SANDRO MATHEUS RAMOS



**TEOREMA CAP**

CASCAVEL

SET / 2022

SANDRO MATHEUS RAMOS

**TEOREMA CAP**

Este trabalho é parte integrante da avaliação bimestral da disciplina de BigData, do 1º bimestre do curso de Engenharia de Software, 8º Período.

Professor: Leandro de Souza

Cascavel, 2022

# **O que é o Teorema CAP**

Proposto pela primeira vez pelo professor Eric A. Brewer, em uma palestra sobre computação distribuída no ano 2000, e posteriormente comprovado por Lynch e Gilbert em 2002, o Teorema de CAP trás que um sistema de banco de dados só pode garantir ao mesmo tempo duas propriedades fazendo com que seja necessário decidir sobre a compensação e prioridade dessas propriedades, que são: Consistência (consisrency), Disponibilidade (availability) e Tolerância à Partição (partition tolerance).

Com o teorema é possível explicar o motivo para que os sistemas NoSQL costumam eximir a consistência dos dados com o intuito de melhorar o desempenho. Porem, o teorema gera uma grande quantidade de discussões, pois sua interpretação vai depender de como as três prioridades são definidas e interpretadas, além do tipo de sistema em consideração.

* Consistência: Independente do nó em que se esteja, será possível ver os mesmos dados. Para que isso aconteça se um dado é gravado em um nó ele deve ser replicado para os demais nós, dessa forma, a consistência dos dados será forte.
* Disponibilidade: Todas requisições estão disponíveis e terão resposta, mesmo que algum nós esteja desativo ou com falha com determinado motivo. Dessa forma, todos os nós em funcionamento no sistema de banco de dados retornaram a resposta mais recente.
* Tolerância à Partição: Aqui o banco de dados é distribuído em diversos servidores independentes, onde a partição é uma quebra de comunicação dentro de um sistema distribuído, uma conexão perdida ou lenta temporariamente. A Tolerância a participação indica que todo o sistema de banco de dados continuará funcionando, com a exceção da parte que estiver fora.

# **Qual a importância do Teorema CAP.**

Na maior parte do tempo sistemas distribuídos operam de forma normal, e erros no sistema são exceções, e quando ocorrem, costumam ser resolvidos rapidamente. Entretanto, com o surgimento do big data e o aumento do volume de dados que é criado diariamente e sua variedade nos bancos distribuídos, torna a probabilidade de erros acontecer mais vezes com o aumento das réplicas.

O CAP torna mais consciente a necessidade de compensação que deve ser realizadas quando os erros acontecem tornando o processo de criação de dados mais ativo.

O teorema faz com que os engenheiros de dados e tenham que decidir qual característica do banco será mais “frágil” uma vez que deverem prezar somente por duas, dessa forma ele podem se perguntar por exemplo, a receita vai ser afetada caso o sistema estiver indisponível? Ou então, se as solicitações receberem respostas imprecisas ocasionalmente?

Por um banco de dado distribuído ser naturalmente tolerável a partição, o engenheiro terá que decidir se o banco será mais consistente ou disponível.

**Exemplo de Banco de dados.**

## MongoBD

É um sistema de gerenciamento de banco de dados (SGBD) NoSQL que armazena dados como documentos BSON (JSON binário), muito usado para big data e aplicativos em tempo real. A SGBD é um armazenamento de dados do tipo CP, então ela resolve partições de rede mantendo a consistência.

Este BD é um sistema de mestre único, o que significa que pode ter um nó primário que recebe todas as operações de gravação e os demais nós são secundários que replicam o nó primário. Quando o nó principal disponível um nó secundário é eleito como novo nó primário para que o cluster volte a ficar disponível.

## Cassandra

Como um banco de dados NoSQL de código aberto, o Cassandra é um banco de coluna ampla que permite armazenar dados em rede distribuída, sendo um AP no Teorema de CAP, sendo assim ele entrega a disponibilidade e tolerância de partição, não conseguindo entregar consistência o tempo todo.

