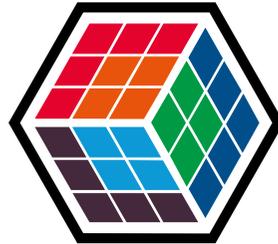


THE DEVELOPER'S CONFERENCE

Trilha – NO SQL

Alessandra Monteiro Martins

Especialista em Governança de TI pela Universidade Católica de Brasília,
Licenciada em Informática pela Universidade do Estado do Amazonas,
Certificações ITIL, COBIT, ISO27002, CTFL, KMPI, Scrum Master, CLF



THE DEVELOPER'S CONFERENCE

**Quebrando Paradigmas, Migrando
SQL para No SQL, Como não
"falir" seu Projeto**

Agenda



- Arquitetura de Dados – Definições e Diferenciações
 - Papéis – DA & DBA
 - Uso Atual – SQL
 - Quebrando Paradigmas
 - Uso Futuro – No SQL
 - Mapeamento de Dados
 - Modelagem de Dados
 - ETL
 - *Arquitetura de Software – Micro Serviços e Bancos NO SQL*
 - *Como não “falir” seu Projeto*
 - *O que não fazer – Não ignore as etapas*

Arquitetura de Dados: Diferenciações



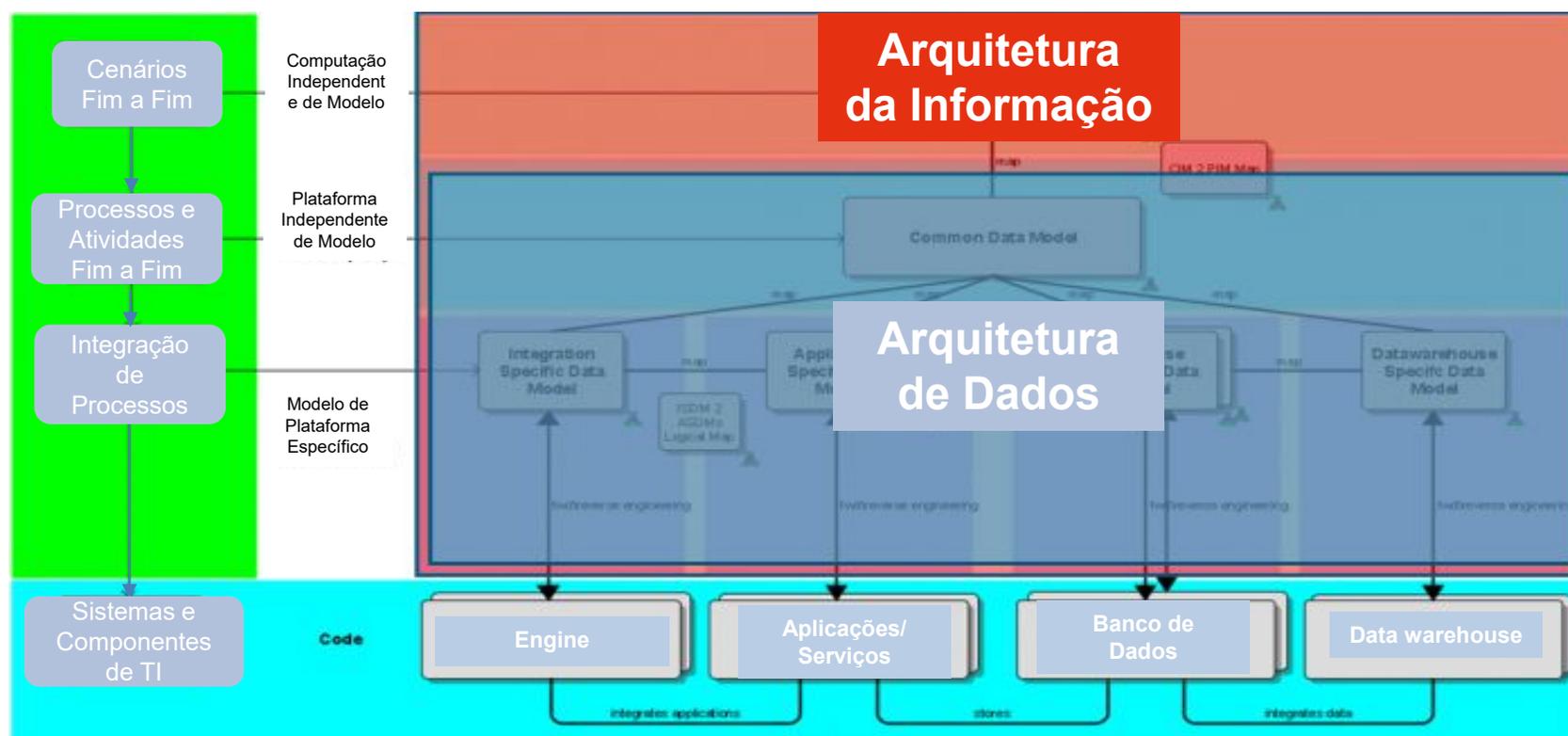
THE
DEVELOPER'S
CONFERENCE

- Dados não são Informações. Logo Arquitetura da Informação não é o mesmo que Arquitetura de Dados, porém, devem trabalhar em conjunto.
- A arquitetura da informação refere-se ao desenvolvimento de programas projetados para inserir, armazenar e analisar informações significativas. A arquitetura de informações interpreta os pontos de dados individuais em informações úteis e significativas.
- A arquitetura de dados é o desenvolvimento de programas que interpretam e armazenam dados. A arquitetura de dados define a coleta, armazenamento e movimentação de dados em uma organização.

Arquitetura de Dados: Diferenciações



THE
DEVELOPER'S
CONFERENCE



Arquitetura de Dados: Definições



THE
DEVELOPER'S
CONFERENCE

- A arquitetura de dados descreve a estrutura de dados utilizada por uma organização e/ou seus aplicativos e contempla descrições de dados - tanto armazenados quanto em movimento, descrições de meios de armazenamento, grupos de dados, itens de dados e modelos de dados de soluções de TI.
- Essencial à concepção da situação futura, a Arquitetura de Dados descreve como os dados são processados, armazenados e utilizados em um determinado sistema (ou conjunto de sistemas). Ela fornece os critérios para as operações de processamento de dados, possibilitando que sejam projetados e também controlados os fluxos de dados no sistema.

Arquitetura de Dados: Definições



Data Architecture



Papéis: DA x DBA



THE
DEVELOPER'S
CONFERENCE

DA

PREOCUPA-SE

Com a definição da situação futura, pelo alinhamento durante o desenvolvimento, e pelo acompanhamento para garantir que melhorias sejam feitas sempre de acordo com as especificações arquitetônicas originais, desde a arquitetura de sistemas de base de dados, modelagem dos dados e administração dos dados de uma organização. Se ocupa de trabalhar os dados como um recurso estratégico da organização, representando-os independentemente dos processos das diferentes unidades que os utilizam.

