

# Curso de DocumentDB

...

Extractta

# O Roteiro Completo: Do Básico ao Avançado



Módulo 1: **Visão Geral e Fundamentos**



Módulo 2: **Administração e Gerenciamento**



Módulo 3: **Segurança e Compliance**

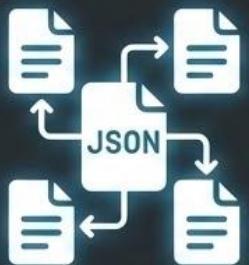


Módulo 4: **Performance e Tuning**



Módulo 5: **Backup e Exportação**

# Pré-requisitos



Conhecimentos  
básicos de bancos  
de dados NoSQL



Familiaridade com  
conceitos de cloud  
computing



Acesso à instância  
fornecida pelo  
instrutor

# Módulo 1 - Visão Geral do AWS DocumentDB

O AWS DocumentDB é um serviço de banco de dados de documentos totalmente gerenciado que oferece compatibilidade com a API do MongoDB.

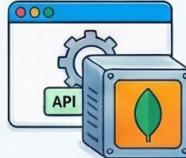
Ele foi projetado para fornecer a flexibilidade e facilidade de uso de bancos de dados de documentos com a confiabilidade, escalabilidade e segurança da AWS.

# DocumentDB vs MongoDB Atlas

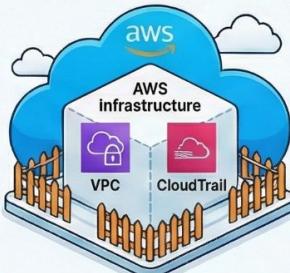


## AWS DocumentDB (A Alternativa Integrada da AWS)

Motor Proprietário que Emula a API do MongoDB



Motor Proprietário que Emula a API do MongoDB  
Não utiliza o motor original do MongoDB, mas é compatível com seus drivers e ferramentas.



Exclusivo do Ecosistema AWS



Gerenciamento de Acesso Dividido  
O AWS IAM gerencia a infraestrutura; usuários e permissões são criados no próprio banco.

## Resumo Comparativo

Característica	AWS DocumentDB	MongoDB Atlas
Origem	Serviço da AWS compatível com API MongoDB	Serviço oficial da MongoDB Inc.
Motor	Implementação proprietária da AWS	Motor nativo do MongoDB
Infraestrutura	Exclusivo da AWS	Multi-cloud (AWS, Azure, GCP)
Gestão	Forte integração com o ecossistema AWS	Plataforma própria e consistente entre nuvens

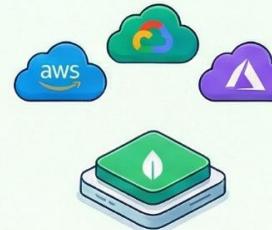


## MongoDB Atlas (A Oferta Oficial e Multi-Cloud)

Motor Nativo e Oficial do MongoDB



Garante acesso às versões mais recentes e a todas as funcionalidades avançadas do MongoDB.



Flexibilidade Multi-Cloud  
Pode ser executado em AWS, Google Cloud (GCP) ou Microsoft Azure de forma consistente.



Utiliza um painel próprio para gerenciar usuários e pode se integrar a provedores de identidade.

## Uma Analogia Simples



AWS DocumentDB é um bolo inspirado na receita famosa  
O sabor é similar, mas feito com técnicas próprias e servido apenas no restaurante da Amazon.



MongoDB Atlas é o bolo da confeitearia original  
Você recebe o produto oficial, com os ingredientes originais e as últimas novidades.

# On-premises vs Gerenciado

**BANCO DE DADOS ON-PREMISES**

- Tarefas críticas e demoradas
- Gestão complexa de infraestrutura
- Falta de agilidade para inovar

**CASA COM MANUTENÇÃO PRÓPRIA**

**CONDOMÍNIO DE LUXO GERENCIADO**

**A analogia perfeita: É como mudar de uma casa onde você conserta tudo sozinho para um condomínio de luxo onde a administração cuida de toda a manutenção estrutural, permitindo que você se concentre em aproveitar o espaço.**

**FOCO NA APLICAÇÃO, NÃO NA INFRAESTRUTURA**

**NUVEM GERENCIADA AWS DocumentDB**

**GESTÃO DE INFRAESTRUTURA SIMPLIFICADA**  
AWS cuida de provisionamento, backup e failover automático.

**AUTOMAÇÃO TOTAL DAS OPERAÇÕES**  
Use AWS CLI, SDKs e Terraform para ambientes reproduíveis e livres de erros.

**ESCALABILIDADE E PROVISIONAMENTO FLEXÍVEL**  
Provisionamento Ágil de Clusters via Console AWS ou programática.

**MONITOREAMENTO AVANÇADO DE PERFORMANCE**  
Monitoramento Avançado de Performance para ajustar o desempenho.

**ALTA DISPONIBILIDADE E PROTEÇÃO DE DADOS**  
Failover Automático e Resiliente  
Backup e Recuperação de Desastres  
Segurança Integrada por Padrão

**Facilite o manejo de grandes quantidades de dados com estratégias integradas de backup e exportação para o Amazon S3.**

**NotebookLM**

# Compatibilidade do DocumentDB



## Versão da API Limitada

Implementa APIs do MongoDB 3.6, 4.0 ou 5.0

MongoDB 3.6  
MongoDB 4.0  
MongoDB 5.0

Versões Mais Recentes (6.0, 7.0+) 

sem recursos mais recentes.



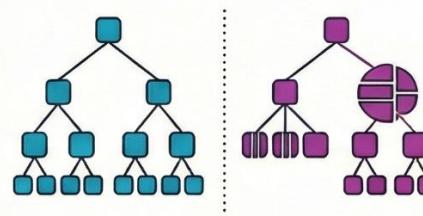
## Suporte Parcial a Operadores e Comandos

Nem todos os operadores de agregação ou comandos administrativos são suportados pela AWS.



## Implementação de Índices Diferente

Índices geoespaciais e de texto podem ser implementados e escalar de forma distinta.



# Console, CLI e SDK

## Console AWS: O Caminho Visual



### Interface Gráfica via Navegador

Ideal para configurações visuais, monitoramento integrado e tarefas administrativas manuais.



## AWS CLI: O Poder da Automação

### Ferramenta de Linha de Comando

Focada na automação de tarefas, scripts de deployment e operações em lote.



## AWS SDKs: Integração Nativa

### Bibliotecas para Integração no Código

Permite que as aplicações interajam programaticamente com os serviços da AWS.



## Resumo Comparativo

Ferramenta	Tipo de Interface	Público-Alvo Principal	Ideal para...
Console	Visual (Navegador)	Iniciantes / Administradores	Configurações手工和 visualização.
CLI	Linha de Comando	DevOps / SysAdmins	Automação rápida e scripts de infraestrutura.
SDK	Código (Bibliotecas)	Desenvolvedores	Integrar serviços AWS diretamente em aplicações.

# Módulo 2 - Administração e Gerenciamento do DocumentDB

# Objetivos

- Provisionar clusters DocumentDB via Console e Terraform
- Configurar políticas de backup e snapshots automáticos
- Implementar e testar failover
- Configurar monitoramento com CloudWatch e EventBridge
- Realizar operações de manutenção e atualizações

# Módulo 2 - Exercício 1: Provisionamento de Clusters

Objetivos:

- Provisionar um cluster DocumentDB via AWS Console
- Provisionar um cluster DocumentDB via Terraform
- Entender as configurações principais de um cluster
- Comparar as duas abordagens

# Parte 1: Provisionamento via AWS Console

Passo 1: Criar Subnet Group

Passo 2: Criar Security Group

Passo 3: Criar o Cluster

Passo 4: Verificar o Cluster

Passo 5: Testar Conexão

# Parte 2: Provisionamento via Terraform

Passo 1: Instalar terraform

Passo 2: Revisar Configuração

Passo 3: Configurar Variáveis

Passo 4: Inicializar Planejar e Aplicar

Passo 5: Verificar Outputs

Passo 6: Testar Conexão

# Validação

/home/\$ID/Curso-documentDB/modulo2-lab/exercicio1-provisionamento/grade\_exercicio1.sh

# Limpeza

```
cd terraform/  
# O terraform destroy usará o .tfstate e as variáveis para remover os recursos corretos  
terraform destroy -auto-approve
```

# Módulo 02 - Exercício 2: Backup e Snapshots Automáticos

## Backup Automático: A Rede de Segurança Continua



### Proteção 24/7 Contínua e Gerenciada

Realiza backups incrementais e automáticos, eliminando a necessidade de intervenção manual.

