

Arquiteturas Big Data

Equipe



Det Teixeira



Kelen Ferreira



Mateus Andrade



Matheus Paz



Thiago Nátalys



Tópicos



Definições / Conceitos / Exemplos



Existência de Padrões Arquiteturais /
Arquiteturas de Referências



Abordagens de apoio ao Design Arquitetural



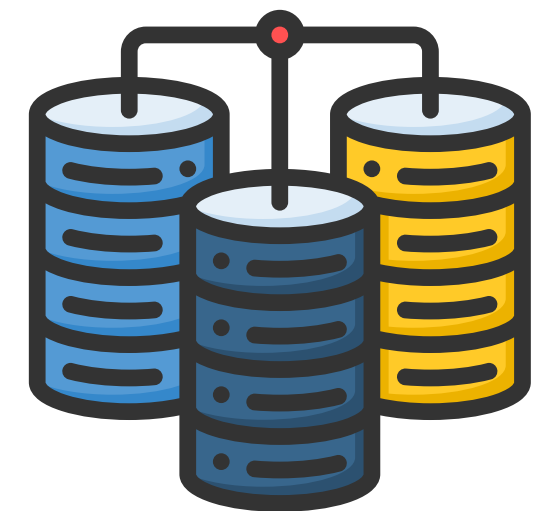
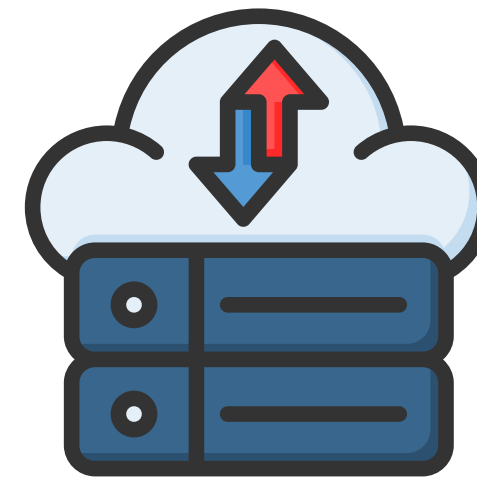
Desafios na Arquitetura de Software /
Engenharia de Software





Definição

Uma **arquitetura de Big Data** foi projetada para lidar com ingestão, processamento e análise de dados grandes ou complexos demais para sistemas de banco de dados tradicionais. O limite no qual as organizações ingressam no campo do Big Data é diferente, dependendo das capacidades dos usuários e de suas ferramentas.





Conceitos

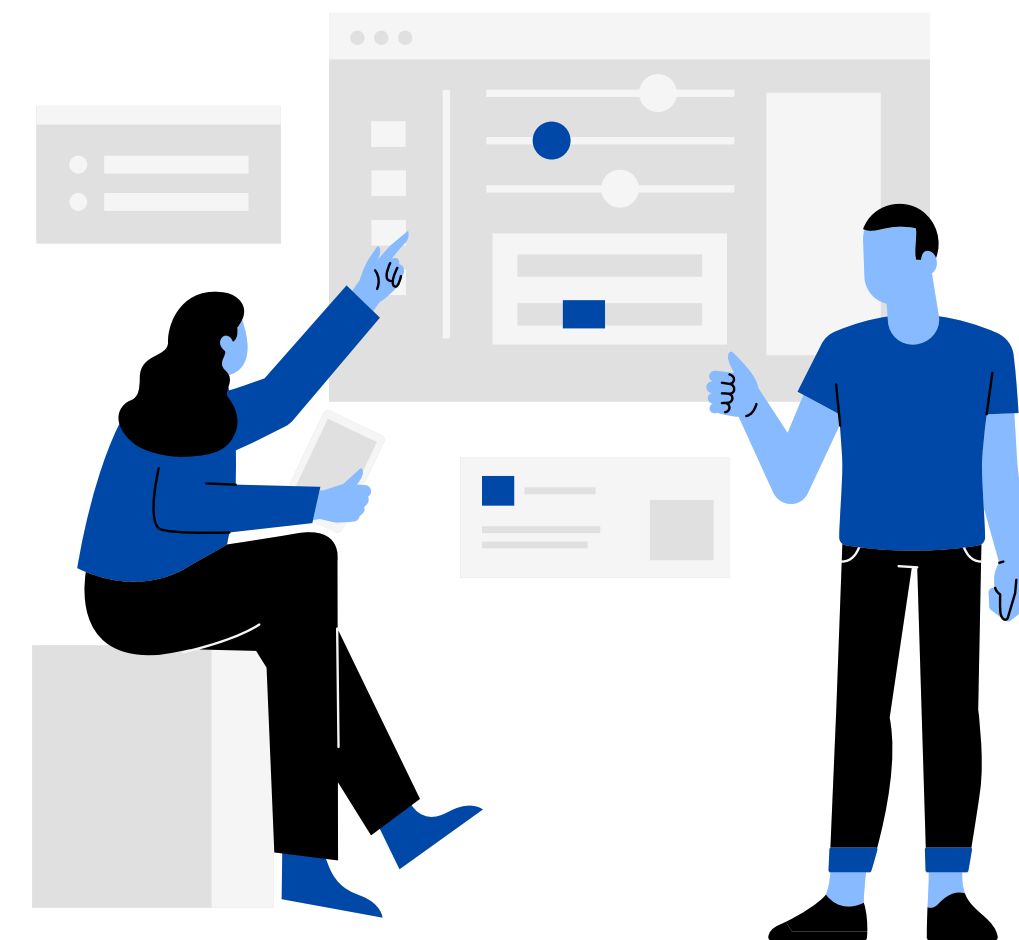
- Escalabilidade
 - Tolerância a falhas
 - Dados distribuídos
 - Processamento distribuído
 - Localidade dos dados
 - 5 V's
- Volume
 - Velocidade
 - Variedade
 - Veracidade
 - Valor