ENVOLVE-SE

O Arquiteto de Dados deve ocupar do projeto das bases de dados e sua estrutura lógica, não se envolvendo diretamente nos aspectos físicos. No que tange a administração está mais especificamente voltada para manutenção dos modelos e dicionários de dados.

DBA

PREOCUPA-SE

Com o gerenciamento, instalação, configuração, atualização e monitoração do banco de dados ou sistema de bancos de dados. Deve zelar pela integridade e saber modificar a estrutura do Banco de Dados. Garantir Acesso e o máximo de desempenho as consultas ao Banco de dados.

ENVOLVE-SE

O DBA deve se ocupar da administração diária dos bancos de dados físicos e seu conteúdo, efetivamente uma administração dos sistemas gestores de Bases de Dados. Realizar atualizações, testes e backup.

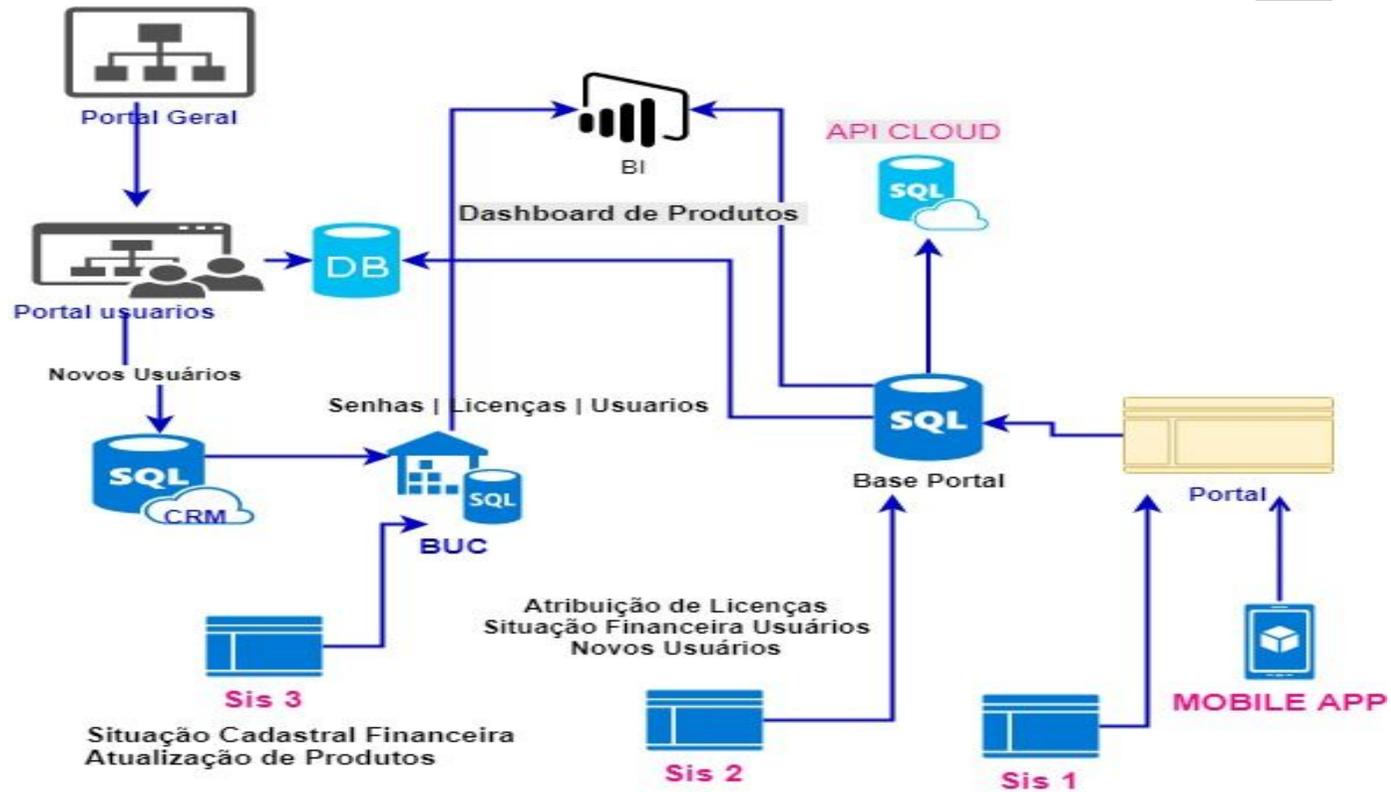
Uso Atual - SQL



- Arquitetura de Dados Tracionais utilizam linguagem Estruturada – SQL, porém, isso causa um certo trabalho quando temos atributos muito específicos ou que sofrem uma atualização ou mudança periódica, por exemplo: uma banco de dados de produtos de uma aplicação de e-commerce, você tem diversos produtos.
- Em uma arquitetura SQL tradicional cada tabela poderá ter varias tabelas auxiliares com PK e FK para gerenciar as ligações e dependências, tudo para poder tratar dos produtos com atributos diferentes, a manutenção e atualização destes bancos pode ser trabalhosa, ou demandar mais tempo causando indisponibilidades.



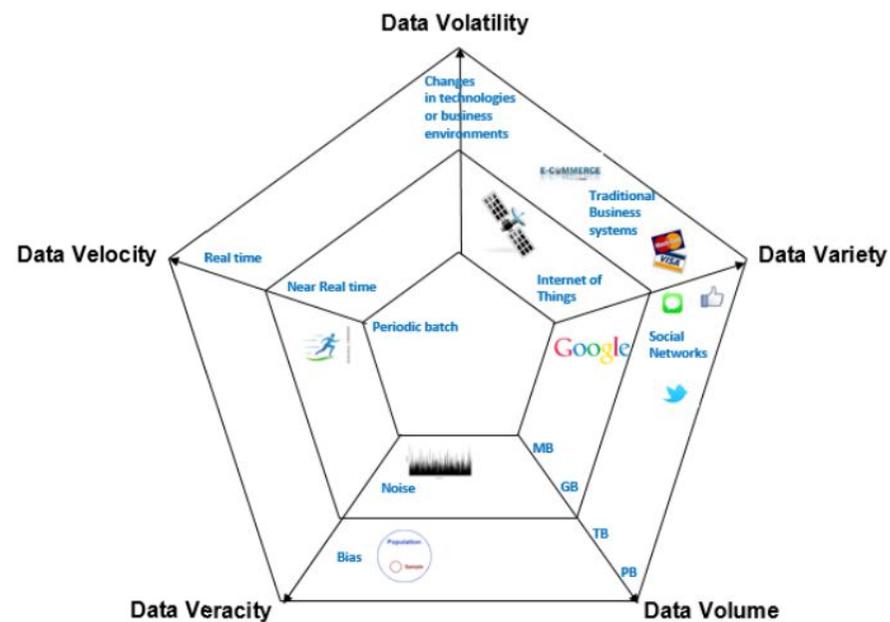
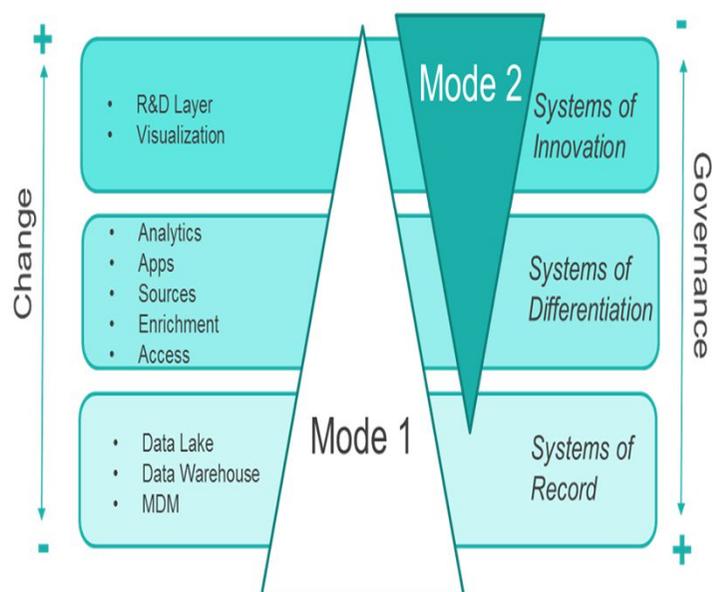
Uso Atual - SQL



