение, горизонтально масштабируется, но СУБД не оптимальна для тяжелой аналитики.

⎯ ClickHouse – аналитическое хранилище (быстрее PostgreSQL при выполнении сложных аналитических запросов), горизонтально масштабируется, снимает аналитическую нагрузку с PostgreSQL, оставляя его для оперативных транзакций и работы с текущими данными.

**2.4. Слой обработки данных**

⎯ Apache Spark – для пакетной и потоковой обработки (параллельно с Flink, гкубокий анализ больших пакетов, истории, перерасчет моделей прогнозирования спроса, суточные отчеты). Вычислительная мощность для больших пакетов данных (batch), богатый набор библиотек (MLlib, Spark SQL).

⎯ Apache Flink– для обработки в реальном времени (например, поведение пользователей, события). Для низколатентной обработки в реальном времени (задержка в милисекунды), что критично для персонализации, динамических цен и мониторинга.

**2.5. Слой аналитики и машинного обучения**

⎯ Jupiter Notebook – для аналитик и ML. Стандарт для аналитиков и дата-сайентистов, поддерживает Python, ML-библиотеки (scikit-learn, TensorFlow, PyTorch).

⎯ Яндекс Даталенс – для визуализации и дашбордов. Бизнес-аналитика (продажи, маркетинг, ABC).

⎯ Grafana. Визуализация метрик из Prometheus (и других источников) в графиках/дашбордах. Мониторинг инфраструктуры / performance / SLA.

**2.6. Слой управления данными**

⎯ Apache Ranger – для централизованного управления политиками безопасности (авторизация, маскировка данных, аудит).

2.7. Слой оркестрации и мониторинга

⎯ Apache Airflow – для оркестрации, объединения данных из разных СУБД.

⎯ Prometheus – для мониторинга и алертинга (система сбора метрик)

Cloud.ru предоставляет объектное хранилище OBS. Это позволяет хранить данные в естественном виде, без предварительной структуризации, и обеспечивает масштабируемость и безопасность. Также для работы с неструктурированными данными в Cloud.ru доступны другие сервисы, например, Data Lakehouse, который включает инструменты для хранения больших объёмов неструктурированных данных и аналитики.

**Безопасность**

1. Шифрование: упомянуто на всех уровнях (Kafka, Fluentd, OBS, PostgreSQL, ClickHouse, Ranger). Обоснование: Обязательное требование PCI DSS и 152-ФЗ для защиты данных как при хранении (encryption at rest), так и при передаче (TLS/SSL).
2. Соответствие 152-ФЗ: Ranger (управление доступом, аудит), шифрование, четкое разделение данных (где хранятся перс. данные), процессы обработки перс. данных, реализуемые в пайплайнах Airflow.
3. Данные карт (PAN) обрабатываются только на стороне банка. То есть сайт или приложение запрашивает платеж у банка, потом перенаправляет на платежную страницу банка, токен возвращается обратно. Ритейлер хранит только id транзакции и сумму. Это соответствует модели SAQ A.

**Отказоустойчивость:**

Tier III ЦОД: обеспечивает физическую основу для доступности 99.99% (менее 53 минут простоя в год). Все критичные компоненты (Kafka, PostgreSQL, сеть) подключены к разным источникам бесперебойного питания. Стойки с GPU для ML-обработки имеют усиленное охлаждение. Датчики температуры на каждой стойке с авторегулировкой. При отказе электричества автоматически включаются дизель-генераторы. Резервные сервера.

**Потянет рост данных**

Год 1: 500 ТБ

Год 2: 750 ТБ (+50%)

Год 3: 1125 ТБ (+50%)

Пиковая нагрузка: 5000 → 7500 → 11250 транз/сек.

Например, уже к 3-ему году PostgreSQL будет недостаточно. Решение: Архивирование старых данных + перенос в S3.

Рекомендации для поддержки роста:

1. Перемещать данные старше 6 мес. в холодное хранилище.
2. Оптимизация ClickHouse: использовать партиционирование по времени.
3. Шардинг PostgreSQL.
4. Автомасштабирование Flink/Spark.

Критические точки:

1. PostgreSQL требует замены/оптимизации к 3-му году.
2. Задержки ClickHouse приблизятся к лимиту 5с.
3. Стоимость хранения вырастет на 89% к 3-му году.

Архитектура масштабируема до 3-го года при своевременной оптимизации. Ключевые действия: архивирование данных, усиление кластера PostgreSQL, мониторинг задержек ClickHouse.

**Выводы:**

1. Объем данных (500ТБ+/год, рост 50%): ClickHouse (горизонтальное масштабирование, сжатие), Cloud.ru (неограниченное масштабирование), Kafka (горизонтальное масштабирование), Spark/Flink (распределенная обработка). Tier III ЦОД - база.
2. Скорость (5000 транз/сек): Kafka (высокая пропускная способность), Flink (миллисекундная обработка потока), ClickHouse (быстрая вставка и запрос).
3. Типы данных: структурированные (60%): PostgreSQL (ядро транзакций), ClickHouse (аналитика), полустуктурированные (30%): PostgreSQL (JSONB для оперативных данных), ClickHouse (для аналитики логов/событий), Kafka (транспорт), неструктурированные (10%): Cloud.ru (оптимально).
4. Анализ поведения в реальном времени: Flink - ключевой инструмент. Обрабатывает поток из Kafka, обогащает, выявляет паттерны с минимальной задержкой.
5. Прогнозирование спроса: Spark (обучение на больших исторических данных в ClickHouse), ML-фреймворки в Jupyter, ClickHouse (быстрое получение обучающих выборок).
6. Доступность 99.99%: Tier III ЦОД (физика), репликация в PostgreSQL и ClickHouse, отказоустойчивые кластеры Kafka/Flink/Spark, Airflow (надежность пайплайнов), Prometheus/Grafana (мониторинг и алертинг).
7. Время отклика <5с: Flink (реал-тайм агрегаты), ClickHouse (быстрые аналит. запросы), оптимизированные запросы к PostgreSQL, CDN для медиа.
8. Безопасность (Шифрование, 152-ФЗ, PCI DSS): Шифрование на всех уровнях (Kafka, Cloud.ru, СУБД, трафик), Apache Ranger (централизованный контроль доступа, маскировка, аудит), изоляция компонентов обработки персон. данных/платежей, процессы соответствия.