Especialista

O **Arquiteto de Big Data** domina as ferramentas de comunicação e negociação para entender os dados obtidos e assessorar o cliente (acompanhando-o durante todo o processo) em sua tomada de decisão. O arquiteto saberá informar os riscos e oportunidades nos resultados alcançados com sua ferramenta.





Exemplos

01. Análise e visualização: vem em primeiro lugar. Aqui os dados são exibidos para exploração e análise usando técnicas estatísticas, algoritmos de análise preditiva, aprendizado de máquina, etc.

02. Governança de dados: concentra-se na integração, governança e segurança de dados. É necessário escolher os dados adequados que permitirão um processamento eficiente, com a qualidade exigida e protegê-los adequadamente, minimizando os riscos de segurança.

03. Armazenamento e processamento: seu foco está no armazenamento dos dados obtidos e seu processamento eficaz de acordo com as necessidades que temos.



Existências de padrões arquiteturais

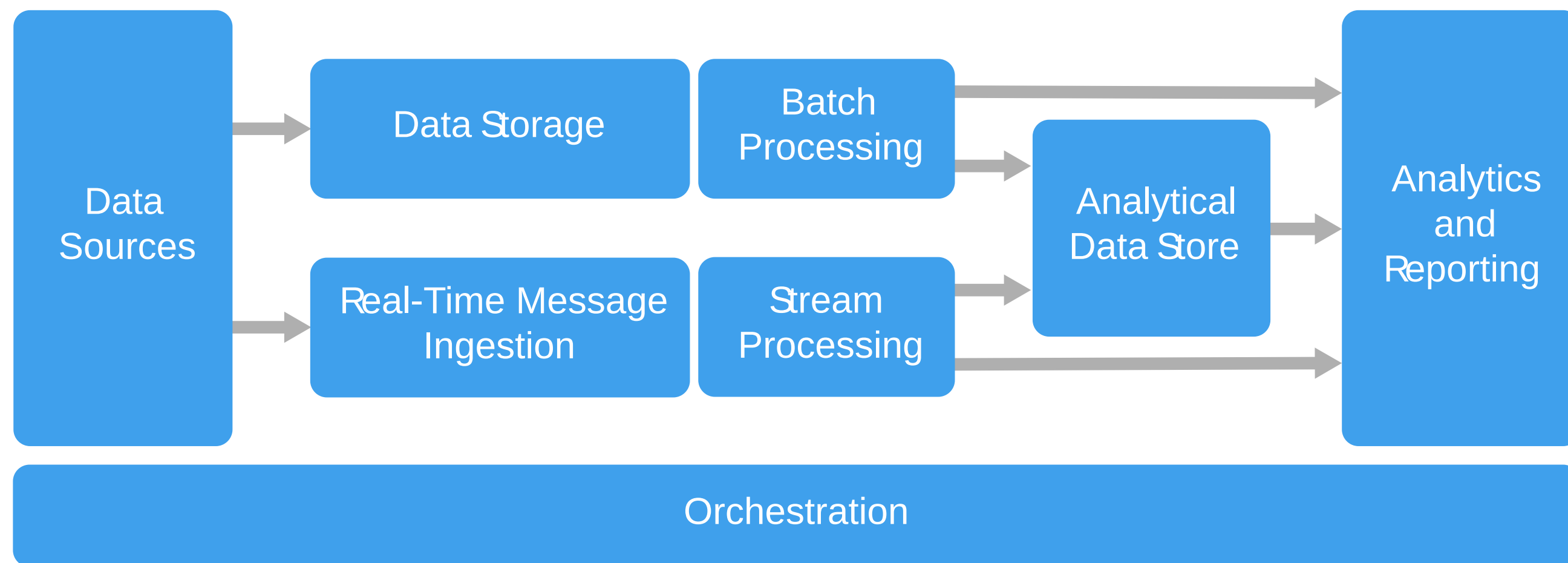
Características desejadas em uma arquitetura de Big Data:

