



Data Science Academy

[www.datascienceacademy.com.br](http://www.datascienceacademy.com.br)

Data Lake – Design, Projeto e Integração

Big Data Storage  
Fornecedores de Storage Para o Hadoop



O Hadoop é uma plataforma de armazenamento e processamento de dados altamente escalável, para processar grandes volumes de dados estruturados e não estruturados. Em grande escala, queremos dizer vários petabytes de dados espalhados por centenas ou milhares de servidores ou nós de armazenamento físico.

O Hadoop, desenvolvido em 2005 e agora uma plataforma de código aberto gerenciada pela Apache Software Foundation, usa um conceito conhecido como MapReduce, que é composto de duas funções separadas.

A etapa **Map** mapeia os dados de entrada e divide-os para processamento em nós em um cluster Hadoop. Esses nós "trabalhadores" podem, por sua vez, dividir ainda mais os dados para processamento. Na etapa **Reduce**, os dados processados são coletados novamente e montados em um formato baseado na consulta original que está sendo executada.

Para lidar com a análise de dados verdadeiramente de grande escala, os desenvolvedores do Hadoop implementaram uma arquitetura de *scale-out*, baseada em muitos servidores físicos de baixo custo com processamento distribuído de consultas de dados durante a operação Map. Sua lógica era habilitar o sistema Hadoop para ser capaz de processar muitas partes de uma consulta em paralelo para reduzir o tempo de execução tanto quanto possível. Isso pode ser contrastado com o design de banco de dados estruturado que parece expandir em um único servidor usando processadores mais rápidos, mais memória e armazenamento compartilhado.

Olhando para a camada de armazenamento, o objetivo do design do Hadoop é executar o processamento distribuído com a mínima latência possível. Isso é obtido executando o processamento da operação **Map** no nó que armazena os dados, um conceito conhecido como localidade de dados. Como resultado, as implementações do Hadoop podem usar unidades SATA conectadas diretamente ao servidor, mantendo assim o custo total do sistema o mais baixo possível.

Para implementar a camada de armazenamento de dados, o Hadoop usa um recurso conhecido como HDFS ou o Hadoop Distributed File System. O HDFS não é um sistema de arquivos no sentido tradicional e geralmente não é montado diretamente para um usuário visualizar (embora existam algumas ferramentas



disponíveis para isso), o que às vezes pode dificultar a compreensão do conceito; talvez seja melhor pensar nisso simplesmente como o armazenamento de dados do Hadoop.

As instâncias do HDFS são divididas em dois componentes: o namenode, que mantém metadados para rastrear o posicionamento de dados físicos no cluster Hadoop e os datanodes, que realmente armazenam os dados.

Você pode executar vários datanodes lógicos em um único servidor, mas uma implementação típica executará apenas um por servidor em um cluster. O HDFS suporta um único espaço de nome do sistema de arquivos, que armazena dados em um formato hierárquico tradicional de diretórios e arquivos. Os dados são divididos em blocos de 128 MB que são espelhados em três partes no cluster para fornecer resiliência. Obviamente, em clusters muito grandes do Hadoop, a falha de um componente ou mesmo a falha total do servidor ocorrerá, portanto, a duplicação de dados em muitos servidores é um requisito importante de design do HDFS.

### **Analisando os principais recursos do Hadoop, como eles se traduzem em armazenamento?**

O suporte ao Hadoop no armazenamento compartilhado não funciona no sentido tradicional, já que a carga de trabalho e a distribuição de armazenamento são inerentes ao Hadoop. No entanto, estamos vendo cada vez mais produtos de diversos fornecedores de armazenamento que suportam o Hadoop nativamente e apontam o caminho para uma possível direção futura no armazenamento de arquiteturas distribuídas em larga escala.

O Hadoop foi projetado para aproximar a computação dos dados e fazer uso de recursos maciços de scale-out. Isso não se ajusta bem às implementações SAN (Storage Area Network) tradicionais, que têm um custo muito mais alto por GB a ser implementado do que o que pode ser obtido com o DAS (Direct-Attached Storage) local.

