

BID DATA –DATABASES

Amazon Athena é o serviço sem servidor que é usado para consultar dados no Amazon S3 usando SQL.
permite a execução de queries no padrão SQL no S3.

Amazon EMR é uma plataforma de **Big Data da AWS**.que permite pré utilizado para criar e gerenciar clusters de servidores de alto desempenho para processamento paralelo
Processar e analisar grandes volumes de dados de forma rápida e eficiente

AWS DMS - Data Migration Service é utilizado para migrar banco de dados mantendo a disponibilidade e consistências dos mesmos.

Amazon Redshift é um banco de dados, mas a sua função é de atuar como um **DW (Data Warehouse)**. **Relação preço-performance até 3x melhor do que outros data warehouses na nuvem**

Amazona MSK (Managed Streaming for Kafka) é um serviço totalmente gerenciado que facilita a criação e a execução de aplicativos criados no Apache Kafka sem a necessidade de gerenciar a infraestrutura.
Permitiria à empresa de serviços financeiros concentrar-se na construção da sua plataforma analítica sem se preocupar com a gestão da infra-estrutura de Kafka, reduzindo assim a sobrecarga operacional

AWS DataSync é um serviço online seguro que automatiza e acelera a movimentação de dados entre serviços de armazenamento on-premises e da AWS

O AWS DataSync o nome já diz, é um serviço para Sync de dados

Amazon QuickSight é o serviço utilizado para gerar Dashboards e gerar insights para tomadas de decisão.

A tabela abaixo ajuda você a entender quando usar diferentes tipos de implantação de banco de dados:

Data Store	Use Case
Database on EC2	<ul style="list-style-type: none">• Need full control over instance and database• Third-party database engine (not available in RDS)
Amazon RDS	<ul style="list-style-type: none">• Need traditional relational database• e.g. Oracle, PostgreSQL, Microsoft SQL, MariaDB, MySQL• Data is well-formed and structured
Amazon DynamoDB	<ul style="list-style-type: none">• NoSQL database• In-memory performance• High I/O needs• Dynamic scaling
Amazon RedShift	<ul style="list-style-type: none">• Data warehouse for large volumes of aggregated data
Amazon ElastiCache	<ul style="list-style-type: none">• Fast temporary storage for small amounts of data• In-memory database
Amazon EMR	<ul style="list-style-type: none">• Analytics workloads using the Hadoop framework

ou seja , de novo ////

DynamoDB: Imagine o DynamoDB como uma prateleira modular na sua cozinha. Ele é um banco de dados NoSQL do tipo documento e chave/valor, ideal para microsserviços autônomos. Sem uma estrutura fixa, é como ter prateleiras que podem se adaptar conforme a necessidade. Ótimo para armazenar itens diversos, mas lembre-se de que, como cada prateleira é independente, consultas complexas podem exigir um pouco mais de planejamento.

- **Quando utilizar:**

- Microsserviços autônomos com necessidade de escalabilidade.

- Aplicações web movimentadas e sistemas de comércio eletrônico.
- **Exemplo de Aplicação:**
 - Aplicações de jogos online que requerem escalabilidade rápida para lidar com picos de usuários.
 - Sistemas de gerenciamento de pedidos em tempo real para comércio eletrônico.
- **Empresas que utilizam:**
 - Amazon.com: Utiliza o DynamoDB para escalabilidade e flexibilidade no gerenciamento de pedidos e recomendações.

Amazon RDS: Pense no Amazon RDS como um assistente pessoal que organiza a sua casa. Não é um banco de dados específico, mas um serviço que facilita a gestão de diversos bancos relacionais como MySQL, PostgreSQL, SQL Server e outros. Similar a um assistente, ele cuida de tarefas diárias como backups e atualizações, ou seja, te dá um conforto maior. Flexível e escalável, adapta-se às suas necessidades, permitindo que você se concentre em outras coisas.

- **Quando utilizar:**
 - Aplicações que já têm familiaridade com bancos de dados relacionais.
 - Projetos que demandam facilidade na gestão e manutenção do banco de dados.

- **Exemplo de Aplicação:**

- Sistemas de gerenciamento de conteúdo que fazem uso extensivo de SQL.
- Aplicações empresariais tradicionais que exigem um banco de dados relacional.

