

Microsoft Data Base Performance Seminar



Odilon Almeida oalmeida@Microsoft.com

Big Data and Data Warehouse Pre-Sales Lead

Microsoft

O MUNDO DOS DADOS ESTA MUDANDO !!!

10x a
cada cinco
anos



85%
relacionados a
novos tipos de
dados

Explosão dos
Dados

Consumerização
de TI

4.3
devices
conectados
por adulto



27%
usando
social
media

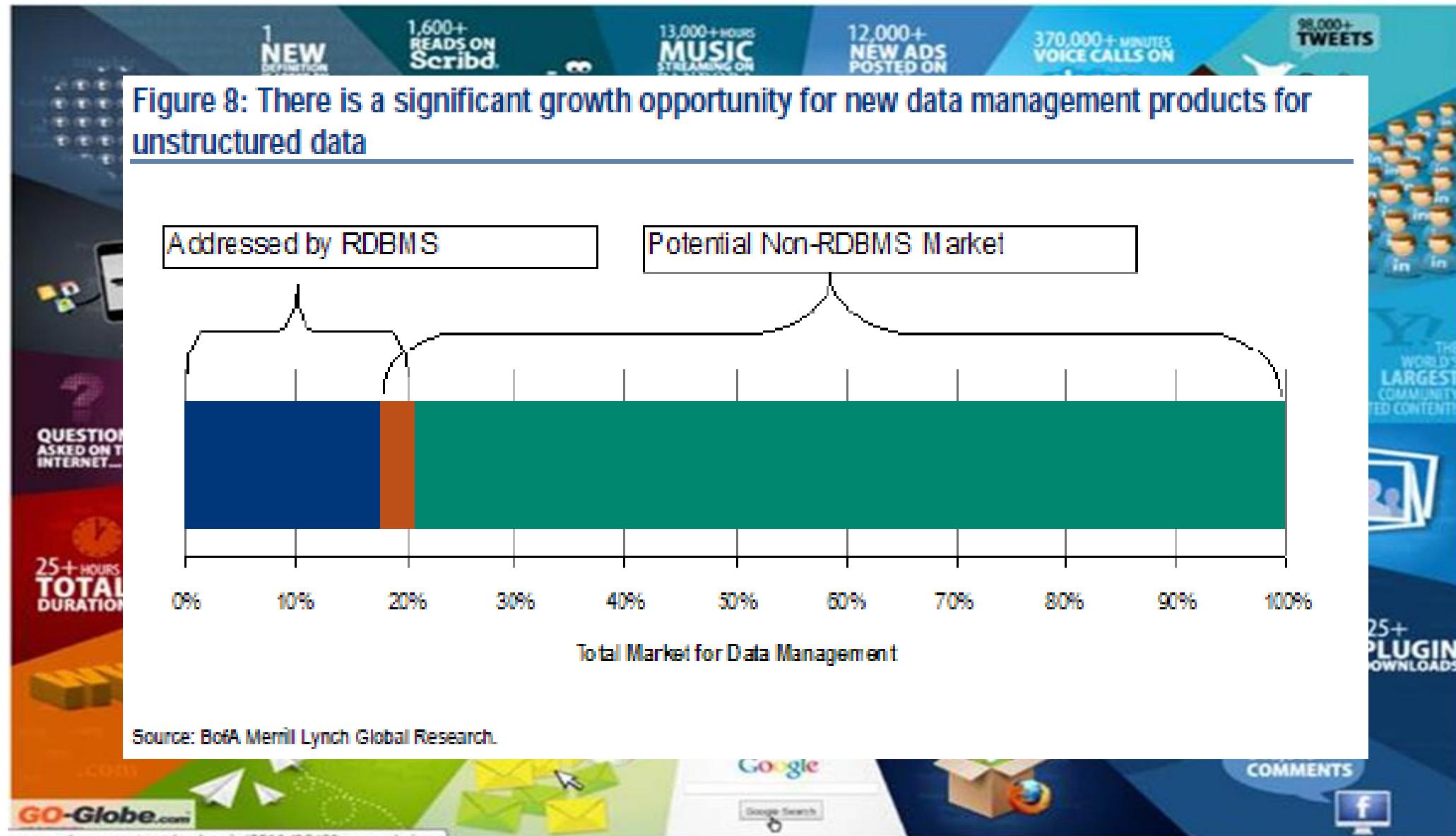
“

Até 2015, organizações que modernizarem seus sistemas de gestão da informação ultrapassarão financeira seus pares em 20 porcento.

– Gartner, Mark Breyer, “Information Management in the 21st Century”

”

INFORMAÇÃO GERADA POR MINUTO !



O QUE É BIG DATA?

A quantidade de dados em nosso mundo **tem crescido muito**, e a **análise de grandes conjuntos de dados**, os chamados **Big Data**, se tornará uma peça chave-base de competição, sustentando novas ondas de crescimento da produtividade, inovação e aumento de consumidores, de acordo com pesquisa realizada pela **McKinsey Global Institute e McKinsey Business Technology Office.**

http://www.mckinsey.com/insights/mgi/research/technology_and_innovation/big_data_the_next_frontier_for_innovation

MUNDO DO BUSINESS INTELLIGENCE HOJE

Apenas 28% dos potenciais usuários de BI possuem acessos adequados

– *Business Intelligence Purchase Drivers and Adoption Rates, Gartner*

32% dos usuários de Excel estão confortáveis com seu uso para “Advanced Analysis”
Ex.: PivotTables

– *Excel Segmentation Research*

31% de todas as iniciativas de BI atendem apenas parte ou não atendem as expectativas originais do Negócio

– *BI Survey 9, BARC*

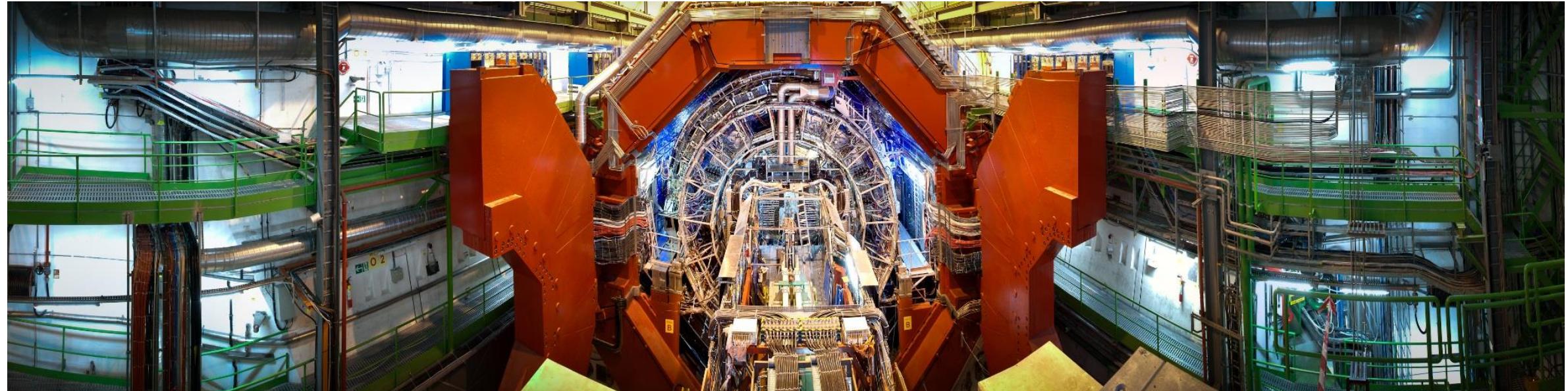
“Profissionais experientes passam muito tempo buscando a informação, sem tempo suficiente para analisá-la.”

“As barreiras entre aqueles que produzem e consomem a informação estão desaparecendo”

“Os Requisitos mudam muito rapidamente para a TI e tradicional tecnologia acompanharem.”

O QUE UM GRANDE VOLUME DE DADOS

Exemplo de utilização para Big Data



O maior Hadron Collider produz
1 PB/sec

[world's largest and highest-energy particle accelerator](#)

POR QUE BIG DATA?

Porque é legal !!!

Falta video neste slide, retirado pelo tamanho !!!!!

BIG DATA



Novas Fontes de Dados



Tipos
de dados não usuais em TI



Grandes Volumes de Dados



Novas Tecnologias

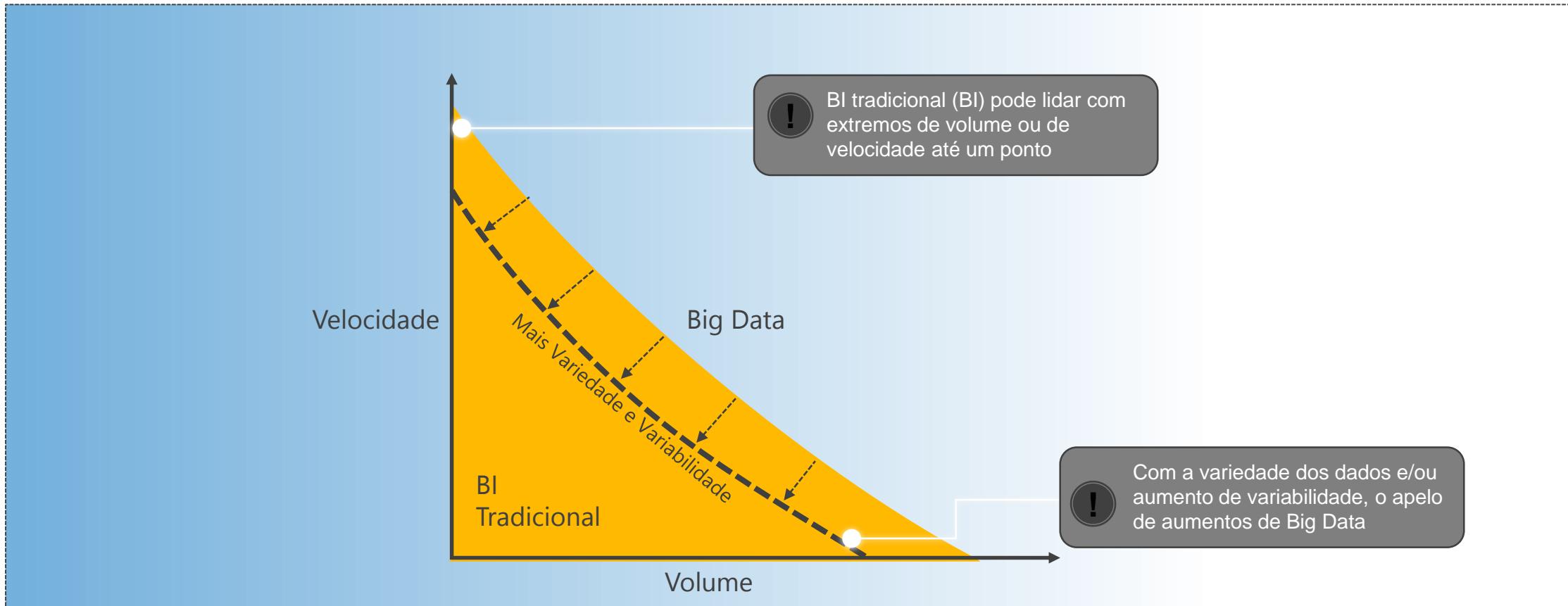


Novos aspectos econômicos

Novas Questões & Novos Insights



CARACTERÍSTICAS DO BIG DATA



Volume – Excede os limites físicos de escalabilidade vertical

Variedade – Muitos formatos diferentes fazem integração trabalhosa

Velocidade – Pequena janela de decisão em relação a taxa de variação dos dados

Variabilidade – Muitas opções ou interpretações de variáveis podem confundir análise

FONTES – BIG DATA

Há inúmeras fontes para Big Data e os diferentes tipos de dados que eles criam, mas podem ser classificados em três categorias aceitas pelo mercado:

- **Dados estruturados:** é o tipo que se encaixaria perfeitamente em um Sistema de Gerenciamento de Banco de Dados Relacional padrão, RDBMS, e prestar-se a esse tipo de processamento;
- **Dados Semi-estruturados:** é o que tem algum nível de comunalidade mas não se encaixa no tipo de dados estruturados;
- **Dados Não-estruturados:** é o tipo que varia em seu conteúdo e pode mudar de entrada para entrada.

