

Disciplina:

BANCOS DE DADOS NoSQL

Professor: Augusto Zadra



1.2 Definição de negócios e motivação

DEFINIÇÃO DE NEGÓCIOS

- Há um foco na criação de plataformas consolidadas e compactas para realizar uma série de ações essenciais em direção a acesso aos dados de forma simplificada e descoberta de conhecimento.
- Porém não há como verificarmos isto sem pensarmos infraestruturas **ideais, dinâmicas e convergentes** para serem eficazes em suas operações levando em conta a série de transformações, como **otimização, racionalização, automação e simplificação** que as infraestruturas de TI passaram no passado recente.

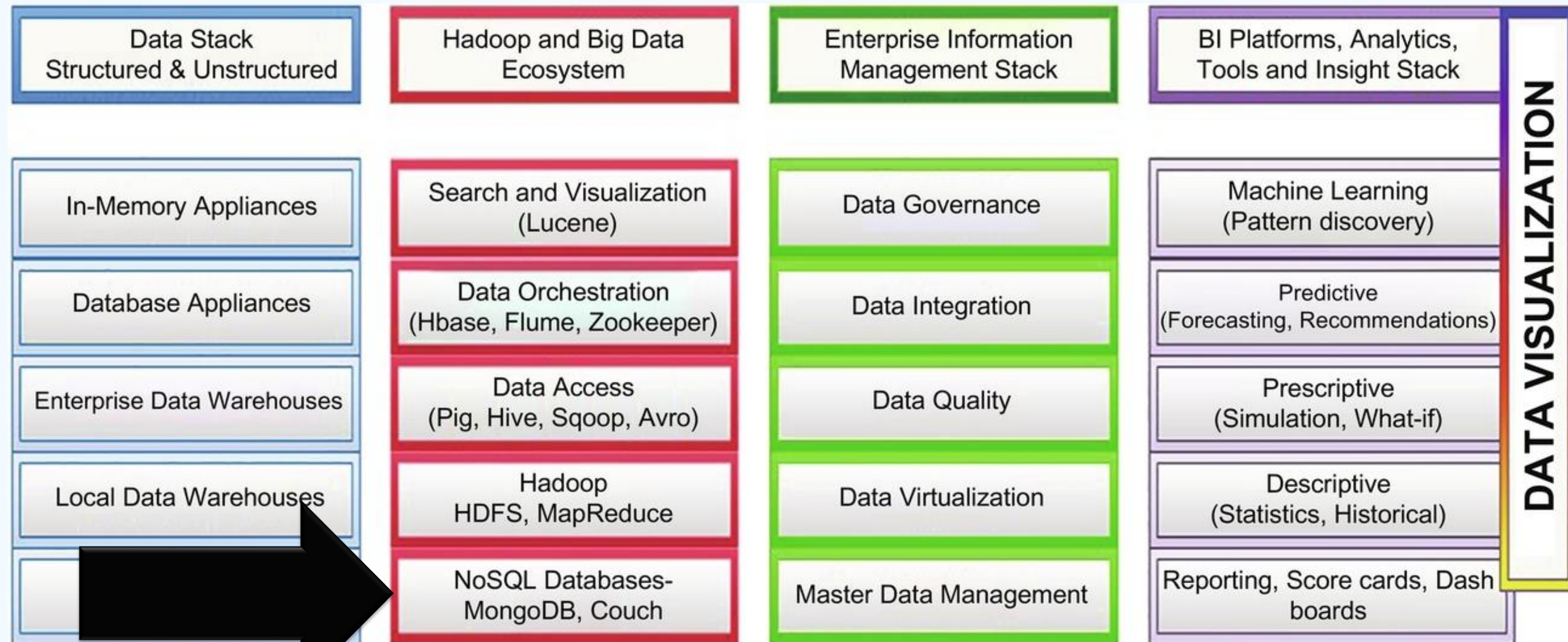
DEFINIÇÃO DE NEGÓCIOS

- A ideia da nuvem vem dominando o pensamento dos analistas de infraestrutura atualmente, à medida que este paradigma possui uma abordagem mais pragmática para alcançar os ideais de otimização da infraestrutura.
- Como consequência desta popularidade, a computação em nuvem com toda a oferta de infraestrutura de TI (servidores, armazenamentos e soluções de rede) está sendo submetida a uma série de tarefas de modernização para capacitá-los a serem baseados em políticas, definidos por software, orientados a serviços, programável, e etc.

