



Olá, aluno(a)!
Seja bem-vindo(a) à aula interativa!

Você entrará na reunião com a câmera e o microfone desligados.

Sua presença será computada através da enquete.
Fique atento(a) e não deixe de respondê-la!

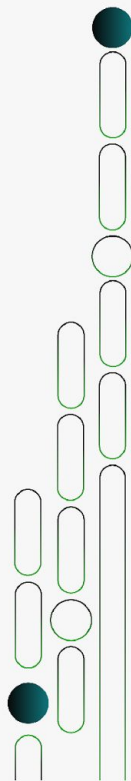


Aprenda com quem faz

Performance & Otimização

Primeira Aula Interativa

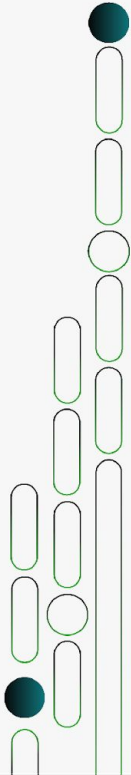
Prof. Adriano G. D. Ferreira





Nesta Aula

- ❑ Sistemas Distribuídos X Sistemas Centralizados
- ❑ Performance x Banco de Dados Distribuído.
- ❑ Arquitetura Distribuída
- ❑ SPOF (Ponto único de Falha).
- ❑ Replica Set x Sharding no MongoDB.
- ❑ Sobre o Trabalho Prático.





Sobre o Professor:

Atuação:

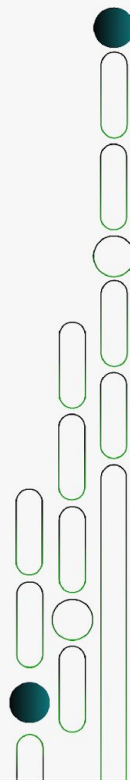
- DBA , Arquiteto de dados e Analista de Dados;
- Professor XP Educação;
- Professor Pitágoras;

Formação:

- Mestre em Administração Estratégica com foco em TI/BI;
- Pós graduado em Banco de Dados com foco SQL Server;
- Graduação em Sistemas de Informação;
- MBA em Ciência de Dados.
- Cursando I2A2 e Pós em Analytics e Ciência de Dados

Empresas:

- Teksid do Brasil (Grupo Fiat);MRV Engenharia;Liga Sistemas; Patrus Transportes; Callink ; Policard; Cedro Technologies, Woli e CloudOpss.





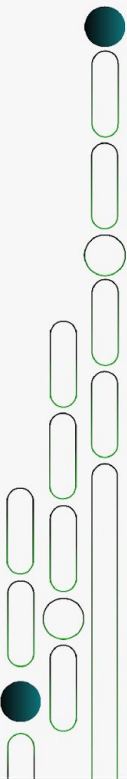
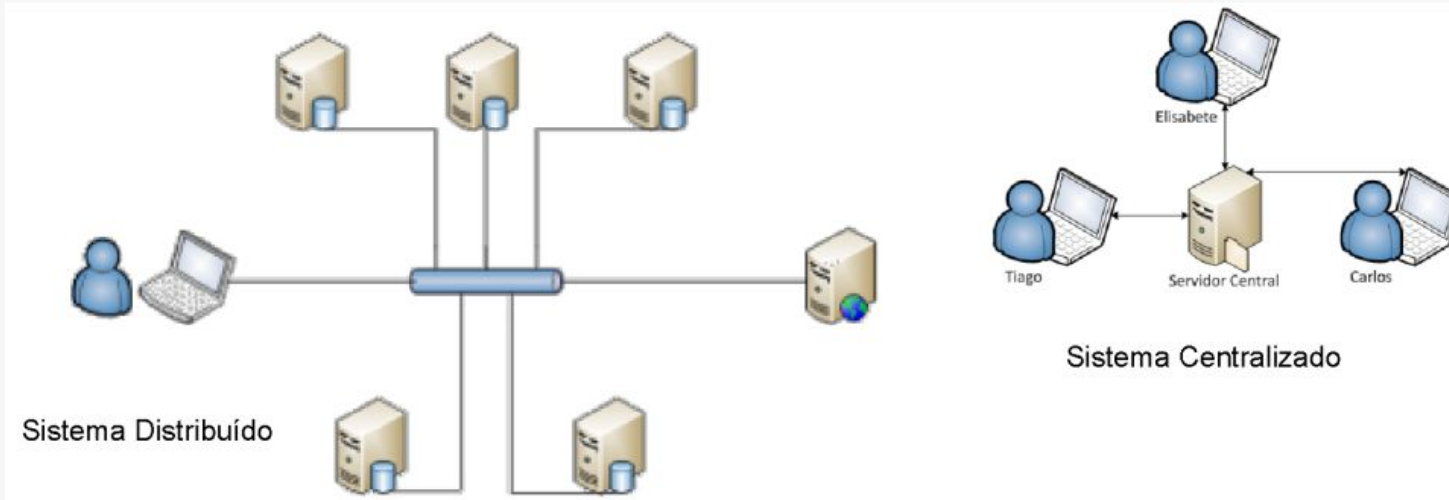
Sobre o Professor: Artigos:



<http://www.devmedia.com.br/space/adriano-geraldo-dias-ferreira>

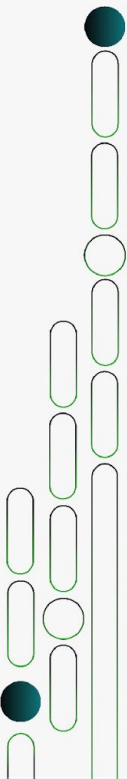
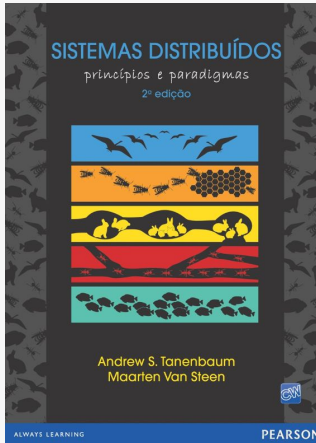


Sistemas Distribuídos X Sistemas Centralizados



Sistemas Distribuídos

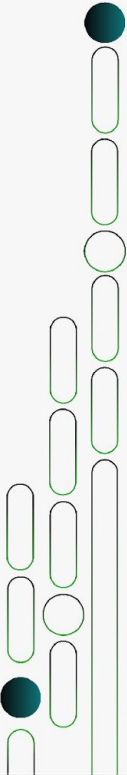
Um **sistema distribuído**, segundo a definição de **Andrew Tanenbaum**, é uma "coleção de computadores independentes entre si que se apresenta ao usuário como um sistema único e coerente"; outra definição, de **George Coulouris** diz: "coleção de computadores autônomos interligados através de uma rede de computadores e equipados com software que permita o compartilhamento dos recursos do sistema: *hardware*, *software* e dados".



Sistemas Distribuídos: Resposta:

1.A primeira e mais importante MOTIVAÇÃO para a utilização de sistemas distribuídos esta relacionado diretamente com a:

- a. Partilha de Processos;
- b. Partilha de Recursos;
- c. Partilha de Componentes;
- d. Partilha de Hardware;
- e. Partilha de Middleware.

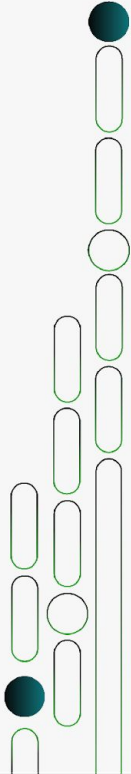




Sistemas Distribuídos: Resposta:

2.A implantação de um sistema distribuído enfrenta muitos desafios, são eles: Abertura, Segurança, Transparência, Escalabilidade, Tolerância a falhas e Concorrência e Heterogeneidade. Cada um desses desafios apresenta obstáculos com várias características pelas quais podemos identificá-los. Custos, Perda de desempenho, Prevenção à falta de recursos e Gargalos no sistema são características de um dos obstáculos enfrentados pelos sistemas distribuídos. Identifique a alternativa que representa o obstáculo que mais se enquadra nessas características.

- a) Segurança.
- b) Transparência.
- c) Escalabilidade.
- d) Concorrência.
- e) Heterogeneidade.



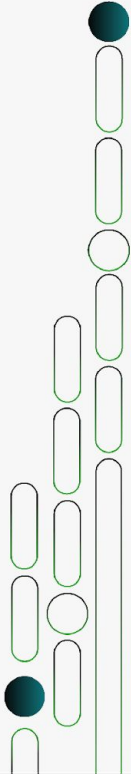


Sistemas Distribuídos: Resposta:

3.Quando planejamos implantar um sistema distribuído, há vários fatores importantes que devem ser levados em consideração. Podemos chamá-los de aspectos de projeto, e, no projeto de sistemas distribuídos, os principais são: Segurança, Escalabilidade, Resiliência e Heterogeneidade.

A capacidade de máquinas com diferentes sistemas operacionais se comunicarem na execução de um sistema se relaciona a qual aspecto de projeto?

- a) Segurança.
- b) Escalabilidade.
- c) Resiliência.
- d) Heterogeneidade.
- e) Transparência.





