

Folheto

# Hadoop de nível empresarial

Extraia o máximo de valor e desempenho do Apache Hadoop e faça escala sem concessões



**Hewlett Packard**  
Enterprise



As organizações se empenham em identificar e materializar o valor do Big Data. Por esse motivo, muitas estão procurando sistemas de análises mais ágeis e capazes. O Apache Hadoop é uma estrutura de software que oferece economia e valor notáveis no armazenamento e no processamento de dados em escala imensa. Muitos já usaram o Hadoop como um repositório de dados para cargas de trabalho simples, mas é possível criar muito mais valor com o Hadoop fazendo um maior uso dos dados na plataforma, para interagir com os dados e revelar novos insights de negócios. Mas existem desafios que limitam esse valor e não permitem que essa oportunidade seja aproveitada. A Hewlett Packard Enterprise desenvolveu uma solução que vence esses desafios através de uma oferta robusta, porém flexível, que fornece uma base centrada nos dados e preparada para o futuro, capaz de fazer escala junto com a evolução das necessidades de seus negócios.

## Os desafios atuais do setor de negócios

Os dados são o combustível da economia das ideias. Porém, para aproveitar todos os seus dados relevantes, você precisa utilizar novos motores para consumir esse combustível e obter melhores resultados em seus negócios. Nos últimos dez anos, o Hadoop se transformou em uma combinação essencial de motor e tanque de combustível, conquistando o apoio de uma massa crítica de desenvolvedores, cientistas de dados e profissionais de TI. Porém, ainda que o Hadoop tenha funcionado bem para alguns visionários importantes, ele se limitava a experimentos científicos, cargas de trabalho simples e pilotos para a grande maioria de seus primeiros adeptos. A Hewlett Packard Enterprise conhece os desafios e as oportunidades apresentadas pelo Hadoop, e pode capacitar sua organização a transformá-lo em um pilar de nível empresarial para sua plataforma centrada nos dados.

60%

de desempenho do HDFS<sup>1</sup>

100%

de seus dados relevantes

Até 8X

mais velocidade de análise<sup>2</sup>

### DESAFIOS DO CLIENTE

#### Incapacidade das tecnologias existentes de lidar com o

**crescimento dos dados:** a maior parte dos dados disponíveis hoje foi criada nos últimos dois anos, muitas vezes a partir de novas fontes e tipos de dados, e as arquiteturas, os bancos de dados e as abordagens existentes para o gerenciamento de dados estão chegando ao seu limite. Uma pesquisa mostra que 41% das empresas entrevistadas afirmam que seus sistemas não são capazes de processar grandes volumes de dados de fontes diferentes.<sup>3</sup>

**Escala de Hadoop sem concessões:** as organizações querem obter mais valor de seus investimentos no Hadoop e escalar seus lagos de dados. Porém, o Hadoop não inclui recursos de análises nem otimização de desempenho capazes de lidar com cargas de trabalho complexas, e a falta de uma segurança robusta com escala aumentada acaba aumentando os riscos.

#### Determinar como obter o valor do Hadoop:

as organizações têm dificuldade em identificar os projetos e casos de uso certos para extrair o maior valor do Hadoop. A estrutura do Hadoop possibilita muitas cargas de trabalho e casos de uso, mas as organizações têm dificuldades em identificar quais executar e avaliar o impacto e o valor para os negócios.

### NECESSIDADES DO CLIENTE

#### Repositório consolidado, escalável e acessível para Big Data:

as organizações precisam de um local centralizado através do Hadoop para armazenar e gerenciar, por um custo acessível, uma ampla variedade de dados, incluindo desde dados de negócios, desprovidos da estrutura rígida dos bancos de dados tradicionais, até a vasta gama de dado estruturados, semiestruturados e não estruturados, utilizando não só o SQL padrão, mas também ferramentas mais avançadas.

#### Gerenciamento de carga de trabalho, desempenho,

**disponibilidade e segurança de nível empresarial:** soluções flexíveis, porém robustas, que aproveitam a infraestrutura e as tecnologias de análises de qualidade comprovada, empregando consultas e resultados de SQL ANSI na quantidade certa de recursos de computação, armazenamento e análises, com base nos casos de uso. Segurança integrada com Hadoop que protege seus dados em trânsito, estacionários e em uso.

#### Um parceiro que entenda seus negócios e tenha experiência

**com ambientes de TI complexos e heterogêneos:** sua organização precisa acessar e priorizar casos de uso de valor comprovado do Hadoop, integrando-os aos aplicativos existentes e ao cenário maior de análises e inteligência de negócios (BI). Os ambientes de TI são difíceis e complexos, e exigem que o parceiro tenha experiência avançada na solução desses desafios.