- Reusabilidade
- Manutenibilidade
- Modularidade
- Performance
- Escalabilidade



Existências de padrões arquiteturais

Componentes de uma arquitetura de Big Data





Existências de padrões arquiteturais

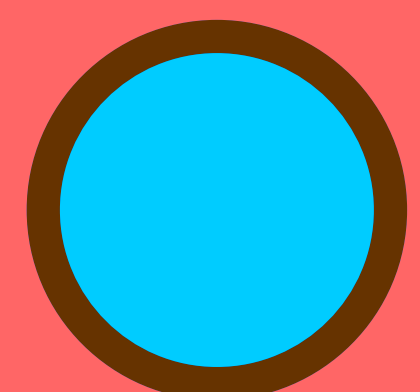
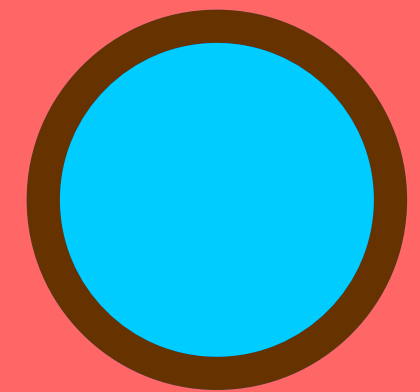
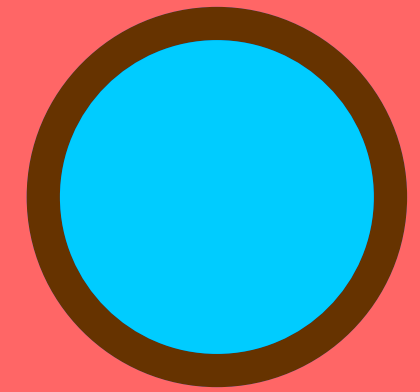
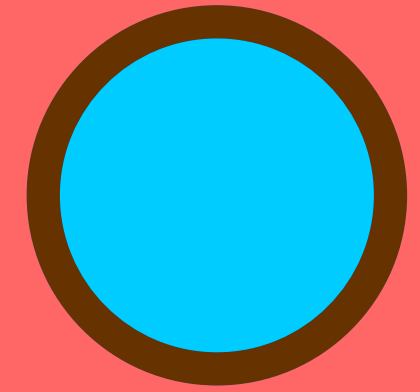
Componentes de uma arquitetura de Big Data:

- Data sources (Fonte de dados)
 - Um ou mais fontes de dados. Exemplos:
 - Armazenamentos de dados de aplicativo.
 - Arquivos estáticos produzidos por aplicativos.
 - Fontes de dados em tempo real;
- Data storage (Armazenamento de dados)
 - Processamento em lotes.
 - Repositório de arquivos distribuídos.
 - Repositório Data Lake.
 - Contêineres de blobs ou Azure Data Lake Store;

Existências de padrões arquiteturais

Componentes de uma arquitetura de Big Data:

- Batch processing (Processamento em lotes)
 - Conjunto de dados muito grandes.
 - Lotes de execução longa.
 - Ler, processar e gravar.
 - U-SQL no Azure Data Lake Analytics;
- Real-time message ingestion (Ingestão de mensagem em tempo real)
 - Fontes em tempo real.
 - Capturar e armazenar mensagens.
 - Armazenamento de dados simples.
 - Removidas para uma pasta para processamento;