Sistemas Distribuídos: Respostas:



1. Partilha de Recursos.

2. Alternativa C.

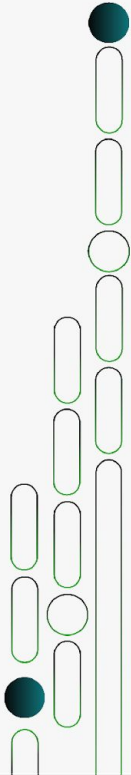
Resposta Comentada:

A resposta correta seria 'Escalabilidade', onde os principais obstáculos encontrados ao aplicar escalabilidade em um sistema distribuído seriam os custos, perda de desempenho, prevenção à falta de recursos e gargalos encontrados nos sistemas.

3. Alternativa D

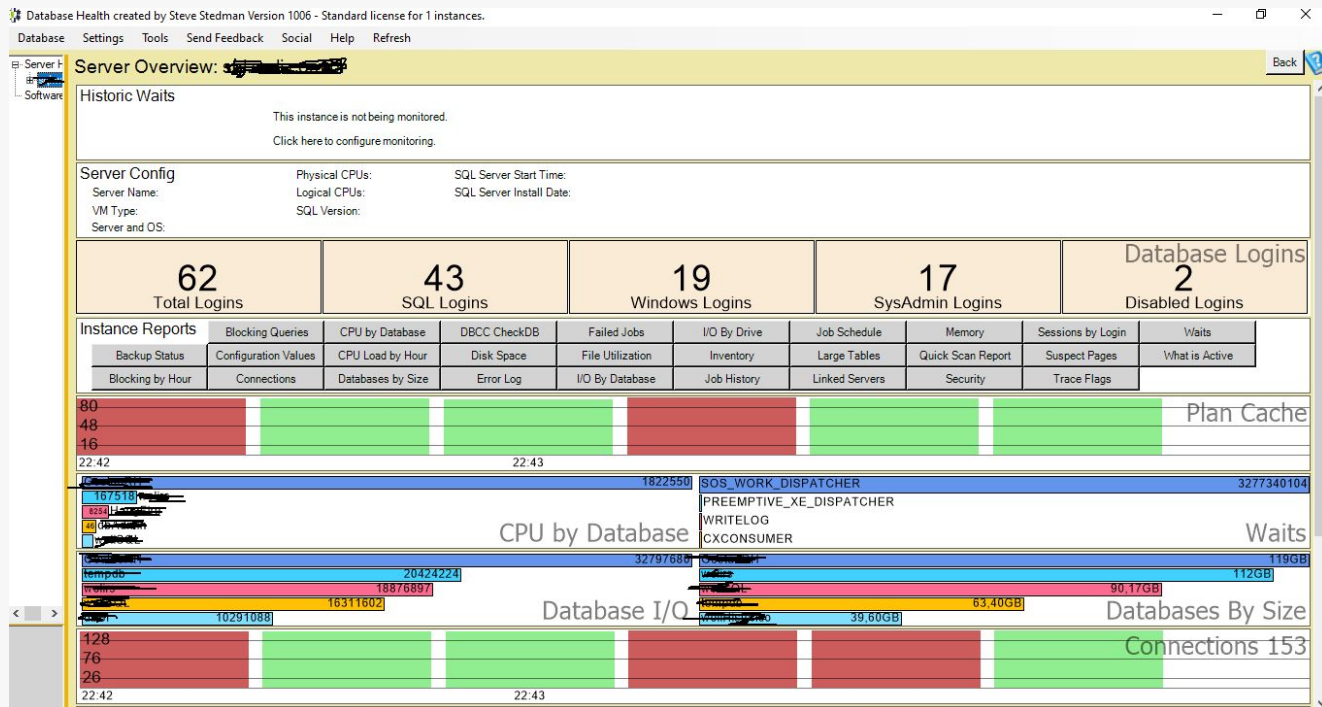
Resposta Comentada:

A resposta correta seria 'Heterogeneidade', pois este aspecto está relacionado à comunicação entre máquinas com diferentes sistemas operacionais e recursos de software para execução de um sistema. Esse aspecto só pode ser atingido com a ajuda da camada de comunicação entre as máquinas (middleware).





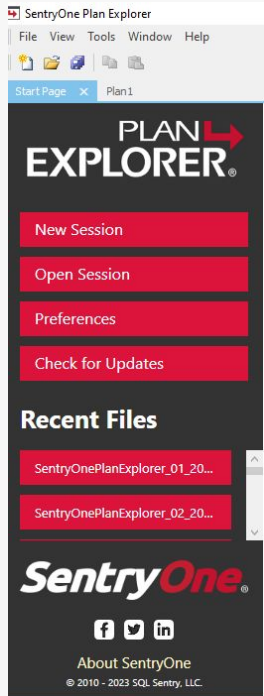
DBs Relacionais: Performance



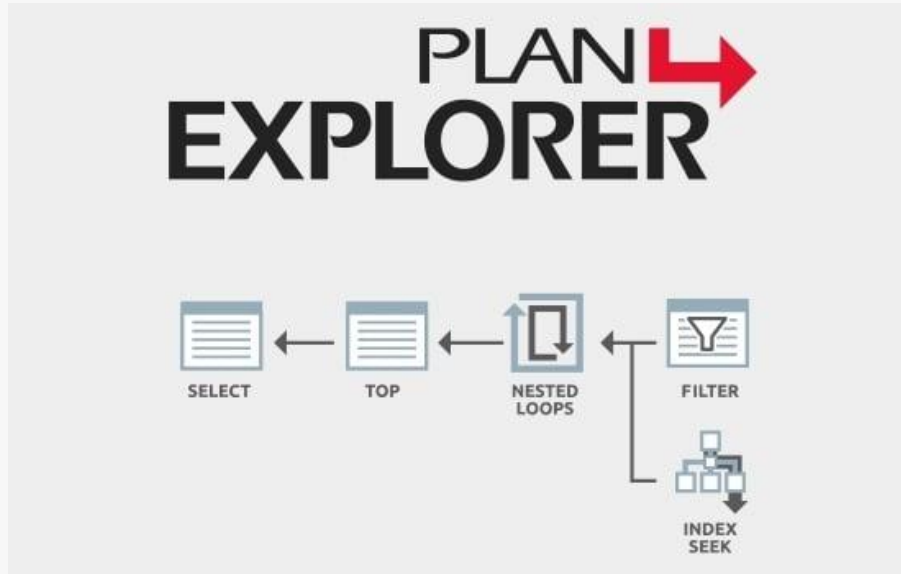
Fonte: <http://databasehealth.com/>



Db's Relacionais: Performance



- Integrado ao Plano de Execução (Execution Plan) do Sql Server.



<https://www.sentryone.com/plan-explorer>



DBs Relacionais: Dicas para otimizar Plano de Query:

Revisar e otimizar os JOINS

- Nested loops apenas para conjuntos pequenos de dados.

Rever operadores SCANS e RID

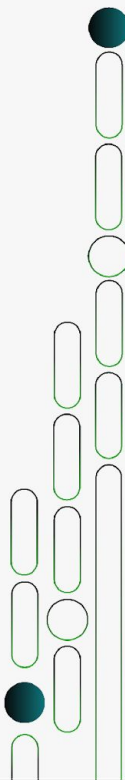
- Removê-los usando índices que contemplem toda a query.

Rever operadores Key Lookups

- Removê-los usando índices que contemplem toda a query.

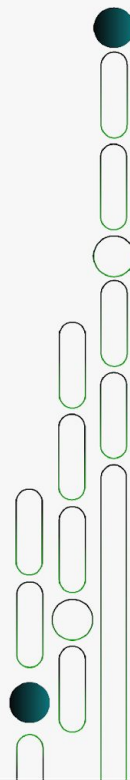
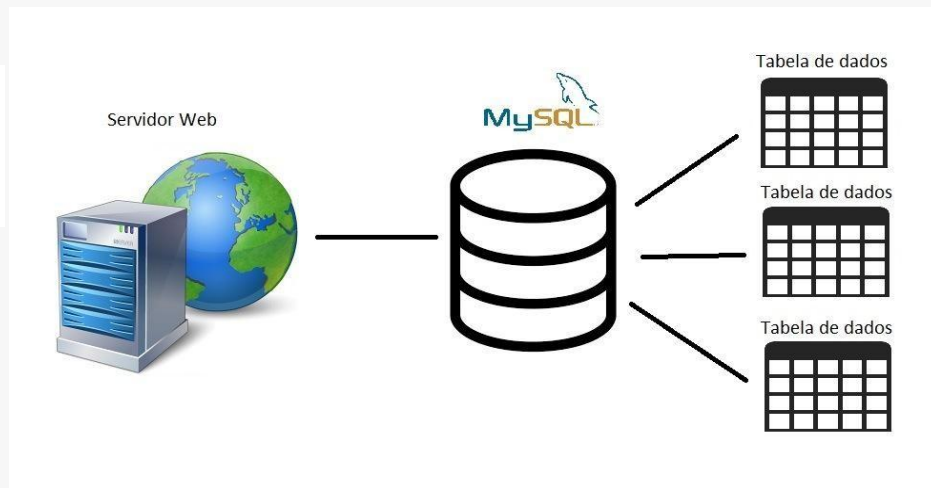
Outras considerações

- Eliminar spools.