Estruturado

Registro de Clientes

Dados de Pontos de Vendas

Inventário

Registros Financeiros

Semi-Estruturado

Web-Logs

Dados de Midia Social

E-Commerce

Dados Geológicos

Não-Estruturado

Fotos

Dados de edição de vídeo

Produtividade

Documentos MS-Office

BIG DATA – ASSUNTO DA MODA

A. Novas fontes e tipos de dados de ***dentro & fora*** das organizações

- “Dados da Internet”, sensores, RFID, dispositivos inteligentes, carros, etc;
- Informação não-estruturada - documentos, web logs, email, tudo que esteja na Web
- Fontes de terceiros confiáveis – provedores a indústria & agregadores, informações governamentais abertas “Open Data”, informações de clima, redes sociais, etc.

B. Inovações Tecnológicas para Explorar **Um Novo Mundo de Informações**

- Storage de baixo custo e capacidade de processamento (Nuvem, On-Site & Híbrida)
- Novos padrões de software para lidar com velocidade e volume, estruturados e não estruturados (In-memory computing, Hadoop, MapReduce, etc)
- Revolução na experiência do usuário – information precise no momento correto
 - Preditivo, analítico, recomendação de dados
 - Não esqueças das respostas automáticas / sistemas tuning próprio e automático

A + B = Vantagem Competitiva

NOVAS QUESTÕES DE NEGÓCIO

Solicitações do negócio numa economia extremamente competitiva



Social & Web
Analytics

Qual é o sentimento da
sociedade sobre minha
marca ou produtos



Live Data Feeds

Como otimizo minha
frota com base em
padrões metereológicos
e de tráfego?



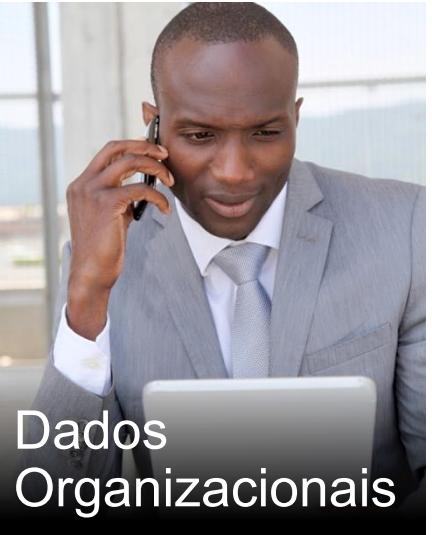
Advanced Analytics

DESCUBRA E CONECTE

Respondendo Novas Questões de Negócio



Dados
Pessoais



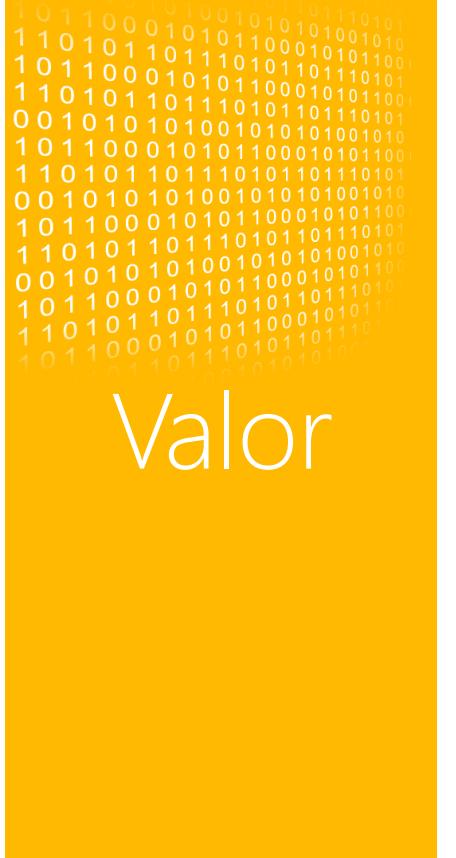
Dados
Organizacionais



Dados da
Comunidade



Dados do
Mundo



BIG DATA É UMA CONVERSA SOBRE NEGÓCIOS

- **Telecom & Media**

- Unir informações de Rede com as informações de clientes para melhor planejamento
- Prevenção de Churn
- Novas ofertas de Marketing
- Percepção da Marca (Social Network)

- **Finanças**

- Análise de Mercado
- Utilização de fontes externas de dados
- Análise de Risco Cliente/Empresa
- Segmentação de Clientes

- **Manufatura**

- P&D
- Produção
- Logistica & distribuição

- **Varejo**

- Supply Chain
- Análises Financeiras - cash flow
- Otimização da Logística
- Análise de Out of Stock

- **Healthcare**

- Redução de tratamentos excessivos e insuficientes
- Planos de tratamentos individuais
- Intervenção médica precisa (antecipada)

- **Setor Público**

- Cobrança de Impostos
- Prevenção de Fraudes
- Planejamento urbano
- Compartilhamento de informações entre departamentos

Estratégia Microsoft para Big Data

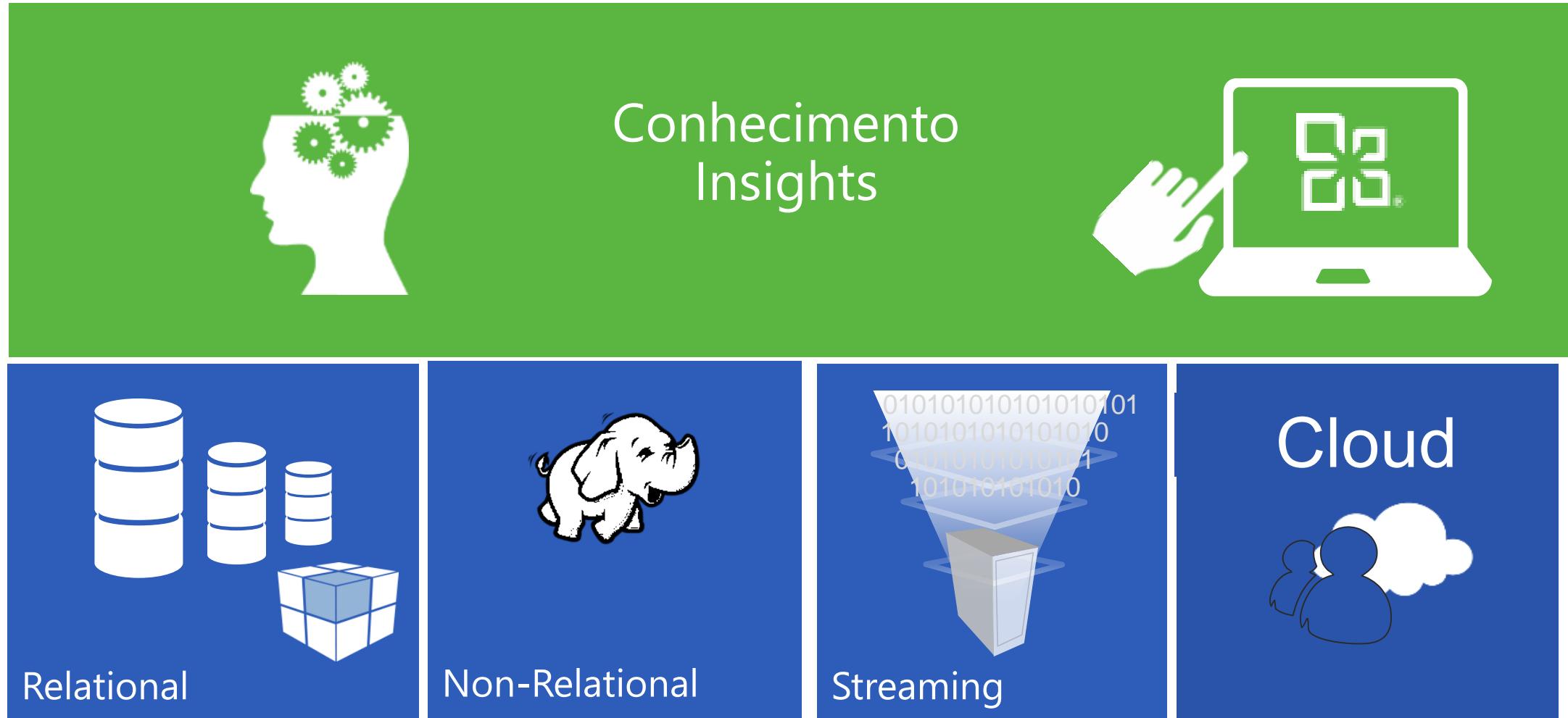


ABORDAGEM COMPLETE PARA BIG DATA



QUALQUER QUANTIDADE, TAMANHO E LOCAL

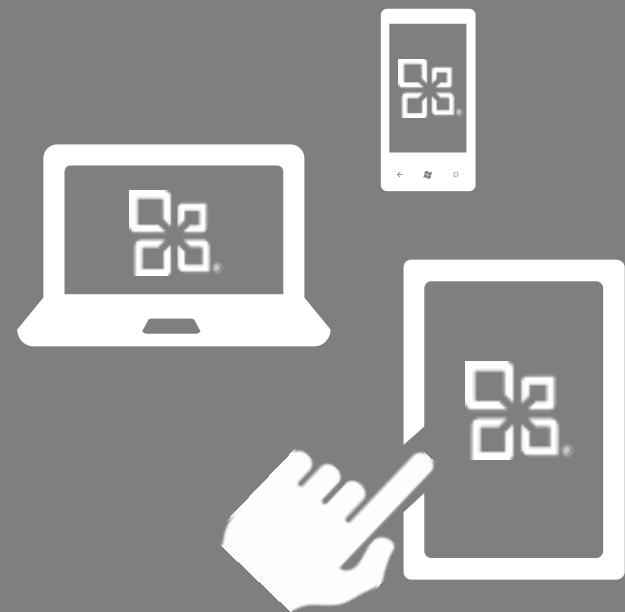
Dados relacional, Semi-estruturado e Não-estruturado