- **Empresas que utilizam:**

- Airbnb: Utiliza Amazon RDS para gerenciamento eficiente de dados de reservas e transações.

Aurora: Agora, o Aurora é como ter um carro esportivo na garagem. É até **5x mais rápido que o MySQL**, oferece segurança robusta e pode lidar com cargas intensas. Imagine um carro de alta performance: elegante, potente, mas talvez um pouco mais do que você realmente precisa. Ideal para projetos que exigem desempenho excepcional, como corridas de velocidade em projetos de análise de dados.

- **Quando utilizar:**

- Projetos que demandam desempenho excepcional, como análises de dados em tempo real.
- Aplicações críticas onde a segurança e replicação robusta são cruciais.

- **Exemplo de Aplicação:**

- Sistemas de análise de grandes volumes de dados em tempo real, como em ambientes financeiros.
- Aplicações que requerem alta disponibilidade e replicação para garantir a integridade dos dados.

- **Empresas que utilizam:**

- Netflix: Utiliza o Amazon Aurora para gerenciar dados críticos, como informações de visualização e preferências de usuários.

Redshift: Se o Redshift fosse um supermercado, seria um centro de análise com caixas ultrarrápidos. Este é um data warehouse poderoso, pronto para integrar com RDS e Aurora. Ao entrar, você percebe que os corredores são amplos e os caixas extremamente eficientes. O Redshift oferece desempenho até 3x superior aos concorrentes, como se estivesse acelerando suas análises. Além disso, tem a capacidade de interagir com seu data lake no S3, uma espécie de área de armazenamento reservada para produtos especiais.

- **Quando utilizar:**

- Análises pesadas de grandes conjuntos de dados.
- Projetos que necessitam de integração com data lakes no Amazon S3.

- **Exemplo de Aplicação:**

- Análise de dados para otimização de cadeias de suprimentos em empresas de varejo.
- Geração de relatórios de business intelligence em tempo real.

- **Empresas que utilizam:**

- Yelp: Utiliza o Amazon Redshift para analisar grandes volumes de dados de avaliações e interações de usuário

ElastiCache: Pense no ElastiCache como ter uma cafeteira na sua mesa de trabalho. Ele armazena dados em cache, proporcionando acesso rápido e eficiente, diminuindo a dependência de acessos frequentes ao banco de dados principal. É como ter uma xícara de café sempre pronta na sua mesa de trabalho, evitando as idas à cozinha a todo momento. Ideal para melhorar o desempenho de aplicações web, mantendo as informações essenciais à mão. Esse cara é muito bom, mas já tive algumas experiências ruins, com microsserviços que implementaram o cache na mesma instância que estavam, e com isso concorreram em memória e CPU, aí a estratégia de dar velocidade pode acabar derrubando seu serviço.

- **Quando utilizar:**

- Melhorar o desempenho de aplicações web através do armazenamento em cache.
- Reduzir a dependência de bancos de dados para operações frequentes de leitura.

- **Exemplo de Aplicação:**

- Aceleração do carregamento de páginas em sites de notícias com conteúdo dinâmico.
- Melhoria na latência de consultas frequentes em aplicativos de análise de dados.

- **Empresas que utilizam:**

- Twitter: Utiliza o Amazon ElastiCache para otimizar a entrega de tweets em tempo real.

Além dessas opções, a AWS oferece soluções especializadas:

- **Amazon DocumentDB:** Imagine como um arquivo organizado para armazenar e consultar documentos, semelhante a um sistema de arquivos eficiente.
- **Amazon Keyspaces:** Pensando em um espaço expansivo para guardar chaves, proporcionando escalabilidade semelhante a um armário de chaves.
- **Amazon Neptune:** Visualize como uma teia de conexões, perfeita para armazenar e explorar relacionamentos complexos, como um mapa interativo.
- **Amazon Timestream:** Como um diário de bordo extremamente organizado, especializado para armazenar e analisar séries temporais, é como ter um diário que organiza eventos ao longo do tempo.
- **Amazon QLDB:** Enxergue como um livro de registros infalível, garantindo a integridade dos dados por meio de registros imutáveis. É como ter um livro contábil que não pode ser alterado.
- Desde a agilidade do DynamoDB até a robustez do Aurora e o poder analítico do Redshift, cada opção atende a necessidades específicas.
- v3