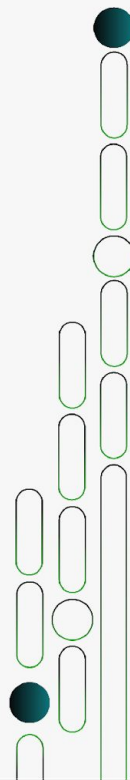
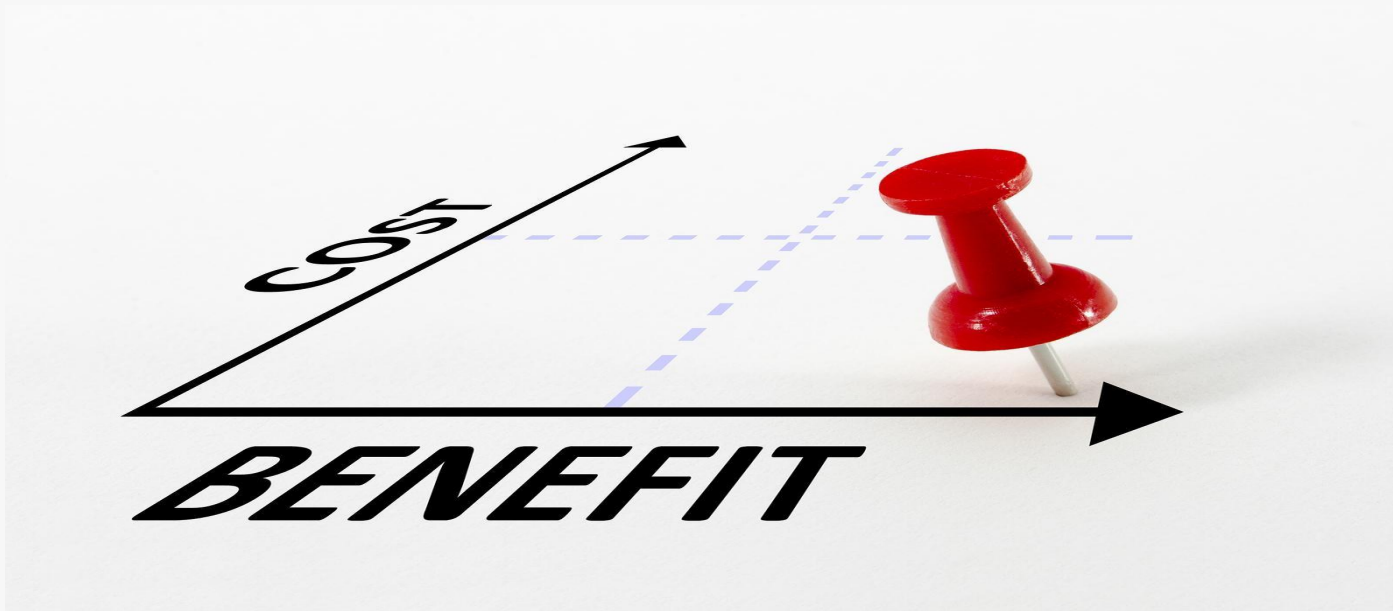


Performance X BD Distribuídos



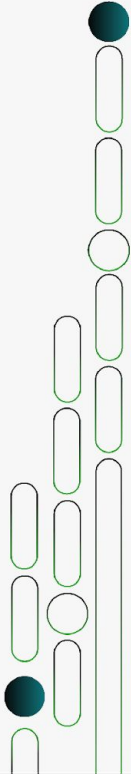
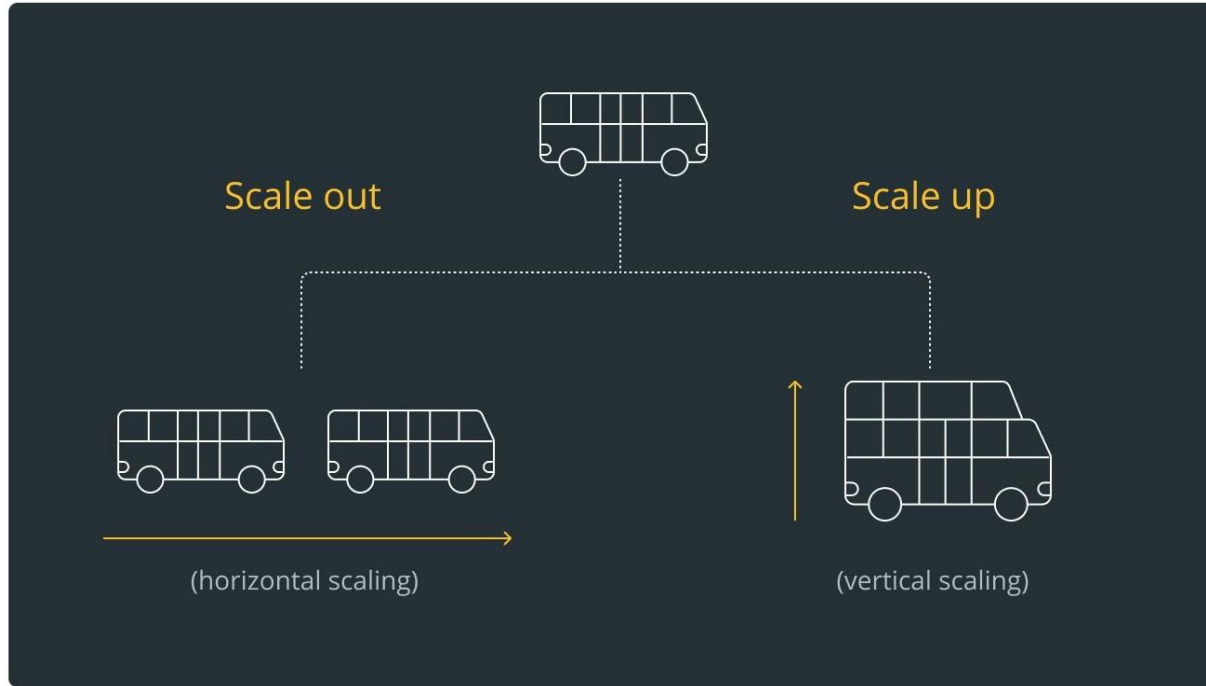


