



# ЛР1

## Вопросы для защиты

### ? 1. Что такое данные?

Данные – это сырые, необработанные факты и цифры, которые не имеют смысла без контекста. Они могут быть собраны из различных источников и в различных формах.

---

### ? 2. Что такое цифровые данные?

Вид данных, доступных пользователю только после компьютерной обработки и хранящиеся в цифровом виде, представленном строками из нулей и единиц, называют цифровыми данными.

---

### ? 3. Что такое информация?

Информация – это данные, которые были обработаны и организованы таким образом, чтобы быть полезными и значимыми для пользователя.

---

### ? 4. Какие типы данных можно выделить?

Данные могут быть классифицированы в зависимости от способа управления данными и их хранения:

- Структурированные данные организуются в строки и столбцы строго определенного формата, хранятся, как правило, с применением системы управления базой данных (DBMS).
- Неструктурированные данные: электронные письма, документы, текстовые файлы.
- Квази-структурированные: данные, которые имеют некоторую структуру, которая может изменяться, но не соответствуют строгим схемам. Соответственно не всегда легко обрабатываются. Они могут содержать элементы, такие как теги или ключи, которые помогают организовать данные. Примеры: JSON, XML, файлы с произвольными данными (например, CSV с переменным количеством колонок).
- Частично структурированные данные: данные, которые содержат как структурированные, так и неструктурированные элементы. Они могут включать фиксированные метаданные, которые облегчают обработку. Примеры: электронные письма (содержат текст, метаданные, такие как отправитель и дата).



## 5. Что такое ЦОД?

В организациях функционируют специализированные помещения или комплекс помещений – дата-центры (ЦОД – центры обработки данных) для централизованной обработки информации всего предприятия.

Инфраструктура дата-центра включает в себя аппаратные компоненты, такие как компьютеры, системы хранения данных, сетевые устройства и резервные источники питания, и программные компоненты, такие как приложения, операционные системы и управляющие программы. Они включают в себя также средства контроля состояния окружающей среды, такие как системы кондиционирования воздуха, противопожарное оборудование и системы вентиляции.



## 6. Характеристики ЦОД?

- Объем (Capacity): для эффективного хранения и обработки большого количества данных операции дата-центра требуют соответствующих ресурсов.
- Доступность (Availability): дата-центры должны обеспечивать доступность запрашиваемой информации.
- Целостность данных (Data integrity): целостность данных означает применение таких механизмов, как коды коррекции ошибок или биты контроля четности, гарантирующих хранение и извлечение данных точно в таком же виде, в котором они были получены.
- Безопасность (Security): в дата-центрах необходимо установить правила, процедуры и надлежащую интеграцию ключевых компонентов с целью предотвращения несанкционированного доступа к информации.
- Масштабируемость (Scalability): развитие компании часто требует развертывания большего количества серверов, установки новых приложений и дополнительных баз данных.
- Производительность (Performance): все основные компоненты дата-центра должны обеспечивать оптимальную производительность в соответствии с необходимыми уровнями обслуживания.
- Управляемость (Manageability): дата-центр должен обеспечивать простое и интегрированное управление всеми своими компонентами. Хорошая управляемость может достигаться путем автоматизации и снижения роли человека (ручного управления) при выполнении стандартных заданий.



## 7. Ключевые процессы управления ЦОД?

К основным функциям управления можно отнести следующие.

- Мониторинг (Monitoring) – непрерывный сбор информации о различных компонентах и службах, запущенных в дата-центре. Мониторинг дата-центра проводится по следующим направлениям: безопасность, производительность, доступность и объем.

- Составление отчетов (Reporting) – периодически проводимая оценка производительности ресурса, его объема и загруженности.
- Предоставление услуг (Provisioning) – процесс обеспечения оборудованием, программами и другими ресурсами, необходимыми для работы дата-центра.



## 8. Этапы развития вычислительных платформ?

1. Мейнфреймы (1950-1960-е годы): огромные, мощные компьютеры, использовавшиеся в крупных организациях для выполнения сложных вычислений.
2. Мини-компьютеры (1960-1970-е годы): более доступные и менее мощные, чем мейнфреймы, компьютеры, предназначенные для использования в малых и средних предприятиях.
3. Персональные компьютеры (1980-е годы и далее): компьютеры, разработанные для индивидуального использования, что сделало вычисления доступными для широких масс.
4. Сети и клиент-серверные архитектуры (1990-е годы): появление локальных сетей (LAN) и архитектур "клиент-сервер", позволяющих компьютерам взаимодействовать и обмениваться данными.
5. Интернет и веб-приложения (1990-е — 2000-е годы): распространение Интернета, развитие веб-приложений и сервисов, работающих в глобальной сети.
6. Мобильные и облачные вычисления (2000-е — настоящее время): появление смартфонов, планшетов и развитие облачных технологий, которые позволяют хранить и обрабатывать данные на удалённых серверах.
7. Искусственный интеллект и машинное обучение (2010-е — будущее): развитие технологий, использующих искусственный интеллект и машинное обучение для анализа данных, автоматизации и создания интеллектуальных систем.

## Ключевые понятия работы

DPE (Disk Processor Enclosure) — это модуль (корпус), который содержит диски (жесткие или твердотельные) и контроллеры, управляющие их работой.

SP (Storage Processor) — это процессор, используемый в системах хранения данных для управления входящими и исходящими запросами к массиву хранения.

LUN (Logical Unit Number) — это логический номер устройства в системе хранения данных, который позволяет серверам обращаться к определенному разделу или блоку данных в массиве хранения. LUN представляет собой абстракцию физического диска или группы дисков, позволяющую администратору эффективно управлять данными.

#### Протоколы FC и File:

1. FC (Fibre Channel) — высокоскоростной сетевой протокол, используемый для передачи данных между серверами и системами хранения в SAN (Storage Area Network).
2. File — протоколы передачи данных на основе файлов, такие как NFS (Network File System) и SMB (Server Message Block), используются для доступа к данным на уровне файлов в сетевых файловых системах (NAS). Эти протоколы позволяют нескольким клиентам получать доступ к файлам на сервере хранения через общую файловую систему.

Пул в системах хранения данных — это совокупность физических или логических дисков, объединенных для предоставления общего пространства хранения.

MultiTier Pool — это пул, состоящий из нескольких уровней хранения (например, SSD, SAS, NL-SAS), где

данные автоматически перемещаются между уровнями в зависимости от их активности.

Fast VP (Fully Automated Storage Tiering for Virtual Pools) — технология, используемая в системах хранения данных, которая позволяет автоматически и динамически перемещать данные между разными уровнями хранения (SSD, SAS, NL-SAS) на основе их активности. Это помогает улучшить производительность, так как более часто используемые данные перемещаются на более быстрые диски, а менее используемые — на более медленные и дешевые.

Initiators (инициаторы) — это устройства или программное обеспечение, которые инициируют запросы на доступ к данным в системах хранения.