DEFINIÇÃO DE NEGÓCIOS

- As infraestruturas de nuvem centralizadas/federadas, virtualizadas, automatizadas, compartilhadas e otimizadas (privadas, públicas e híbridas) vem sendo o foco e os negócios apontam cada vez mais nesta direção.
- Os Ambientes de TI estão sendo preparados para a era do **BIG DATA** e as redes orientadas a software são as infraestruturas de comunicação mais procuradas para transmissão e processamento de dados.
- Vamos visualizar assim os componentes relevantes na simplificação e agilidade para o acesso aos dados.

DEFINIÇÃO DE NEGÓCIOS



Fonte: HURSON, MEMON, 2018. Disponível em: <
<https://learning.oreilly.com/api/v2/epubs/urn:orm:book:9780128137871/files/images/S0065245818300020/f01-01-9780128137864.jpg> >

DEFINIÇÃO DE NEGÓCIOS

- O sucesso de qualquer tecnologia deve ser decidido com base no número de aplicativos de missão crítica que ela poderia criar e manter.
- A aplicabilidade ou empregabilidade do novo paradigma para tantos domínios de aplicação quanto possível é o principal fator de decisão para sua jornada de sucesso.
- Como grandes *insights* estão se tornando obrigatórios para vários segmentos da indústria, há um escopo maior para aplicativos de **BIG DATA**.

Definição de negócios

Motivação

- De forma direta, a motivação para esta tecnologia, como já dissemos algumas vezes é a enorme quantidade de dados, proveniente de várias fontes diferentes, complexos, dinâmicos, e que demandam tecnologias e técnicas específicas para armazenamento, análise e visualização.
- Bancos de dados relacionais escalam, mas quanto maior o tamanho, mais custoso – sob a ótica de consumo de recursos e por conseguinte sob a ótica financeira.

Tipos e modelos de transações

Conceitos

- Faz-se necessário uma apresentação dos modelos de transações e sua ambientação no mosaico tecnológico desta nova realidade.
- Introduziu-se os conceitos básicos dos modelos ACID e BASE porém vamos expandir agora o teorema CAP e apresentar a comparação entre os modelos de transações.

Tipos e modelos de transações

Conceitos

- O controle de transações é importante em ambientes de computação distribuída nos quesitos desempenho e consistência.
- Normalmente, um dos dois tipos de modelos de controle de transação são utilizados o **ACID**, usado em bancos de dados relacionais, e **BASE**, encontrado em muitos sistemas NoSQL.

Tipos e modelos de transações

Conceitos

- Importante reforçar que os sistemas relacionais e NoSQL possuem a capacidade de criar esses controles e a diferença entre esses modelos está na quantidade de esforço exigida pelos desenvolvedores de aplicativos e na localização (camada) dos controles transacionais.
- A diferença entre esses modelos está na quantidade de esforço exigida pelos desenvolvedores de aplicativos e na localização (camada) dos controles transacionais.

Tipos e modelos de transações

ACID

- Iniciemos então com a conceituação de ACID:
- **Atomicidade:** significa que em uma transação envolvendo duas ou mais partes de informações discretas, ou a transação será executada totalmente ou não será executada, garantindo assim que as transações sejam atômica.
- **Consistência:** é quando uma transação cria um novo estado válido dos dados ou, em caso de falha, retorna todos os dados ao seu estado anterior ao início da transação.