Por não ter um nó principal, Cassandra precisa estar disponível de maneira continua. Por uma consistência eventual, esse banco de dados fornece a funcionalidade de “reparação” para ajudar a recuperar quando uma informação fica inconsistente entre os nós.

## Redis

O Redis, um sistema de gerenciamento de banco de dados NoSQL de código aberto que oferece um conjunto de estruturas versáteis de dados na memória que permite fácil criação de aplicações personalizadas, tendo como principal uso o cache, gerenciamento de sessões e classificações

Assim como o banco de dados MongoBD, o Redis é um sistema de gerenciamento de banco de dados do tipo CP dentro do Teorema CAP, entregando uma maior consistência dos dados ao invés da disponibilidade,

## PostgreSQL

O sistema de gerenciamento de banco de dados relacional gratuito Postgre possui o código fonte aberto sendo o SGBD que possui o conjunto mais complexo de recursos, sendo um banco objeto-relacional.

Apesar dos bancos relacionais terem substituído os modelos hierárquico e de redes, sua implementação se torna difícil em situação que alta demanda de dados como em aplicações web distribuídas. Em relação ao Teorema de CAP, por ser um banco relacional o Postgre é do tipo CA onde ele consegue oferecer a consistência e disponibilidade dos dados, não possuindo partições.

# **Conclusão**

Com os sistemas distribuídos é possível atingir um nível de capacidade e disponibilidade de computação que não estavam disponíveis no passado, como em bancos relacionais muito usado ainda.

Possuem alto desempenho, menor latência e estão em atividade constante em quase 100% das vezes em todo o mundo, porém, são mais complexos e suscetíveis a falhas devidos sua distribuição em vários nós de informação que podem ou não ter disponibilidade constante ou consistente, levando os engenheiros a precisar escolher qual propriedade é mais relevante para o projeto que estão envolvidos.

# **Referências**

MAC0439 – Laboratório de Bancos de Dados. SGBDs Distribuídos e o Teorema CAP. Disponível em: < <https://edisciplinas.usp.br/pluginfile.php/5214218/mod_resource/content/1/mac439_aula6_teorema_cap.pdf> >. Acesso em: 09 de setembro de 2022.

REIS, Fabio dos. Conceitos de bancos de Dados – O Teorema CAP. Boson Treinamentos em Ciência e Tecnologia. Disponível em < <http://www.bosontreinamentos.com.br/bancos-de-dados/conceitos-de-bancos-de-dados-o-teorema-cap/> >. Acesso em: Acesso em: 09 de setembro de 2022.

MENDES, Bruno L. Oque é teorema de CAP e por que é importante na engenharia de dados. Linkedin. Disponível em: < [https://www.linkedin.com/pulse/oque-%C3%A9-teorema-de-cap-e-por-que-importante-na-dados-bruno-l-mendes/?originalSubdomain=pt](https://www.linkedin.com/pulse/oque-é-teorema-de-cap-e-por-que-importante-na-dados-bruno-l-mendes/?originalSubdomain=pt) >. Acesso em: Acesso em: 09 de setembro de 2022.

EDUCATION, IBM Cloud. Teorema CAP. IBM. Disponível em: < <https://www.ibm.com/br-pt/cloud/learn/cap-theorem> >. Acesso em: 09 de setembro de 2022.

RABELO, Eduardo. Arquitetura de Software: O Teorema CAP. DEV. Disponível em: < <https://dev.to/oieduardorabelo/arquitetura-de-software-o-teorema-cap-3m18> >. Acesso em: 09 de setembro de 2022.

AWS Amazon. O que é o Redis?. AWS Amazon. Disponível em: < <https://aws.amazon.com/pt/elasticache/what-is-redis/> >. Acesso em: 09 de setembro de 2022.

LEITE, Gleidson Sobreira. Análise Comparativa do Teorema CAP Entre Bancos de

Dados NoSQL e Bancos de Dados Relacionais. Docplayer. Disponível em: < <https://docplayer.com.br/1711115-Analise-comparativa-do-teorema-cap-entre-bancos-de-dados-nosql-e-bancos-de-dados-relacionais.html> >. Acesso em: 9 de setembro de 2022.