



МИНОБРНАУКИ РОССИИ

Федеральное государственное бюджетное образовательное учреждение
высшего образования

«МИРЭА – Российский технологический университет»

РТУ МИРЭА

Институт информационных технологий (ИИТ)
Кафедра информационных технологий в атомной энергетике (ИТАЭ)

ОТЧЕТ ПО САМОСТОЯТЕЛЬНОЙ РАБОТЕ №2
по дисциплине «Информационно-технологическая инфраструктура
организаций атомной отрасли»

Студент группы

ИКБО-50-23 Павлов Н.С.

(подпись)

Старший преподаватель

Нежданов И.В.

(подпись)

Москва 2025 г.

СОДЕРЖАНИЕ

1. ПОСТАНОВКА ЗАДАЧИ	3
2 ТЕОРЕТИЧЕСКАЯ ЧАСТЬ	4
2.1 ОСНОВНЫЕ ТИПЫ ДАННЫХ В СИСТЕМАХ ХРАНЕНИЯ	4
2.2 RAID-МАССИВЫ.....	4
3 ПРАКТИЧЕСКАЯ ЧАСТЬ	6
3.1 ПРИМЕРЫ РЕШЕНИЙ ДЛЯ ХРАНЕНИЯ ДАННЫХ	6
3.2 СРАВНИТЕЛЬНЫЙ АНАЛИЗ RAID	7
4 ЗАКЛЮЧЕНИЕ	11

1. ПОСТАНОВКА ЗАДАЧИ

Изучить основные типы данных, их применение в системах хранения, а также ознакомиться с принципами работы RAID-массивов, их разновидностями и особенностями использования.

2 ТЕОРЕТИЧЕСКАЯ ЧАСТЬ

2.1 ОСНОВНЫЕ ТИПЫ ДАННЫХ В СИСТЕМАХ ХРАНЕНИЯ

В современных системах хранения информации данные классифицируются на три основных типа: структурированные, неструктурированные и полуструктурированные. Каждый из них имеет свои особенности и области применения.

Структурированные данные представляют собой информацию, организованную в строгом формате, например, в виде таблиц или баз данных. Такие данные легко обрабатываются с помощью SQL-запросов и используются в системах, где важна точность и скорость доступа. Примерами являются реляционные базы данных (MySQL, PostgreSQL) и таблицы Excel. Их преимущество — высокая скорость обработки, а недостаток — сложность масштабирования при больших объемах.

Неструктурированные данные не имеют четкой структуры и включают в себя мультимедийные файлы, текстовые документы, изображения и видео. Они составляют значительную часть современных данных и хранятся в файловых системах или специализированных хранилищах, таких как Hadoop. Преимущество — гибкость хранения, недостаток — сложность обработки и анализа.

Полуструктурированные данные занимают промежуточное положение. Они не имеют жесткой схемы, но содержат метки или теги, которые позволяют их систематизировать. Примеры — JSON, XML и лог-файлы. Такие данные часто используются в веб-приложениях и API. Их преимущество — адаптивность, а недостаток — необходимость дополнительной обработки для извлечения информации.

2.2 RAID-МАССИВЫ

RAID (Redundant Array of Independent Disks) — это технология объединения нескольких дисков в единый массив для повышения производительности, надежности или их комбинации. RAID используется для

обеспечения отказоустойчивости, увеличения скорости чтения/записи и эффективного использования дискового пространства.

Уровни RAID и их характеристики:

1. RAID 0 (страйпинг): данные разделяются между дисками, что увеличивает скорость работы. Однако отсутствие избыточности делает этот уровень ненадежным — выход одного диска из строя приводит к потере всех данных.

2. RAID 1 (зеркалирование): данные дублируются на двух дисках, обеспечивая высокую надежность. Недостаток — высокие затраты на хранение, так как используется только половина емкости.

3. RAID 5: использует распределенную четность для восстановления данных при отказе одного диска. Оптимален для баланса между производительностью и надежностью, но требует значительных вычислительных ресурсов.

4. RAID 6: аналогичен RAID 5, но с двойной четностью, что позволяет пережить отказ двух дисков. Подходит для критически важных систем, но имеет более низкую скорость записи.

5. RAID 10: комбинация RAID 1 и RAID 0, обеспечивающая высокую производительность и надежность. Недостаток — высокая стоимость из-за дублирования дисков.

3 ПРАКТИЧЕСКАЯ ЧАСТЬ

3.1 ПРИМЕРЫ РЕШЕНИЙ ДЛЯ ХРАНЕНИЯ ДАННЫХ

PostgreSQL — одна из самых мощных и гибких реляционных СУБД с открытым исходным кодом. Она поддерживает сложные запросы, транзакции, индексы, а также расширения для работы с JSON, геоданными и полнотекстовым поиском.

Преимущества:

- Полная поддержка ACID-транзакций, что критично для финансовых систем и других приложений, где важна целостность данных.
- Гибкость: помимо стандартных SQL-функций, PostgreSQL позволяет хранить полуструктурированные данные в формате JSONB.
- Масштабируемость за счет репликации и шардирования.

Недостатки:

- Требуется грамотной настройки для больших нагрузок.
- Менее эффективен, чем специализированные NoSQL-решения, при работе с очень большими массивами данных.

Amazon S3 — это облачное хранилище, предназначенное для хранения больших объемов неструктурированных данных, таких как изображения, видео, логи и резервные копии.

Преимущества:

- Практически неограниченная масштабируемость.
- Высокая надежность и доступность (99,99% SLA).
- Интеграция с другими сервисами AWS (Lambda, Athena, Redshift).