Performance X DB Distribuído





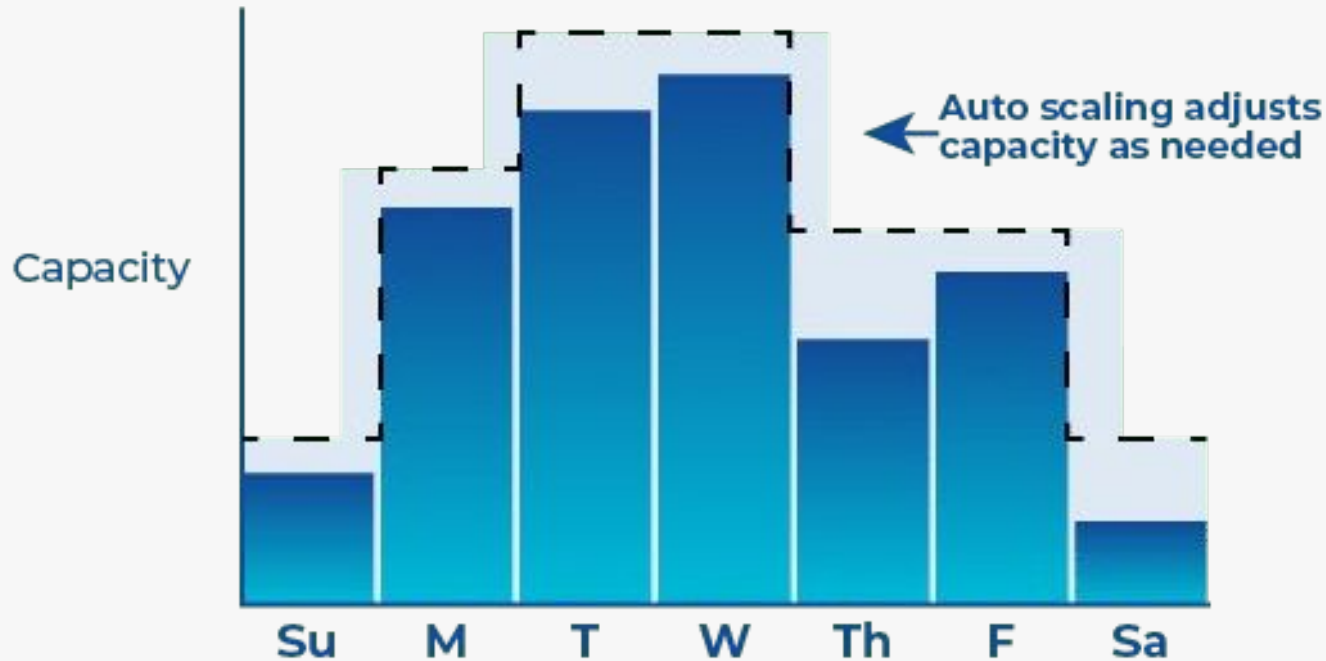
Horizontal X Vertical





Computação Elástica

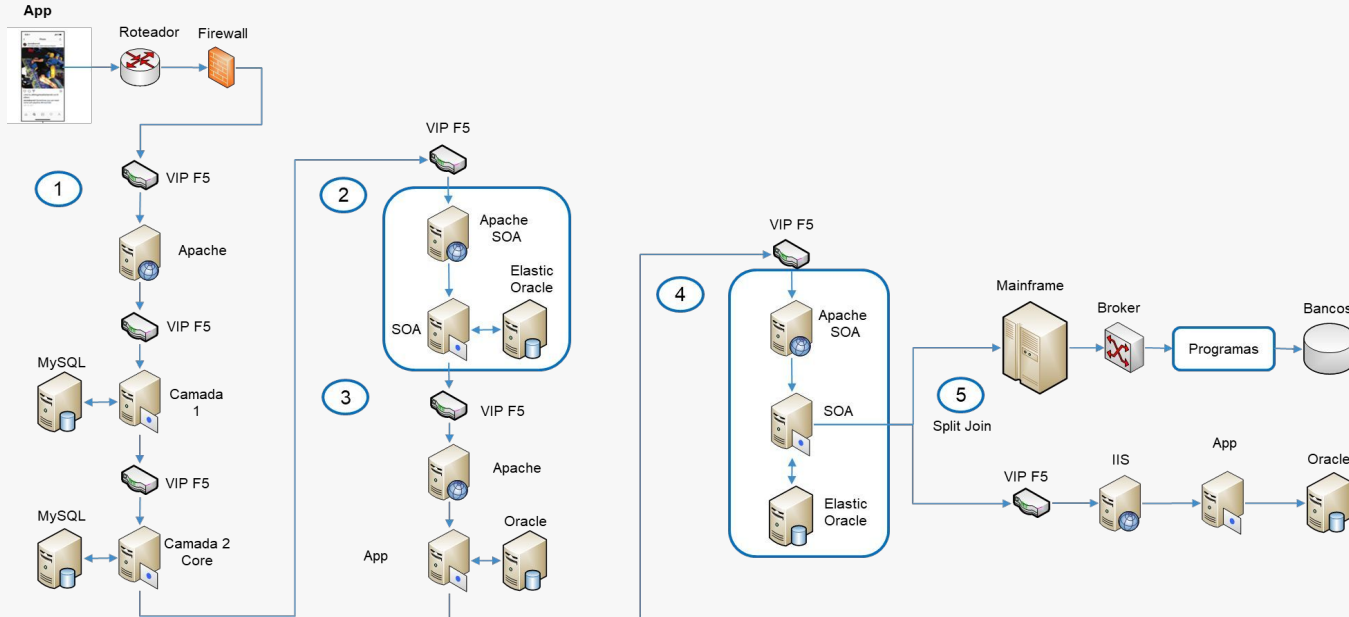
- Para atender às exigências dinâmicas, sem se preocupar com o planejamento de capacidade e picos de utilização.

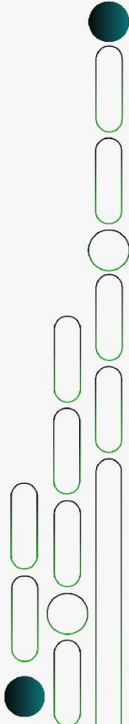




Performance X DB Distribuído

- **Heterogeneidade:** “Uma faca de 2 gumes” □ cuidado com o acúmulo de tecnologias.





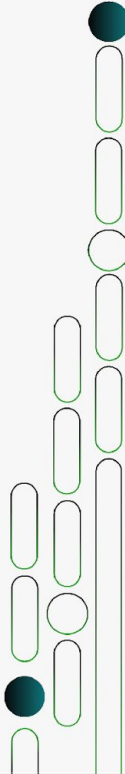
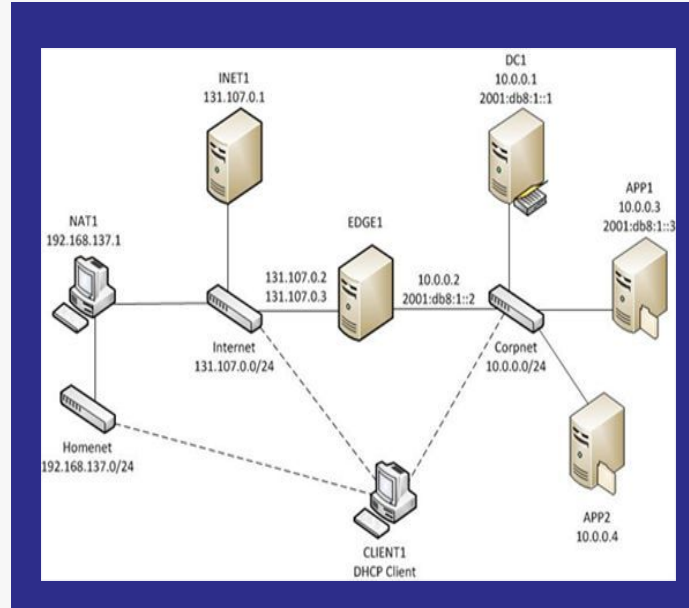
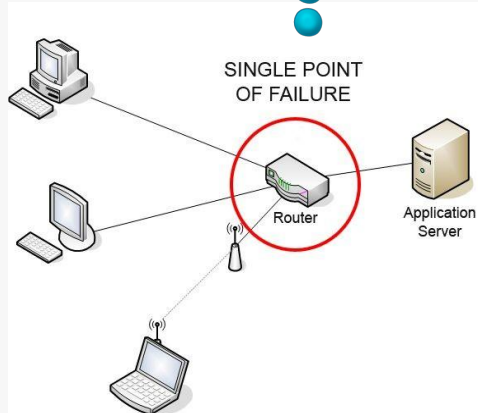
<https://neo4j.com/>



SPOF:

▪ Single Point Of Failure: ponto único de falha.

- Para designar um local num sistema que, caso falhe, provoca a falha de todo o sistema.
- É possível eliminar?
- 100%



SGBD e SGBDD

Banco de dados
centralizado / distribuído
Multiplataforma
Escrito em C++



Lançado em 02/2009

Inicialmente 10gen.
Atualmente ©MongoDB Inc.



www.mongodb.com

NOSQL

"Not Only SQL"

Não relacional

Orientado a Documento



{JSON}

Esquema Dinâmico
Document Database
Formato *JSON* (field-value)
Armazenado como *BSON*

Gratuito e Open Source

Edição *Community Server*
*Edição *Enterprise*
(subscrição com suporte)

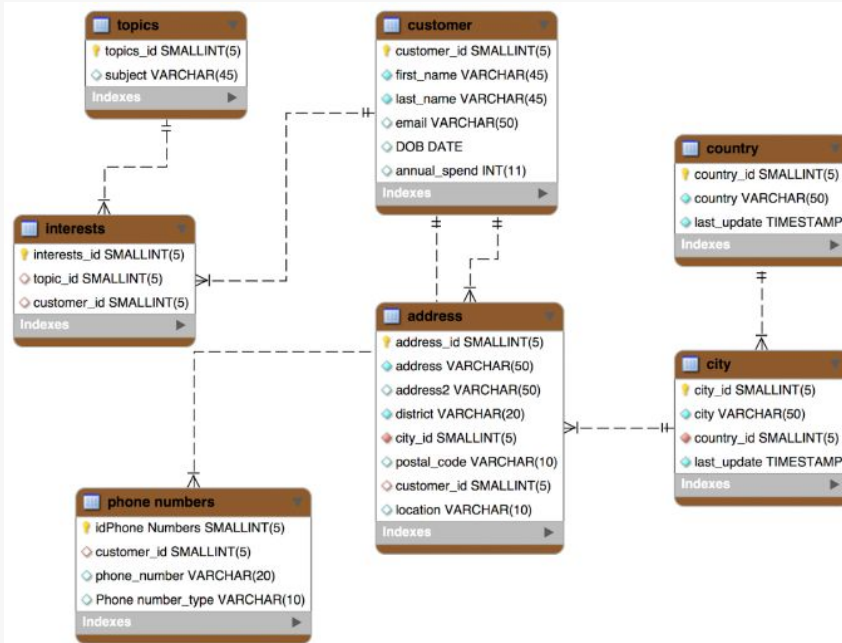


Universidade MongoDB

Cursos gratuitos e online, em
inglês, com certificado
<https://university.mongodb.com>



