



Když se řekne

„HADOOP“

Ondřej Dolák

Hadoop je jeden z top projektů Apache Software Foundation. Můžeme si představit, že se jedná o výpočetní prostředí postavené nad distribuovaným souborovým systémem, který byl navržen speciálně pro velmi rozsáhlé operace s daty. Hadoop byl inspirován prací Googlu na vlastním souborovém systému (GFS) a MapReduce programovacím paradigma. MapReduce není nová koncepce, přesto byl Hadoop první systém, který ho udělal prakticky použitelným pro mnohem širší soubor úloh.

Na rozdíl od transakčních systémů je Hadoop navržen tak, aby procházel velké soubory dat a prováděl nad nimi výpočty prostřednictvím vysoce škálovatelného, distribuovaného, dávkového systému. Klíčem škálovatelnosti Hadoopu je model, kde jsou výpočetní funkce posílány k datům, na rozdíl od běžného směru, kdy data putují k funkcím. Hadoop však není o odezvách v řádu milisekund, realtime warehousingu nebo o transakční rychlosti, ale o hledání informací v tak velkém množství dat, které je běžnými prostředky technicky i ekonomicky nemožné.

Hadoop – složení

Hadoop je nejčastěji vnímán jako systém sestávající ze dvou komponent – systému souborů (Hadoop Distributed File System) a programovacího paradigma (MapReduce). Jednou z klíčových vlastností Hadoopu, na kterých je postaven, je redundance. Nejen že jsou data uložena redundantně na více místech po celém clusteru, ale programovací model je takový, že poruchy jsou očekávány a vyřešeny automaticky tím, že části programu jsou spuštěny na jiných serverech v clusteru. Díky této redundanci je možné distribuovat data a přidružené funkce i přes velmi velké clustery, kde je možné očekávat chybu uzlu s velmi vysokou pravděpodobností.

Mimo tyto hlavní stavební kameny existuje celá řada souvisejících projektů, které poskytují další nástroje pro práci s daty v Hadoopu. Některé z nich zde uvádíme:

Nástroje pro vývoj

- **MapReduce** – psaní paralelních spustitelných programů se ukázalo v průběhu let jako velmi náročný úkol vyžadující různé specializované schopnosti. MapReduce

poskytuje programátorům schopnost produkovat paralelní distribuované aplikace mnohem snadněji tím, že stačí napsat pouze funkce `map()` a `reduce()`, které řeší logiku specifického problému, zatímco framework MapReduce se automaticky stará o řízení distribuovaných serverů, paralelizaci úloh, řízení veškeré komunikace a datových přenosů mezi různými částmi systému a zajištění redundance v případě selhání.

- **Hive** – jazyk vyvinutý ve Facebooku umožňující psát dotazy v Hive Query Language (HQL), což je obdoba klasického SQL. HQL dotaz je pak Hive službou přeložen na MapReduce úlohy. Oproti SQL je zde řada funkčních omezení (operace pouze pro čtení, vysoká latence vyplývající z podstaty MapReduce atd.).
- **Pig** – původně vyvinut společností Yahoo!, aby se lidé mohli soustředit více na analýzu dat a strávit méně času psaním MapReduce úloh. Pig se skládá ze dvou částí: první je jazyk sám, nazývá se PigLatin a druhým je běhové prostředí, ve kterém jsou programy v PigLatin jazyce překládány na sérii MapReduce úloh.
- **Jaql** – funkční, deklarativní dotazovací jazyk pro JavaScript Object Notation (JSON) umožňující vybrat, spojovat, agregovat a filtrovat data uložená v HDFS. Kód v Jaql je opět převáděn na MapReduce úlohy.
- **Mahout** – škálovatelná knihovna pro strojové učení. Jsou v ní implementovány algoritmy pro clustering, klasifikaci a kolaborativního filtrování optimalizované pro běh v prostředí Hadoopu.