## Proposta de valor

A Hewlett Packard Enterprise entende do que você precisa não só para sobreviver, mas também para prosperar na economia das ideias. Ideias precisam ser convertidas em experimentos bem-sucedidos e, em seguida, em soluções de negócios de escala total. Essa conversão em dois estágios lida com a necessidade de conversão de dados em insights e, em seguida, com a conversão desses insights em resultados melhores para os negócios. A Hewlett Packard Enterprise pode ser sua parceira e ajudá-lo a descobrir o valor de seus dados. Nós podemos ajudá-lo a construir uma base centrada nos dados para transformar esses dados em insights e, em seguida transformar esses insights em ótimos resultados de negócios, na velocidade dos negócios, incorporando o Hadoop como um pilar de sua solução de análise de Big Data.

<sup>1</sup> **Arquitetura de referência para Big Data**  
HPE: uma abordagem moderna, 4AA5-6141PTL, outubro de 2015, Rev. 2

<sup>2</sup> Estudo interno da HPE.

<sup>3</sup> "Lucrando com a promessa do Big Data", documento da PWC, janeiro de 2013.

**Experiência com o Hadoop:** a Hewlett Packard Enterprise otimizou todos os aspectos do ambiente Hadoop para que nossos clientes possam operar vários aplicativos e consolidar vários armazenamentos de dados em um único sistema. As arquiteturas de referência da HPE foram desenvolvidas para apoiar a adaptação a novas demandas, permitindo que você escale computação e armazenamento de maneira independente, quebrando a corrente padrão de armazenamento e computação de nó local do Hadoop, e ainda otimize as arquiteturas tradicionais. Esses avanços permitem que você extraia o máximo do YARN, Spark, Kafka, Mahout e, obviamente, dos módulos básicos: Hive, HBase, MapReduce e Hadoop Database File System (HDFS). Além disso, a Hewlett Packard Enterprise tem software líder do setor para oferecer funcionalidades adicionais além do Hadoop através do Vertica para SQL em Apache Hadoop, do Vertica Enterprise Edition para análises preditivas e de diversos pacotes de governança de informações, gerenciamento e segurança de dados para ajudar a garantir recursos completos de nível empresarial.

Embora o Hadoop tenha se popularizado como uma tecnologia de código aberto viável no mercado de Big Data e análises, como mostra o cenário geral de revolução dos dados, vários desafios consideráveis surgiram. Em escala e com os devidos níveis de segurança e confiabilidade, as implementações do Hadoop apresentam muita complexidade de planejamento e implantação, além de dificuldades no gerenciamento de longo prazo. No momento, faltam talentos em Hadoop no mercado. Os desafios trazidos por essa falta de talentos são agravados ainda mais pela contínua evolução e pelo amadurecimento da pilha de tecnologia do Hadoop, que impõem níveis mais altos de dificuldade e incerteza.

A Hewlett Packard Enterprise tem uma abordagem comprovada de quatro etapas para a implementação de sistemas de análises ou dados sofisticados: descoberta, desenvolvimento, integração e implementação. Essa abordagem já ajudou um conjunto global de clientes, grandes e pequenos, a otimizar o valor do Hadoop para a obtenção de ótimos resultados de negócios.

**Talento em análises com alcance global:** a Hewlett Packard Enterprise implementa soluções de missão crítica para os clientes através de nossa presença global de mais de 3.500 especialistas em análises, com média de 18 anos de experiência na área de análises, distribuídos por nove centros analíticos de excelência em quatro continentes.

**Soluções integradas:** nosso poderoso portfólio de produtos de padrões abertos, arquiteturas abertas e serviços, somado à nossa experiência avançada, vai ajudá-lo a integrar e incorporar análises a processos e fluxos de trabalho padrão, contribuindo para que você transforme insights em ações sem ficar preso a soluções proprietárias inflexíveis, que oferecem retornos cada vez menores.