### Recuperação Precisa no Tempo (PITR)

Permite restaurar o cluster para qualquer segundo dentro do período de retenção (1-35 dias).

### Como uma Câmera de Segurança

Grava tudo continuamente e apaga as gravações antigas após o fim do período de retenção.

## Snapshot Manual: A Foto Instantânea Sob Demanda



### "Foto" do Banco de Dados em um Ponto Específico

Você decide quando capturar o estado exato dos dados, criando um backup sob demanda.

### Essencial para Momentos Críticos e Auditoria

Ideal para antes de atualizações, migrações ou para manter cópias de conformidade.

### Persistência Total e Compartilhamento

Não expiram automaticamente e podem ser compartilhados entre diferentes contas da AWS.



## A Estratégia Completa: Por que Usar Ambos?

### Proteção Máxima com uma Estratégia Dupla

A combinação dos dois garante um plano eficaz de recuperação de desastres (Disaster Recovery).

### Automático para Falhas Diárias, Manual para Marcos Históricos

Proteja-se contra falhas do dia a dia e preserve o histórico de eventos críticos.

# Objetivos

- Entender políticas de backup automático do DocumentDB
- Criar snapshots manuais
- Restaurar clusters a partir de snapshots
- Configurar janelas de backup
- Gerenciar retenção de backups

# Exercício 2: Backup e Snapshots Automáticos

Escolha uma estratégia, via console ou via cli

Parte 1: Configurar Backup Automático

Parte 2: Criar Snapshot Manual

Parte 3: Restaurar a partir de Snapshot

Parte 4: Point-in-Time Recovery (PITR)

Parte 5: Gerenciar Snapshots

# Validação

/home/\$ID/Curso-documentDB/modulo2-lab/exercicio2-backup-snapshots/grade\_exercicio2.sh

# Limpeza

Passo 1: Deletar todas as instâncias do cluster restaurado

Passo 2: Deletar cluster restaurado

Passo 3: Deletar todas as instâncias do cluster restaurado PITR

Passo 4: Deletar cluster restaurado PITR

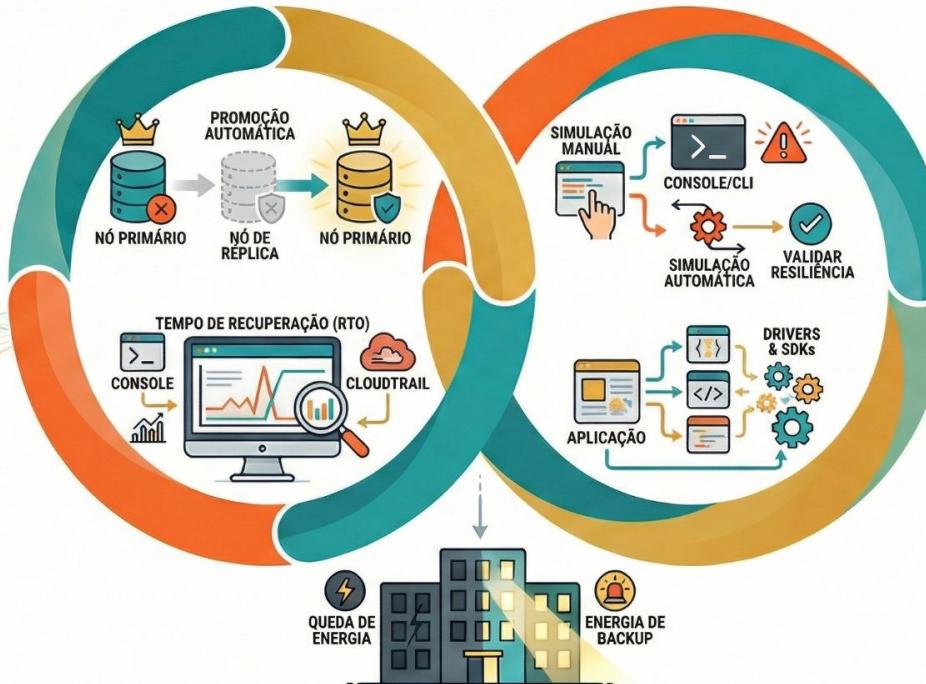
Passo 5: Deletar snapshot manual

Passo 6: Deletar snapshot manual da console

# Módulo 2 - Exercício 3: Gerenciamento de Failover

## 1. Entenda o Conceito de Failover

Uma réplica assume como primária para minimizar o tempo de inatividade do banco.



### Analogia: O Sistema de Luz de Emergência

Failover é o sistema que mantém seu banco de dados funcionando durante uma "queda de energia".

# O que é Failover

## Conceitos Fundamentais



### O que é Failover?

É a comutação automática para um sistema reserva quando o principal falha.



### Objetivo Principal: Continuidade

Reducir ou eliminar completamente o impacto de falhas para os usuários.

### Failover vs. Transição (Switchover)



Failover:  
Automático



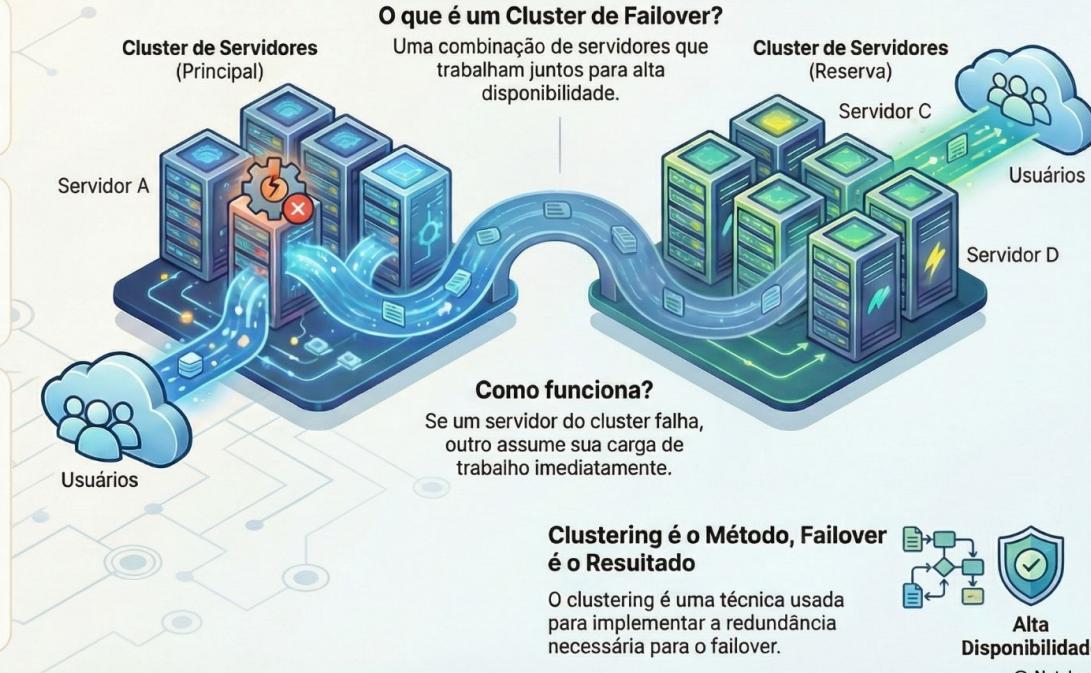
Transição (Switchover):  
Manual

A transição é um processo semelhante, mas requer intervenção humana para ser iniciada.

## O Cluster de Failover: A Implementação

### O que é um Cluster de Failover?

Uma combinação de servidores que trabalham juntos para alta disponibilidade.



### Como funciona?

Se um servidor do cluster falha, outro assume sua carga de trabalho imediatamente.

### Clustering é o Método, Failover é o Resultado

O clustering é uma técnica usada para implementar a redundância necessária para o failover.