Ukládání dat a správa metadat

- **HDFS** – Hadoop Distributed File System (HDFS) je distribuovaný souborový

systém navržený pro provoz na běžném hardwaru. Má mnoho stejných vlastností jako další distribuované souborové systémy, ale zároveň má některé unikátní vlastnosti. HDFS je především vysoce odolný proti chybám a je určen k nasazení na komoditním hardwaru. HDFS poskytuje vysokou propustnost v přístupu k aplikačním datům a je vhodný pro aplikace, které obsahují velké datové sady.

- **Cassandra** – nejedná se o souborový systém, ale o NoSQL (klíč-hodnota) úložiště. Cassandra je vhodnou alternativou k HDFS v aplikacích, které vyžadují rychlý přístup k datům z/do Hadoopu.
- **HBase** – sloupcově orientovaný databázový systém, který běží nad HDFS, a který je vhodný pro řídké soubory dat. Na rozdíl od relačních databázových systémů, HBase nepodporuje strukturovaný dotazovací jazyk (jako např. SQL), ve skutečnosti HBase není ani relační úložiště dat. HBase aplikace jsou psány v Javě stejně jako typické MapReduce aplikace.
- **HCatalog** – systém pro správu tabulek a úložišť v Hadoop, který umožňuje uživatelům s různými systémy pro zpracování dat (Pig, MapReduce a Hive) snadněji číst a zapisovat data v Hadoopu. HCatalog poskytuje uživatelům relační abstrakci nad daty uloženými v HDFS a unifikuje přístup k datům v různých formátech.

Řízení

- **Zookeeper** – koordinační služba pro distribuované aplikace. Vystavuje služby pro správu konfigurace, pojmenování, synchronizace a skupinové služby.
- **Oozie** – systém pro správu Hadoop jobů (workflow scheduler).

Distributor	Apache	Cloudera	DataStax	EMC	Hortonworks	IBM	Intel	MapR	Microsoft	
Název distribuce	Hadoop	CDH4	Enterprise	Greenplum HD	Data Platform	InfoSphere BigInsights	Distr. for Hadoop	M5 Edition	HDInsight Server (beta)	
Verze distribuce	1.0.4	4.2.0	3.0	1.2	1.2.3.1	2.0	2.3	2.1.2	Technology Preview	
Verze Hadoop	1.0.4	2.0.0	1.0.4	1.0.3	1.1.2	1.0.3	1.0.3	0.20.2	HDP Win 1.1.0 (beta)	
Integrované subprojekty	Cassandra	x	x							
	Chukwa	x								
	Flume	x	x			x	x	x		
	HBase	x	x		x	x	x	x		
	HDFS	x	x	x	x	x	x	x		
	HCatalog	x	x			x	x	x	x	
	Hive	x	x	x	x	x	x	x	x	
	Jaql						x			
	Mahout	x	x	x	x	x	x	x		
	MapReduce	x	x	x	x	x	x	x		
	Oozie	x	x			x	x	x	x	
	Pig	x	x	x	x	x	x	x	x	
Sqoop	x	x	x		x	x	x	x		
Zookeeper	x	x		x	x	x	x			
Správa	Clusterů	Apache Ambari	Cloudera Manager	DataStax OpsCenter	Greenplum Command Center	Apache Ambari	BigInsights Console	Intel Manager	MapR Control System	HDInsight Dashboard / Ambari
	RT/LL data management		Impala (beta)				x	x		
Integrace	POSIX příst.						až ve v2.1		x	
	SQL přístup		Impala (beta)		až v Pivotal HD	Teradata SQL-H	až ve v2.1			PolyBase
Masazení	Vysoká dostupnost	ve v2.0 alpha	x			x	až ve v2.1		x	
	Certifikace pro OS Appliance	Cross platform	RHEL, CentOS, Oracle L., SLES, Ubuntu, Debian +Oracle	RHEL, CentOS, Oracle L., SLES, Debian	RHEL, CentOS	RHEL, CentOS, SLES, WIN (beta) +Teradata	RHEL, SLES	RHEL, CentOS, Oracle L., SLES	RHEL, CentOS, SLES, Ubuntu	Windows