Existências de padrões arquiteturais

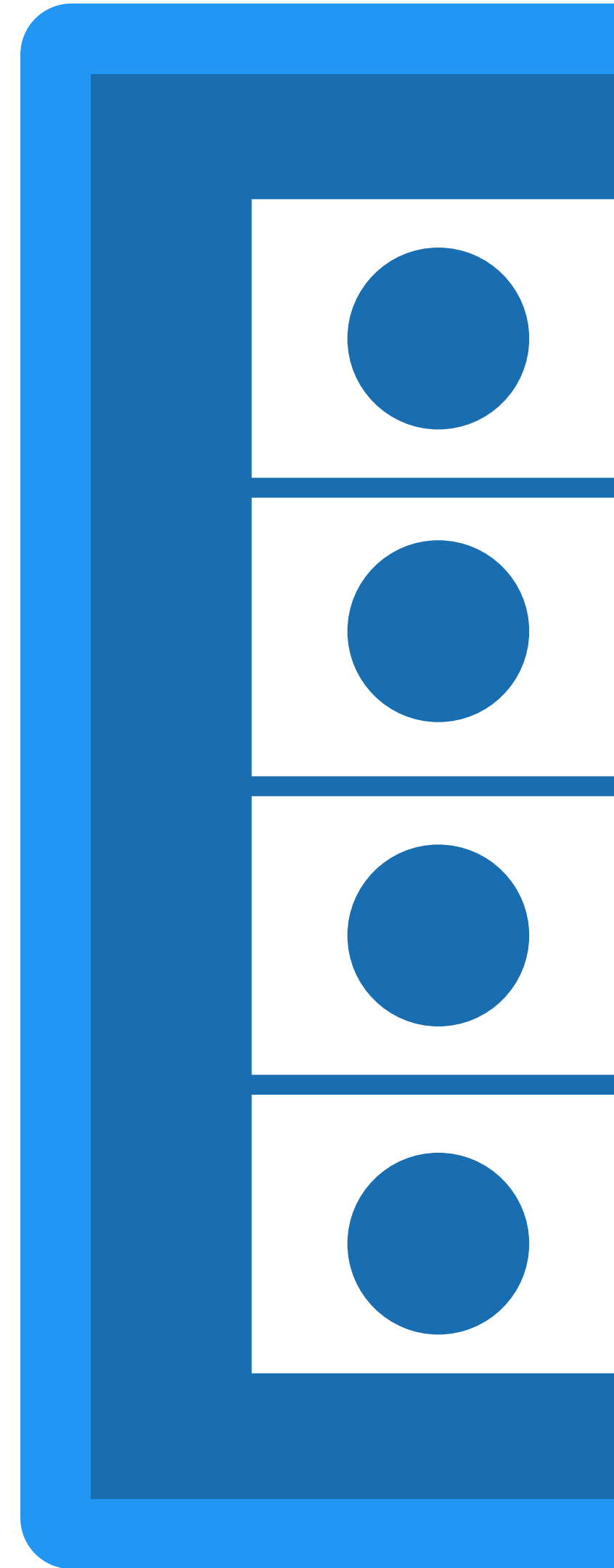
Componentes de uma arquitetura de Big Data:

- **Stream processing (Processamento de fluxo)**
 - Capturar mensagens em tempo real.
 - Processar filtrando, agregando e preparando os dados.
 - Coletor de saída.
 - Azure Stream Analytics (Consultas SQL, fluxos não associados);
- **Analytical data store (Armazenamento de dados analíticos)**
 - Prepara e fornece dados processados em um formato estruturado.
 - Data warehouse relacional estilo Kimball.
 - NoSQL de baixa latência (Alternativa).
 - Azure Synapse Analytics (Dados em larga escala);