Bancos de dados NoSQL

Propriedades dos bancos de dados NoSQL

ACID

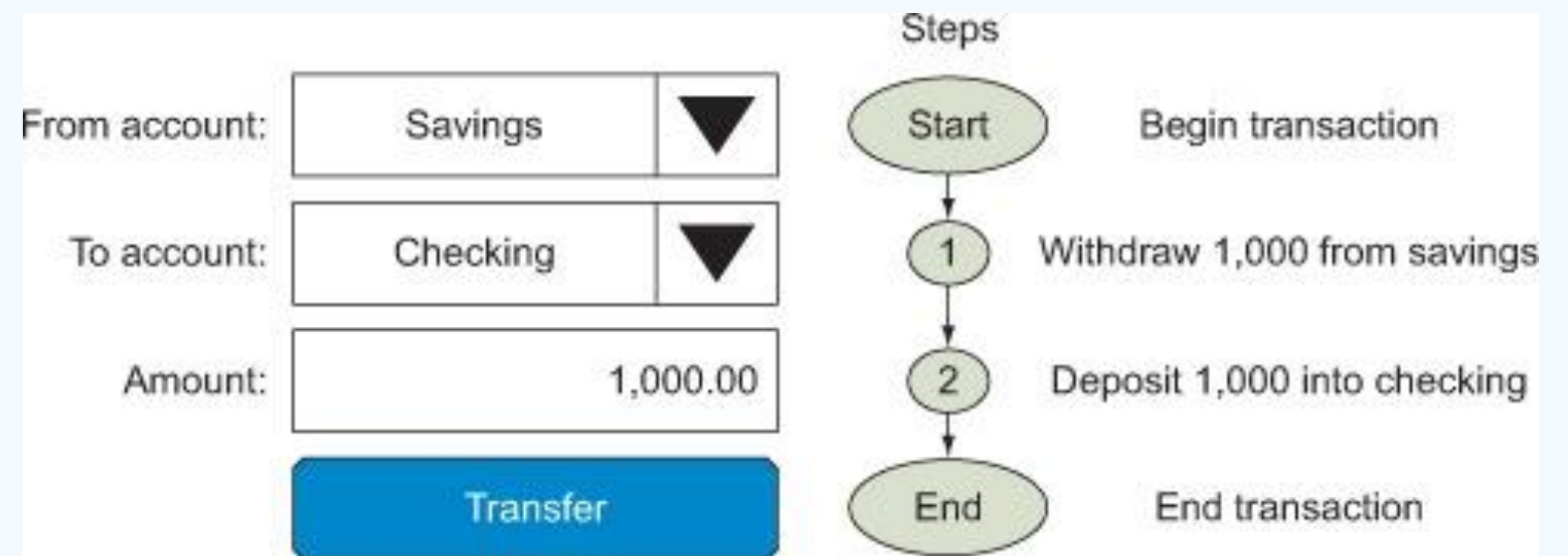
- **Isolamento:** significa que uma transação em andamento, mas ainda não validada, deve permanecer isolada de qualquer outra operação externa, ou seja, garante-se que a transação não será interferida por nenhuma outra transação concorrente.
- **Durabilidade:** indica que dados validados são registrados pelo sistema de tal forma que, mesmo no caso de uma falha ou reinício do sistema, os dados estão disponíveis em seu estado correto.

Bancos de dados NoSQL

Propriedades dos bancos de dados NoSQL

ACID

- O melhor exemplo de transação ACID é a bancária onde há os marcadores para início e fim da análise.



Bancos de dados NoSQL

Propriedades dos bancos de dados NoSQL

BASE

- Uma alternativa ao modelo de transação ACID é o BASE que apresenta os seguintes conceitos:
 - ✓ **Basic Availability:** permite que os sistemas fiquem temporariamente inconsistentes para que as transações sejam gerenciáveis.
 - ✓ **Soft-state:** reconhece que alguma imprecisão é temporariamente permitida e os dados mudam enquanto são usados reduzindo os recursos consumidos.

Bancos de dados NoSQL

Propriedades dos bancos de dados NoSQL

BASE

- ✓ **Eventual consistency:** significa que, eventualmente, quando toda a lógica de serviço é executada, o sistema é deixado em um estado consistente.
- ✓ **Soft-state:** reconhece que alguma imprecisão é temporariamente permitida e os dados mudam enquanto são usados reduzindo os recursos consumidos.

Bancos de dados NoSQL

Propriedades dos bancos de dados NoSQL

BASE

- Os sites que usam as recursos de carrinho de compras e checkout são exemplo de como utiliza-se o conceito de transações BASE.
- Eles consideram as transações de forma diferente e assim, a emissão de relatórios inconsistentes por alguns minutos é menos importante do que algo que o impede de aceitar um pedido, porque se você bloquear um pedido, você perde um cliente.

Bancos de dados NoSQL

Propriedades dos bancos de dados NoSQL

BASE

- A forma de organização do modelo BASE é notável pois o objetivo principal é permitir que novos dados sejam armazenados, mesmo com o risco de ficarem fora de sincronia por um curto período de tempo.
- Para que isto aconteça ele relaxa as regras e permitem a execução de relatórios, mesmo que nem todas as partes do banco de dados estejam sincronizadas.
- Eles presumem que, eventualmente, todos os sistemas serão atualizados e se tornarão consistentes.