Representação Tabular

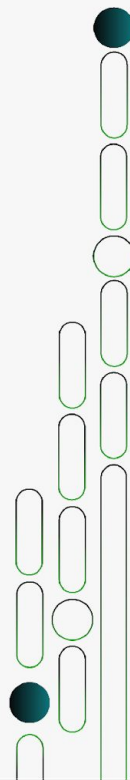
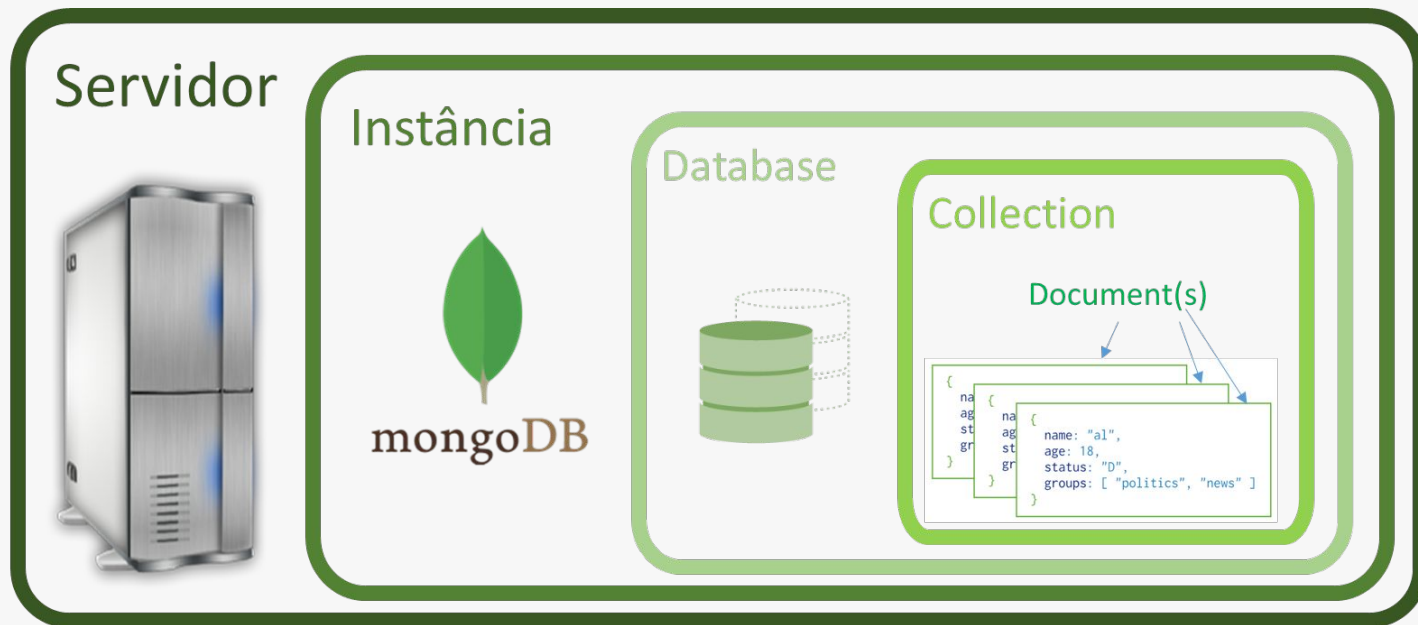


Representação JSON

```
{
  "_id":
    ObjectId("5ad88534e3632e1a35a58d00"),
  "name": {
    "first": "John",
    "last": "Doe" },
  "address": [
    { "location": "work",
      "address": {
        "street": "16 Hatfields",
        "city": "London",
        "postal_code": "SE1 8DJ"},
      "geo": { "type": "Point", "coord": [
        51.5065752,-0.109081]}},
    { ... }
  ],
  "phone": [
    { "location": "work",
      "number": "+44-1234567890"},
    { ... }
  ],
  "dob": ISODate("1977-04-01T05:00:00Z"),
  "retirement_fund":
    NumberDecimal("1292815.75")
}
```



Macroarquitetura





Macroarquitetura

- **Documento (*Document*)** □ um registro:
 - Conceito similar ao de tupla do mundo relacional;
 - Composto por **pares de campos (*fields*) e valores (*values*)**;

```
{  
  name: "sue",  
  age: 26,  
  status: "A",  
  groups: [ "news", "sports" ]  
}
```

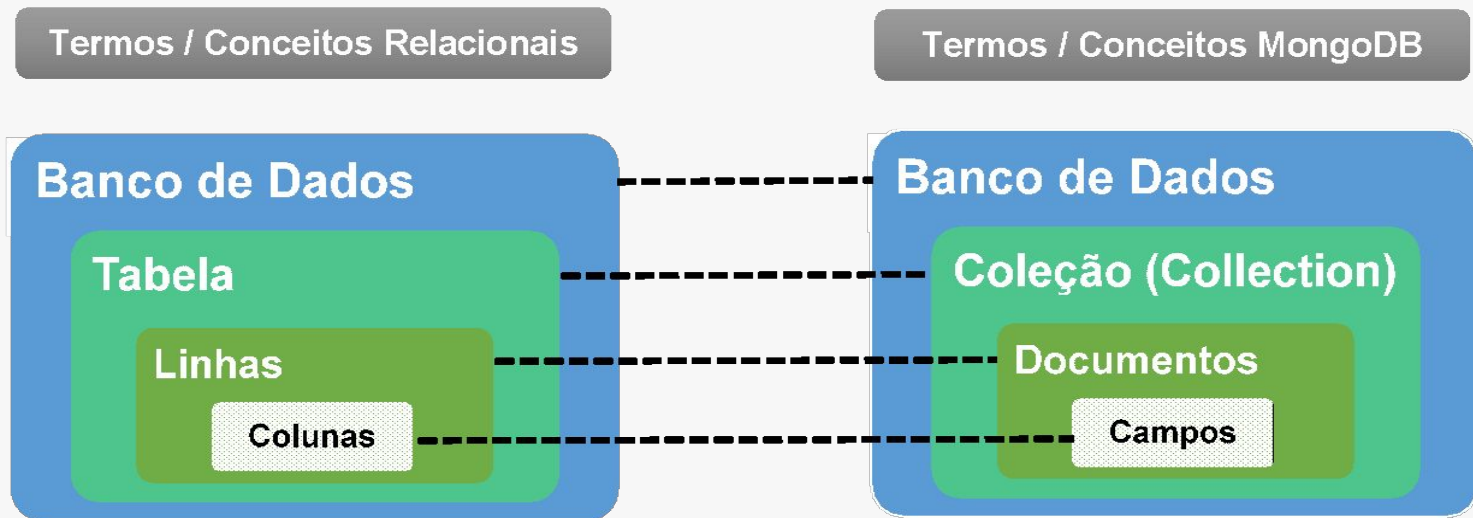
← field: value
← field: value
← field: value
← field: value

- Formato *JSON* □ armazenado como *BSON*;
- Tamanho máximo □ 16 MB (pode ser expandido usando-se GridFS).



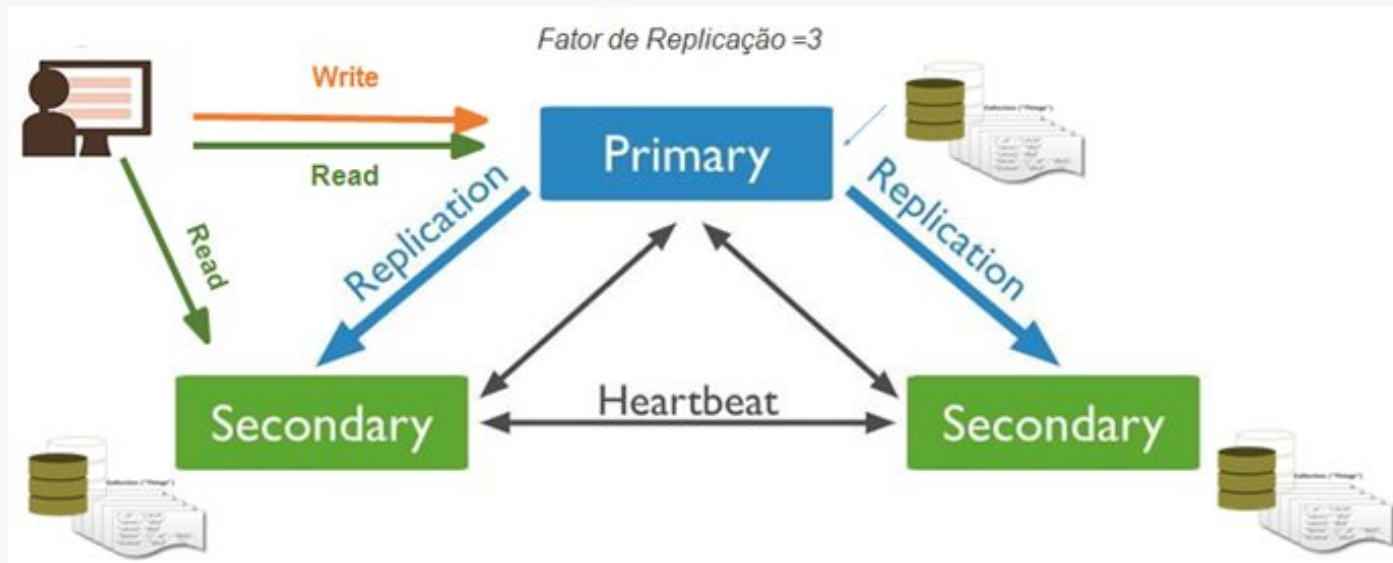
Macroarquitetura

- Analogia dos termos / conceitos Relacionais x MongoDB:





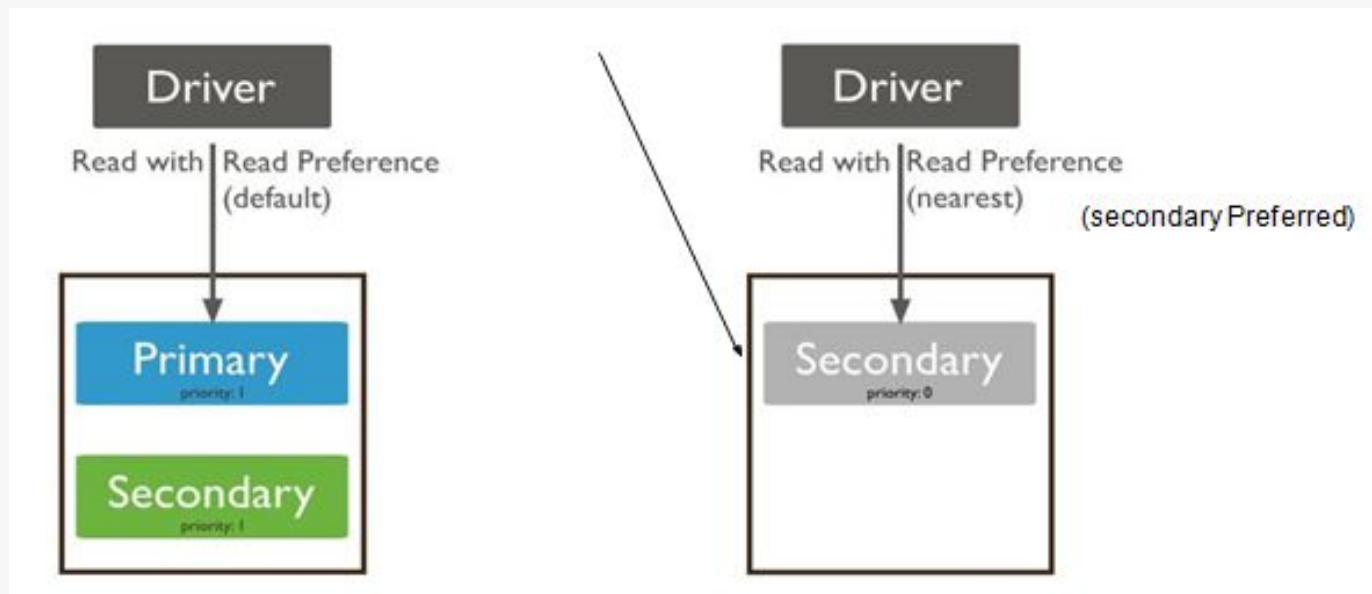
MongoDb Replica Set





MongoDb Replica Set

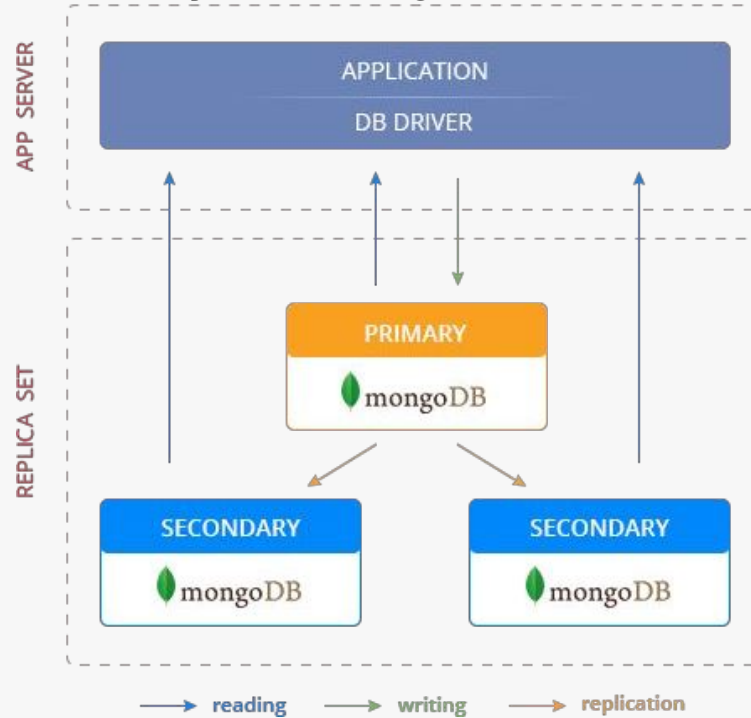
- Leitura nas secundárias via parâmetro na string de conexão.





MongoDb Replica Set

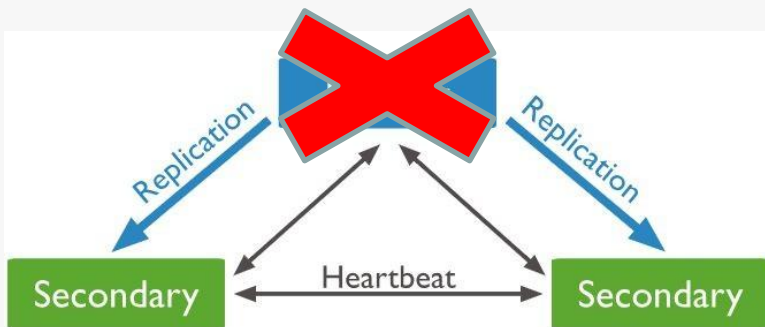
- Distribuição da carga de leituras.





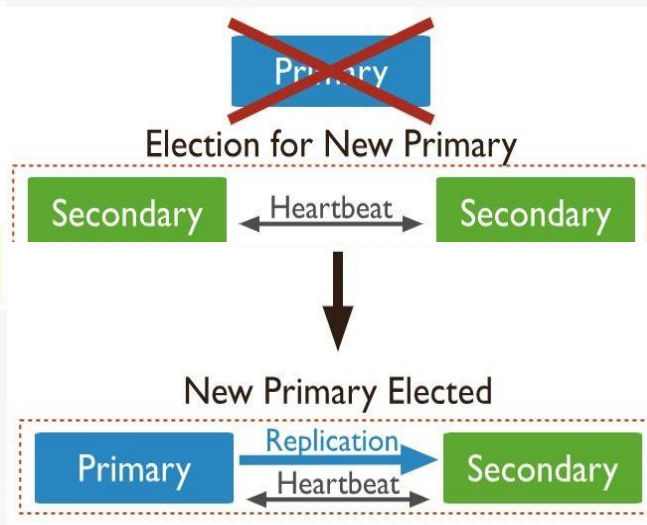
MongoDb Replica Set

- Alta disponibilidade e tolerância a falhas de hardware:
 - **Eleição de uma nova réplica primária** em caso de falhas.



! Quantidade de Réplicas:

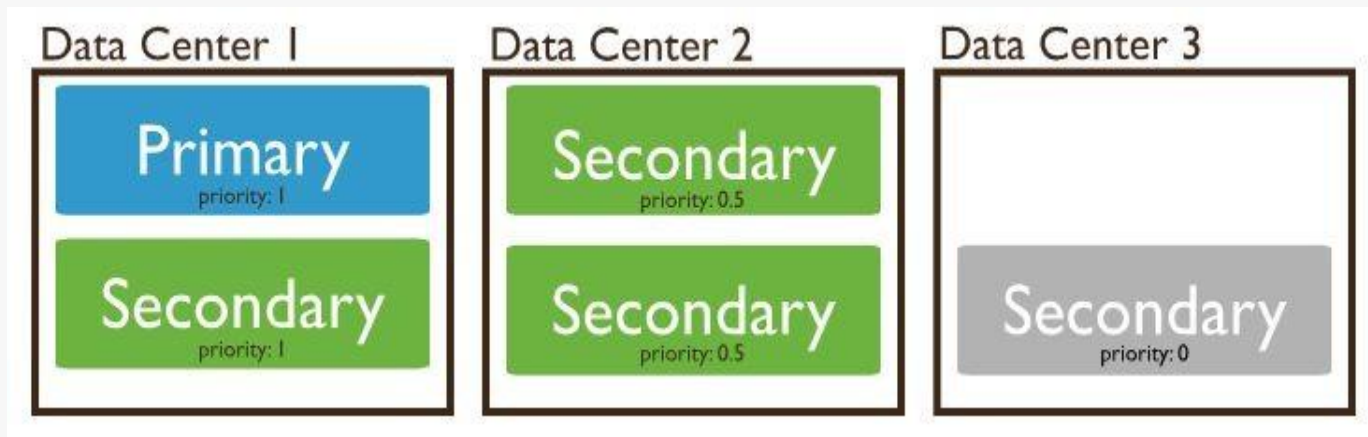
- Ímpar: sem empate;
- Par: usar árbitro para desempate.