Existências de padrões arquiteturais

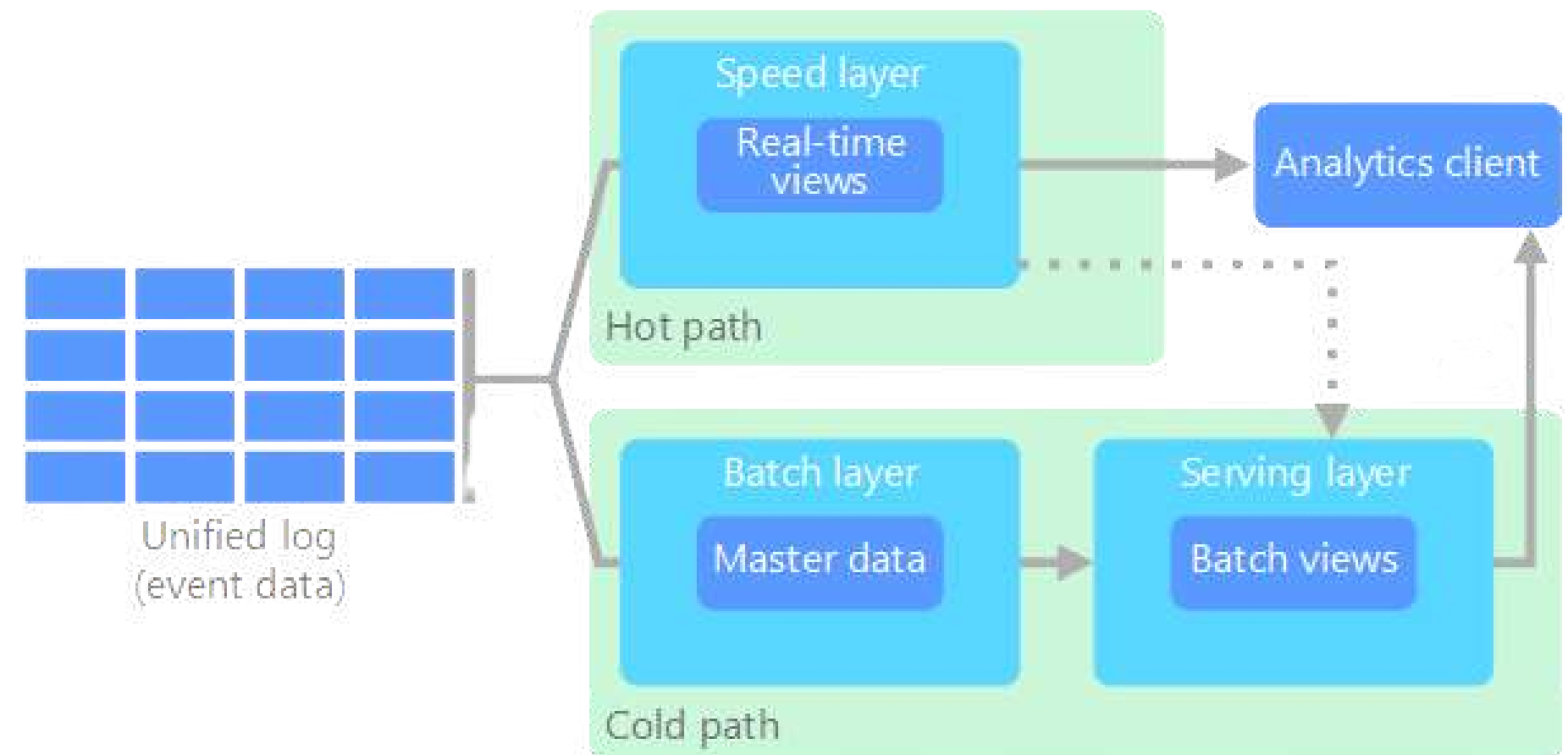
Componentes de uma arquitetura de Big Data:

- **Analytics and reporting (Análise e relatórios)**
 - Camada de modelagem de dados como modelo de dados tabular no Azure Analysis Services.
 - Pode dar suporte a Business Intelligence de autoentendimento, usando as tecnologias de modelagem e visualização do Power BI/Excel;
- **Orchestration (Orquestração)**
 - Forma de automatizar os fluxos de trabalho (Azure Data Factory);



Arquiteturas de Referências

- 1.0 Arquitetura Kappa
- 1.1 MapReduce
- 1.2 Latência
- 1.3 Lambda traz dois caminhos
 - 1.3.1 Camada de lote
 - 1.3.2 Camada de velocidade