Bancos de dados NoSQL

Propriedades dos bancos de dados NoSQL

Teorema CAP

- O teorema CAP é outra base os bancos de dados NoSQL. Este teorema afirma que é impossível para qualquer sistema distribuído fornecer simultaneamente todos os três recursos:
 - ✓ **Consistência:** todos os nós veem os mesmos dados ao mesmo tempo.
 - ✓ **Disponibilidade:** garantia de resposta em cada solicitação sobre conclusão bem-sucedida ou falha.

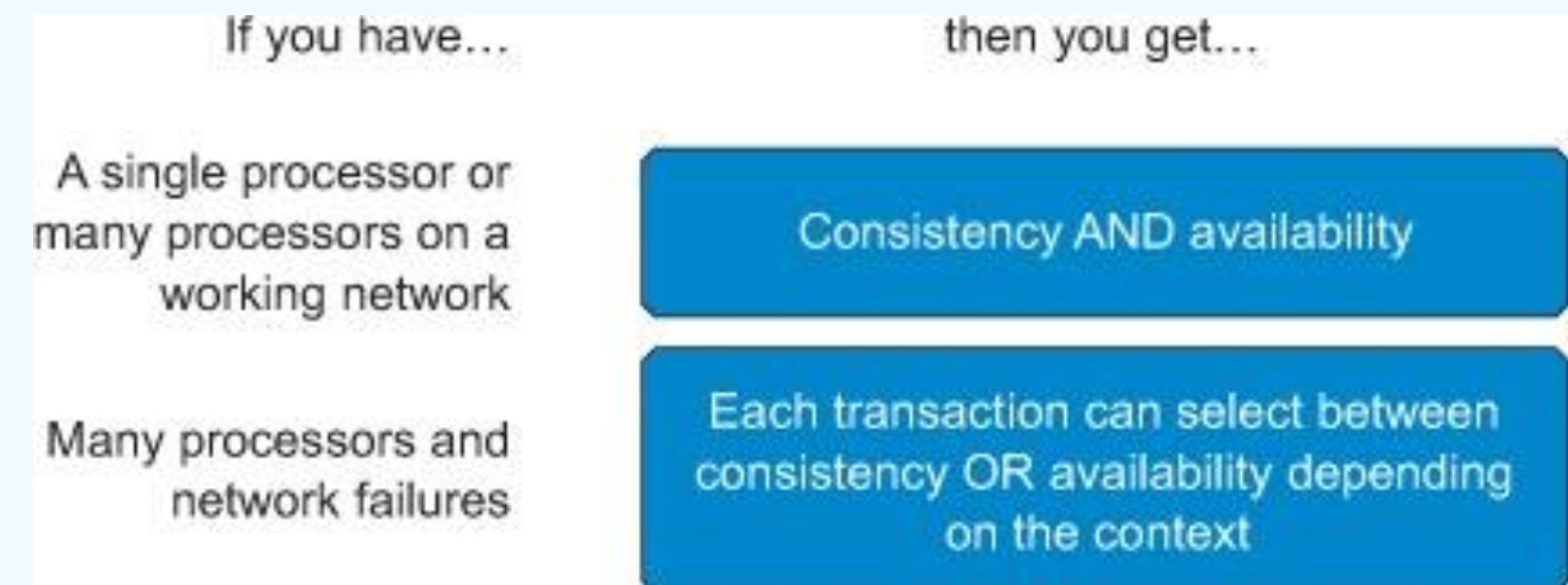
Bancos de dados NoSQL

Propriedades dos bancos de dados NoSQL

CAP Teorema

✓ **Tolerância de partição:** o sistema continua a operar apesar da perda arbitrária de mensagens ou da falha de uma parte do sistema.