Alta  
Disponibilidade

# Exercício 3 - Parte 1: Configurar Ambiente de Teste

Passo 1: Verificar Cluster

Passo 2: Identificar a instância primária atual

# Exercício 3

Parte 1: Configurar Ambiente de Teste

Parte 2: Failover Manual

Parte 3: Simular Failover direcionado (você escolhe a nova RW)

Parte 4: Medir Tempo de Recuperação (RTO)

Parte 5: Aplicação Resiliente a Failover

# Validação

/home/\$ID/Curso-documentDB/modulo2-lab/exercicio3-failover/grade\_exercicio3.sh

# Módulo 2 - Exercício 4: Monitoramento com CloudWatch, EventBridge e SNS

## Fluxo 1: Alertas Baseados em Métricas



## Fluxo 2: Alertas Baseados em Eventos



# Exercício 4

Parte 1: Criar Dashboard no CloudWatch

Parte 2: Configurar Alarmes

Parte 3: EventBridge para Eventos do Cluster

# Validação

/home/\$ID/Curso-documentDB/modulo2-lab/exercicio4-monitoramento/grade\_exercicio4.sh

# Limpeza

Passo 1: Deletar alarmes

Passo 2: Deletar regra EventBridge

Passo 3: Deletar dashboard

Passo 4: Deletar tópico SNS

# Módulo 2 - Exercício 5: Operações de Manutenção e Atualizações

Parte 1: Configurar Janela de Manutenção

Parte 2: Upgrade de Versão (não executar)

Parte 3: Modificar Instâncias

Parte 4: Modificar Parâmetros

Parte 5: Rollback (não executar)

# Validação

/home/\$ID/Curso-documentDB/modulo2-lab/exercicio5-manutencao/grade\_exercicio5.sh

# Limpeza

Passo 1: Deletar snapshots de teste

Passo 2: Deletar parameter group customizado

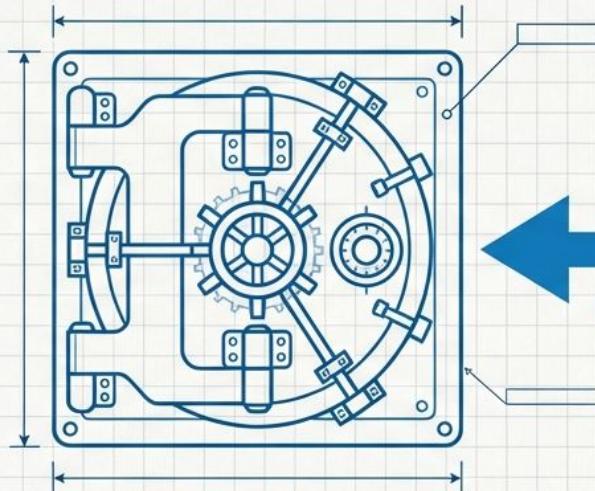
# Módulo 3 - Segurança e Compliance do DocumentDB

# Objetivos

- Implementar autenticação nativa de banco de dados
- Configurar integração segura com VPC, subnets e security groups
- Estabelecer controle de acesso com TLS e roles de privilégios mínimos
- Habilitar auditoria completa com CloudTrail e CloudWatch Logs

## A Analogia Central: Seu Banco de Dados é um Cofre de Banco

Para entender a segurança do DocumentDB, imagine que você está projetando um banco. Cada componente da AWS tem um papel físico e tangível na proteção do seu ativo mais importante: os dados.



O Ativo Valioso  
(Seus Dados)



Vamos desenhar a planta  
desta fortaleza.



A Estrutura de Proteção  
(Sua Nuvem AWS)

# Passo 1: Construindo o Perímetro (A VPC)

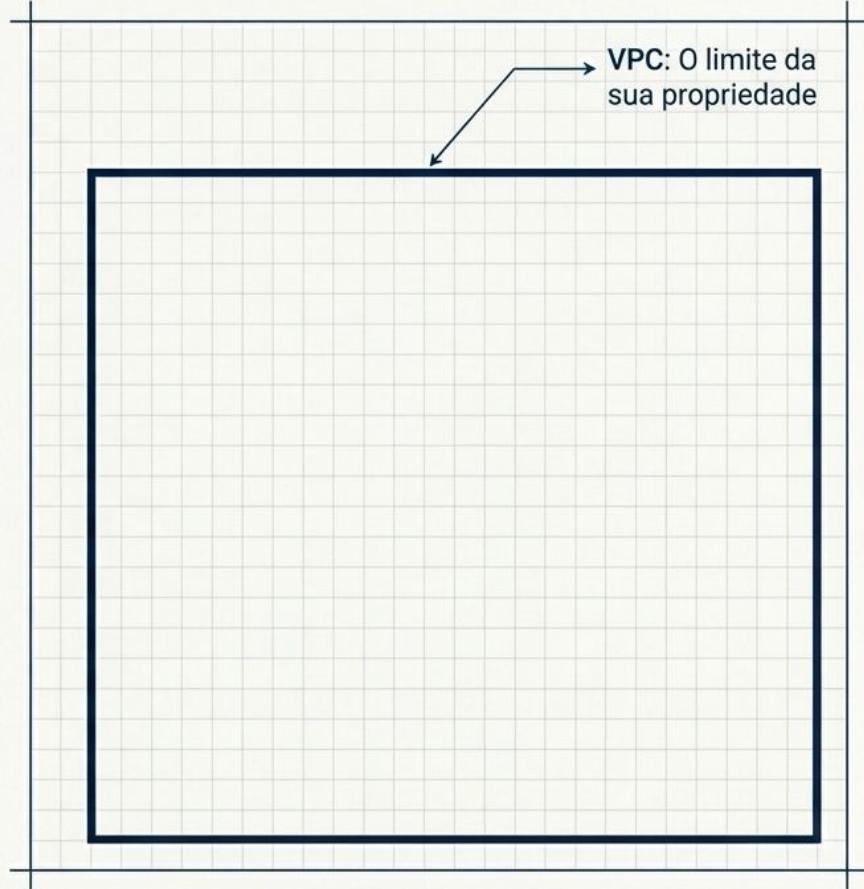
## A Metáfora

O Terreno e o Prédio do Banco. É a sua propriedade privada dentro do vasto bairro da AWS. Define os limites de tudo o que está dentro.