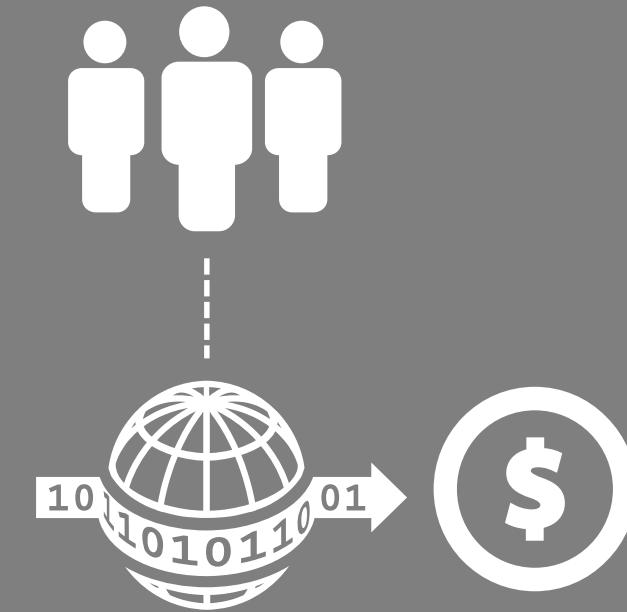
GERAÇÃO DE CONHECIMENTO / INSIGHTS

Plataforma de Business Intelligence da Microsoft

Insights Profundos,
Onde quer que você
esteja



Conectados aos
Dados do Mundo

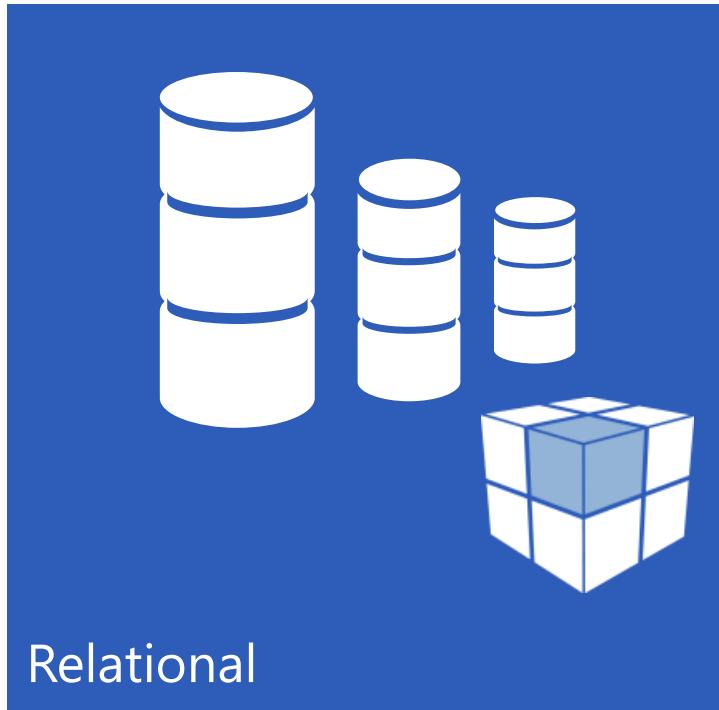


Qualquer dado,
De qualquer tamanho e
formato, Em qualquer lugar

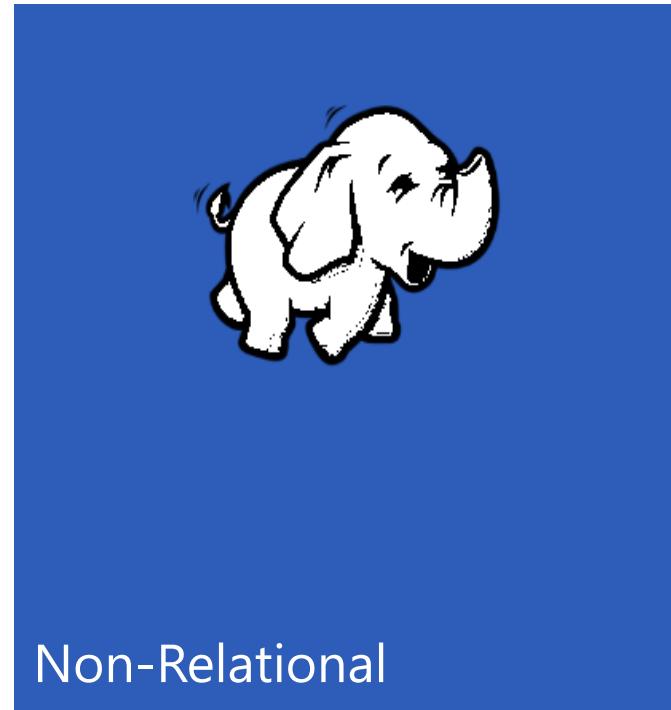


INTEGRANDO O USO MASSIVO DE INFORMAÇÕES

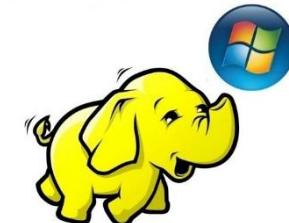
Data Base Performance



Relational



Non-Relational



VISÃO DE DATA WAREHOUSING



Escalabilidade Massiva
Custo Baixo



Hardware padrão de
Mercado com
independência de
fornecedores

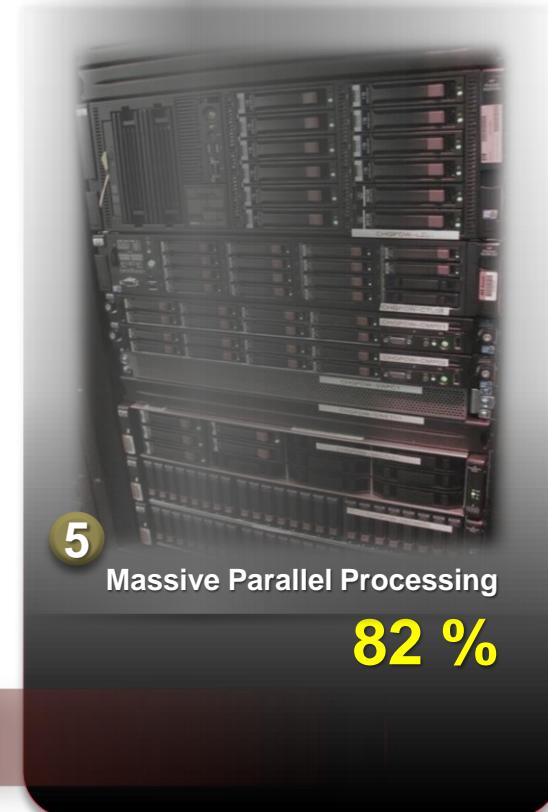
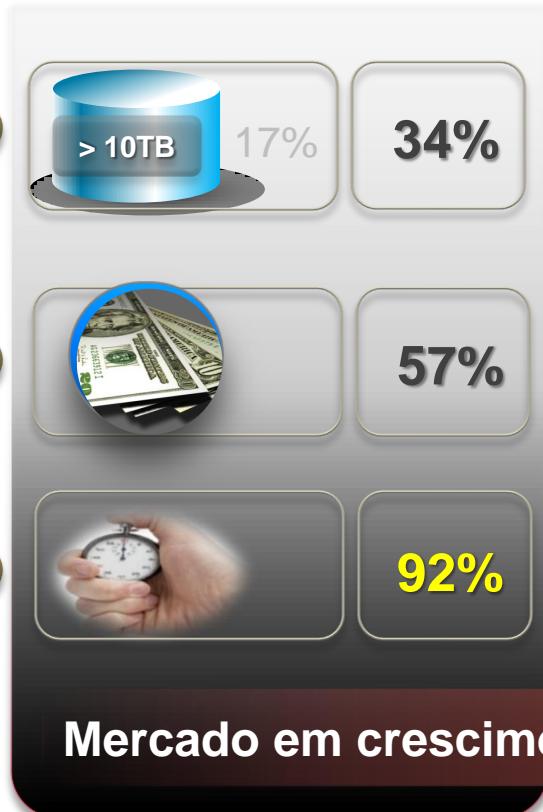


Melhorar a agilidade dos
negócios

Tornando o banco de dados SQL Server mais rápido e mais acessível para clientes de todos os tamanhos.

DESAFIOS DO DATAWAREHOUSE

- 1 Explosão de Dados – número de empresas com mais de 10TB – 2 X em 3 anos
- 2 Clientes necessitam reduzir custos
- 3 Mais clientes planejam fazer real-time analytics
- 4 Busca por **Appliances** para suportar ambientes de Data Warehouse
- 5 Movimento para arquitetura **MPP** – clientes planejam migrar para MPP nos próximos 3 anos



SOLUÇÕES MICROSOFT PARA DATA WAREHOUSING

Tier 1 offerings



Microsoft®
SQL Server® 2008 R2
Enterprise



Microsoft®
SQL Server® 2008 R2
Fast Track Data Warehouse



Microsoft®
SQL Server® 2008 R2
Parallel Data Warehouse

Plataforma escalável e confiável

Ideal para data marts ou pequenas bases analíticas

Somente Software

Scale-Up DW

ARQUITETURA SMP VERSUS MPP

SMP (Symmetric multiprocessing)	MPP (Massively Parallel Processing)
<p>Na computação, multiprocessamento simétrico (SMP) envolve uma arquitetura de hardware de computador, onde dois ou mais processadores idênticos estão ligados a uma memória principal comum única e são controlados por uma única instância de sistema operacional.</p>	<p>A computação massivamente paralela, refere-se à utilização de um grande número de processadores (ou computadores separados) para executar um conjunto de coordenadas de cálculos ou execução de tarefas em paralelo.</p>
<p>Todas as CPUs compartilham a mesma memória, discos e rede</p>	<p>Cada CPU controla seus próprios recursos para uso exclusivo incluindo memória discos e rede (I/O)</p>
<p>Processamento de consultas ocorre inteiramente dentro de uma instância de um banco de dados</p>	<ul style="list-style-type: none">• Arquitetura MPP partitiona tabelas através de múltiplos nós de processamento• Cada nó de processamento possui de forma DEDICADA, CPU, Memoria e Storage, rodando sua PRÓPRIA INSTÂNCIA DE SQL Server, numa arquitetura paralela “shared-nothing”

ARQUITETURA SMP VERSUS MPP

SMP (Symmetric multiprocessing)	MPP (Massively Parallel Processing)
Múltiplas CPUs, dependem do threading das aplicações para executar processos individuais de forma concorrente	Usa múltiplas CPUs trabalhando em paralelo para a execução de um programa ou tarefa
Assim que atingir o máximo de utilização dos recursos disponíveis, terá que crescer através de um outro HW equivalente e dividir a carga de trabalho entre dois sistemas SMP	Arquiteturas MPP permitem a escalabilidade linear, através da adição de novas unidades de processamento e armazenamento ao ambiente existente, reduzindo o impacto do crescimento dos mesmos
Tradicionalmente, a maioria das implementações do SQL Server foram SMP	SQL Server agora capaz de tirar proveito da arquitetura MPP através do PDW

SQL SERVER PARALLEL DATA WAREHOUSE



- **Data Warehouse Appliance MPP Tier 1**
 - Alta escalabilidade, de dezenas a centenas de Terabytes
 - Alta performance através de uma arquitetura **MPP** e um optimizador inteligente para resolução de queries
- **Flexibilidade e Opção**
 - Múltiplos fornecedores de hardware
 - Possibilidades de configurações com Appliances de $\frac{1}{2}$ até o 4 Racks
- **Solução completa e abrangente**
 - Solução completa de data warehouse, suportando a implementação de DWs corporativos e data marts
 - Integração completa com portfolio de Microsoft Business Intelligence (BI)
 - Conjunto abrangente de ferramentas para BI, ETL, MDM, e data streaming (Stream Insight)



RAZÕES PARA UTILIZAR O PDW

 <p>PDW = PERFORMANCE + ESCALABILIDADE</p>	<p>Crescimento significativo de dados</p>	<p>Alta latência no processamento de informações</p>
<p>O processamento de cubos analíticos está lento</p>	<p>Alta concorrência e cargas de trabalho mistas (mixed workload)</p>	<p>Necessidade de crescimento da infra de DW atual</p>
<p>Necessidade de carga de dados simultânea a consultas</p>		<p>Ambiente de DW atual é muito caro</p>

PARALLEL DATA WAREHOUSE



Microsoft®
SQL Server® 2008 R2
Parallel Data Warehouse



Rack de Controle



Microsoft®
SQL Server® 2008 R2
Parallel Data Warehouse



Rack/s de Dados

BENEFÍCIOS DO PDW

Alta disponibilidade nativa

Control Node

Recebe as solicitações de trabalho e é responsável pela criação do Plano de Query



Management Server

Gerenciamento do Sistema, Atualizações, Monitotamento HPC, Status, Networking, Console de Administração