Vysvětlivky: integrované subprojekty – v rámci příslušné distribuce Hadoopu; real time / low latency data management – podpora zpracování dat v reálném čase; vysoká dostupnost – žádná porouchaná komponenta nesmí způsobit výpadek systému; POSIX přístup – kompatibilní POSIX rozhraní pro HDFS; SQL přístup – nativní SQL dotazovací rozhraní pro Hadoop; appliance – možnost dodání komplexního optimalizovaného řešení včetně hardwarových komponent

Agregování a získávání dat

- **Sqoop** – nástroj určený pro efektivní přenos hromadných dat mezi Hadoop a strukturovanými datovými sklady, jako jsou relační databáze.
- **Chukwa** – systém sběru dat pro správu rozsáhlých distribuovaných systémů. Je postaven na HDFS a MapReduce frameworku. Obsahuje flexibilní a výkonné nástroje pro zobrazování, sledování a analýzu výsledků ze shromážděných dat.
- **Flume** – služba pro distribuované přenosy, shromažďování a agregování velkého množství dat z datových logů.

Kupujeme Hadoop

Pokud uvažujeme o pořízení Hadoopu, máme v podstatě na výběr dvě možnosti: cloud a on-premises. Obecně porovnání těchto možností by vyžadovalo minimálně samostatný článek, ale pro srovnání Hadoop distribucí nám může stačit informace, že většina cloudových řešení využívá distribuce od dodavatelů on-premises. Samotný Hadoop a jeho komponenty lze

pořídit zdarma, jelikož se jedná o open-source produkt, který zároveň nemá specifické požadavky na hardware. Nicméně tato varianta přesouvá veškerou tíhu instalace, integrace a provozu na vlastní zdroje firmy, což vzhledem ke značné komplexitě a „surovosti“ některých komponent nemusí být realizovatelné. Na tento problém zareagovala většina velkých IT firem i nových startupů a nabízejí vlastní distribuce Hadoopu, které by tyto nedostatky měly řešit. Některé distribuce pak řeší i problémy, které jsou v samotném Hadoopu nedostupné, jako například vysoká dostupnost nebo POSIX přístup k HDFS. Tabulka poskytuje přehled distribucí a jejich vlastností k březnu 2013.

Používáme Hadoop

Jak již bylo zmíněno, primární síla Hadoopu je v nákladově efektivní škálovatelnosti s komoditním hardwarem. Poskytuje podporu pro zpracování všech typů dat, ať strukturovaných, semistrukturovaných nebo nestrukturovaných, a jeho otevřená rozšiřitelnost umožňuje

vývojářům rozšířit jej pro širokou škálu aplikací. Úloha Hadoopu ovšem není nahradit stávající systémy. Naopak Hadoop rozšiřuje jejich schopnosti tím, že umožňuje zpracování velkého objemu dat mimo ně, takže stávající systémy se mohou soustředit pouze na to, co je jejich primární úlohou. Jako jeden z příkladů lze uvést nahrávání dat do klasického relačního datového skladu. Datový sklad vychází většinou z předpokladů, na základě kterých jsou vstupní data filtrována a transformována. Pokud ovšem tyto předpoklady v průběhu času měníme, dostáváme se do situace, kdy již nedisponujeme původními surovými daty, protože jejich ukládání v relační databázi je neefektivní. V případě, že bychom surová data ukládali v Hadoopu a pak z nich teprve plnili datový sklad, lze pak všechny předpoklady ověřovat přímo na nich a zároveň je libovolně měnit, protože vždy máme k dispozici bez omezení kompletní sadu dat. ■

Autor je teamleader pro oblast big data ve společnosti Sophia Solutions.