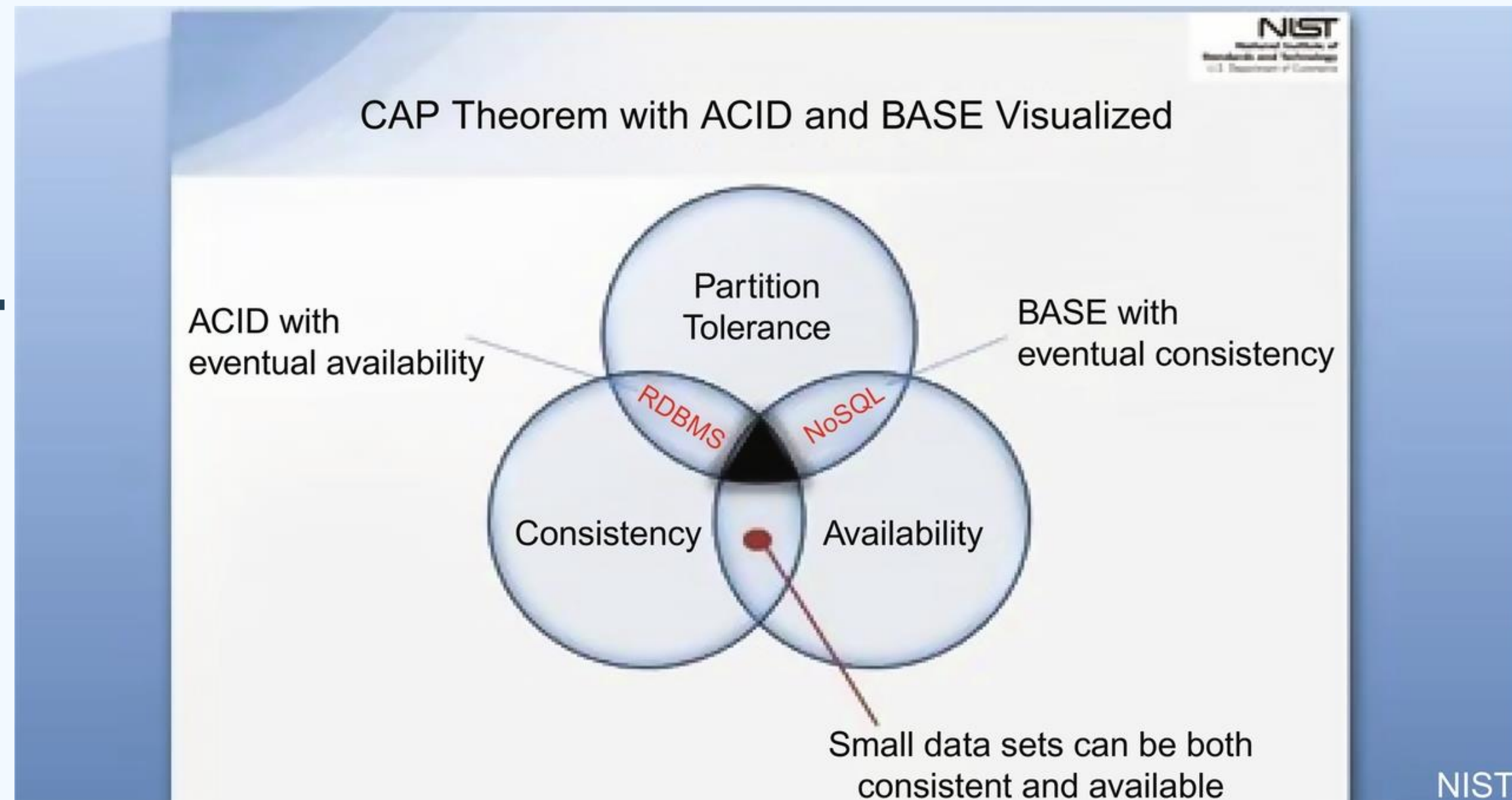
- Ou seja, qualquer sistema distribuído garante quaisquer dois dos três declarados posteriormente.