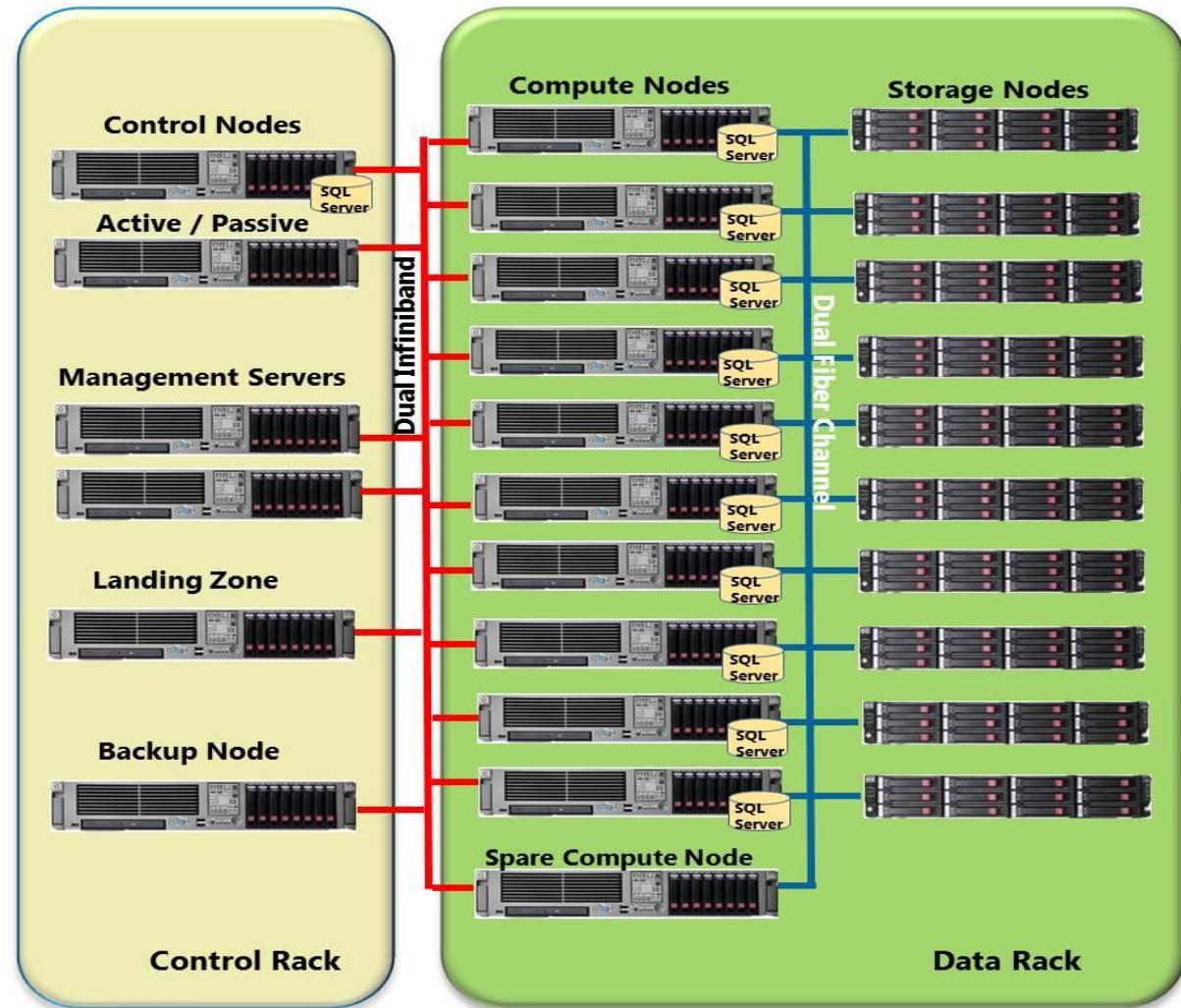
Landing Zone

Staging Area, Carga em Paralelo, testado com +1TB/hora. LZ com até 10TB de storage



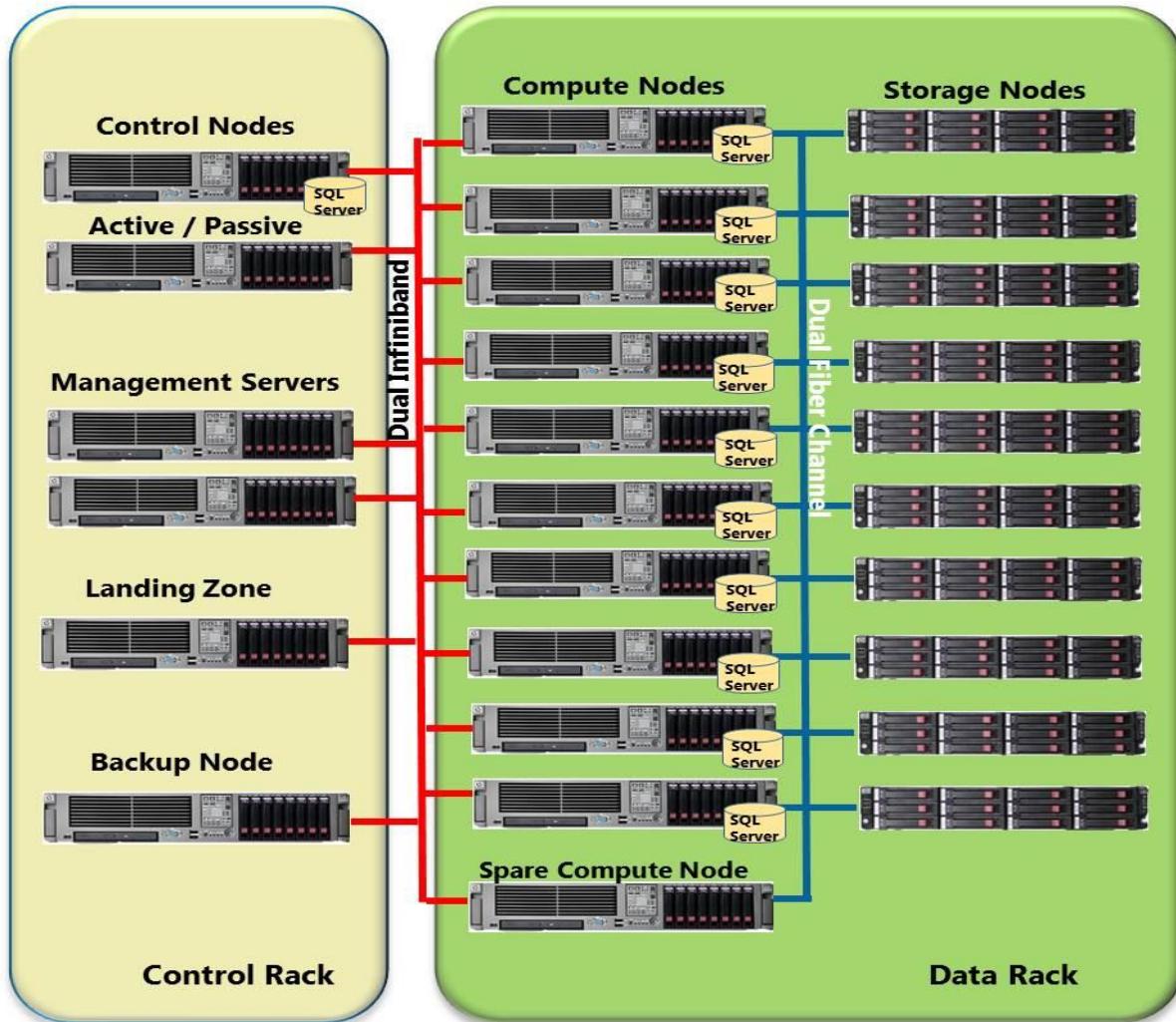
Backup Node

Configurável p back-up completo da base instalada.
Backup de até 4-6 TB an hour (teste)



Alta Disponibilidade não representa um custo EXTRA no PDW

BENEFÍCIOS DO PDW



Cada **Compute Node** é completa, Instância independente de SQL server.
Arquitetura MPP native.,
12 cores, 96GB Ram, por nó

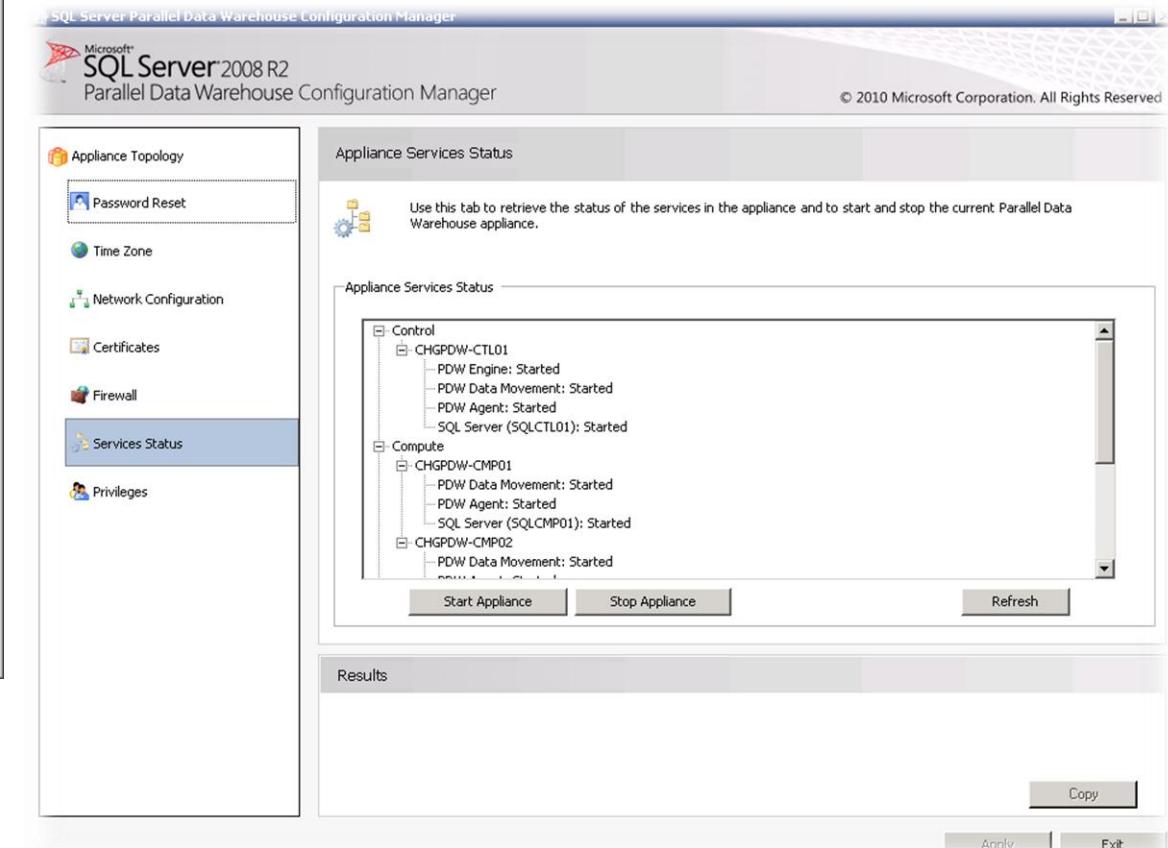
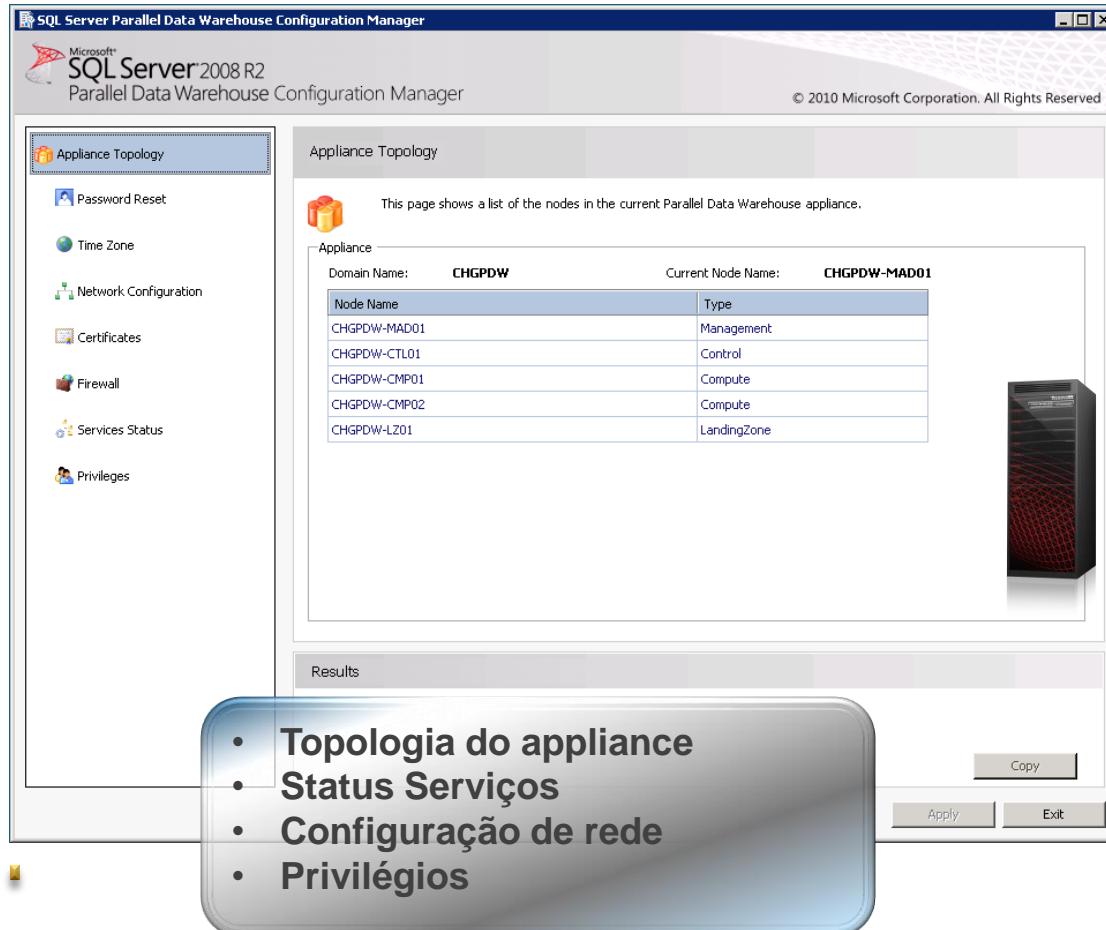
Data Rack:
960 GB RAM
120 Cores