Quebrando Paradigmas



Gartner's Bimodal IT & Modern Data Architecture



Uso Futuro - NO SQL

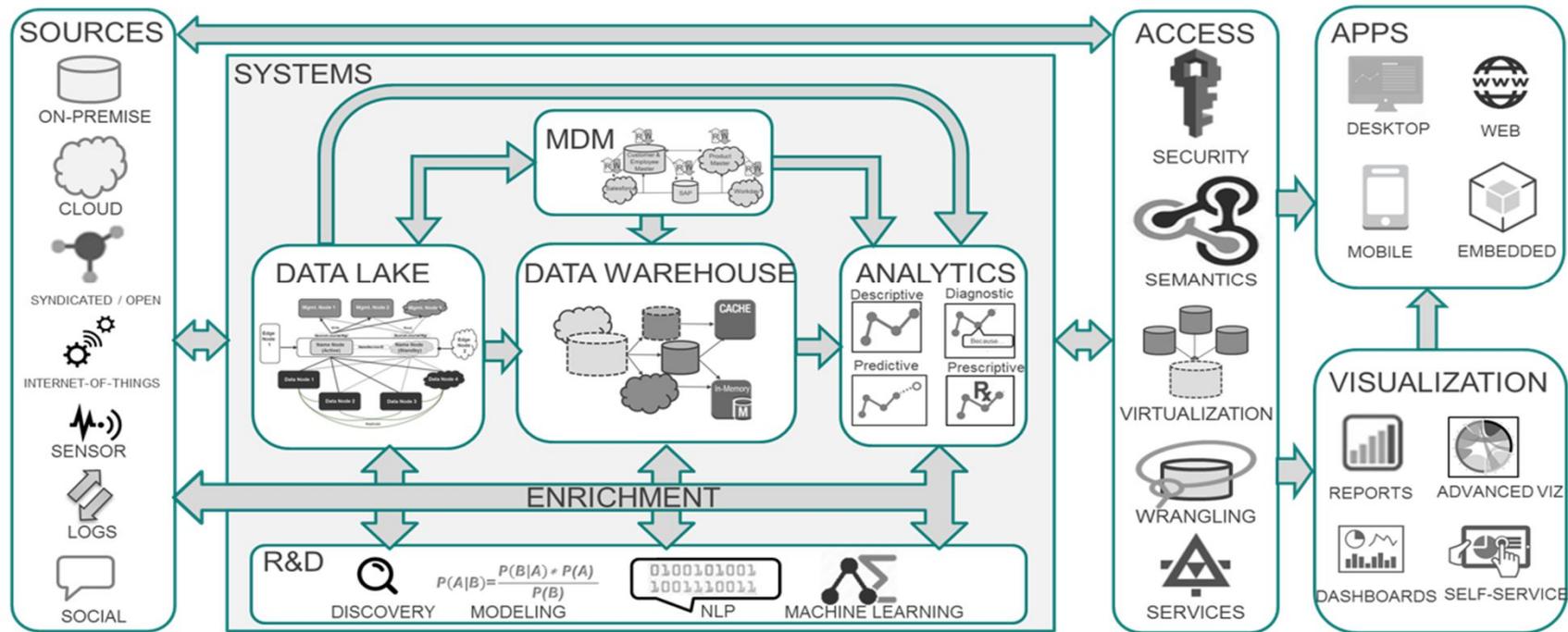


- Com a demanda crescente por informações para tomada de decisão cada vez mais rápidas, pela propagação da Internet das Coisas, Big Data, Machine Learning, cada vez mais cresce a demanda pela coleta, tratamento, processamento, saída e compartilhamento de dados estruturados e não estruturados.
- Pensando no volume das informações, quantidade de tipos e fontes, não é possível, mais pensar em um modelo linear, apenas estruturado, assim, o No SQL se apresenta como uma solução para o uso atual e futuro de dados, sendo essencial pensar numa arquitetura “Bimodal” ou híbrida.