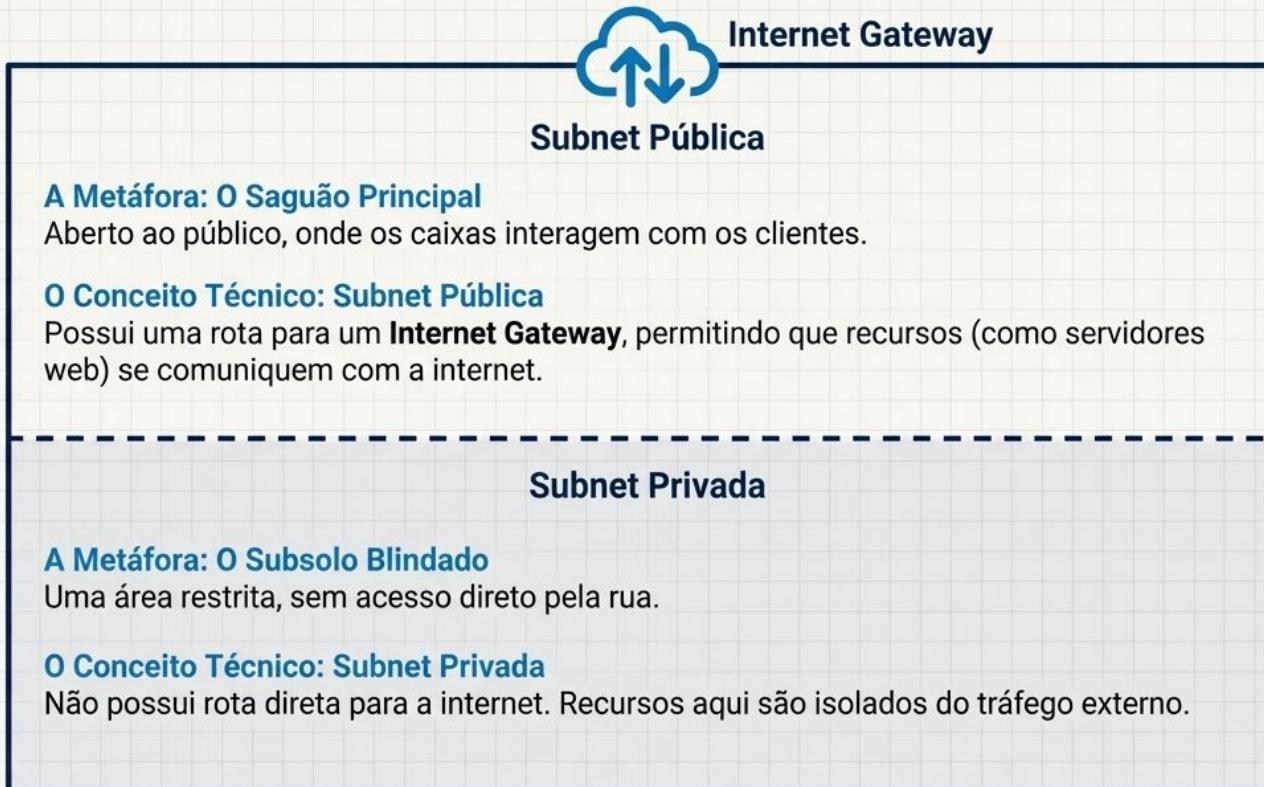
## O Conceito Técnico

### Amazon Virtual Private Cloud (VPC).

Uma rede virtual logicamente isolada na nuvem AWS. Todos os seus recursos, incluindo o cluster DocumentDB, residirão dentro desta rede. É a primeira e mais fundamental camada de isolamento.



## Passo 2: Separando o Saguão do Subsolo Seguro (As Subnets)



## Passo 3: Posicionando o Cofre (O Cluster DocumentDB)

O cluster DocumentDB é projetado para ser um recurso de rede interno. Sua segurança começa com o isolamento total.

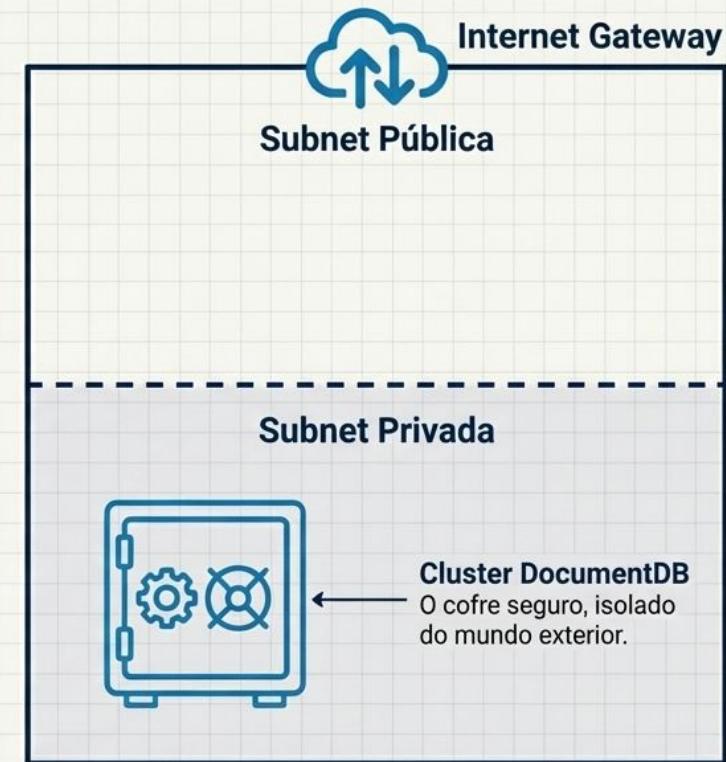
### A Metáfora

O cofre não tem janelas para a rua. Ele reside na sala blindada do subsolo (a Subnet Privada), completamente “escondido” do mundo exterior.

### O Conceito Técnico

O cluster DocumentDB é implantado exclusivamente em uma **Subnet Privada**. Ele não possui um IP público, eliminando um vetor inteiro de ataques da internet.

Essa prática implementa o ‘princípio do menor privilégio’ desde a concepção da arquitetura.



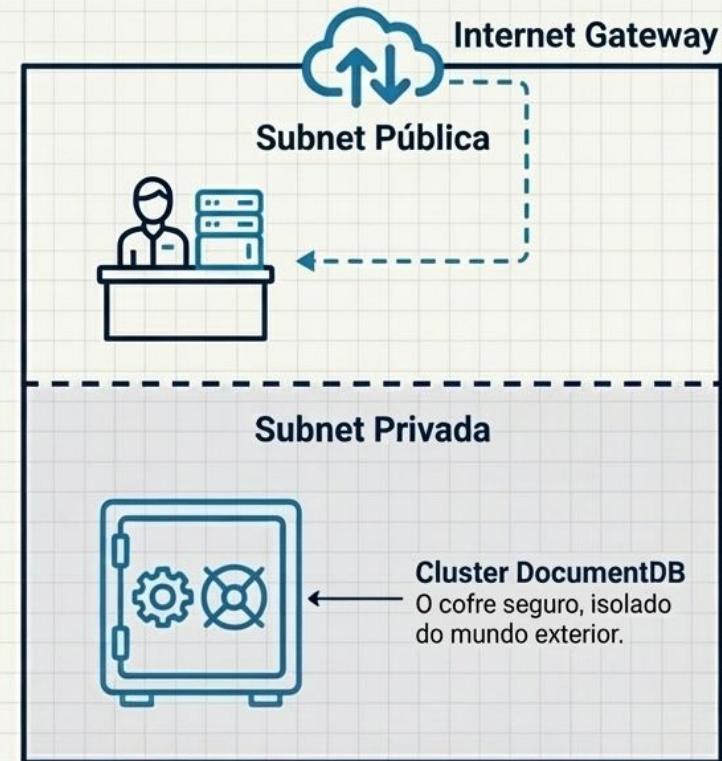
## Passo 4: Os Intermediários (A Aplicação na Subnet Pública)

### A Metáfora

Os Caixas no Saguão. Eles interagem com o público (internet), mas para acessar o dinheiro (dados), eles usam uma porta interna e segura para o subsolo.

### O Conceito Técnico

Uma **instância EC2** (servindo uma aplicação web, por exemplo) é colocada na **Subnet Pública**. Ela pode receber tráfego do **Internet Gateway**, processar requisições e então se comunicar internamente com o DocumentDB na Subnet Privada.



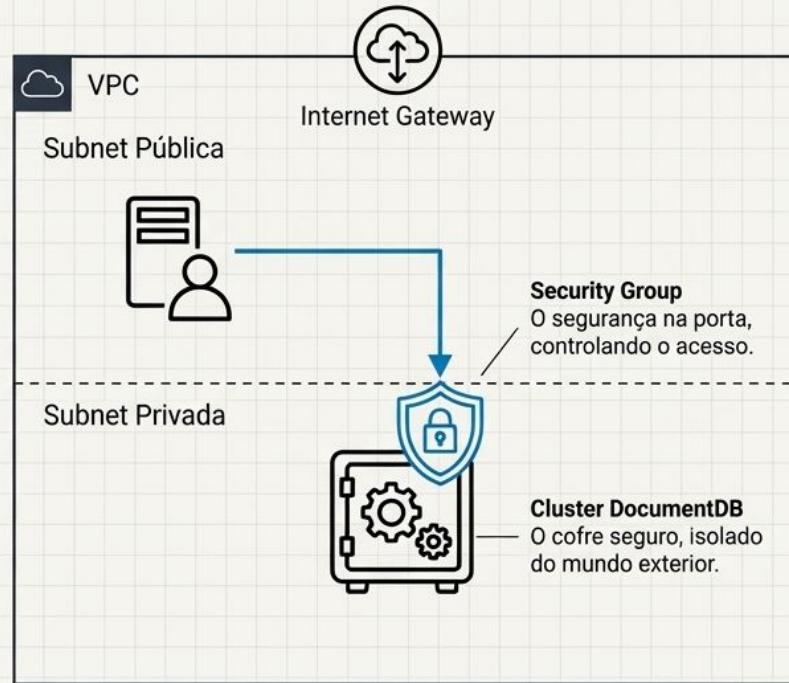
# Passo 5: O Controle de Acesso (O Security Group)

## A Metáfora

O **Segurança na Porta do Cofre**. Ele não deixa qualquer um passar. Ele tem uma lista específica e só libera a entrada para quem tem o crachá correto (a instância EC2 da aplicação) e que está tentando entrar pela porta designada (a porta do banco de dados).