Bancos de dados NoSQL

Propriedades dos bancos de dados NoSQL

CAP Teorema

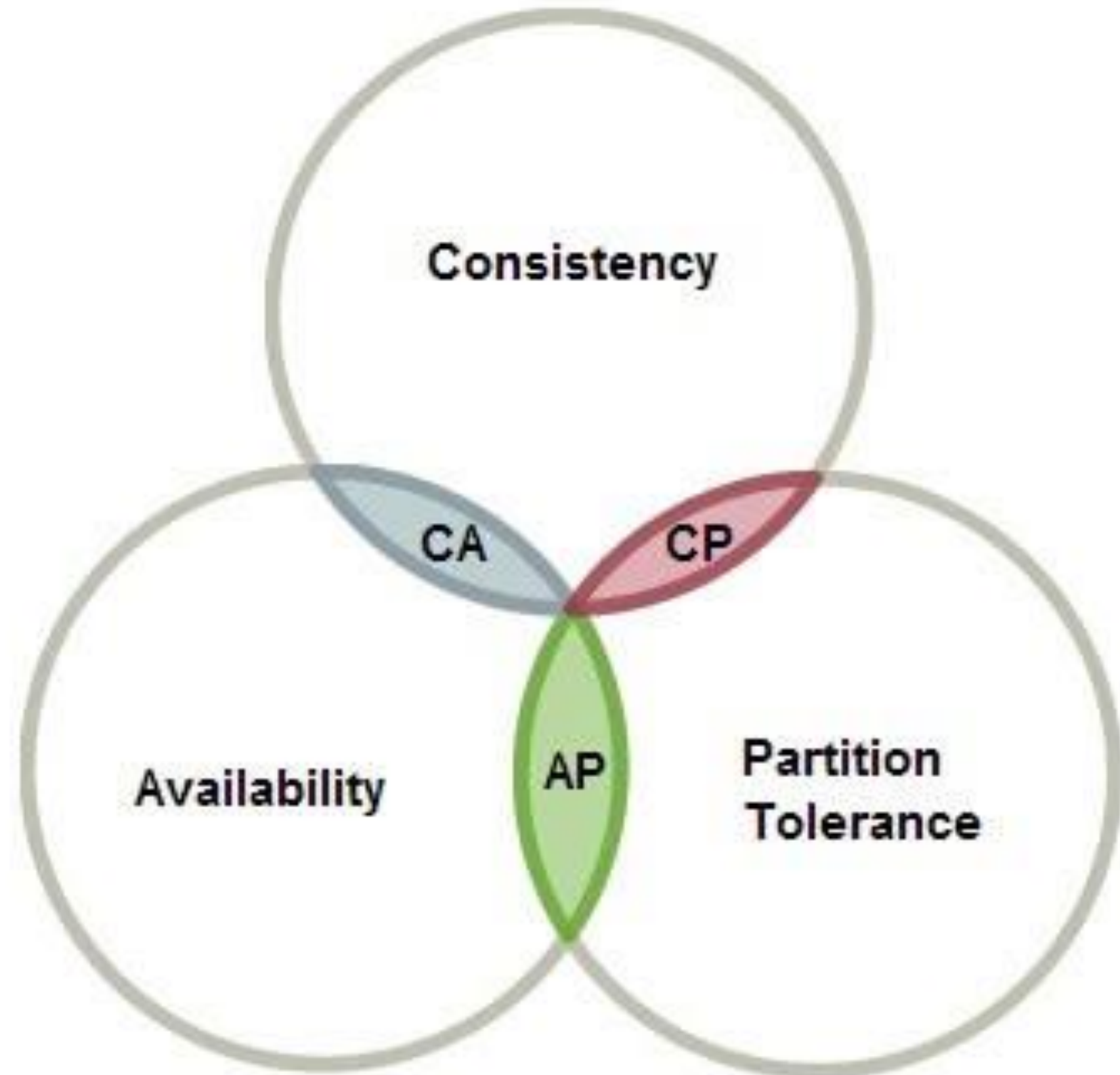


Fonte: NIST, 2021. Disponível em: <
https://bigdata.nist.gov/_uploadfiles/M0065_v1_4451775754.pdf >