Capacidade de dados do Rack:
Mais de 150TB

Scan Rate médio:
Até 3.6 TB/min/Data Rack

Alta Disponibilidade não representa um custo EXTRA no PDW

GESTÃO E CONFIGURAÇÃO



C:\Program Files\Microsoft SQL Server Parallel Data Warehouse\100\dwconfig.exe

CONSOLE DE ADMINISTRAÇÃO

The screenshot shows the Microsoft SQL Server 2008 R2 Parallel Data Warehouse Administration Console. The top navigation bar includes links for Home, Sessions, Queries, Loads, Backups/Restores, Appliance State, Alerts, Locks, and Dashboard. The main interface features four key sections: Review Queries (Active: 5, Queued: 0), Review Loads (Active: 0, Queued: 0), Review Backups (Backups: 0, Restores: 0), and Review Appliance State. Below these are four icons representing data, time, storage, and hardware. The bottom left displays the page footer with the date (10/31/2010 4:43:18 AM) and version (Microsoft SQL Server 2008 R2 (RTM) - Parallel Data Warehouse Edition v10.0.2091.0). The bottom right shows a detailed table of appliance states for various nodes.

Node Name	Type	IP Address	Passive	Failed Over	Cluster	Cooling	Network	Power Supply	Processing	Software	Storage External	Storage Internal
CHOPDW-SQLCTP01	Control	172.16.130.240	✓	✓	✓	✓	✓	✓	✓	✓	NA	✓
CHOPDW-SQLCMP01	Compute	172.16.130.1	✓	✓	✓	✓	✓	✓	✓	✓	NA	✓
CHOPDW-SQLCMP02	Compute	172.16.130.2	✓	✓	✓	✓	✓	✓	✓	✓	NA	✓
CHOPDW-SQLCMP02	Compute	172.16.130.2	✓	✓	✓	✓	✓	✓	✓	✓	NA	✓
CHOPDW-SQLCMP02	Compute	172.16.130.2	✓	✓	✓	✓	✓	✓	✓	✓	NA	✓
CHOPDW-SQLCMP02	Compute	172.16.130.2	✓	✓	✓	✓	✓	✓	✓	✓	NA	✓
CHOPDW-SQLCMP02	Compute	172.16.130.2	✓	✓	✓	✓	✓	✓	✓	✓	NA	✓
CHOPDW-SQLCMP02	Compute	172.16.130.2	✓	✓	✓	✓	✓	✓	✓	✓	NA	✓
CHOPDW-SQLCMP02	Compute	172.16.130.2	✓	✓	✓	✓	✓	✓	✓	✓	NA	✓
CHOPDW-SQLCMP02	Compute	172.16.130.2	✓	✓	✓	✓	✓	✓	✓	✓	NA	✓
CHOPDW-LZ01	LandingZone	172.16.128.230	NA	✓	✓	✓	✓	✓	✓	✓	NA	✓
CHOPDW-MAD01	Management	172.16.128.250	NA	✓	✓	✓	✓	✓	✓	✓	NA	✓

https://controlnodeipaddress

- Dashboard
- Queries Activity
- Load Activity
- Backup and Restore
- Active Locks
- Active Sessions
- Alerts
- Appliance State

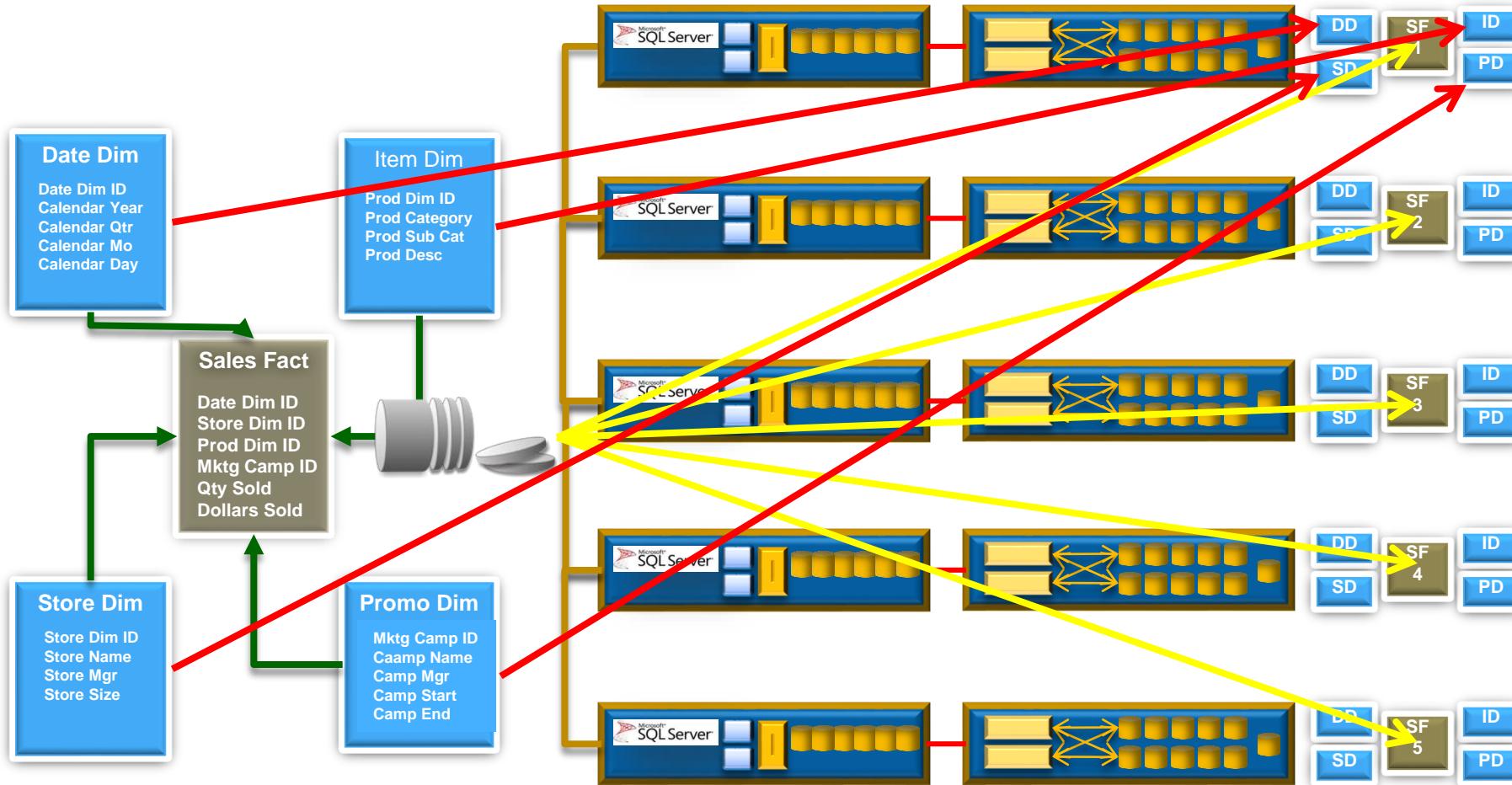
DISTRIBUIÇÃO DE DADOS

Utilização da Arquitetura MPP

- **Replicados:** A estrutura de tabela que existe como uma cópia completa replicada dentro de cada Storage de cada Compute Node do PDW
- **Distribuídos:** A estrutura de tabela que é distribuída, através de Hashing, uniformemente distribuida por todos os Nós de Processamento (Compute Nodes) do PDW. Cada distribuição completa é uma tabela física independente dentro do Banco de Dados.
- **Shared Nothing:** A capacidade de projetar um schema de distribuição e replicação de tabelas para reduzir/minimizar a movimentação de dados entre os nós de processamento.
 - Pequenas quantidades de dados poderão ser processadas mais eficientemente se armazenadas completas (replicação).
 - O Scan de grandes volumes de dados é muito mais eficiente quando utiliza o benefício do paralelismo da arquitetura MPP.