Certamente não é prático considerar o uso de Fibre Channel em implantações do HDFS devido ao custo da implementação em termos de

adaptadores de barramento de host (HBA) e portas SAN. Além disso, o HDFS é projetado para atender a dados de fluxo contínuo (Streaming de dados), já que as transações do Hadoop geralmente gravam dados uma vez no cluster e, em seguida, o leem várias vezes. Isso funciona bem com unidades SATA conectadas diretamente, mas não tão bem com ambientes de armazenamento compartilhado, em que o mesmo disco físico subjacente é usado para suportar o cluster do Hadoop.

No entanto, existem cenários em que as soluções de armazenamento existentes podem ser usadas. Por exemplo, os nós do Hadoop podem ser implantados com armazenamento hierárquico para obter desempenho adicional de discos flash SSD e PCIe.

A classificação por níveis pode ser obtida usando software de código aberto, como o Flashcache, desenvolvido pelo Facebook, que mantém ativamente blocos de dados usados em armazenamento de alto desempenho. Haveria um custo associado à implantação do flash em servidores, e isso precisaria ser considerado em relação à eficiência de cada nó, mas o acréscimo de flash pode tornar os nós capazes de processar mais consultas.

A outra opção é observar os fornecedores de armazenamento que fornecem suporte nativo ao HDFS em seus produtos existentes, especialmente aqueles que são soluções de scale-out e, portanto, têm a mesma natureza distribuída em seu próprio design.

O uso de uma solução de armazenamento pronta para HDFS pode fornecer vários benefícios. Em primeiro lugar, a computação e o armazenamento podem ser escalonados independentemente, e não dentro da capacidade fixa de um nó. Há também a opção de fornecer dados mais rápidos e visualizar o conteúdo dos arquivos diretamente no cluster. Os dados e a capacidade podem ser compartilhados entre várias instâncias do Hadoop e um nível maior de proteção em torno dos metadados do HDFS pode ser fornecido. Em uma implementação padrão do Hadoop, o namenode é um ponto único de falha, mas pode ser replicado manualmente.

Embora o HDFS forneça recursos para recuperação automatizada de dados e verificação de integridade, uma solução de armazenamento pronta para HDFS



pode descarregar esse trabalho (deixando a instância do Hadoop capaz de processar mais) e reduzir a necessidade de manter três cópias de dados em uma instância.

## **Fornecedores de Armazenamento Para o Hadoop**

A plataforma NAS scale-out Isilon da EMC fornece suporte nativo para o Hadoop e adiciona recursos extras, incluindo um namenode distribuído, proteção de dados por meio de snapshots e backups NDMP e suporte multiprotocolo. Um cluster Isilon pode ser usado para várias cargas de trabalho, tornando-se uma boa maneira de avaliar as soluções do Hadoop sem implantações caras em larga escala.

O Cleversafe suporta o Hadoop como um substituto para a camada de armazenamento HDFS. Sua tecnologia dispersa os dados em vários dispositivos de armazenamento, proporcionando resiliência e desempenho adicionais mesmo além das fronteiras geográficas. À medida que o número de nós de armazenamento em uma solução Cleversafe aumenta, a resiliência também aumenta sem exigir capacidade adicional, tornando-a uma solução de armazenamento eficiente.

O Open Solution for Hadoop da NetApp é baseado no appliance de armazenamento direto E2660 e em um controlador FAS2040 que gerencia o namenode. O E2660 fornece proteção RAID aos dados, reduzindo a necessidade de reter várias cópias em uma instância do Hadoop e se conecta por meio de SAS de 6 Gbps a até quatro servidores físicos que atuam como datanodes.

A Hitachi Data Systems (HDS) tem uma arquitetura de referência para o Hadoop baseada na distribuição Cloudera Hadoop. Ele usa o produto de servidor HDS CR220S com unidades SATA de 3 TB integradas, mas não usa nenhum dos principais produtos de armazenamento do HDS. Essa também é a abordagem adotada pela HP com sua solução AppSystem for Apache Hadoop, que também usa armazenamento de conexão direta na implementação.



Existem outras soluções em desenvolvimento que visam substituir o HDFS, incluindo o suporte Lustre da Intel, o GPFS da IBM e soluções de código aberto, incluindo o Ceph e o Cassandra. Além da distribuição da Intel, essas soluções ainda não são suportadas diretamente pelos fornecedores.

Recomendamos você visitar os sites dos fornecedores citados para obter mais detalhes sobre as soluções de hardware para o Hadoop.