Bancos de dados NoSQL

Propriedades dos bancos de dados NoSQL

CAP Teorema

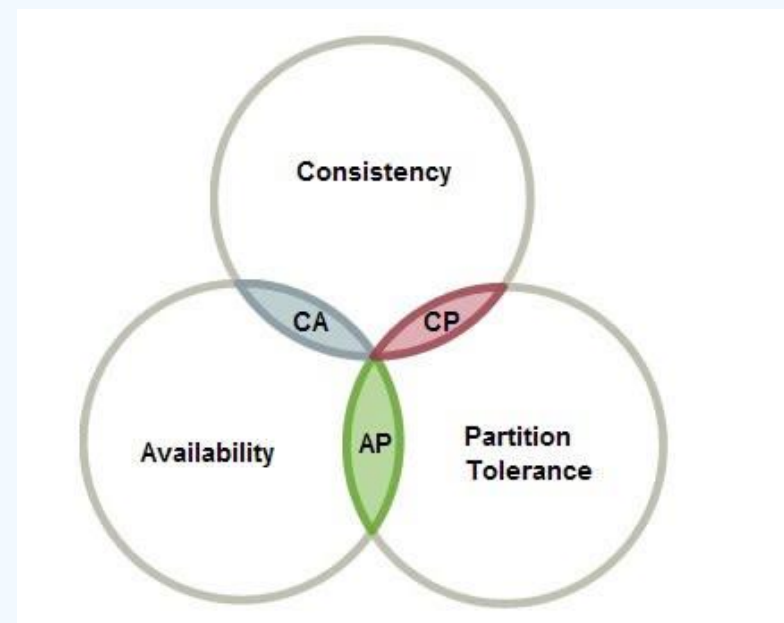


Bancos de dados NoSQL

Propriedades dos bancos de dados NoSQL

CAP Teorema: SISTEMAS CA (Consistency/Availability)

- Os sistemas com consistência forte e alta disponibilidade (alta disponibilidade de um nó apenas) não sabem lidar com a possível falha de uma partição.
- Caso ocorra o sistema inteiro pode ficar indisponível até o membro do cluster voltar.
- Exemplo: configurações clássicas de bancos relacionais.

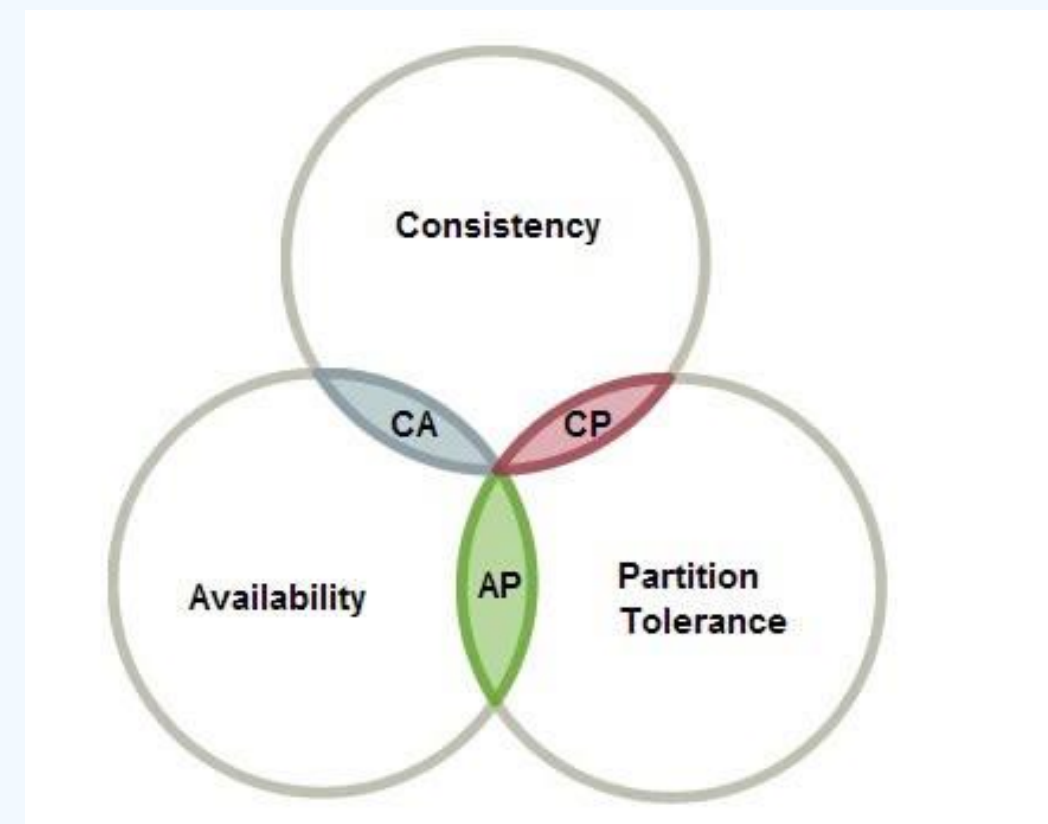


Bancos de dados NoSQL

Propriedades dos bancos de dados NoSQL

CAP Teorema: SISTEMAS CP (Consistency/Partition Tolerancy)

- Para sistemas que precisam da consistência forte e tolerância a particionamento é necessário abrir a mão da disponibilidade (um pouco).
- **Exemplo:** BigTable, HBase ou MongoDB entre vários outros.