DISTRIBUIÇÃO DE DADOS NO PDW

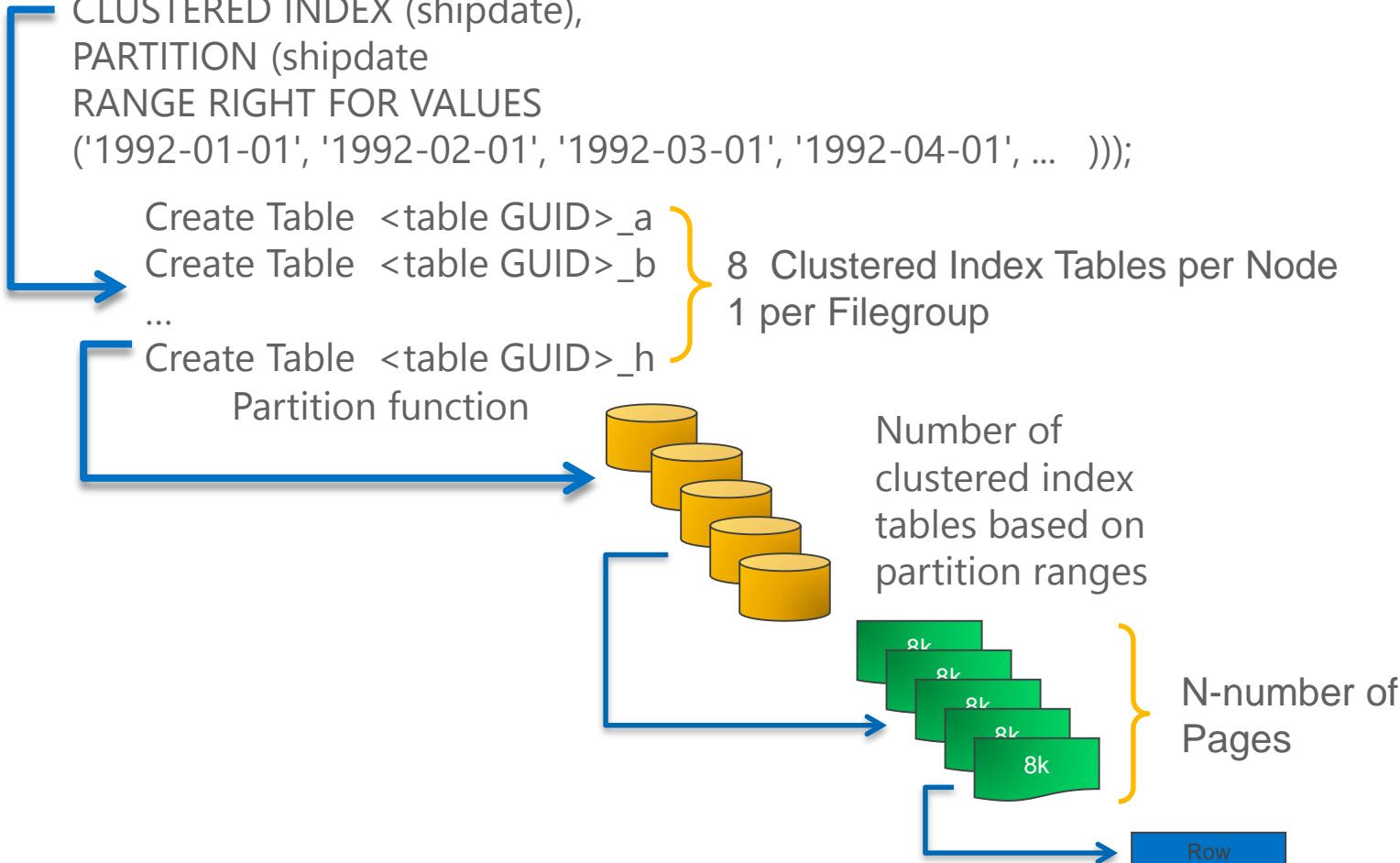
Utilização da Arquitetura MPP



SIMPLICIDADE

Comando CREATE TABLE de uma tabela distribuída

```
CREATE TABLE myTable  
(column Defs)  
WITH ( DISTRIBUTION = HASH (id),  
CLUSTERED INDEX (shipdate),  
PARTITION (shipdate  
RANGE RIGHT FOR VALUES  
('1992-01-01', '1992-02-01', '1992-03-01', '1992-04-01', ... )));
```



SOLUÇÃO COMPLETA – BUSINESS ANALITYCS

Integração NATIVA com soluções de BI Microsoft

Integração com ferramentas de BI - SQL Server

- SQL Server Integration Services (ETL) ter o PDW como destino
- SQL Server Analysis Services (OLAP) ter o PDW como fonte
- SQL Server Reporting Services
- Integração com PowerPivot

Ferramentas Complementares SQL Server

Master Data Services (para MDM)

StreamInsight (para streaming de dados)



Acelerado com In-memory

Tecnologia xVelocity In-Memory embutida na plataforma



Desktop



Data Mart



Data Warehouse

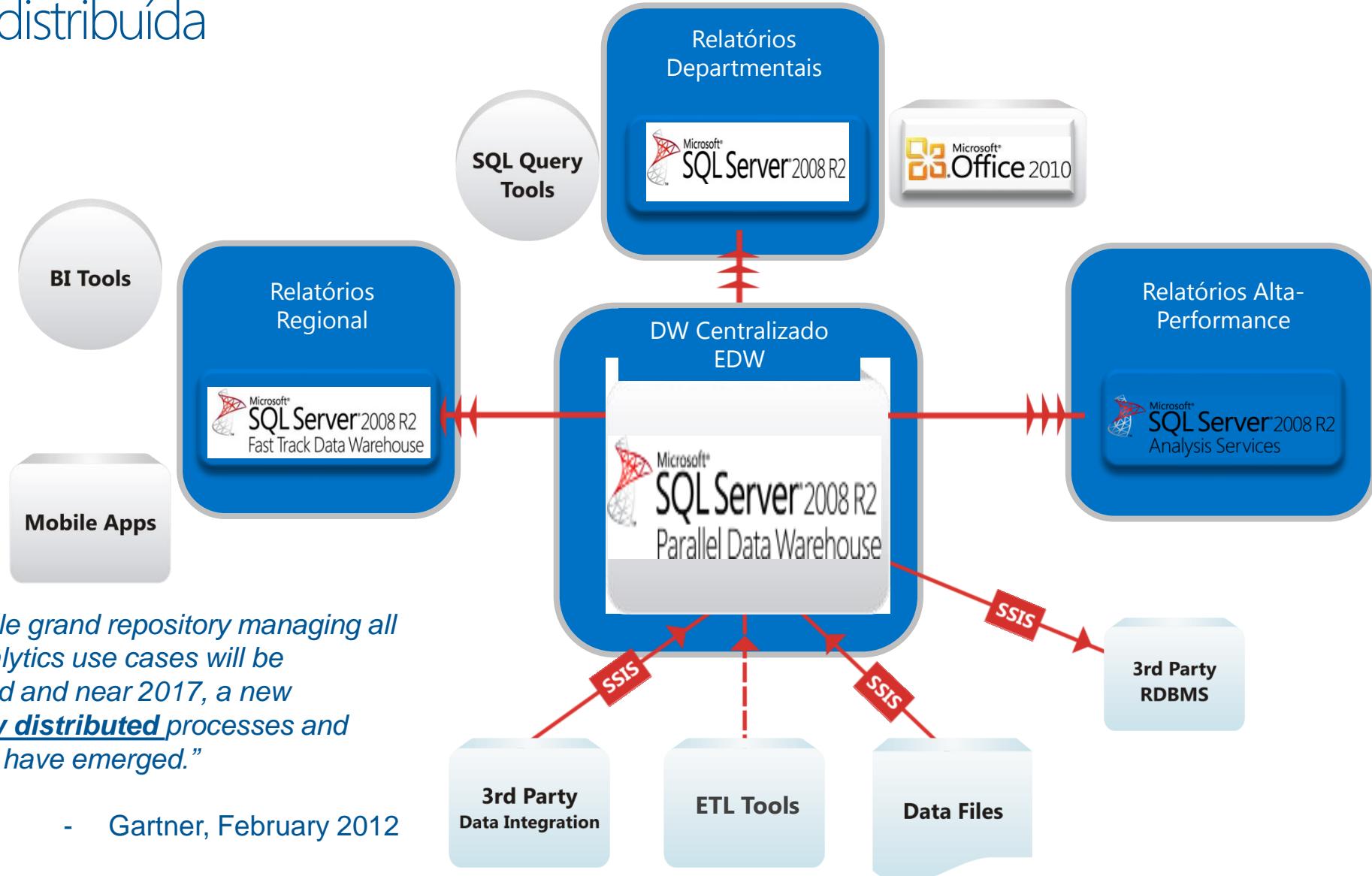
xVelocity In-Memory

“ Ao reestruturar nosso warehouse com o xVelocity, a carga dos dados e a produção de reportes estão significativamente mais rápidos. Um reporte grande que costumava levar **17 minutos** agora acontece em **apenas 3 segundos**. ”

- Wolfgang Kutschera
Team Leader of Database Engineering, bwin.party

DATA WAREHOUSE

Arquitetura distribuída



"The concept of a single grand repository managing all the information for analytics use cases will be increasingly challenged and near 2017, a new infrastructure of highly distributed processes and information assets will have emerged."