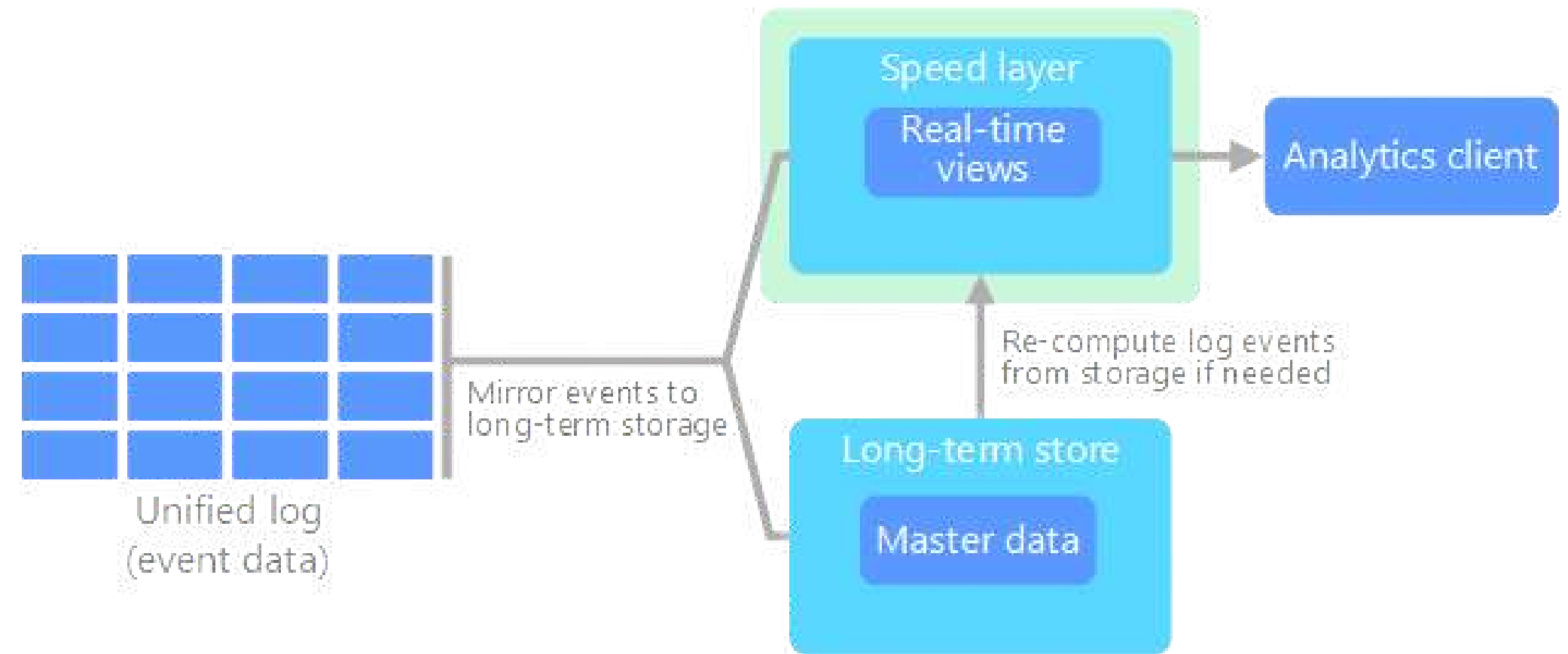


Arquiteturas de Referências

2.0 Arquitetura Kappa

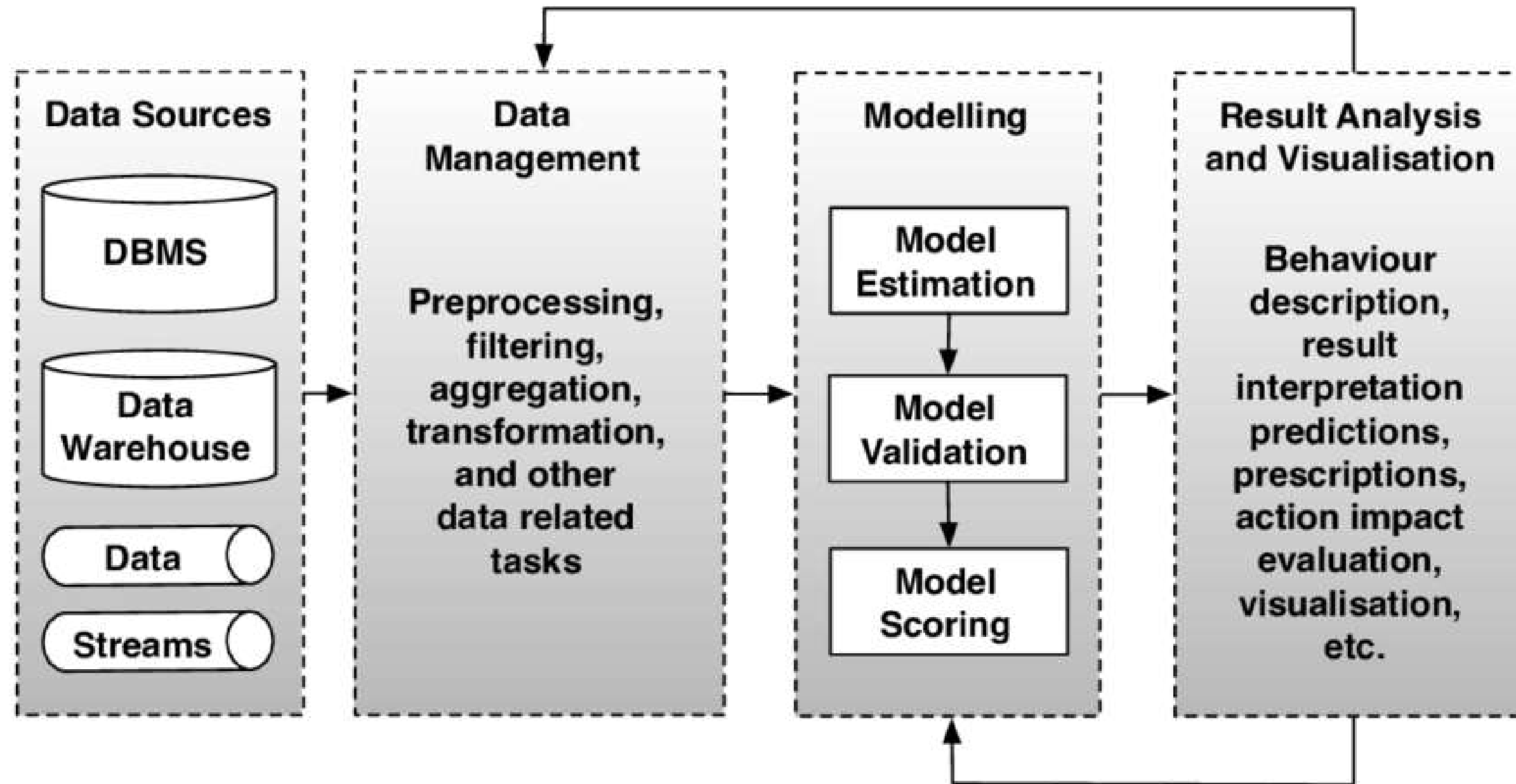
2.1 Alternativa da Lambda

2.2 Único caminho



Abordagens de Apoio

Fluxo de trabalho em Big Data(Processo)



Abordagens de Apoio

Ferramentas



Armazenamento e
Processamento

Análises de Dados e
Estatísticas

Visualização e
Relatórios



Desafios Enfrentados:

01

- Lidar com imenso volume de dados e a sobrecarga dos sistemas de processamento e armazenamento.

02

- Lidar com a complexidade e a diversidade dos dados.

03

- Garantir a segurança e privacidade dos dados.

04

- Lidar com a escalabilidade dos dados.

A close-up photograph of a person's hand holding a silver pen, pointing at a network diagram on a document. The diagram shows various nodes and connections, with labels like 'TOWN HOME', 'CONDO', and 'TRATTORIA' visible. The background is blurred, showing a laptop screen.

Referências

MARTINEKUAN. Arquiteturas de Big Data - Azure Architecture Center. Microsoft.com. Disponível em: <<https://learn.microsoft.com/pt-br/azure/architecture/data-guide/big-data/>>. Acesso em: 09 abr. 2023.

Assunção, M. D., Calheiros, R. N., Bianchi, S., Netto, M. A., & Buyya, R. (2015). Big Data computing and clouds: Trends and future directions. Journal of parallel and distributed computing, 79, 3-15.

DE, Analista. 5 V's do Big Data: como se aplicam no seu negócio. Cortex-intelligence.com. Disponível em: <<https://www.cortex-intelligence.com/blog/os-5-vs-do-big-data>>. Acesso em: 08 abr. 2023.



Agradecemos
a Atenção!