Usó Futuro - NO SQL



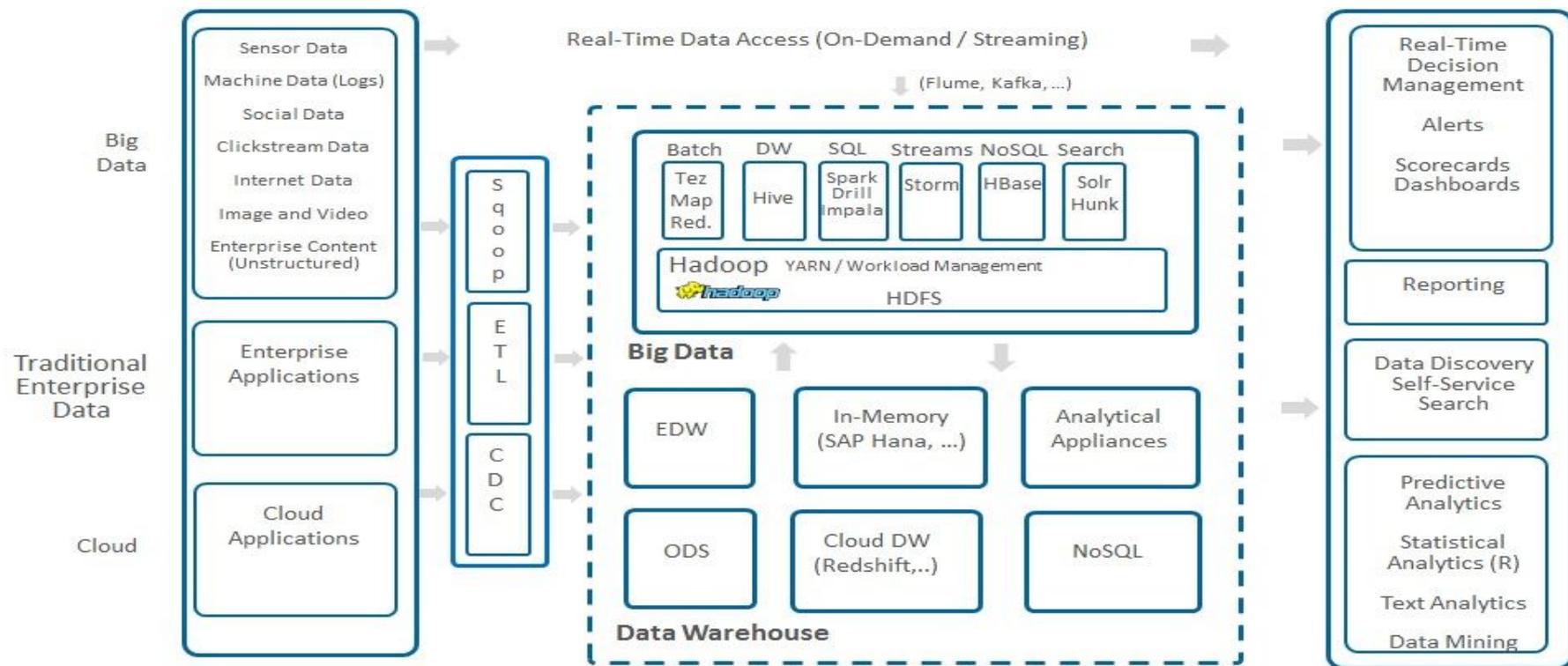
Modern Data Architecture



Usó Futuro - NO SQL



THE DEVELOPER'S CONFERENCE



Mapeamento de Dados



- 1º Passo para não “Falir” seu projeto: ouça as pessoas envolvidas (key users) do sistema atual ou que vão utilizar o sistema a ser construído;
- Verifique a documentação do sistema, caso exista, mesmo que desatualizada;
- Mapeie os processos que as pessoas executam e com base neles verifique a documentação do sistema;
- Então crie o mapeamento dos dados, sim, para cobrir o ciclo de vida do dado é preciso saber do input ao processamento e output.
- O gerenciamento do ciclo de vida de dados refere-se aos processos automatizados ou não que enviam dados de um estágio para o próximo ao longo de sua vida útil até que, finalmente, se tornem obsoletos e sejam excluídos de um banco de dados.
- O gerenciamento do ciclo de vida das informações analisa questões como se um dado é ou não útil e, se sim, como?

Mapeamento de Dados



THE
DEVELOPER'S
CONFERENCE

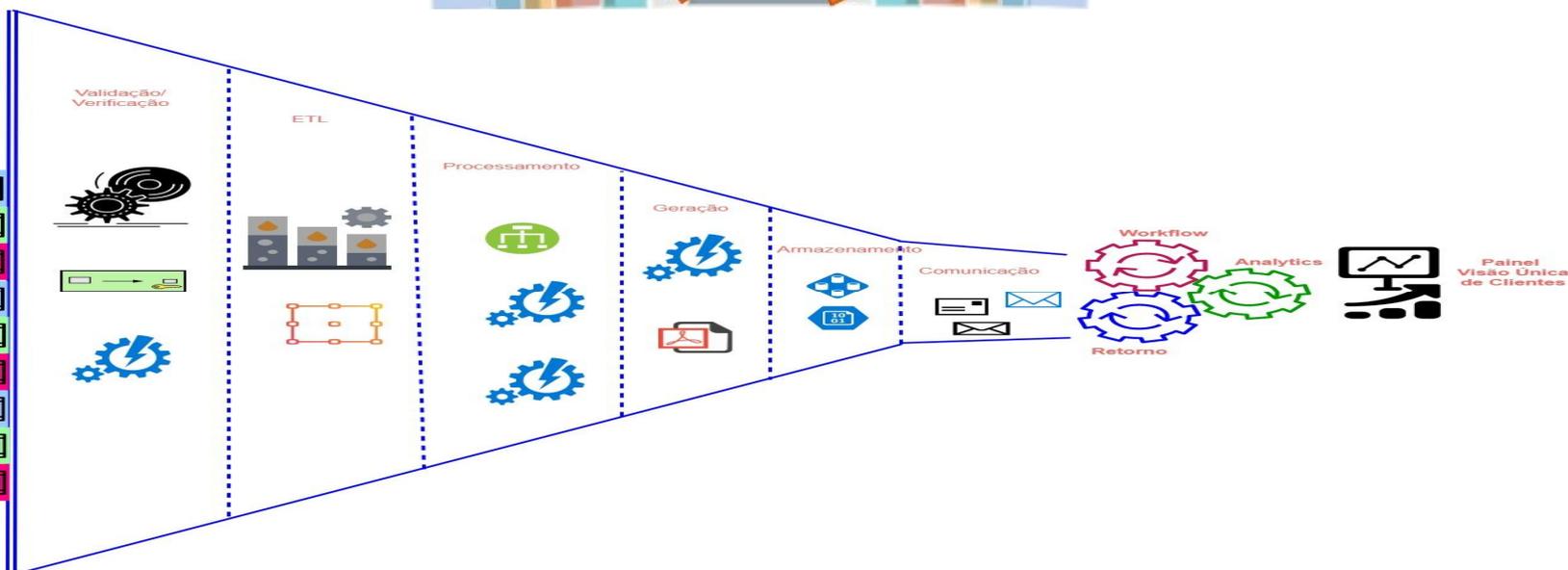
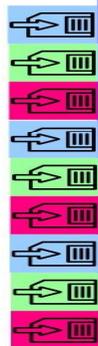


Recursos:

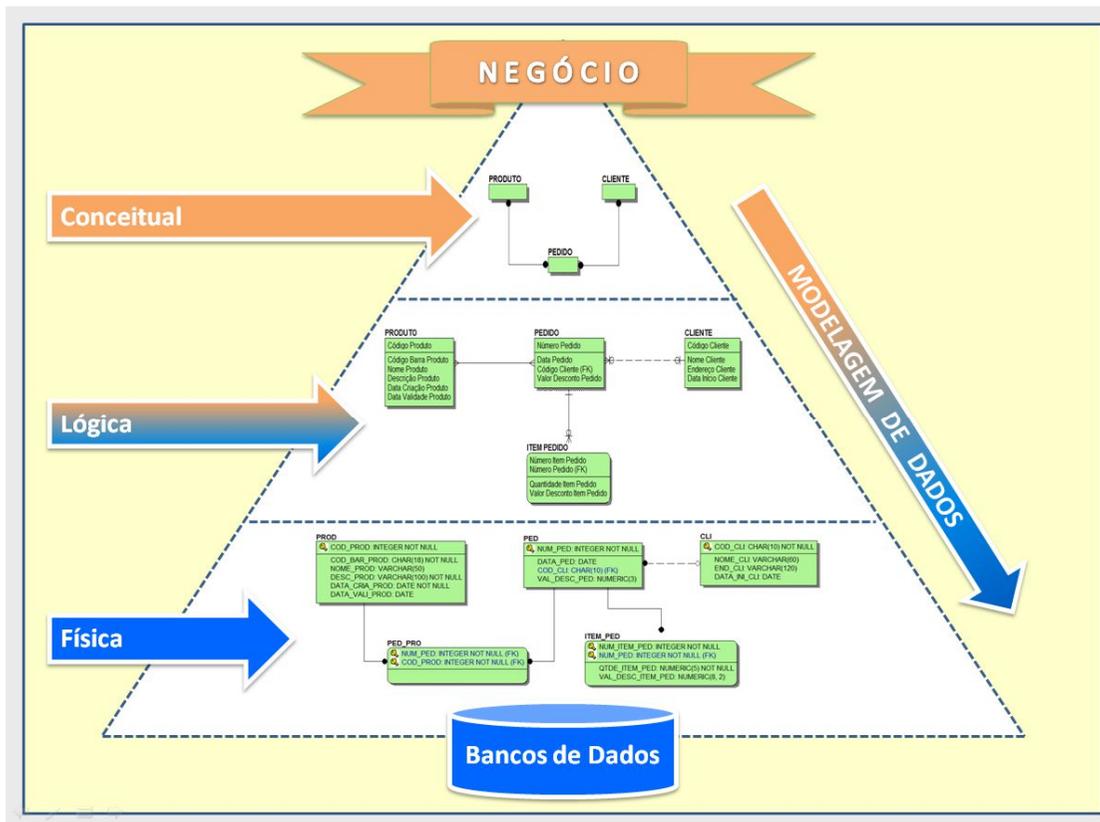
Arquivos de

Entrada de Dados:

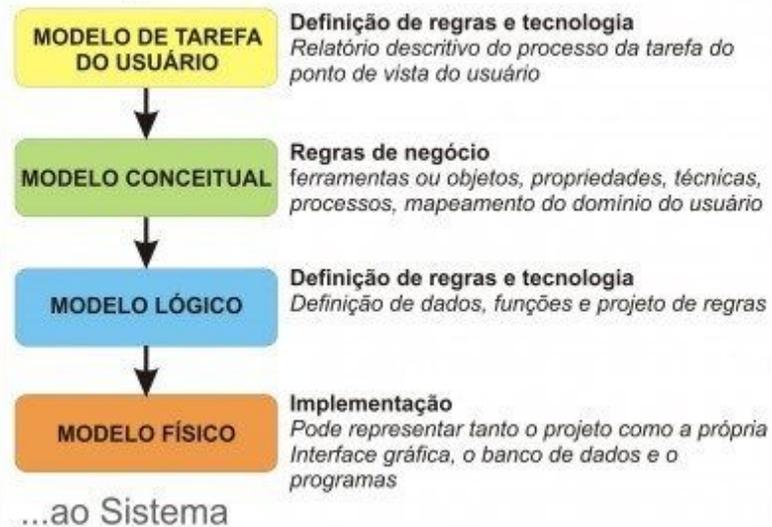
- PDF
- XLS
- CSV
- XML
- JSON
- SPOLL
- FAT
- TXT
- ZIP



Modelagem de Dados



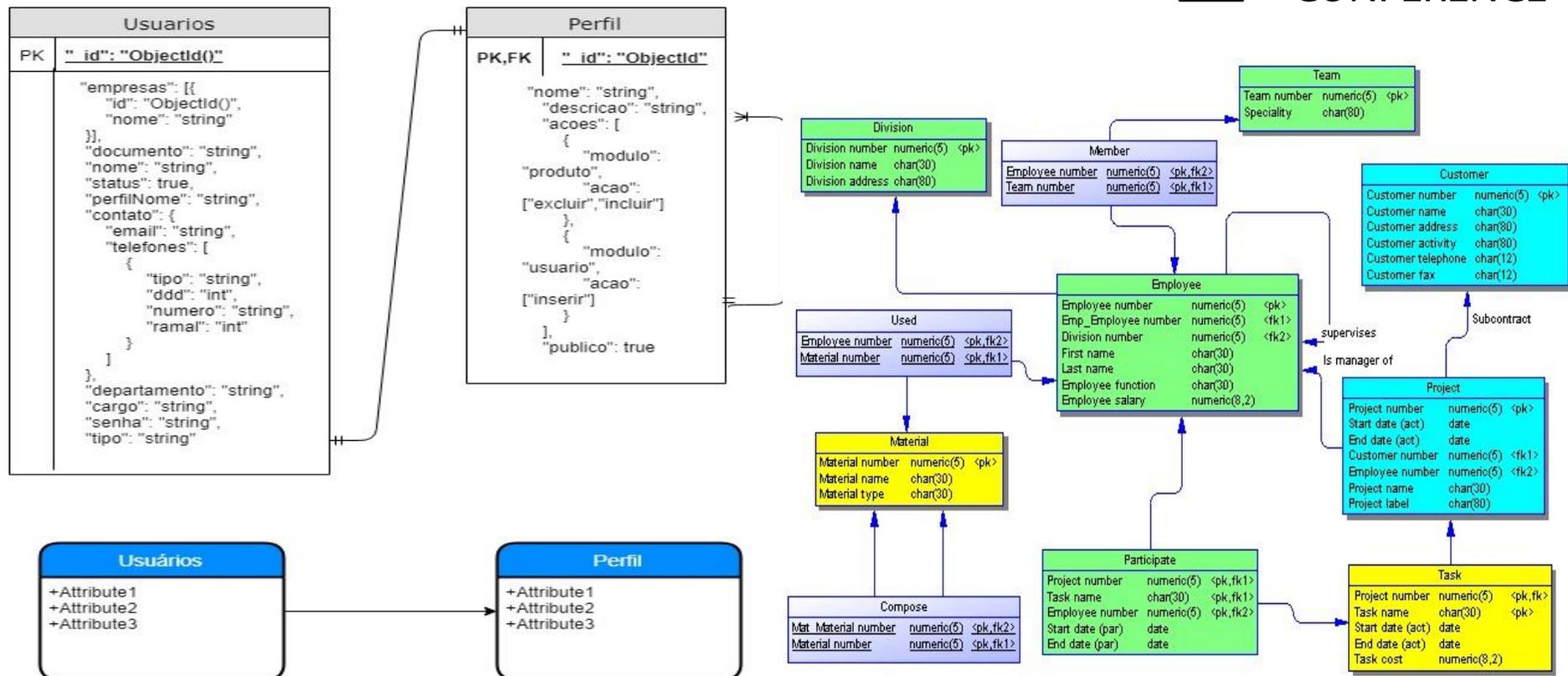
Do Negócio...