MongoDb Replica Set

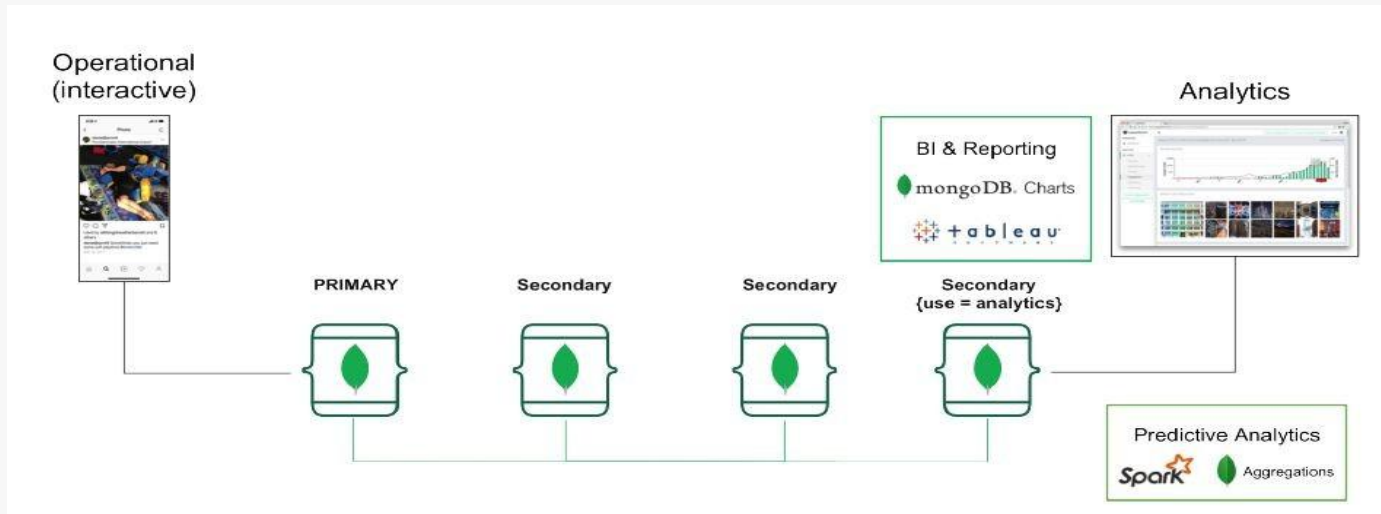
- Distribuição geográfica de dados;
- Ambientes de disaster recovery (DR).





MongoDb Replica Set

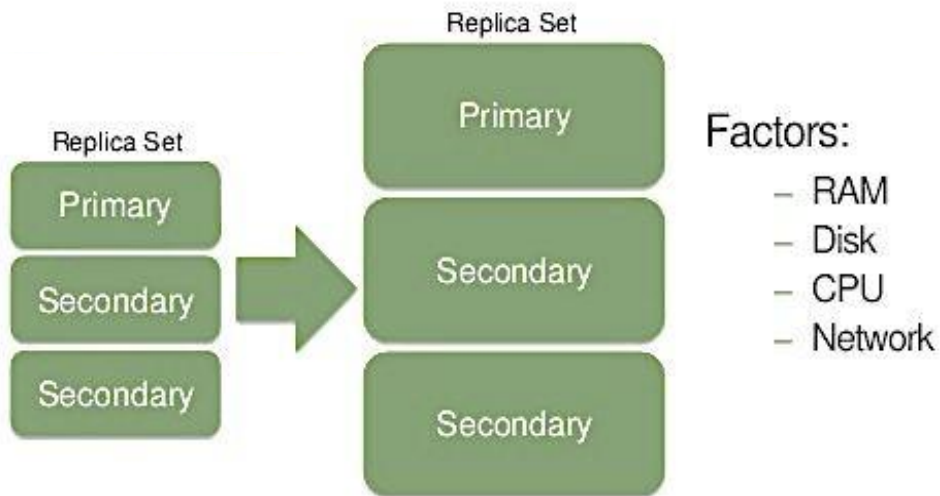
- Coexistência de workloads distintos, sem impactos para o sistema transacional.





MongoDb Replica Set

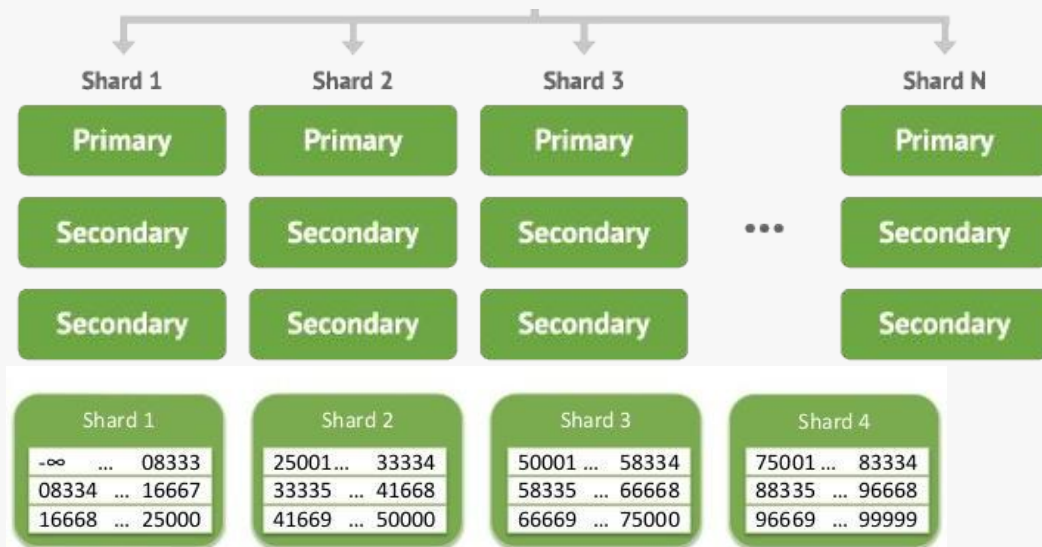
- Escalabilidade vertical.





MongoDb Sharding

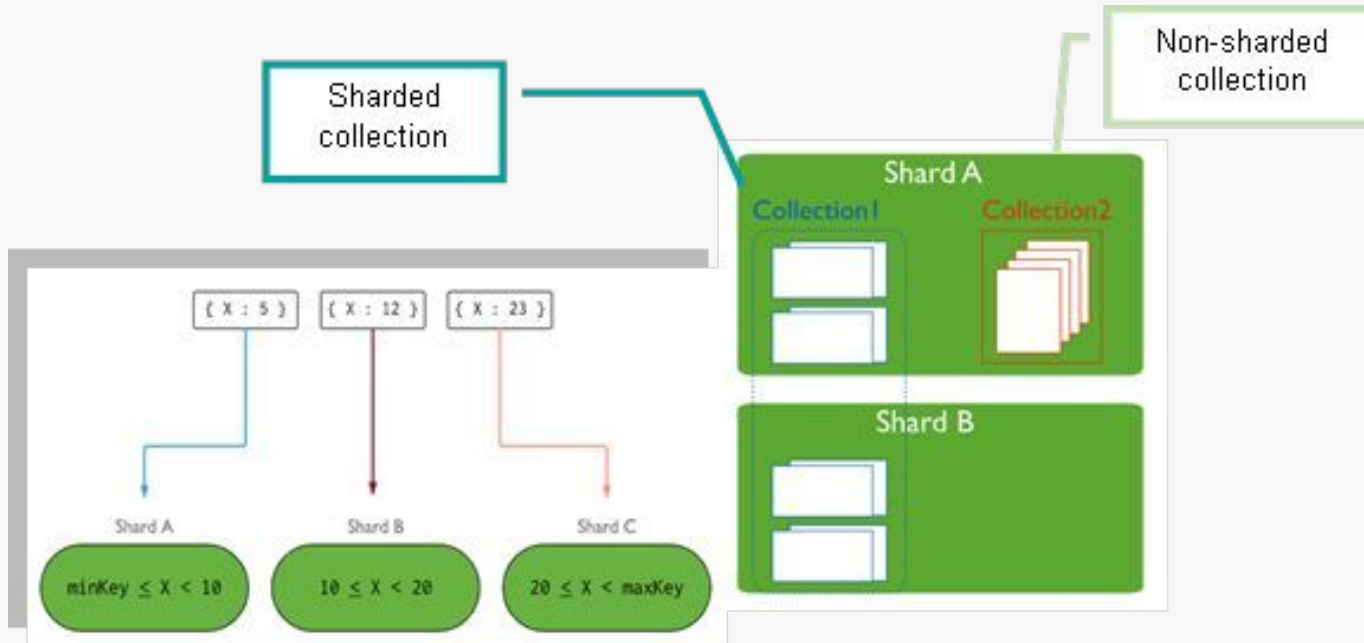
- Usa técnicas de particionamento horizontal (sharding);
- Leitura e escrita em todos nós; Escalabilidade: horizontal e vertical;
- Alta disponibilidade: [(Replicação dos shards (a partir da versão 3.2))].





MongoDb Sharding

- **Sharded Collections** □ são divididas em fragmentos (*shards*);
- **Shard Key** □ chave de particionamento de cada collection.