## O Conceito Técnico

O **Security Group** atua como um firewall virtual para o seu cluster DocumentDB. Ele controla o tráfego de entrada. A regra fundamental é: permitir tráfego na porta do DocumentDB (TCP 27017) **APENAS** da fonte específica que é o Security Group da sua instância EC2.



# Um Olhar Detalhado nas Regras do Guarda

Security Groups são *stateful*. Se você permite uma conexão de entrada, o tráfego de retorno correspondente é automaticamente permitido. A configuração incorreta é uma das causas mais comuns de falha de conexão.

## Exemplo de Regra de Entrada (Inbound Rule)

Tipo	Protocolo	Intervalo de Portas	Origem	Descrição
Custom TCP	TCP	27017	`sg-xxxxxxxx` (O ID do Security Group da sua instância EC2)	Permite acesso ao DocumentDB a partir da camada de aplicação.



### Dica do Especialista

No laboratório do curso, esses Security Groups são criados automaticamente para garantir a conexão correta entre a instância EC2 do aluno e o cluster. Validar esses grupos é um passo essencial no **troubleshooting de conexão**.

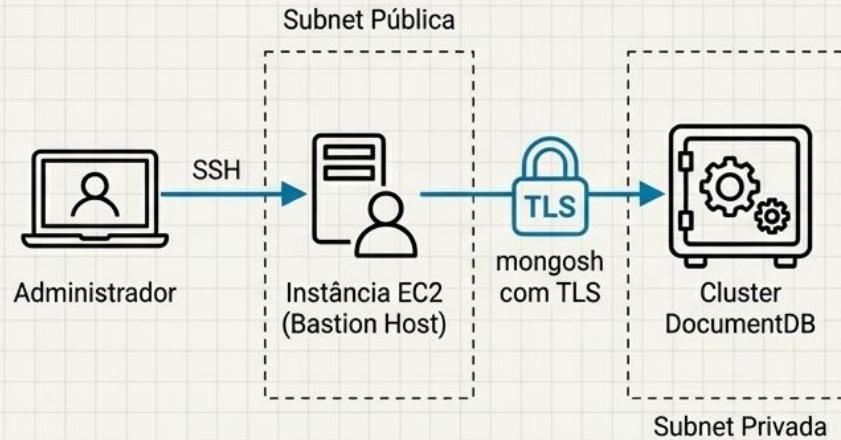
# A Entrada do Administrador: Acesso Seguro via Bastion Host

## A Metáfora

O gerente do banco não usa a porta da frente para acessar o cofre. Ele entra no prédio e usa uma chave mestra e um corredor seguro para chegar ao subsolo.

## O Conceito Técnico

Para acesso administrativo, você se conecta (via SSH) à instância EC2 na Subnet Pública, que funciona como um **Bastion Host** (ou jump box). A partir dela, você usa o **MongoDB Shell** (`mongosh`) para se conectar ao endpoint do cluster DocumentDB. Essa conexão interna exige o uso de **TLS obrigatório**, criptografando os dados em trânsito.



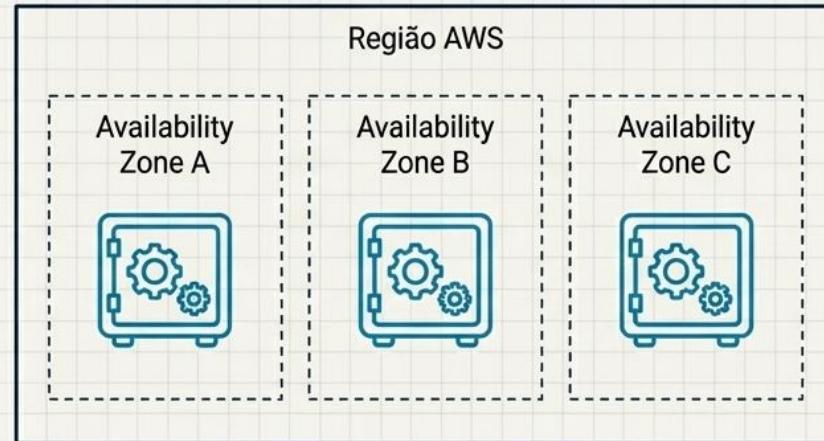
# Plano de Contingência: E se um Prédio Tiver Problemas?

## A Metáfora

Um banco sério não guarda todo o seu dinheiro em um único cofre em um único prédio. Ele possui cofres idênticos em filiais diferentes, geograficamente separadas. Se um prédio for comprometido (por um incêndio, por exemplo), o outro assume as operações instantaneamente.

## O Conceito Técnico

**Arquitetura Multi-AZ (Multi-Availability Zone).** O DocumentDB distribui automaticamente os nós do seu cluster em diferentes **Availability Zones (AZs)** dentro de uma região AWS. As AZs são locais fisicamente distintos com infraestrutura independente.



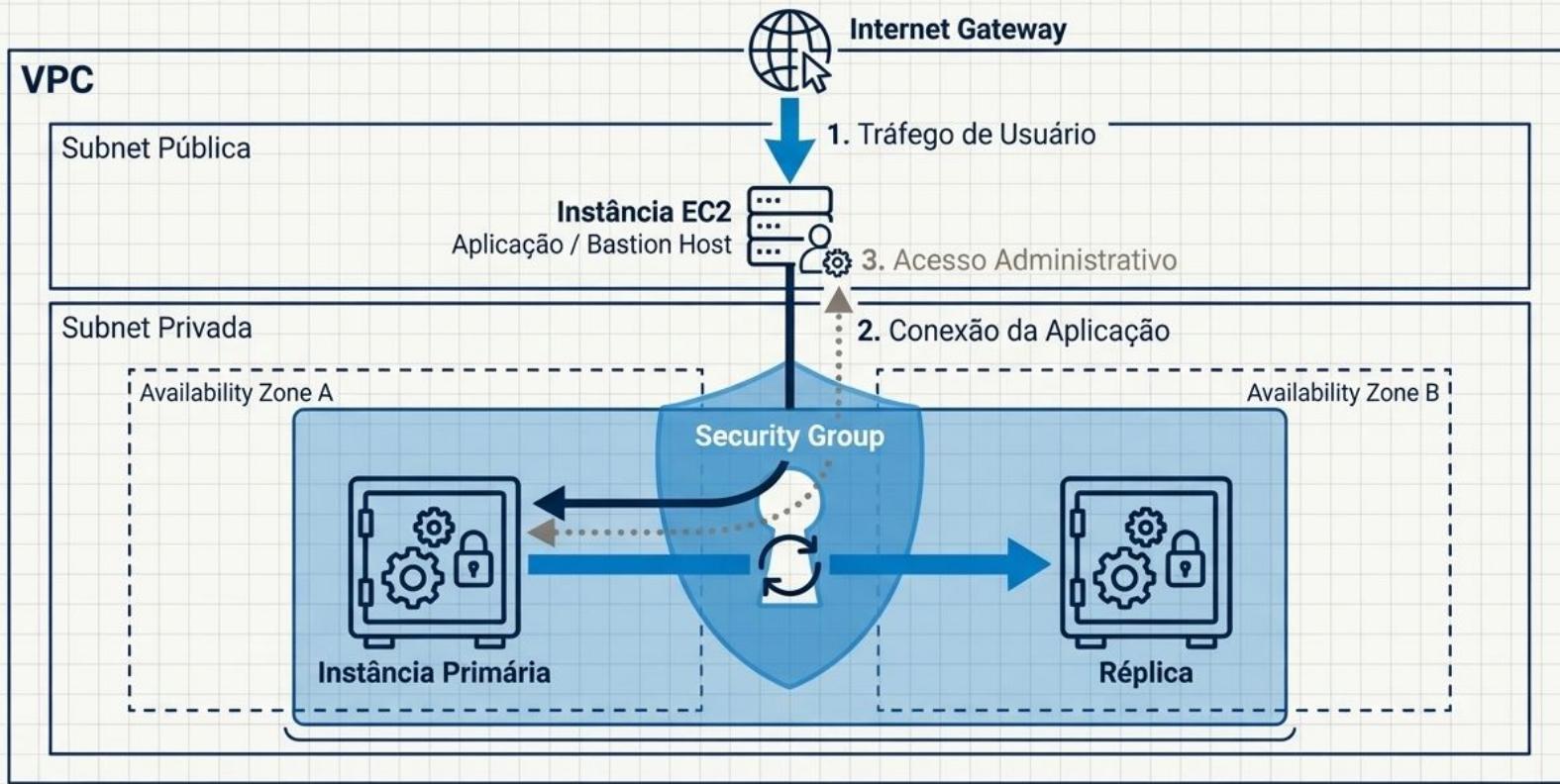
# A Mecânica da Resiliência: Failover Automático

Dentro da arquitetura Multi-AZ, os nós têm **papéis diferentes** para garantir alta disponibilidade para leitura e escrita.