Bancos de dados NoSQL

Propriedades dos bancos de dados NoSQL

CAP Teorema: SISTEMAS AP (Availability/Partition Tolerancy)

- São sistemas que jamais podem ficar offline e para ter alta disponibilidade mesmo com tolerância a particionamento é preciso prejudicar um pouco a consistência.
- Esses sistemas aceitam escritas sempre e tentam sincronizar os dados em algum momento depois (*read-consistency*).
- **Exemplo:** aqui são Amazon Dynamo, Cassandra ou Riak.

Bancos de dados NoSQL

Propriedades dos bancos de dados NoSQL

ACID X BASE

- Bancos de dados SQL cumprem as propriedades **ACID** (atomicidade, consistência, isolamento e durabilidade) e acomodam apenas dados estruturados.
- Quando a quantidade de dados se torna gigantesca, as interações enfrentam enormes dificuldades por aplicativos transacionais.
- Foram projetados para gerenciar dados estruturados em campos, linhas e colunas gerenciáveis, como datas, números de previdência social, endereços correspondentes e valores de transação.

Bancos de dados NoSQL

Propriedades dos bancos de dados NoSQL

SQL X NoSQL

- Na maioria dos bancos de dados NoSQL possuem a característica de serem sem esquema (*schemaless*) o que permite ostensivamente a gravação de dados multiestruturados.
- Isso funciona especialmente bem para documentos e metadados associados a uma variedade de tipos de dados não estruturados.
- BASE determina que o banco de dados irá, em algum ponto, classificar e indexar o conteúdo para melhorar a localização/descoberta de dados ou informações contidas no texto ou no objeto.

Bancos de dados NoSQL

Propriedades dos bancos de dados NoSQL

SQL X NoSQL

- A diferença importante e considerável é que as tecnologias relacionais têm esquemas rígidos, enquanto os modelos NoSQL são sem esquemas.
- Alterar o esquema uma vez que os dados são inseridos possui alta complexidade e é frequentemente evitado.
- No entanto, no BIG DATA os desenvolvedores de aplicativos precisam incorporar constante e rapidamente novos tipos de dados para enriquecer seus aplicativos e este é o caminho que estamos seguindo.

QUANDO UTILIZAR SQL OU NoSQL

UTILIZE BANCOS RELACIONAIS	UTILIZE NoSQL
Aplicações forem centralizadas	Aplicações descentralizadas (Web, IoT, BIG Data)
Não houver requisitos de altíssima disponibilidade 99,9999999999(10)	Não puder haver interrupção no funcionamento da gravação de dados
Geração de dados em velocidade média (há períodos de “descanso”)	Geração de dados em alta velocidade (sensores)
Poucas fontes geradoras de dados	Muitas fontes geradoras de dados
Dados forem estruturados	Dados forem semi ou não-estruturados
Transações complexas	Transações simples
Armazenamento de dados com médio volume	Alto volume de armazenamento

*Now, we'll advance to
more concepts and rating NoSQL DB!*

REFERÊNCIAS BIBLIOGRÁFICAS

NIST. Tackling Big Data. Disponível em: < https://bigdata.nist.gov/_uploadfiles/M0065_v1_4451775754.pdf >. Acesso em 16 de nov. de 2021.

DIANA, M. de;; **GEROSA**, M. A. NOSQL na Web 2.0: Um estudo comparativo de bancos não-relacionais para armazenamento de dados web 2.0. In: Workshop de Teses e Dissertações em BD - WTDB, 9., 2010, Belo Horizonte. Anais... Belo Horizonte: SBC, 2010.

LI, Y.;; **MANOHARAN**, S. A performance comparison of SQL and NoSQL databases. In: IEEE Pacific Rim Conference on Communications, Computers and Signal Processing - PACRIM, 14., 2013, Victoria, B.C., Canadá. Proceedings... IEEE, 2013.

LÓSCIO, B. F.;; **OLIVEIRA**, H. R. de;; **PONTES**, J. C. de S. NoSQL no desenvolvimento de aplicações web colaborativas. In: Simpósio Brasileiro de Sistemas Colaborativos – SBSC, 8., 2011, Paraty (RJ). Anais... Paraty: SBC, 2011.

VIEIRA, M. R. et al. Bancos de Dados NoSQL: conceitos, ferramentas, linguagens e estudos de casos no contexto de Big Data. In: Simpósio Brasileiro de Bancos de Dados, 27., 2012, São Paulo. Anais... São Paulo: SBC, 2012.