Modelagem de Dados



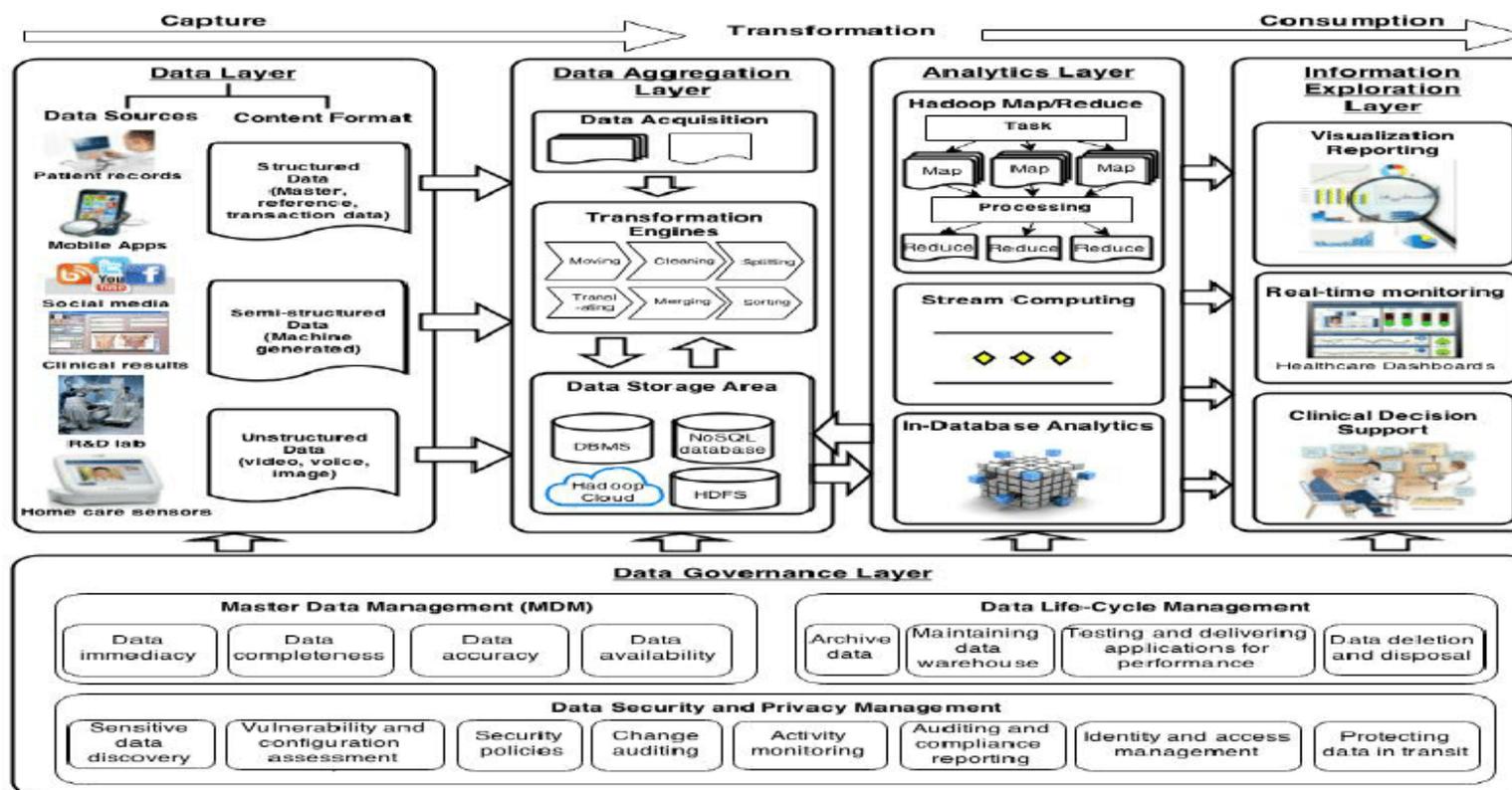
THE DEVELOPER'S CONFERENCE



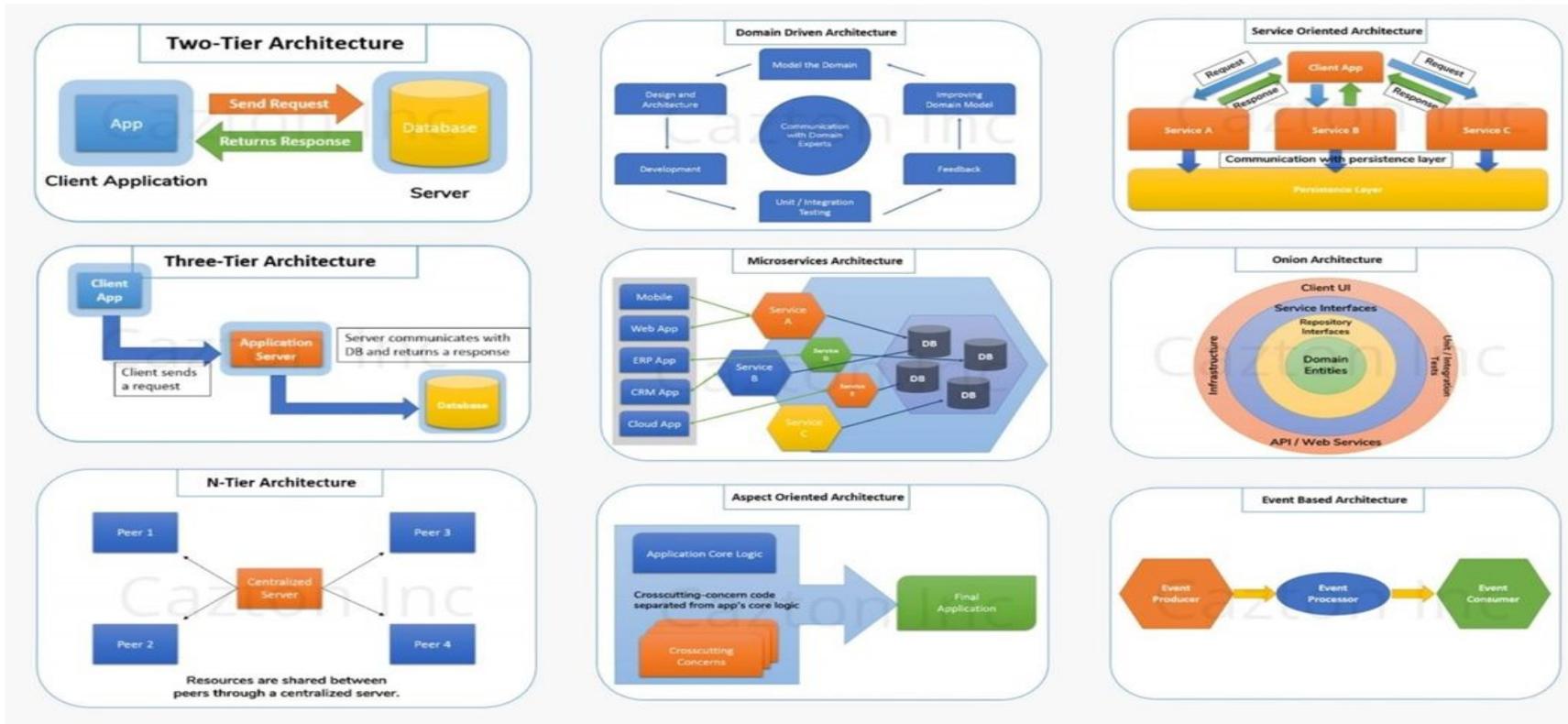
ETL – Extração Transformação Carga



THE DEVELOPER'S CONFERENCE



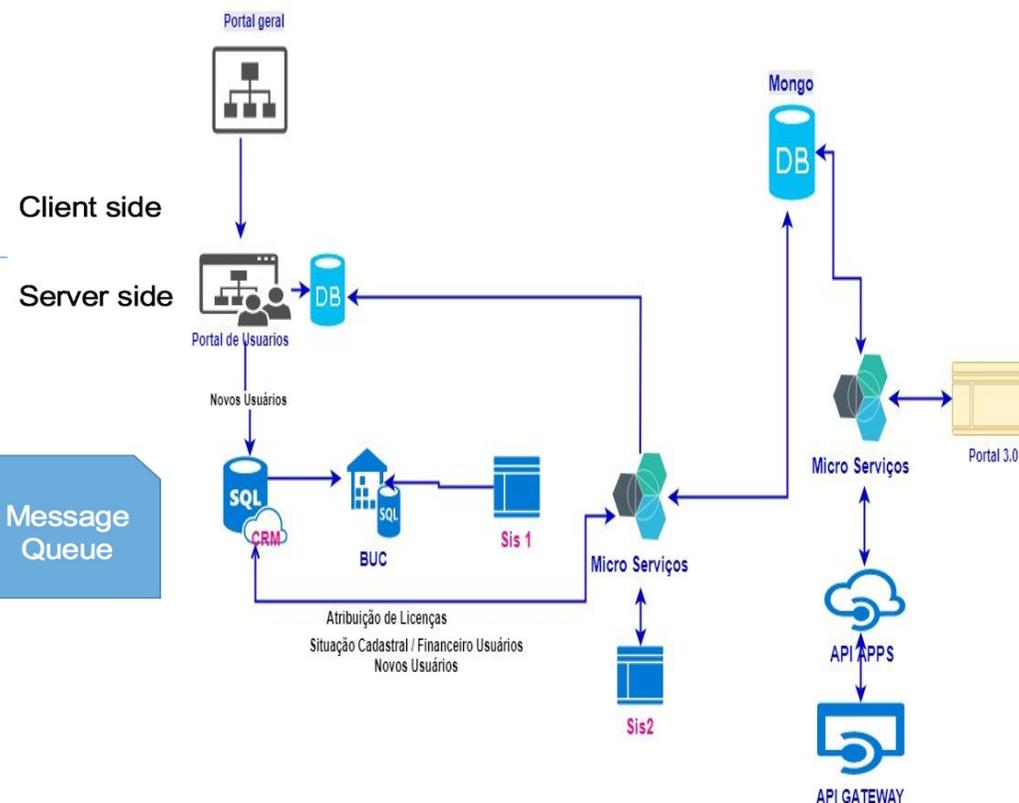
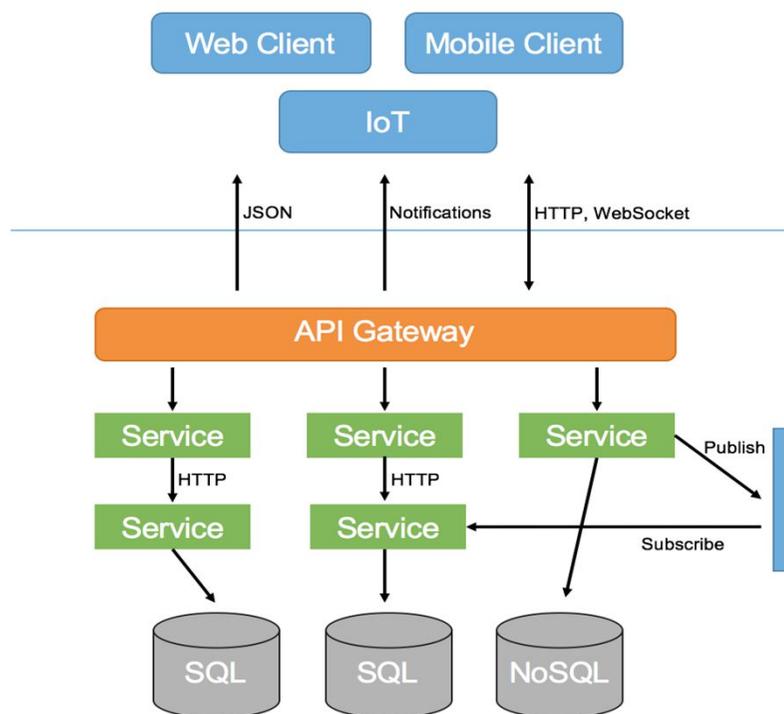
Arquitetura de Software



Arquitetura de Software – Micro Serviços e Bancos NO SQL



THE DEVELOPER'S CONFERENCE



Como Não “Falir” seu Projeto



- Planejamento: Não construa sem as pessoas certas, é preciso planejar, desenhar a arquitetura de software e dados, e alinhar as expectativas de ambas, para que uma má escolha do banco de dados não interfira no custo\$, performance ou desempenho da aplicação que está sendo construída.
- Adote um padrão de Design de Desnormalização - armazene alguns dados de uma tabela filha em cada tabela relacionada. Apenas deve-se armazenar os dados que serão mais consultados, para que o banco de dados não fique muito inchado. Quando você consultar a tabela pai, recebe de volta todos os dados relacionados, pois são armazenados em cada registro na tabela pai.
- Envolve as pessoas corretas, os profissionais com as habilidades necessárias, Arquiteto e Engenheiros de software e de Dados.
- Estratégia de Migração: Modele, Extraia, Migre em etapas, crie o dicionário de dados, Teste a Carga, a Performance, Crie índices, mas somente o necessário, se for preciso mantenha algumas operações transacionais, se for uma questão de segurança do Core business.