**O Processo de Failover:** Se a AZ do nó primário falhar, o DocumentDB promove automaticamente uma das réplicas em outra AZ para se tornar a nova primária. A aplicação é redirecionada para o novo endpoint, garantindo a **continuidade do negócio** com o mínimo de interrupção.

# A Planta Completa da Fortaleza



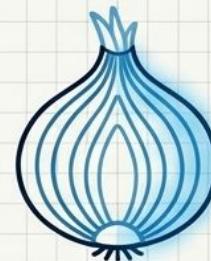
# Os Princípios de Design por Trás da Arquitetura

A arquitetura que construímos não é apenas uma coleção de serviços; é a implementação de princípios de segurança fundamentais.



## Princípio do Menor Privilégio

O cluster não possui IP público. O acesso é negado por padrão e liberado apenas por regras explícitas no Security Group.



## Defesa em Profundidade

Múltiplas camadas de segurança (VPC, Subnets, Security Groups, TLS) garantem que a falha de uma única camada não comprometa o sistema.



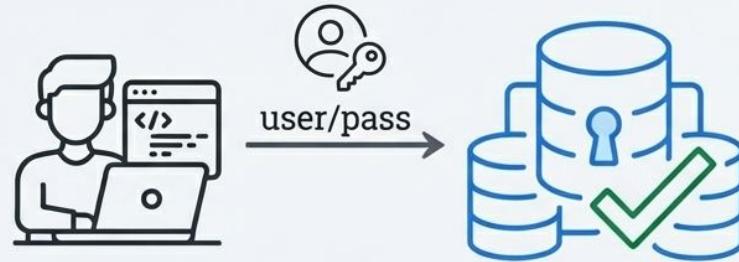
## Alta Disponibilidade por Design

A arquitetura Multi-AZ com failover automático garante que o sistema seja resiliente a falhas de infraestrutura.

# Módulo 3 - Exercício 1: Autenticação Nativa de Banco de Dados

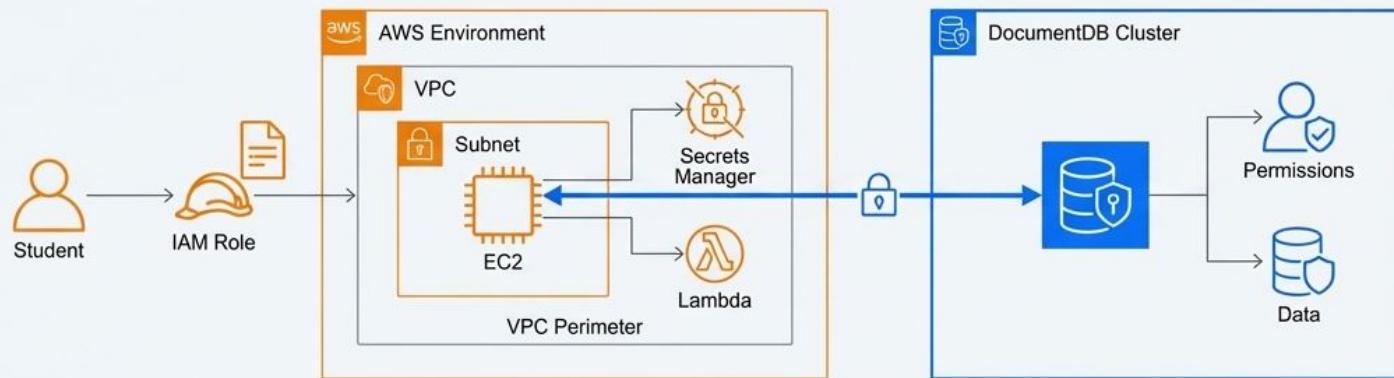
## Como a Autenticação Nativa Funciona na Prática

- **Criação:** Usuários e roles são criados diretamente no cluster do DocumentDB.
- **Acesso:** A conexão ocorre sempre por usuário e senha internos do banco.
- **Validação:** A autenticação é feita pelo próprio DocumentDB, sem nenhuma integração com o IAM.
- **Ferramentas:** Utiliza drivers padrão de mercado e ferramentas como o `mongosh`.



# Módulo 3 - Exercício 1: Autenticação Nativa de Banco de Dados

Autenticação Nativa e IAM não são concorrentes, são parceiros.  
Uma segurança robusta no DocumentDB exige a aplicação de ambas as camadas.



A **Autenticação Nativa** protege os dados *por dentro*, controlando quem pode ver e modificar o quê.



O **IAM** protege o ambiente *por fora*, controlando quem pode provisionar, acessar e gerenciar a infraestrutura.

# Módulo 3 - Exercício 1: Autenticação Nativa de Banco de Dados

Passo 1: Obter Informações de Conexão do Cluster

Passo 2: Baixar Certificado SSL/TLS

Passo 3: Conectar ao Cluster como Usuário Mestre

Passo 4: Criar Base de Dados e Coleções de Teste

Passo 5: Criar Usuários com Diferentes Roles

Passo 6: Testar Autenticação dos Novos Usuários

# Validação

/home/\$ID/Curso-documentDB/modulo3-lab/exercicio1-autenticacao-nativa/grade\_exercicio1.sh

# Módulo 3 - Exercício 2: Integração com VPC, Subnets e Security Groups



# Exercício 2: Integração com VPC, Subnets e Security Groups

Passo 1: Identificar Tipos de Subnets na VPC

Passo 2: Analisar Configuração Atual de Rede

Passo 3: Analisar Security Groups Atuais

Passo 4: Criar Security Groups para Aplicação Cliente

Passo 5: Configurar Regras Restritivas no Security Group do DocumentDB

Passo 6: Testar Conectividade e isolamento

# Validação

/home/\$ID/Curso-documentDB/modulo3-lab/exercicio2-integracao-rede/grade\_exercicio2.sh

# Módulo 3 - Exercício 3: Auditoria



No ambiente AWS, a auditoria de ações de gerenciamento no DocumentDB é realizada primariamente pelo **AWS CloudTrail**. Ele funciona como o sistema de vigilância e registro de eventos para a sua infraestrutura.



**Registra Atividades:** Captura todas as chamadas de API e eventos de gerenciamento no seu cluster.



**Garante Rastreabilidade:** Fornece um histórico detalhado para investigações de segurança e auditorias de conformidade.

# Módulo 3 - Exercício 3: Auditoria

Os logs do CloudTrail são estruturados para responder precisamente às questões essenciais para qualquer auditoria.



## Quem?

(Identifica o usuário, role ou serviço do IAM que fez a chamada).



## O quê?

(Registra a ação específica ou chamada de API executada, como ModifyDBCluster).



## Quando?

(Fornece o timestamp exato do evento).



## De onde?

(Captura o endereço IP de origem da requisição).

# Módulo 3 - Exercício 3: Auditoria

Os logs do CloudTrail são estruturados para responder precisamente às questões essenciais para qualquer auditoria.



## Quem?

(Identifica o usuário, role ou serviço do IAM que fez a chamada).