Недостатки:

- Не подходит для частых изменений данных (лучше для архивного хранения).
- Нет встроенных механизмов для сложных запросов (только поиск по метаданным).

3.2 СРАВНИТЕЛЬНЫЙ АНАЛИЗ RAID

Таблица 1 – Сравнение уровней RAID

	RAID 0	RAID 1	RAID 5	RAID 6	RAID 10
Отказоустойчивость	Нет	1 диск	1 диск	2 диска	До 50% дисков
Скорость чтения	Очень высокая	Средняя	Высокая	Средняя	Очень высокая
Скорость записи	Очень высокая	Низкая	Средняя	Низкая	Высокая
Стоимость	Низкая	Высокая	Умеренная	Высокая	Очень высокая

Таблица 2 – Сравнение программные и аппаратные решения для создания RAID

	Аппаратный RAID	Программный RAID	FakeRAID (BIOS RAID)
Производительность	Высокая	Средняя	Ниже аппаратного
Нагрузка на CPU	Нет	Высокая	Умеренная
Отказоустойчивость	Высокая	Зависит от ОС	Средняя
Стоимость	Высокая	Бесплатно	Низкая
Гибкость	Низкая	Высокая	Средняя
Поддержка загрузки	Да	Частично	Да
Примеры	LSI MegaRAID, Dell PERC	mdadm (Linux), Storage Spaces (Windows)	Intel RST, AMD RAIDXpert2

RAID-массивы применяются для повышения производительности, отказоустойчивости или баланса между этими параметрами. Выбор уровня RAID зависит от типа нагрузки, критичности данных и бюджета. Рассмотрим практические примеры использования.

1. RAID 0 (Stripe) – Максимальная скорость, нулевая надежность

Где используется:

- Геймерские ПК и рабочие станции – ускорение загрузки игр и работы с тяжелыми приложениями (видеомонтаж, 3D-рендеринг).
- Кэш-серверы и временные данные – хранение логов, буферов, которые не жалко потерять.
- Высокочастотные транзакции – в системах, где важна скорость записи, а данные дублируются иначе (например, в распределенных системах).

Пример:

Системный диск для стримера – два NVMe SSD в RAID 0 ускоряют загрузку игр и запись стримов, но при отказе одного диска все данные теряются.

Почему не RAID 5/10?

Не нужна избыточность, важна только скорость.

2. RAID 1 (Mirror) – Надежность с минимальной избыточностью

Где используется:

- Корпоративные серверы – загрузочные диски ОС, гипервизоров (ESXi, Proxmox).
- Файловые серверы малого бизнеса – документы, бухгалтерия, где важна сохранность данных.
- Банковские терминалы и АТМ – защита от сбоев диска без сложных настроек.

Пример:

Сервер Active Directory – два диска в RAID 1 обеспечивают отказоустойчивость: если один выйдет из строя, сервер продолжит работу.

Почему не RAID 5/6?

Избыточность 50% – это дорого для больших массивов, но приемлемо для системных дисков.

3. RAID 5 (Stripe + Parity) – Баланс скорости и надежности

Где используется:

- Файловые хранилища и NAS (Synology, QNAP) – медиатеки, общие документы.
- Веб-серверы – базы данных (MySQL, PostgreSQL) с умеренной нагрузкой.
- Системы видеонаблюдения – запись потокового видео с защитой от сбоя диска.

Пример:

Медиасервер Plex – 4 диска по 4 ТБ в RAID 5 дают 12 ТБ полезного пространства и защиту от отказа одного диска.

Почему не RAID 6/10?

RAID 5 дешевле, но при отказ двух дисков данные теряются. Для более критичных систем выбирают RAID 6.

4. RAID 6 (Stripe + Double Parity) – Защита от двойного сбоя

Где используется:

- **Серверы архивного хранения** – резервные копии, медицинские данные.
- **Облачные хранилища** – Backblaze, Amazon Glacier (используют модификации RAID 6).
- **Видеоархивы и СМИ** – долгосрочное хранение контента.

Пример:

- **Сервер медицинской клиники** – 6 дисков в RAID 6 позволяют пережить отказ **любых** двух без потери данных.

Почему не RAID 5?

Современные диски большие (8+ ТБ), и при восстановлении RAID 5 высок риск второго сбоя.

5. RAID 10 (1+0) – Скорость + максимальная надежность

Где используется:

- **Высоконагруженные базы данных** (Oracle, MongoDB) – где важны и скорость, и отказоустойчивость.
- **Виртуализация** (VMware, Hyper-V) – диски для виртуальных машин.
- **Торговые площадки** (например, биржи) – минимальные задержки при транзакциях.

Пример:

Сервер 1С:Предприятие – 8 дисков в RAID 10 (4 пары зеркал + страйпинг) обеспечивают скорость и живучесть даже при отказе нескольких дисков (но только по одному в каждой паре).

Почему не RAID 5/6?

RAID 10 быстрее пишет (не нужно считать parity), а RAID 6 медленнее из-за двойной контрольной суммы.

4 ЗАКЛЮЧЕНИЕ

В ходе выполнения задания были изучены основные типы данных и их применение в системах хранения, а также принципы работы RAID-массивов. Были рассмотрены преимущества и недостатки различных уровней RAID, что позволило понять их оптимальное использование в реальных сценариях. Выполнение практической части помогло закрепить теоретические знания и научиться сравнивать решения для хранения данных. Результаты работы показали, что выбор типа данных и уровня RAID зависит от конкретных требований к производительности, надежности и стоимости.