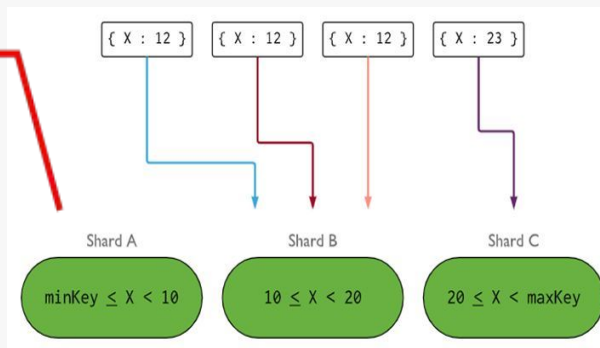




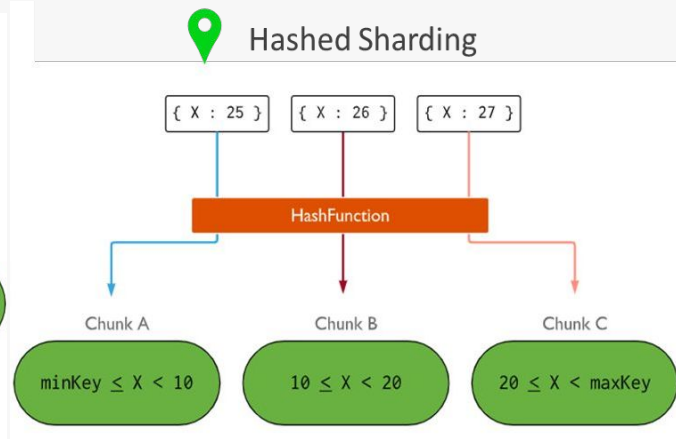
MongoDb Sharding

- Shard Key pode ser de dois tipos:
 - **Hash:** valores gerados e gerenciados pela própria engine do MongoDB;
 - **Range:** permitido definir os limites de valores (faixas) para cada shard.

Desbalanceado



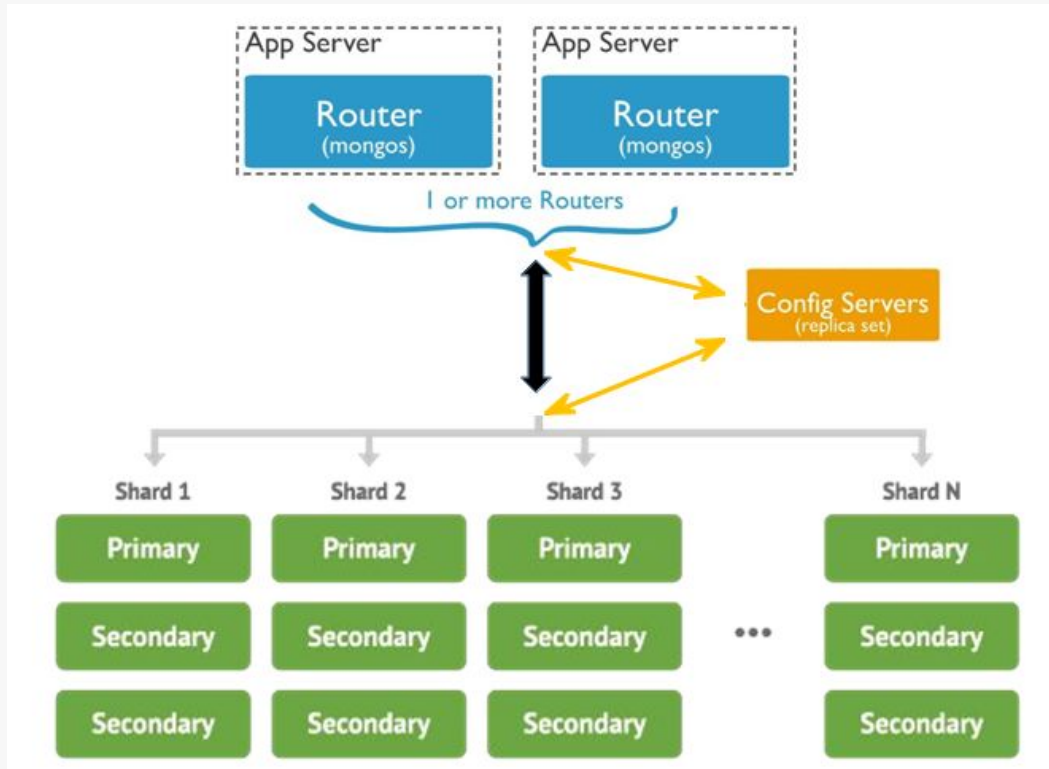
Ranged Sharding



Hashed Sharding



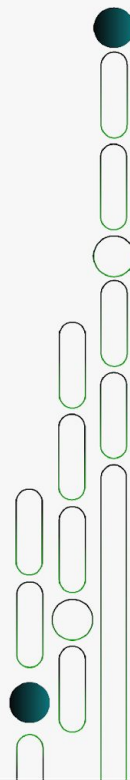
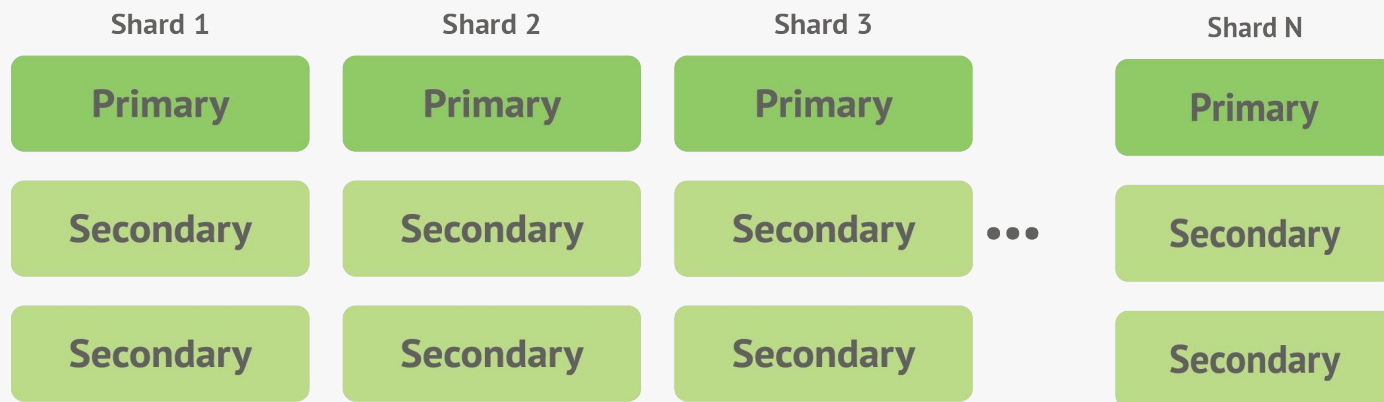
MongoDb Sharding





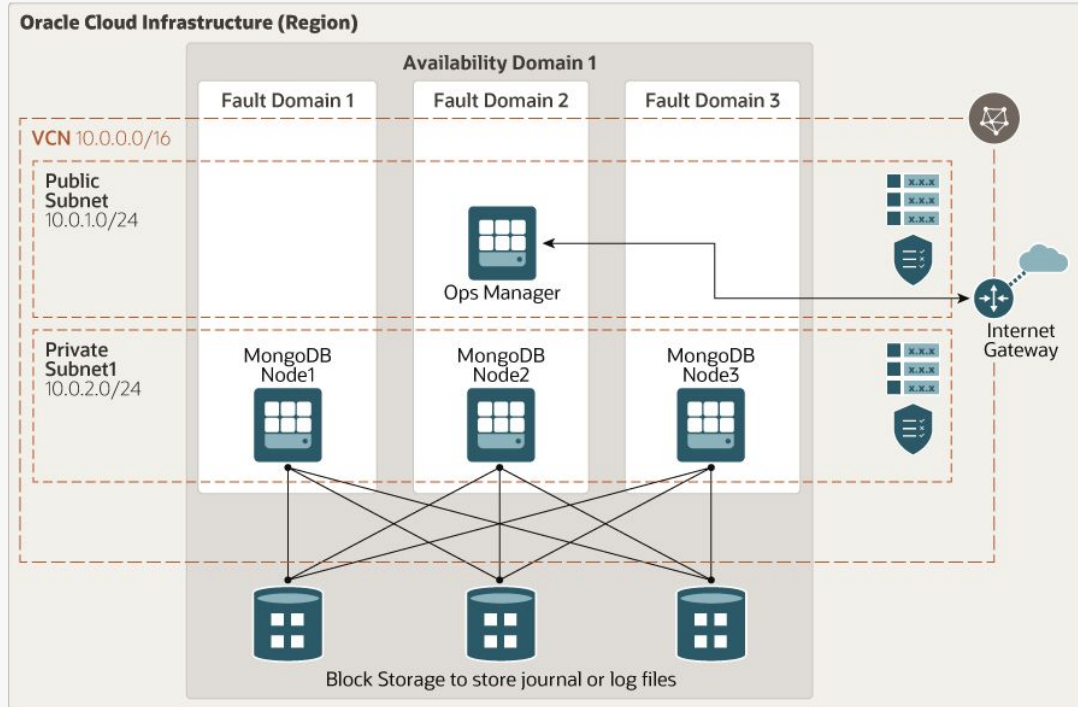
MongoDb Sharding

- Escalabilidade horizontal com sharding.





Sobre o Trabalho Prático





Adriano G. Dias Ferreira

DBA Sr | Data Administrator | Data
Architect | Cloud Computing | Data
Quality | Data Analyst | BI | DW |
Coordenador e Professor de
Graduação e Pós-Graduação

Linked-in: <https://www.linkedin.com/in/adrianogdiasferreira>

Lattes: <http://lattes.cnpq.br/3393239069676217>

Email: adrianogdf@gmail.com

Obrigado
pela
participação!

Bons estudos!
Sucesso!