## O quê?

(Registra a ação específica ou chamada de API executada, como `ModifyDBCluster`).



## Quando?

(Fornece o timestamp exato do evento).



## De onde?

(Captura o endereço IP de origem da requisição).

# Módulo 3 - Exercício 3: Auditoria

A auditoria transcende o compliance. É a base da governança proativa e da segurança operacional.



- **Visão Integrada:** A segurança do DocumentDB depende da tríade inseparável de Autenticação, Rede e Auditoria.



- **Ferramenta Essencial:** AWS CloudTrail é o mecanismo central para registrar todas as ações e garantir a rastreabilidade completa.



- **Valor Estratégico:** Logs de auditoria são cruciais não apenas para conformidade, mas para investigações de segurança, análise de eventos críticos e troubleshooting ágil.

# Módulo 3 - Exercício 3: Auditoria

Passo 1: Habilitar a Auditoria de Eventos do DocumentDB

Passo 2: Exportar Logs para o CloudWatch Logs

Passo 3: Gerar Atividade no Banco para Produzir Logs

Passo 4: Inspecionar os Logs no CloudWatch

# Módulo 4 - Performance e Tuning do DocumentDB

# Objetivos

- Implementar monitoramento avançado de performance com métricas personalizadas
- Analisar planos de execução e otimizar índices suportados pelo DocumentDB
- Identificar gargalos de performance através de análise detalhada
- Aplicar estratégias de indexação para diferentes tipos de queries

# Você aprendeu a dirigir. Agora, vamos abrir o capô.

Nos módulos anteriores, você dominou os fundamentos do DocumentDB. No **Módulo 4**, focamos em um único objetivo: extrair a máxima eficiência de cada consulta. Não se trata apenas de fazer funcionar, mas de fazer voar.

“O Módulo 4 do curso é focado exclusivamente em Performance e Tuning, com o objetivo de transformar a maneira como o banco de dados processa informações para garantir a máxima eficiência.”



# O Kit de Ferramentas do Especialista em Performance

Para alcançar a alta performance, você precisa das ferramentas certas. O Módulo 4 equipa você com um arsenal completo para diagnosticar e resolver qualquer gargalo.



## Análise de Métricas Customizadas

Entenda o comportamento do cluster sob carga.



## Otimização de Índices e Queries

Refine consultas complexas para reduzir latência.



## Troubleshooting Avançado

Identifique e corrija gargalos em ambientes reais.



## Domínio do Plano de Execução

A ferramenta essencial para entender exatamente como o DocumentDB processa suas queries.

# Afinal, o que é o Plano de Execução?

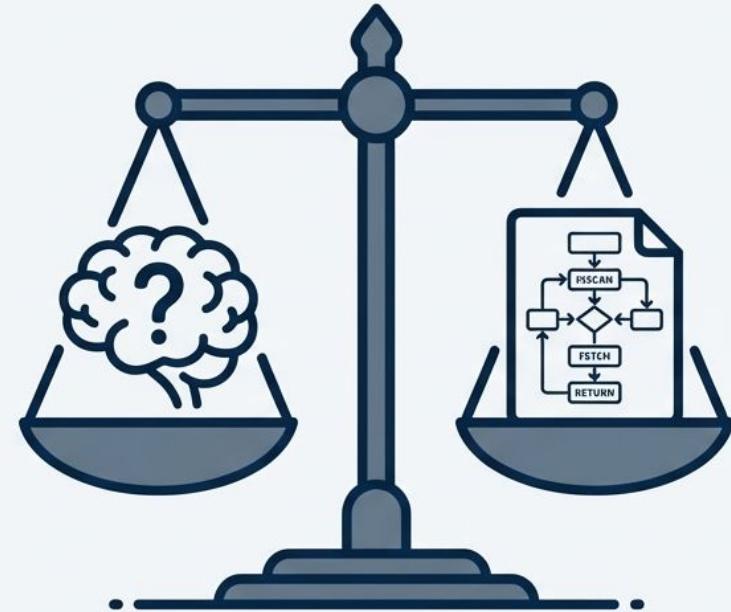
É o mapa passo a passo, o GPS interno que o DocumentDB cria para encontrar e retornar os dados que você solicitou. Ele detalha cada etapa da jornada da sua query.



# Sua Única Fonte da Verdade

Esqueça as suposições. O plano de execução não mente. Ele mostra o que o banco de dados *\*realmente\** fez, não o que você *\*achava\** que ele faria.

- Revela o “Como”:** Mostra a estratégia exata usada (ex: usou um índice? escaneou a coleção inteira?).
- Identifica Ineficiências:** Aponta com precisão onde o tempo está sendo gasto.
- Valida suas Otimizações:** Permite que você veja o impacto real da criação de um novo índice.



# O Caminho da Força Bruta: **COLLSCAN**

Collection Scan

O COLLSCAN ocorre quando o DocumentDB não tem um índice para ajudar na busca. Para encontrar os documentos, ele é forçado a inspecionar cada um dos documentos da sua coleção, um por um.

## A Biblioteca Sem Catálogo

Imagine um bibliotecário que precisa encontrar uma única citação em uma biblioteca gigantesca, mas não há um índice de livros. A única opção é pegar cada livro da prateleira, abrir e folhear página por página.



○ Extremamente Lento

○ Alto Custo de CPU e I/O

○ Não escala com o crescimento dos dados

# A Rota Inteligente: IXSCAN

## Index Scan

O IXSCAN é o resultado de uma query que utiliza um índice. O banco de dados usa o índice como um **atalho para ir diretamente** aos documentos que correspondem aos critérios da busca, ignorando todos os outros.

### A Biblioteca Organizada

O mesmo bibliotecário, na mesma biblioteca, mas agora com um **catálogo de fichas** (o índice). Ele consulta o **catálogo**, encontra a **localização exata do livro e da página**, e vai **direto ao ponto**.



Extremamente Rápido

Baixo Custo de CPU e I/O

Excelente Escalabilidade

# `COLLSCAN` vs. `IXSCAN`: A Batalha pela Performance

## COLLSCAN



- **Método:** Força Bruta (Lê toda a coleção)
- **Velocidade:** Lenta
- **Custo de Recursos:** Alto
- **Escalabilidade:** Ruim (Piora drasticamente com mais dados)

## IXSCAN



- **Método:** Acesso Direto (Usa um índice)
- **Velocidade:** Rápida
- **Custo de Recursos:** Baixo
- **Escalabilidade:** Excelente (Mantém a performance com mais dados)

A sua principal missão em otimização de queries é transformar um `COLLSCAN` em um `IXSCAN`.

# Estudo de Caso: O Diagnóstico da Query Lenta

Escenário: Uma consulta simples para buscar pedidos com status "processando" em uma coleção com milhões de documentos.

## A Query

```
db.orders.find({ "status": "processando" })
```

## O Plano de Execução (Simplificado)

```
{
  "queryPlanner": {
    "winningPlan": {
      "stage": "COLLSCAN",
      "direction": "forward"
    }
  }
}
```

## Resultado

**Tempo de Execução:** 2.850 ms

**Documentos Examinados:** 5.000.000

**Diagnóstico:** `COLLSCAN` identificado. A query está lendo toda a coleção.

# A Correção: Uma Linha de Comando

Com o plano de execução apontando um `COLLSCAN` no campo `status`, a solução é clara: precisamos de um índice nesse campo.



## A Solução

```
// Criando o índice no campo 'status'
```

```
db.orders.createIndex({ "status": 1 })
```

# O Resultado: Performance Transformada

Executando a mesma query após a criação do índice.

## A Query (Inalterada)

```
db.orders.find({  
  "status": "processando"  
})
```

## O Novo Plano de Execução

```
{  
  "queryPlanner": {  
    "winningPlan": {  
      "stage": "IXSCAN",  
      "keysExamined": 1250  
    }  
  }  
}
```

### Novo Resultado

Tempo de Execução: 3 ms

Documentos Examinados: 1.250 (apenas os relevantes)

**99.9%**  
**MAIS**  
**RÁPIDO**

# Principais Lições

**1.**

**O Plano de Execução é a fonte da verdade.**

Sempre comece sua análise por ele.