1200+  
profissionais globais  
de análises

18+  
anos de experiência  
em análises

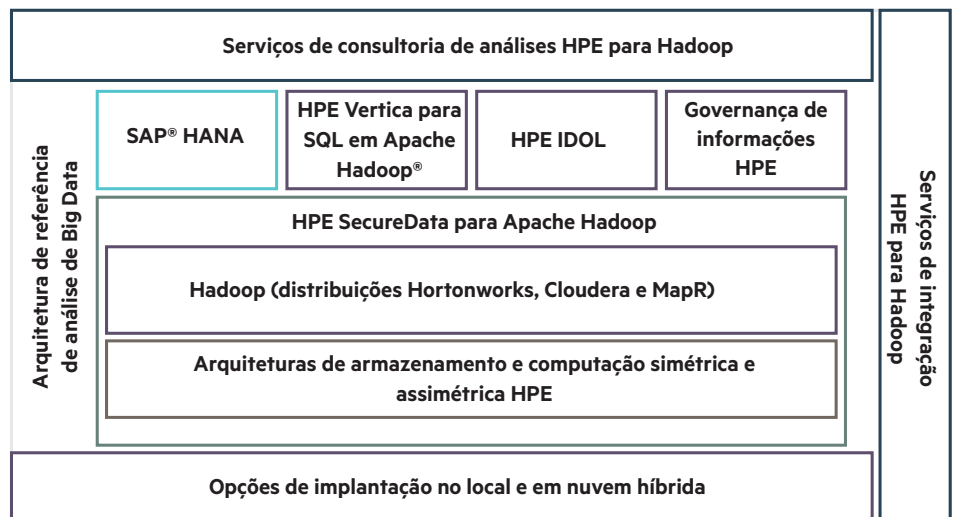


Figura 1: Visão geral da solução

## Hadoop de nível empresarial

Há um consenso cada vez maior de que uma base centrada em dados para lidar com Big Data na faixa de TB-PB em todos os tipos de Big Data (dados de negócios tradicionais, humanos e de máquina) certamente terá o Hadoop como componente central. O HDFS será um destino de lago de dados para dados extraídos de bancos de dados existentes baseados em esquemas, mas também será o destino de lago de dados para fontes de dados novas, menos estruturadas e mais variadas, incluindo mídia social, logs de máquina e outras fontes novas de Big Data. A Hewlett Packard Enterprise trabalha com base no código aberto para tornar possível um Hadoop de alta confiança.

