

3.12.2013



FAKULTA
APLIKOVANÝCH VĚD
ZÁPADOČESKÉ
UNIVERZITY
V PLZNI

Hadoop a HDFS

<KIV>

KATEDRA INFORMATIKY
A VÝPOČETNÍ TECHNIKY

Bc. Milan Nikl

Co je Hadoop:


- Open-Source Framework
- Vyvíjený Apache Software Foundation
- Pro ukládání a zpracovávání velkých objemů dat
- Big Data – trojrozměrný růst dat (3V)
 - Objem (Volume) – množství dat se zvyšuje
 - Typ (Variety) – roste různorodost
 - Rychlost (Velocity) – zrychluje se vytváření dat – roste potřeba rychlé analýzy (real time)



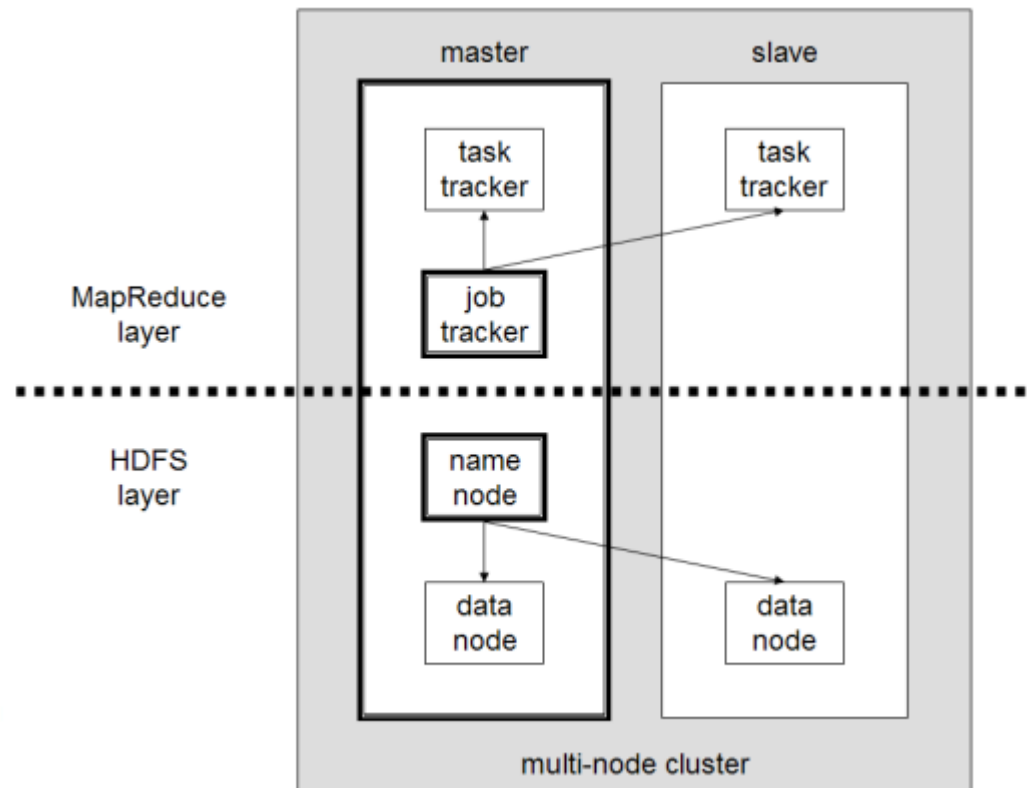


Vývoj Hadoopu:

- Vyvíjen od r. 2006
 - 04.09.2007 – verze **0.14.1**
 - 27.12.2011 – verze **1.0.0**
 - 15.10.2013 – verze **2.2.0**

 - Odvozený od Google File System (GFS)
 - Založený na programovacím modelu MapReduce
- 


Struktura Hadoopu:

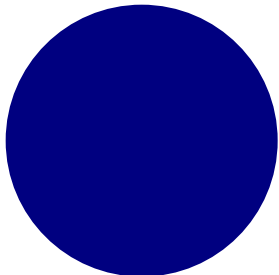




Komponenty Hadoopu:


1. MapReduce - přehled

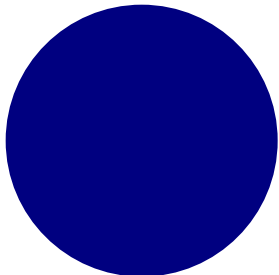
- Framework zajišťující distribuci a paralelizaci úloh
 - Dělí komplexní úlohy (Job) na dílčí (Task)
 - Dále obstarává veškerou komunikaci a datové přenosy
 - Řeší redundantní ukládání aktuálně zpracovávaných dat a s tím související problematiku přidávání a odebírání uzlů
 - Sort dat v objemu PB za pár hodin
- 



Komponenty Hadoopu:


1. MapReduce - princip

- Snaha je dostat výpočet co nejbližší datům – minimalizuje zbytečné přenosy
 - JobTracker (Master)
 - Přijímá úkoly od klienta
 - Rozděluje tasky, monitoruje, obnovuje
 - TaskTracker (Slave)
 - vykoná zadanou úlohu a vrátí odpověď
 - Každý task se spouští na nové JVM
 - Z jednotlivých odpovědí slave uzlů se sestavuje odpověď na původní dotaz
- 



Komponenty Hadoopu:


1. MapReduce - princip

- Uživatel řeší pouze co se s daty děje, nikoli jak se to dělá. Tedy:
 - Funkci Map()
 - Vybere z dat relevantní údaje
 - Dvojice <Klíč, hodnota>
 - Funkci Reduce()
 - Nad vybranými údaji provádí nějaký výpočet
 - Obvykle nějaká forma analýzy, statistiky, atd.
 - Vrátí opět <Klíč, hodnota> ale jiný typ – např. výsledek statistiky
- 



Komponenty Hadoopu:

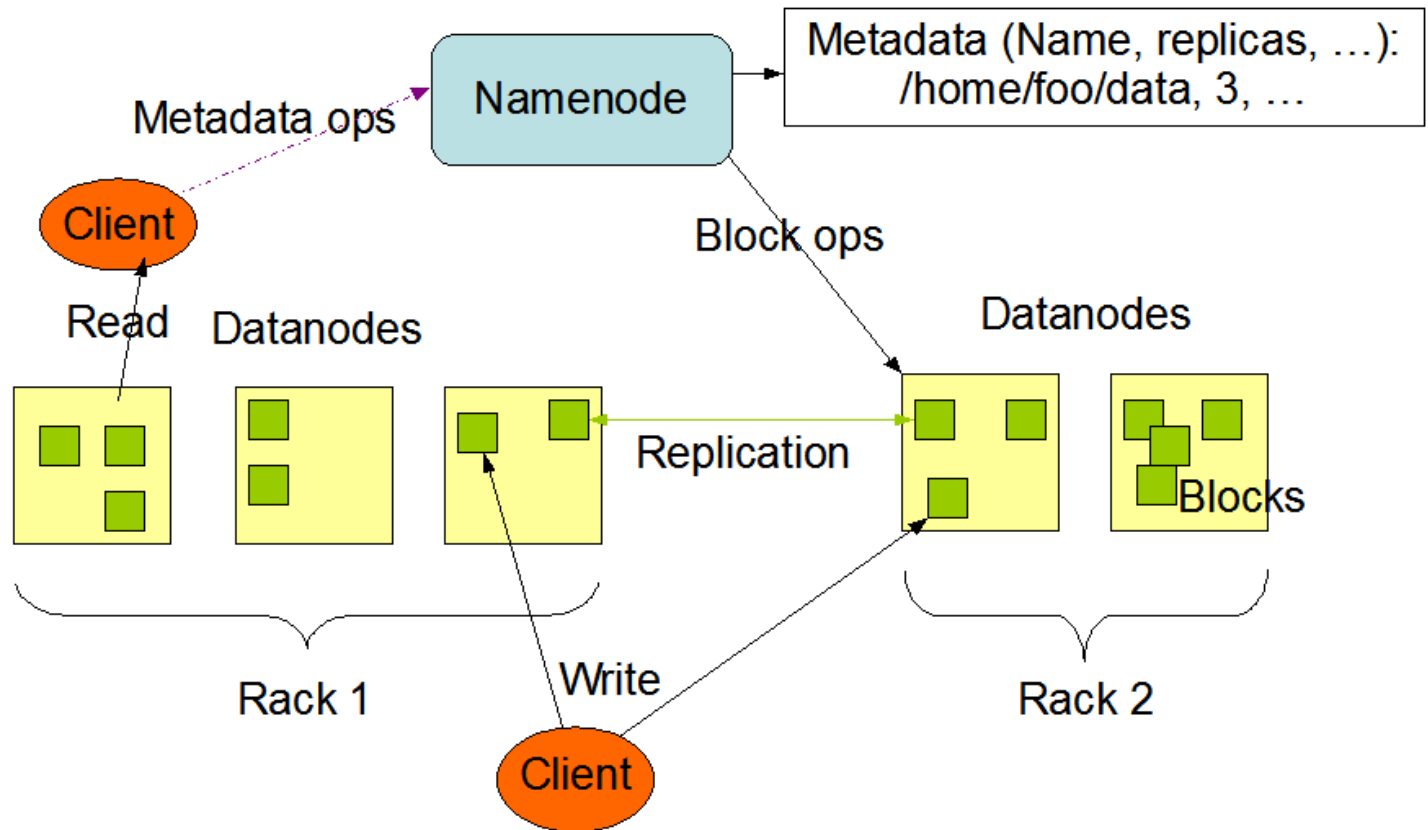
2. HDFS - vlastnosti

- Distribuovaný, škálovatelný souborový systém
 - Vysoce odolný proti chybám
 - Velká míra redundance – nevyžaduje RAID
 - Navržený pro použití na běžném HW
 - Naprogramovaný v Javě
 - Využívá TCP/IP pro komunikaci
 - Pro práci vzdálené volání procedur (RPC)
- 

Komponenty Hadoopu:

2. HDFS - struktura


HDFS Architecture





Komponenty Hadoopu:


2. HDFS - struktura

- Řídící prvek - Namenode (Master)
 - Spravuje FS namespace, mapuje bloky 64MB
 - Řídí požadavky klientů
 - Otevření, uzavření, přejmenování souboru
 - Obsahuje všechna metadata
 - Pracující prvky – Datanode (Slave)
 - Zařizuje práci s bloky podle instrukcí
 - Čtení, zápis, vytváření a mazání bloků
 - Zajišťuje replikaci bloků
 - Standard jsou 3 repliky na blok
 - HeartBeat, BlockReport
- 



Komponenty Hadoopu:


2. HDFS – Bezpečnost

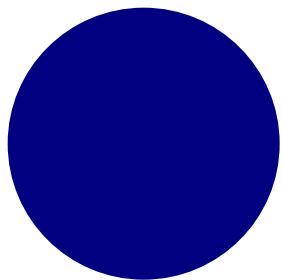
- Namenode zaznamenává všechny změny do EditLogu
 - Vytváření souborů, ale i změna replikace
 - Popis celého FS namespace uložen v souboru FsImage na Namenode
 - Checkpoint
 - Namenode načte EditLog a FsImage
 - Provede změny popsané v EditLogu
 - Vytvoří nové verze souborů
 - Zatím pouze při startu
 - Snapshoty
- 



Komponenty Hadoopu:

2. HDFS – omezení a výhody

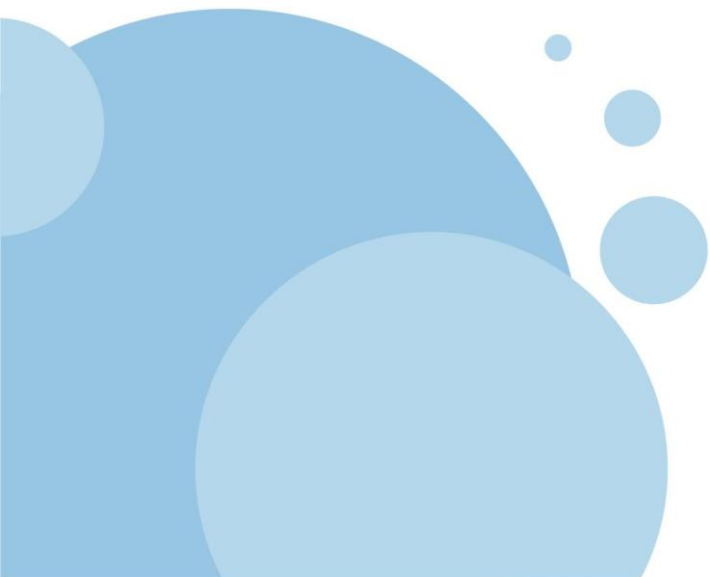
- Nejedná se o klasický FS – fyzický mount pod OS není možný
 - Jde o použití FUSE
 - Zaměřený na čtení dat – u většiny se předpokládá, že se nebudou měnit
 - Nevhodný pro potřeby vícenásobného současného zápisu souborů
 - /trash – možnost obnovení souborů (6h)
 - Používá Java API, ale pomocí Thrift API lze používat i v jiných jazycích
- 



Komponenty Hadoopu:

3. Hive

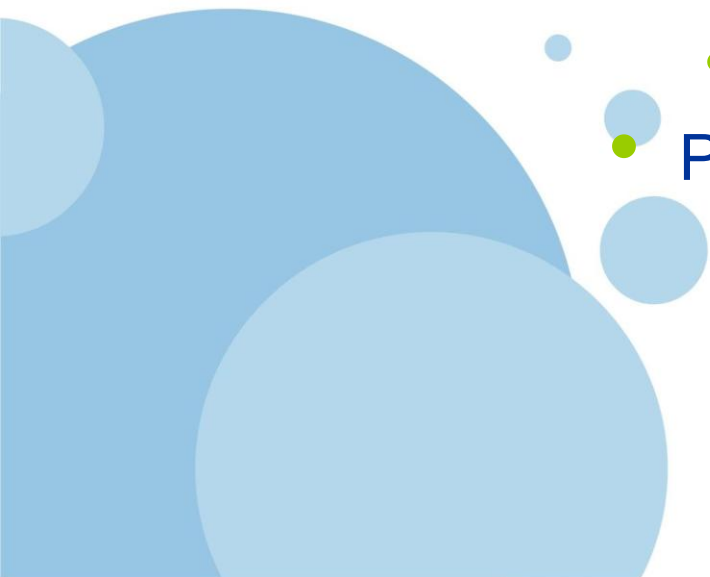
- Použití SQL-like dotazů nad daty v HDFS
- Vlastní jazyk HQL
- Převádí dotazy na MapReduce úlohy
- Oproti SQL značně omezen (pouze čtení, vyšší latence)
- Vytvořen ve Facebooku





Komponenty Hadoopu:

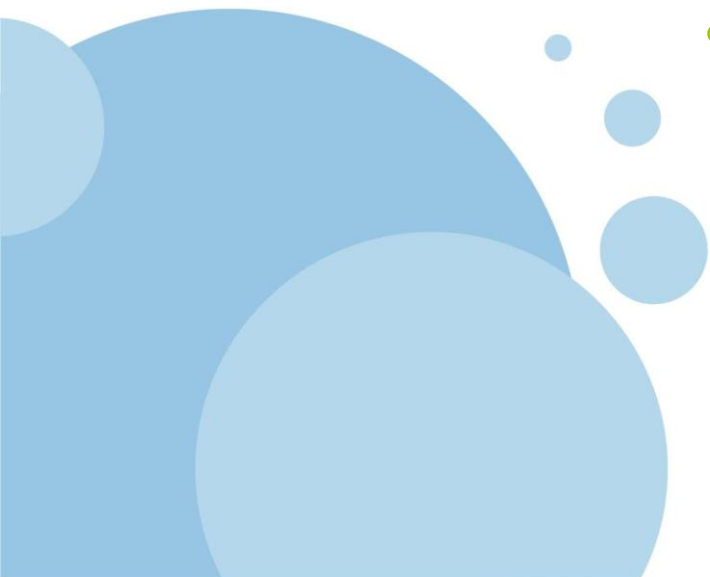
4. Pig

- Vyvinut společností Yahoo!
 - Pro jednoduché analyzování rozsáhlých dat – paralelizaci neřeší uživatel
 - Používá vlastní jazyk Pig Latin
 - Zaměřen na jednoduchost psaní kódu
 - Struktura kódu umožňuje dobrou čitelnost, udržitelnost a snadnou optimalizaci
 - Snadno rozšiřitelný
 - Převádí Pig Latin na MapReduce úlohy
- 



Komponenty Hadoopu:


5. Mahout

- Knihovna pro strojové učení
 - Základní typy úloh:
 - **Recommendation mining** – analyzuje chování uživatele a provádí doporučení
 - **Clustering** – seskupuje obsahově podobné dokumenty
 - **Klasifikace** neoznačených dokumentů
 - **Frequent itemset mining** – analyzuje výskyt určitých objektů – které položky se nejčastěji vyskytují společně – např. nákupní košík
- 



Komponenty Hadoopu:


6. Cassandra

- Alternativa k HDFS
 - NoSQL databázový systém
 - Využívá decentralizaci a replikaci dat
 - Odolnost proti poruchám a výpadkům
 - Výkonný, přesto přizpůsobitelný systém
 - Adaptabilní – snadná rozšiřitelnost
 - V praxi používá např. Ebay, Twitter
 - Největší Cluster 300 TB na 400 strojích
- 




Komponenty Hadoopu:

7. HBase

- Sloupcově orientovaný přístup
 - Vhodné pro náhodný Read/Write přístup k Big Datům
 - Nadstavba nad HDFS, která umožňuje práci s rozsáhlými tabulkami (mld x mil)
 - Není relační databázový systém, ani nepodporuje SQL
 - Postavena na Javě, psána podobně jako MapReduce
- 



Podporované FS:

- Amazon S3 filesystem – používané u clusterů založených na Amazon Cloudu
 - FTP file servery – ukládání dat na vzdálené servery
 - HTTP a HTTPS file systémy – pouze pro čtení
 - Teoreticky provozovatelný na jakémkoli distribuovaném FS podporujícím *file:// URL*
 - Hrozí ale ztráta výkonnosti
- 




Významní uživatelé:

- Ebay
 - 532 Nodů v clusteru, 8x532 jader, 5,3 PB
 - Facebook
 - 1100 Nodů, 8800 jader, 12 PB
 - 300 Nodů, 2400 jader, 3 PB
 - Spotify
 - 690 Nodů, 8280 jader, 28 PB
 - Yahoo!
 - Přes 40000 počítačů
 - 4500 Nodů, 36000 jader, 18 PB
 - Přes 60% tvoří Pig
- 



Použití Hadoop:

- Nativní řešení pouze pro linuxové distribuce (nejčastěji Red Hat, CentOS)
 - Na windows přes Cygwin
 - Příp. přes Wmware

 - Základní verze od ASF
 - X
 - Enterprise řešení různých firem (IBM, HortonWorks, Cloudera, Oracle, ...)
- 



Závěrem:

Děkuji za pozornost.





Zdroje:

- <http://hadoop.apache.org/>
 - <http://www.linuxexpres.cz/software/kdyz-se-rekne-hadoop>
 - <http://www.systemonline.cz/clanky/big-data.htm>
 - <http://www-01.ibm.com/software/data/infosphere/biginsights/>
 - <http://en.wikipedia.org/wiki/Hadoop> a další odkazy zde nalezené
- 