**2.**

**`COLLSCAN` é o inimigo da escalabilidade.**

Evite-o a todo custo. Ele é um sinal de alerta de que um índice está faltando.

**3.**

**Índices são a chave para a performance (**IXSCAN**).**

A otimização mais eficaz começa com a estratégia de indexação correta.

# Módulo 4



## Análise de Métricas

Criação de métricas personalizadas, análise do comportamento do cluster, identificação de padrões sob carga.



## Planos de Execução

Interpretação de planos de execução, entendimento do processamento interno de consultas, diagnóstico de ineficiências.



## Otimização Prática

Otimização de índices, reescrita e refinamento de queries complexas, redução de latência e consumo de I/O.



## Troubleshooting

Identificação de gargalos em tempo real, aplicação de correções, estratégias para ambientes de produção.

# Módulo 4 - Exercício 1: Métricas Avançadas e Monitoramento de Performance

Passo 1: Preparar Ambiente

Passo 2: Script de Demonstração de Métricas

Passo 3: Dashboard de Performance Avançado

Passo 4: Conceitos de Performance Baseline

Passo 5: Conceitos de Monitoramento Contínuo

Passo 6: Validação e Testes

# Validação

/home/\$ID/Curso-documentDB/modulo4-lab/exercicio1-metricas-avancadas/grade\_exercicio1.sh

# Limpeza

Passo 1: Deletar dashboard

# Módulo 4 - Exercício 2: Análise de Planos de Execução e Otimização de Índices

Passo 1: Preparação do Ambiente e Dados de Teste

Passo 2: Análise de Planos de Execução Básicos

Passo 3: Estratégias de Indexação Avançadas

Passo 4: Análise Avançada com explain()

Passo 5: Otimização Baseada em Análise

Passo 6: Monitoramento de Performance de Índices

Passo 7: Validação das Otimizações

# Validação

/home/\$ID/Curso-documentDB/modulo4-lab/exercicio1-metricas-avancadas/grade\_exercicio1.sh

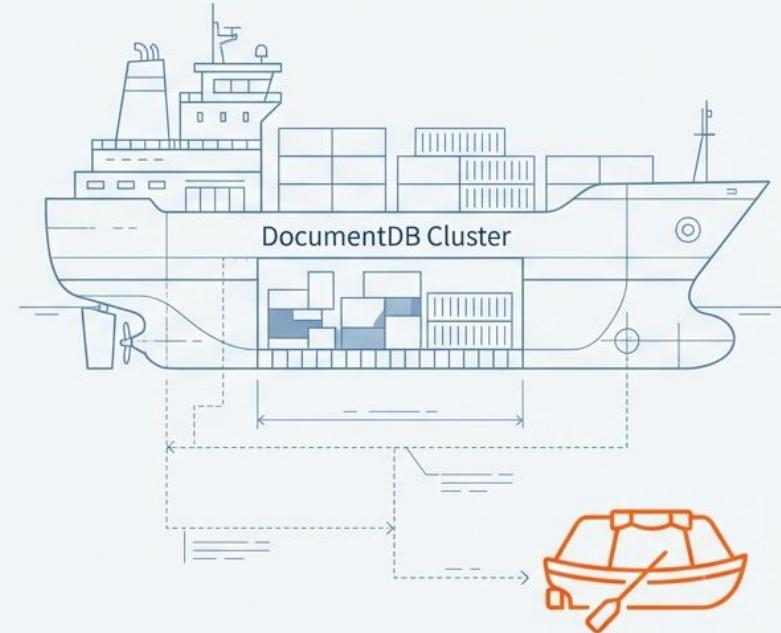
# Módulo 5 - Backup e Exportação de Dados

# Objetivos

- Implementar backup completo e incremental do DocumentDB para S3
- Configurar políticas de retenção no S3 para compliance básico
- Testar procedimentos de restore e validação de integridade
- Estabelecer rotinas manuais de backup operacional

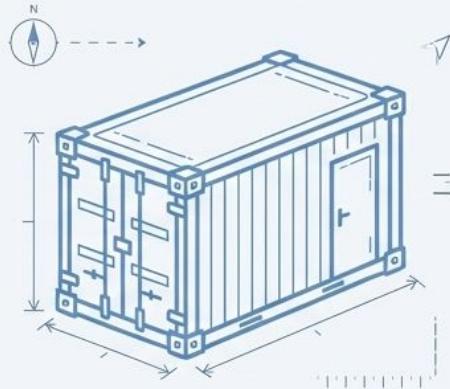
# Módulo 5: Seu Plano de Resiliência para DocumentDB

Garantindo a continuidade da sua operação quando a navegação não sai como o esperado.



*Este módulo não é sobre backups. É sobre a certeza de que, em uma emergência, você tem um plano de evacuação testado e botes salva-vidas prontos para zarpar.*

# Construindo seu Plano de Resiliência em 3 Pilares Fundamentais



## Estratégias de Backup e Exportação

Onde e como armazenar com segurança seus dados mais valiosos.



## Restore e Validação

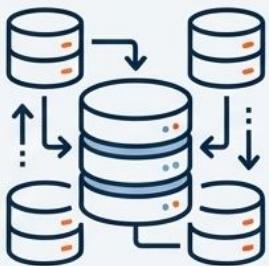
Como garantir que seu plano de resgate realmente funciona sob pressão.



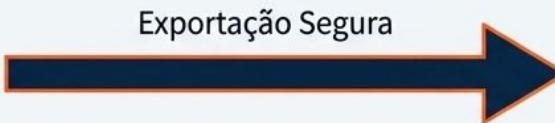
## Ferramentas e Integração

O arsenal tecnológico à sua disposição para executar o plano com precisão.

# Amazon S3: Seu Porto Seguro para Arquivamento de Dados



DocumentDB Cluster



Amazon S3:  
Armazenamento de Objetos

O módulo ensina o processo de exportar dados do DocumentDB diretamente para o **Amazon S3**. Pense no S3 como um cofre externo, imune a qualquer problema que possa ocorrer com o “navio” principal (seu cluster). É a base para qualquer estratégia de arquivamento, conformidade ou recuperação.

- ✓ Exportação direta do cluster para o S3.
- ✓ Garante a durabilidade e a disponibilidade dos dados fora do ambiente primário.
- ✓ Fundamental para cenários de arquivamento de longo prazo e conformidade.

# O Processo de Restore e a Validação da Integridade dos Dados

Ter um backup não é o suficiente. Você precisa da certeza absoluta de que pode usá-lo. O módulo cobre os procedimentos técnicos para restaurar os dados e, mais importante, ensina a realizar **testes de restore** para validar a integridade dos dados recuperados.

**Key Question Answered:** Como garantir que o banco de dados volte exatamente ao estado esperado após uma falha crítica?



- Executar o procedimento técnico de restore a partir de um backup.
- Conduzir testes de validação para verificar a integridade e consistência dos dados.
- Assegurar que a aplicação se conecta e opera normalmente com o banco de dados restaurado.

# De Testes de Restore a uma Estratégia Completa de Disaster Recovery



As habilidades aprendidas neste pilar são a base para um dos objetivos mais críticos de qualquer operação: "Implementar alta disponibilidade e **disaster recovery**". Não se trata apenas de recuperar dados, mas de garantir a continuidade do negócio em face de uma falha catastrófica. Este módulo fornece o conhecimento prático para cumprir essa promessa.

# **Resiliência Não é Reagir ao Desastre. É Navegar com a Certeza da Preparação.**

O verdadeiro objetivo deste módulo é mudar a mentalidade de reativa para proativa. É a diferença entre esperar a tempestade e ter a confiança de que seu navio, sua tripulação e sua carga preciosa estão seguros, não importa o que o oceano reserve.



# Módulo 5 - Exercício 1: Backup de Dados para S3

Passo 1: Configuração do Ambiente de Backup

Passo 2: Criar Bucket S3 para Backups

Passo 3: Baixar Certificado SSL e Verificar Ferramentas

Passo 4: Backup Completo (Full Backup)

Passo 5: Comprimir e Enviar para S3

Passo 6: Backup Incremental

Passo 7: Restore e Validação

# Validação

/home/\$ID/Curso-documentDB/modulo5-lab/exercicio-backup-s3/grade\_exercicio\_backup.sh