Тема 5. Однопроцессные и многопроцессные базы данных.

База данных — это упорядоченный набор структурированной информации или данных, которые обычно хранятся в электронном виде в компьютерной системе. База данных обычно управляется системой управления базами данных (СУБД). Данные вместе с СУБД, а также приложения, которые с ними связаны, называются системой баз данных, или, для краткости, просто базой данных.

Однопроцессные базы данных - это системы управления базами данных, которые используют один процессор для обработки всех операций с базой данных. Они обычно используются для небольших баз данных, где не требуется высокая производительность или масштабируемость. Однопроцессные СУБД обычно проще в использовании и настройке, чем многопроцессорные системы, но они могут не справляться с большими объемами данных или высокой нагрузкой на систему.

Среда однопроцессной базы данных состоит из памяти, диска и единственного процессора. Она может называться по-разному, например, автономной базой данных, базой данных клиент/сервер, последовательной базой данных, однопроцессорной системой и одноузловой или непараллельной средой.

База данных в этой среде обслуживает потребности отдела или небольшого предприятия, где данными и системными ресурсами (в том числе единственным процессором) управляет один менеджер баз данных.

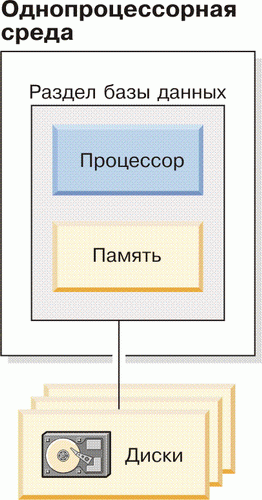


Рисунок 1 Однопроцессная среда

Многопроцессорные базы данных используют несколько процессоров для обработки операций с базой данных, что позволяет увеличить производительность и масштабируемость системы. Многопроцессорные СУБД обычно используются для крупных баз данных с высокой нагрузкой на систему, таких как интернет-магазины или социальные сети. Они могут обрабатывать большое количество запросов одновременно и могут масштабироваться для обработки еще большего количества запросов. Однако многопроцессорные СУБД могут быть сложнее в настройке и использовании, чем однопроцессорные системы.

Любая технология предназначена, прежде всего, для решения конкретных проблем, с которыми она справляется лучше других аналогов.

Зачастую в одном ИТ-проекте необходимо использовать возможности всех категорий этих разнородных хранилищ, что несет дополнительные накладные расходы. Разработчикам приходится писать код приложений, обращающихся к этим хранилищам, используя разные драйверы и коннекторы, архитектор ломает голову над средствами их интеграции, выбирая между оркестрацией и хореографией при взаимодействии нескольких сервисов, а системные администраторы должны поддерживать все это многообразие технологий.

Чтобы уменьшить сложность архитектуры данных, в 2010-х годах возникли мультимодельные СУБД, объединяющие несколько моделей хранения данных в одном решении. Не стоит путать мультимодельные СУБД с технологиями, поддерживающими несколько моделей хранения данных. Например, PostgreSQL и Greenplum изначально представляют собой классические реляционные базы, но с помощью расширений Apache AGE и MADlib могут работать с графами. Также эти СУБД поддерживают тип данных JSON, что характерно, прежде всего, для документо-ориентированных баз данных. Впрочем, под капотом этих дополнительных возможностей по-прежнему работает первичная модель хранения данных, например, в виде таблиц, связанных по внешнему ключу. При этом производительность дополнительных сценариев использования вторичных моделей данных будет намного ниже чем в специализированных решениях. В частности, [здесь](https://bigdataschool.ru/blog/graph-database-under-the-hood.html) мы разбирали, почему нативные графовые базы данных намного быстрее реляционных и документо-ориентированных решений в задачах графовой аналитики.

Считается, что по-настоящему мультимодельные СУБД лишены таких недостатков, будучи изначально спроектированы без привязки к конкретной модели данных. Сегодня наиболее известными из таких хранилищ считаются ArangoDB, OrientDB и Cosmos DB.