- Gartner, February 2012

BENEFÍCIOS DO PDW

Alta Performance



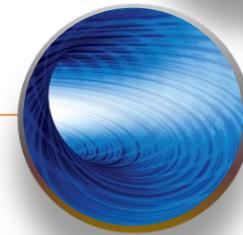
Opção de Platformas de HW



Baixo Custo



Alta Escalabilidade

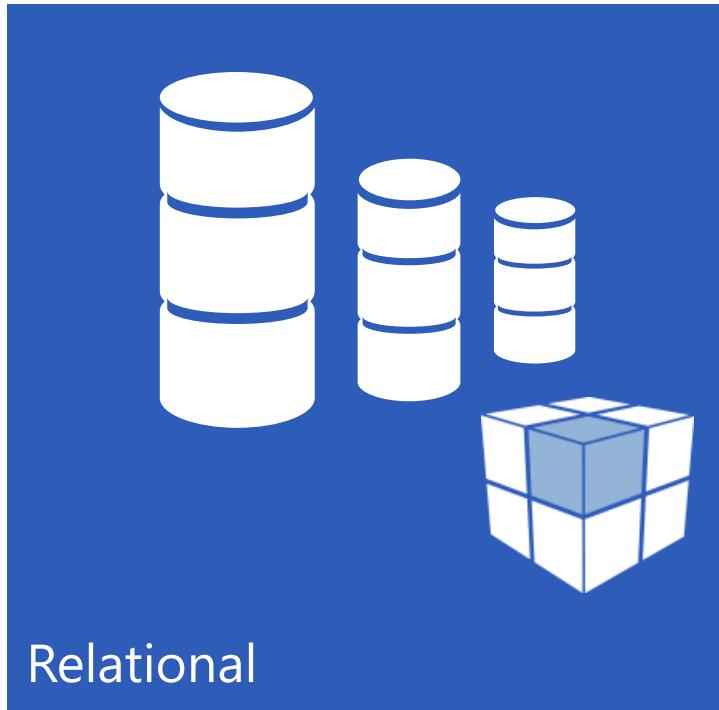


Redução do Risco

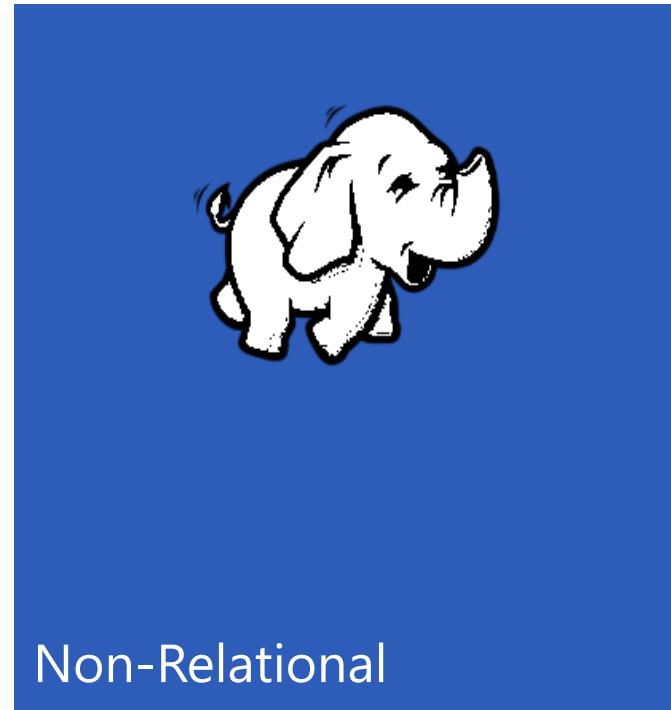


INTEGRANDO O USO MASSIVO DE INFORMAÇÕES

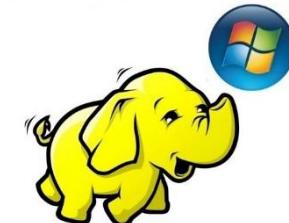
Data Base Performance



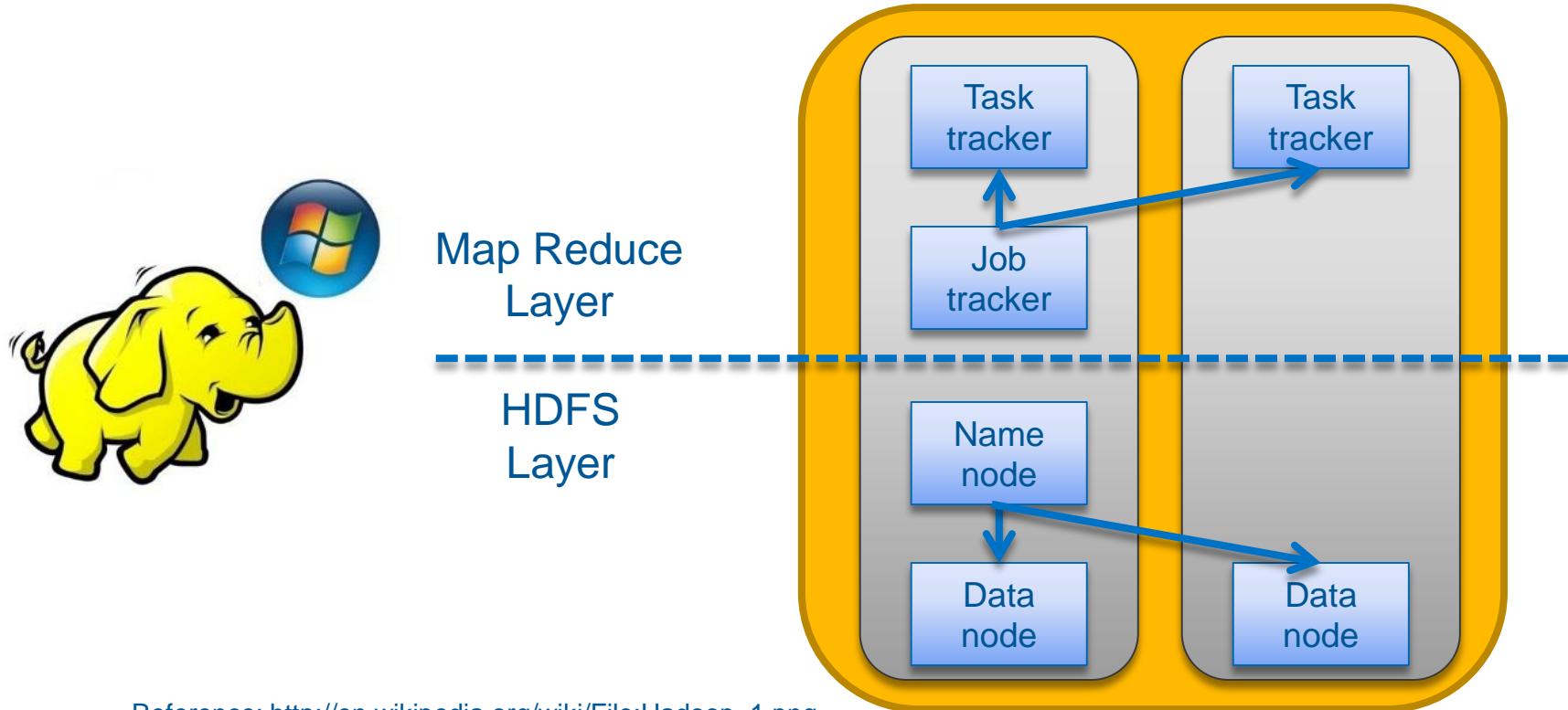
Relational



Non-Relational



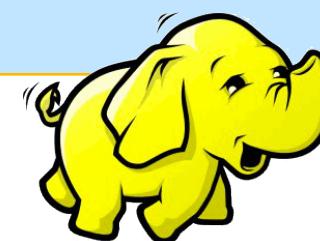
ARQUITETURA DO HADOOP



RDBMS VERSUS MAPREDUCE (HADOOP)

Referencia: Tom White's Hadoop: The Definitive Guide

	RDBMS Tradicional	Big Data MapReduce
Volume	Gigabytes (Terabytes)	Petabytes (Hexabytes)
Acesso	Interativo e Batch	Batch
Atualizações	Leitura e Muitas Gravações	Uma gravação e muitas leituras
Estrutura	Modelo Estático	Modelo Dinâmico
Integridade	High (ACID)	Low (BASE)
Escalabilidade	Não Linear	Linear
DBA Ratio	1:40	1:3000



Big Data Announcements at Strata Conference, New York Oct 24 2012

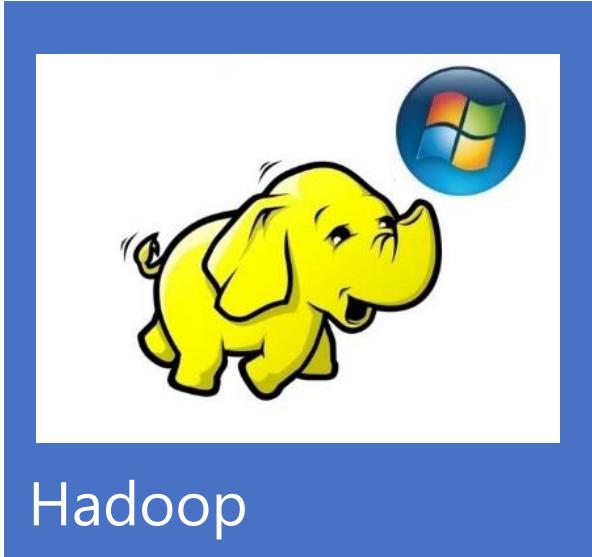
Tomorrow at Strata Conference + Hadoop World, New York, New York, we are making three announcements Big Data, expanding on Microsoft's commitment to help customers manage and gain insights from any data, any size, anywhere. Specifically, we will announce:

1. Availability of the first Community Technology Preview (**CTP**) of **Microsoft HDInsight Server** (formally known internally as Hadoop on Windows).
2. A **Preview of Windows Azure HDInsight Service** (formally known internally as Hadoop on Azure).
3. An **expanded partnership with Hortonworks**, a commercial vendor of Hadoop. This expanded partnership will give customers access to an enterprise-ready distribution of Hadoop with 100% compatibility with Apache Hadoop.

HDInsight shipped by Microsoft, will offer the following benefits to our customers:

1. **100% Apache Hadoop:** HDInsight builds on Hortonworks Data Platform (HDP) which is fully compatible with apache Hadoop.
2. **Enterprise-ready Hadoop:** HDInsight offers high security through integration with ActiveDirectory, and simplified manageability through integration with System Center
3. Simplicity and manageability of Windows: Through smart packaging, HDInsight simplifies the setup and provisioning of Hadoop
4. **Insights for all:** through integration with Microsoft BI tools. Customers can use award winning BI tools like PowerPivot and Power View to gain insights on all their data, including unstructured data on Hadoop.

VOLUME E VARIEDADE



Microsoft SQL Server

Windows Azure™

HDInsight Server (Hadoop com Windows Server)
HDInsight Server (Hadoop com Windows Azure)
Lançamento CTP 24 Outubro de 2012

- Segurança padrão empresarial – classe A, Alta Disponibilidade e Gerenciamento
- Suporte a Virtualização
- Integração native com ferramentas de BI Microsoft
- Experiências similares em instalações Privadas e na Nuvem



Welcome

Sign in

The Developer Preview for the Apache Hadoop-based Services for Windows Azure is available by [invitation](#).

[Sign in](#) 

Microsoft + Big Data

Unlock business insights from all your structured and unstructured data, including large volumes of data not previously activated, with Microsoft's Big Data solution. Microsoft's end-to-end roadmap for Big Data embraces Apache Hadoop™ by distributing enterprise class, Hadoop-based solutions on both Windows Server and Windows Azure.

Windows Azure

The [Windows Azure platform](#) includes a cloud services operating system and a set of developer services, which provide the functionality to build applications that span from consumer Web to enterprise scenarios. Windows Azure and SQL Azure are the key components of the Windows Azure platform.

Windows Azure provides developers

Apache Hadoop

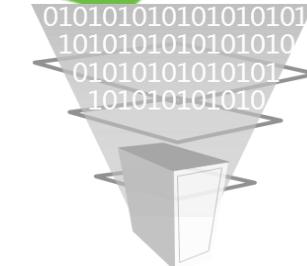
The [Apache Hadoop software library](#) is a framework that allows for the distributed processing of large data sets across clusters of computers by using a simple programming model. It is designed to scale up from single servers to thousands of machines, each offering local computation and storage.

Rather than rely on hardware to deliver

VELOCIDADE



Análises na instalação privada de dados streaming em tempo real



Análise de dados em streaming projetadas para a Nuvem



Microsoft®
SQL Server®
StreamInsight

Processamento de Complex Event Processing (CEP) com StreamInsight (On-premise)

 **Windows Azure™**
SQL StreamInsight

Processamento de eventos complexos na Nuvem com Windows Azure SQL StreamInsight

EXTENDENDO SEU DW COM HADOOP

Benefícios

Integração com Soluções Corporativas de BI



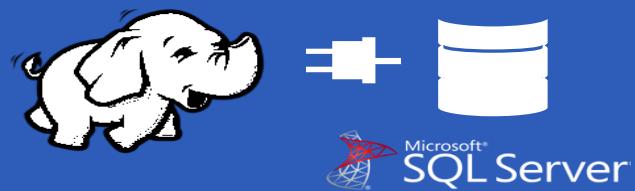
Integração com Data Warehouses Corporativas da Microsoft



Conhecimento e Insights profundos de dados estruturados e não-estruturados



Funcionalidade



Conection do Microsoft SQL Server para Apache Hadoop através do SQOOP (SQL to Hadoop)



Conection do SQL Server Parallel Data Warehouse para Apache Hadoop através do SQOOP

E SUPORTE A DADOS NÃO ESTRUTURADOS

Através do hadoop no seu site ou na Nuvem

Benefícios



Extendendo seu Data Warehouse com Big Data



Plataforma elástica para Big Data na Nuvem



Plataforma para Big Data Enterprise-class no seu Data Center

Funcionalidade



Conector do Hadoop para SQL Server



Serviços Hadoop na plataforma
HDInsight Service
(Windows Azure)



HDInsight Server,
Distribuição Hadoop Hortonworks em Windows Server

BIG DATA SEM BI NÃO É RELEVANTE

Entregamos conhecimento para todos com ferramentas familiares

Benefícios



Interação e análise de dados não-estruturados no Hadoop

Funcionalidade



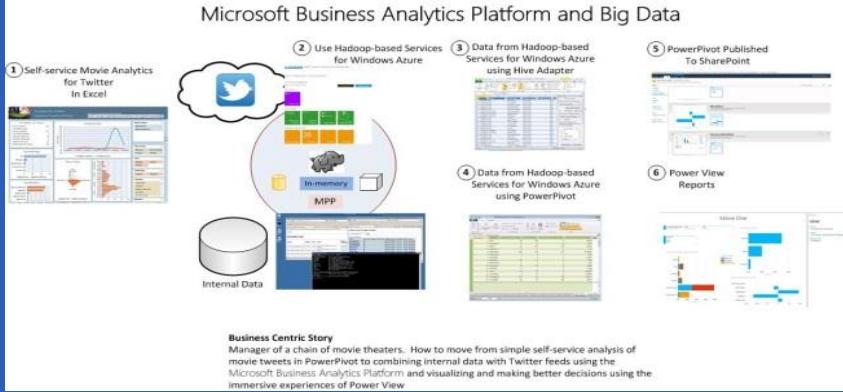
Add-in do Hive para Microsoft Excel

A screenshot of the Microsoft Excel ribbon showing the 'Data' tab selected. A 'Hive Query' dialog box is open, displaying the following configuration:

- Hive Connection: MyHiveData
- Hive Objects - Tables/Views: hivesampletable [Table]
- Columns: A table showing four columns: clientid, querytime, market, and devicenplatform. All four columns have checkboxes checked under 'Column Name'.