COMPONENTES DA SOLUÇÃO	DESCRIÇÃO
<b>Solução HPE como serviço para Hadoop</b>	A solução HPE como serviço para Hadoop é uma plataforma de Big Data de nível empresarial, baseada em código aberto. Trata-se de uma implantação como serviço, baseada em nuvem; um modelo de consumo, somado às nossas soluções hospedadas e locais. Nossa solução oferece aos clientes uma plataforma Hadoop configurada na nuvem do HPE Helion, sem os altos custos iniciais, os altos riscos de implementação, o peso sobre os recursos de TI e software de que você não precisa, permitindo que você se concentre em atingir suas metas de negócios. Disponível em opções de cargas de trabalho em camadas, a solução é flexível e capaz de se adaptar a mudanças, reduzindo os riscos da tecnologia se tornar obsoleta. A solução HPE como serviço para Hadoop pode ser estendida com a adição da capacidade de banco de dados de análise do HPE Vertica para lidar com grandes conjuntos de dados praticamente em tempo real, e inclui o IDOL para processar todos os tipos de dados não estruturados.
<b>Serviços de consultoria de análises HPE</b>	Os serviços de consultoria de análises HPE podem ajudá-lo a determinar o valor de seus dados em horas ou dias, em vez de semanas ou meses. Em seguida, nós o ajudaremos a identificar os componentes apropriados de uma base centrada nos dados, como operacionalizar as análises usando essa base no ponto da ação de seu processo de negócios e ajudá-lo a implementá-la. Vamos ajudar sua organização a determinar a combinação ideal de investimentos existentes em BI/depósito de dados empresariais (EDW), a equilibrar os investimentos em tempo real X análise de lotes e a usar o Hadoop como um lago de dados e uma plataforma de análises.
<b>Hadoop</b>	O Hadoop é uma estrutura de software Apache modular e de código aberto para gerenciamento e análises de Big Data, capaz de escalar para petabytes de dados em hardware de padrões abertos.
<b>Serviços de integração do Hadoop de base centrada em dados</b>	Os serviços de descoberta e transformação HPE para Hadoop vão ajudá-lo a desenvolver uma abordagem formal e estruturada para a descoberta de dados e análises, incluindo uma metodologia estruturada; usar ferramentas especializadas de visualização e compartilhamento de dados como parte de uma plataforma integrada; aprimorar as habilidades e os métodos de colaboração de sua equipe; e apresentar práticas recomendadas, simultaneamente ambiciosas e de custo acessível, para apoiar a experimentação. Os serviços de integração HPE para Hadoop vão ajudá-lo a otimizar e modernizar sua infraestrutura e lidar com seus requisitos adicionais de sistema, levando alta confiança à sua plataforma Hadoop.
<b>HPE IDOL 10</b>	Com o HPE Intelligent Data Operating Layer (IDOL), você pode acessar, analisar, compreender e agir sobre petabytes de informações humanas armazenadas em seu lago de dados do HDFS (contextualmente e em vários idiomas), a partir de praticamente qualquer fonte, incluindo tanto a nuvem quanto fontes locais, com 500 funções de análises para aplicação em texto, imagens, áudio, vídeos, mídia social e dados estruturados em bancos de dados. O IDOL possui conectores prontos para uso com mais de 1.000 formatos padrão de arquivos e 400 interfaces de bancos de dados e aplicativos. O IDOL pode ser usado com o Hadoop para categorizar, indexar e extrair sentido de seu lago de dados Hadoop.
<b>HPE Vertica para SQL em Apache Hadoop</b>	O HPE Vertica para SQL em Apache Hadoop oferece uma solução empresarial mais madura para a realização de consultas SQL em seus dados Hadoop, oferecendo sintaxe SQL ANSI completa em MapR, Hortonworks e Cloudera. Com o HPE Vertica para SQL em Apache Hadoop, seus analistas de negócios, cientistas de dados e desenvolvedores contam com integração e acesso imediatos às suas ferramentas tradicionais de BI e às mais recentes ferramentas de visualização e análises padrão do setor. O HPE Vertica oferece análises avançadas, prontas para uso empresarial, com suporte a SQL ANSI completo, conformidade ACID (atomicidade, consistência, isolamento e durabilidade), tipos de dados complexos e outros recursos que só estão disponíveis através de nossa implementação de SQL em Hadoop. Ele oferece otimizações como compactação, armazenamento em coluna e projeções, proporcionando níveis de desempenho que ultrapassam em muito o SQL padrão em Hadoop.
<b>Arquiteturas de referência HPE para Hadoop</b>	Para reduzir o risco e otimizar recursos humanos, financeiros e de data center, é essencial que você empregue a arquitetura certa para apoiar seus casos de uso específicos do Hadoop: Hive, SQL, processamento na memória do Spark, dados de fluxo Kafka, aprendizagem de máquina Mahout, etc. A Hewlett Packard Enterprise tem inúmeras arquiteturas de referência para suporte à otimização de cargas de trabalho para Hadoop baseadas em combinações de servidores Moonshot, HPE Apollo séries 2000 e 4000 Gen9 e ProLiant XL, equilibrando requisitos de computação e armazenamento para clusters Hadoop com as três distribuições de Hadoop: MapR, Hortonworks e Cloudera. A arquitetura de referência Minotaur amplia as arquiteturas de referência exclusivas do Hadoop, incorporando SAP HANA, HPE Vertica e HPE IDOL.
<b>HPE Financial Services</b>	Equilibrar seu novo lago de dados Hadoop e as análises adicionais com o ambiente e a infraestrutura que você já possui exige flexibilidade financeira. O financiamento pode ajudá-lo a responder e se adaptar às mudanças ao longo do processo de implementação, promover a inovação e alinhar os custos ao uso. Além disso, o HPE Financial Services pode ajudá-lo a reduzir os riscos de infrações normativas e danos à sua reputação associados ao descarte de seu equipamento legado, através dos serviços do HPE Asset Recovery.
<b>HPE SecureData</b>	O HPE SecureData é uma abordagem de segurança centrada em dados para o Hadoop. Ele criptografa e tokeniza os dados, desidentificando-os o mais próximo possível da fonte e transformando elementos de dados confidenciais em equivalentes utilizáveis, porém desidentificados, que mantêm o formato, o comportamento e o sentido. Essa forma protegida de dados pode ser usada em seguida em aplicativos, mecanismos de análises, transferências de dados e armazenamentos de dados, podendo ser pronta e seguramente reidentificada pelos aplicativos e usuários que precisarem dos dados. O HPE SecureData protege uma quantidade praticamente ilimitada de tipos de dados no nível dos dados.

Otimize sua estratégia de investimento em TI com novas maneiras de adquirir, usar e pagar por tecnologia, em sintonia com suas metas de negócios e transformação.

[hpe.com/solutions/hpefinancialservices](http://hpe.com/solutions/hpefinancialservices)



Inscreva-se para receber atualizações

## O que você deve fazer agora?

**Agende um workshop de transformação.** Converse com seu representante HPE sobre como inscrever-se em um workshop de transformação.

Saiba mais em  
[hpe.com/us/en/solutions/empower-data-driven.html](http://hpe.com/us/en/solutions/empower-data-driven.html)

© Copyright 2016 Hewlett Packard Enterprise Development LP. As informações contidas neste documento estão sujeitas a alterações sem aviso prévio. As únicas garantias para produtos e serviços da Hewlett Packard Enterprise são as definidas nas declarações de garantia expressa que acompanham tais produtos e serviços. Nada aqui contido deve ser interpretado como constituindo garantia adicional. A Hewlett Packard Enterprise não se responsabiliza por omissões, erros técnicos ou erros editoriais contidos neste documento.

SAP é marca comercial ou marca comercial registrada da SAP SE na Alemanha e em vários outros países. Apache Hadoop e Hadoop são marcas comerciais ou marcas comerciais registradas da Apache Software Foundation nos Estados Unidos e em outros países.

4AA6-3745PTL, maio de 2016, Rev. 2