O que não fazer: não ignore as etapas



THE
DEVELOPER'S
CONFERENCE

- Arquitetura de Dados deve estar bem alinhada a Arquitetura de Software
- Não deixe o Banco no “Default”, ou pior não deixe como uma responsabilidade do desenvolvedor, já que se trata de NO SQL, e com um JSON você cria uma collection.
- Defina uma chave composta

The Common NoSQL Data Models



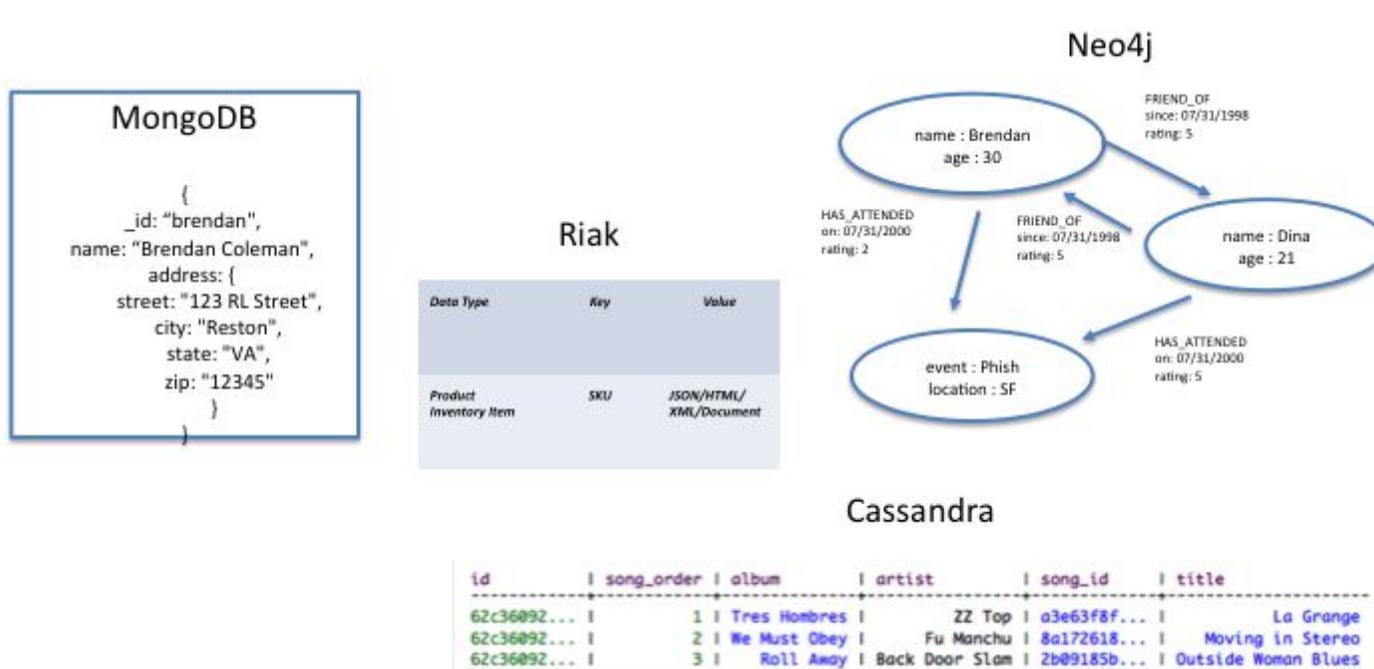
JSON documents are popular because they easily model:

- Hierarchical/nested data
- Evolving data
- Varying data

Defina um Schema – Nada de Default



THE
DEVELOPER'S
CONFERENCE

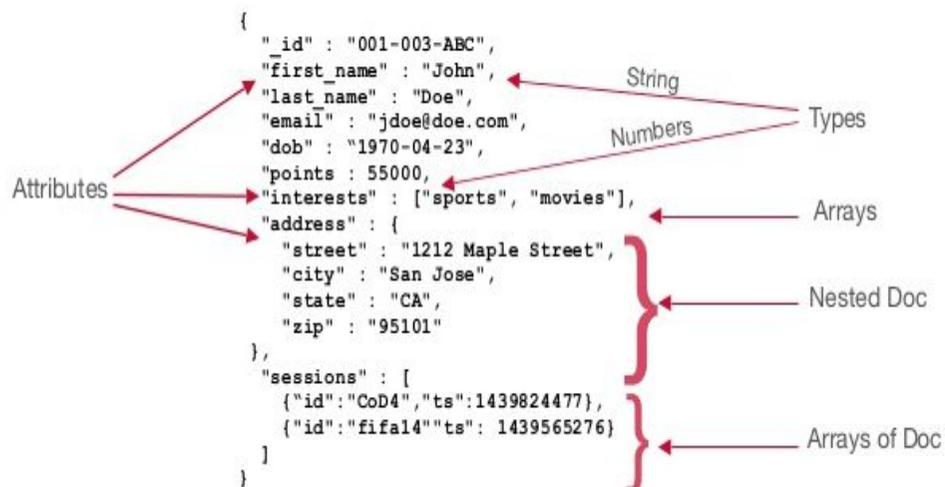


Defina um Schema – Nada de Default



THE
DEVELOPER'S
CONFERENCE

JSON Document : Flexible Schema



Product Catalog - NoSQL/Document

```
{
  "_id" : "rp-prod132546",
  "name" : "Marvel T2 Athena",
  "brand" : "Pinarello",
  "category" : "bike",
  "type" : "Road Bike",
  "price" : 2949.99,
  "size" : "55cm",
  "wheel_size" : "700c",
  "frameset" : {
    "frame" : "Carbon Torayaca",
    "fork" : "Onda 2V C"
  },
  "groupset" : {
    "chainset" : "Camp. Athena 50/34",
    "brake" : "Camp."
  },
  "wheelset" : {
    "wheels" : "Camp. Zonda",
    "tyres" : "Vittoria Pro"
  }
}
```



products
.findById("rp-prod132546")

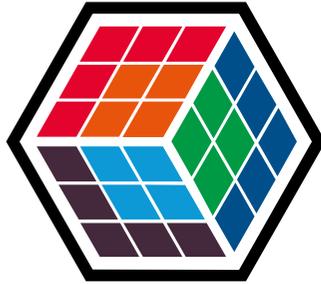
Store the product "as a business object"

To get a single product

Referências:



- https://pt.wikipedia.org/wiki/Arquitetura_de_dados
- <https://www.bmc.com/blogs/data-architecture-vs-information-architecture/>
- <http://geekswithblogs.net/SabotsShell/archive/2006/07/11/84708.aspx>
- <https://www.ariscommunity.com/users/koiv/2010-08-05-information-architecture-big-picture>
- <https://br.hortonworks.com/blog/enterprise-hadoop-journey-data-lake/>
- <http://www.datadansandler.com/>
- BIG DATA: POTENTIAL, CHALLENGES, AND STATISTICAL IMPLICATIONS, IMF Staff Discussion Note; SDN/17/06/2017
- https://en.wikipedia.org/wiki/Operational_data_store
- <http://www.datamodelling.com.br/biblioteca/conceitos/>
- https://www.researchgate.net/figure/The-best-practice-of-big-data-architecture-in-a-health-care-organization_fig1_266613537
- <https://www.devmedia.com.br/modelagem-de-dados-conceitual-construindo-pontes-entre-dados-e-negocios/30597>
- <http://www.codingthearchitecture.com/presentations/sa2008-why-software-projects-fail/>
- <https://highlyscalable.wordpress.com/2012/03/01/nosql-data-modeling-techniques/>



THE DEVELOPER'S CONFERENCE