CASO YAHOO

Ferramentas de BI Microsoft



Cubo de 24



Hadoop - Distribuição



Caso de uso:

- Volume de dados extremamente grande da dados não-estruturados de Web log analysis
- Análises Ad hoc de Web Logs não estruturados para protótipação de perfís.
- Hadoop carrega Cubo em Analysis Services com 24TB

BIG DATA REQUER ABORDAGEM HOLÍSTICA



Microsoft



TRADITIONAL

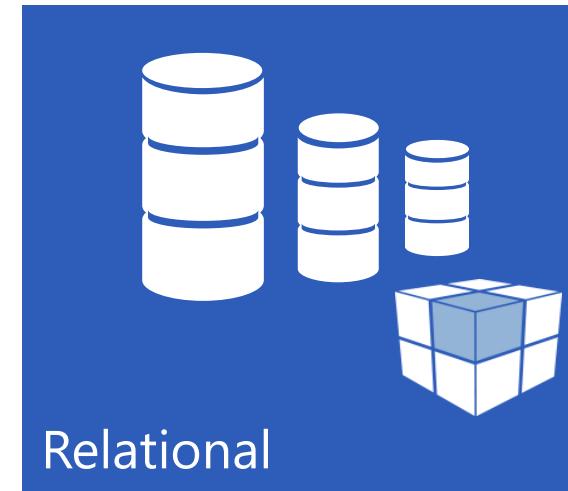
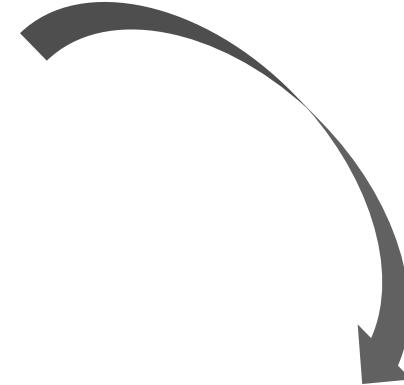
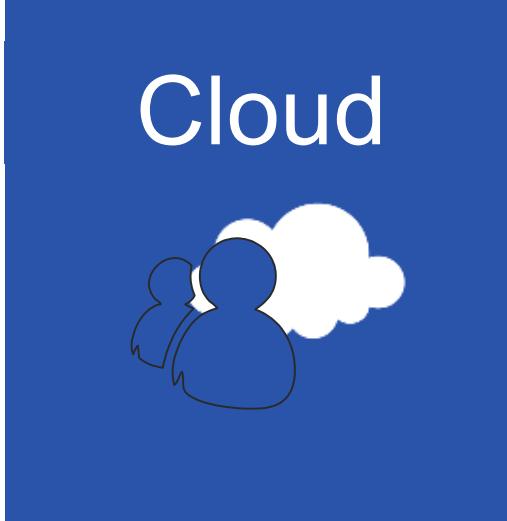
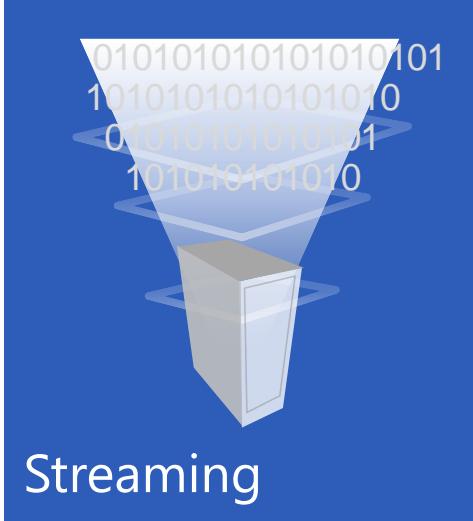
Relational Database Management System

NEW

Petabyte-Scale Services

PROCESSAMENTO DE EVENTOS

Windows AZURE StreamInsight



Relational



Non-Relational

PROCESSAMENTO DE EVENTOS NA NUVEM

StreamInsight on Windows Azure



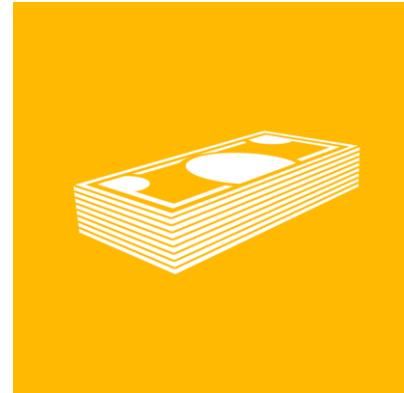
Dados de eventos estão
na Nuvem

Dados não são locais!



Dados de eventos
globalmente distribuidos

Trazer o processamento
para os dados, não os dados
Para o processamento!



Redução de
TCO

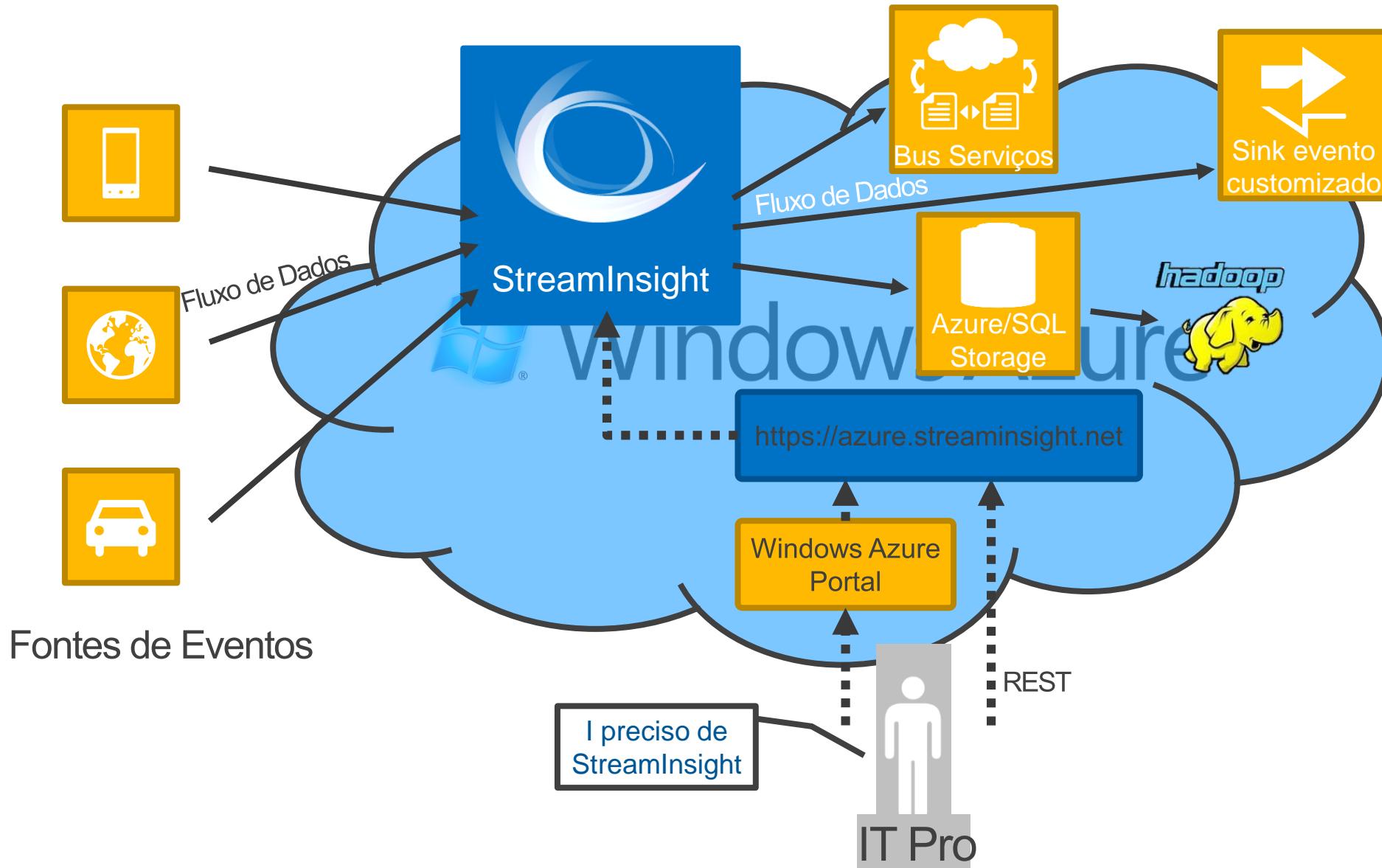


Elasticidade
scale-out



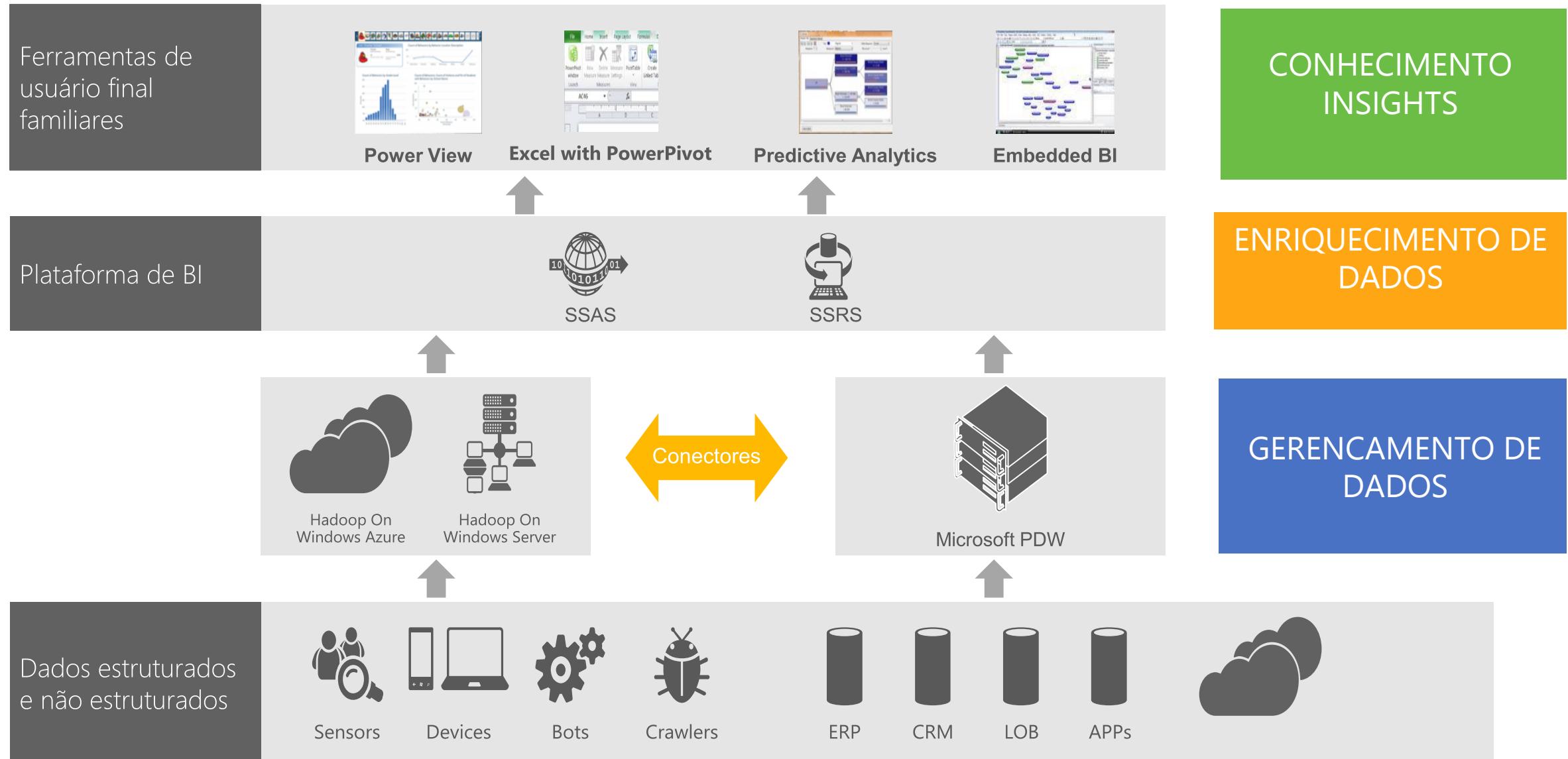
Serviço,
Não
Infraestrutura

O QUE É WINDOWS AZURE STREAMINSIGHT?



ARQUITETURA DE REFERÊNCIA

Integração Dados Estruturados e Não estruturados



Microsoft[®]