

HDInsight: Hadoop și Hive

Conf.dr. Cristian Kevorchian
Facultatea de matematică și informatică

cristian.kevorchian@unibuc.ro

Big Data

MapReduce, Streaming, ML(Machine Learning),
MPP(Massively Parallel Processing)



"Scale Out" pentru
"Pay As You Go"



Schema pentru
Read Nu Write



BASE NU ACID

Prea Mare, Complex, sau Scump pentru mediile IT
exploatate curent

Noi probleme



Tehnologii Big Data

Relational / Baze de Date Structurate (SQL Server și PDW)

Enterprise BI și Instrumente Analitice (Office, Analysis Services)

Multi-Dimensional și Tabular / Baze de Date Structurate (Servicii de analiză)

Hadoop MapReduce Stocare Distribuită & Procesare Date Multi-Structurate (HDInsight)

Surse de Date Multiple

Procese Orientate Eveniment

SSBI (Self Service BI) și Instrumente Analitice (Office, Analysis Services)

Cloud (Azure) Flexibilitate+ Opțiuni On-Premises

Apache Hadoop, Hortonworks, Hdinsight

Orchestrare de tehnologii



Apache Hadoop



Hortonworks



Microsoft

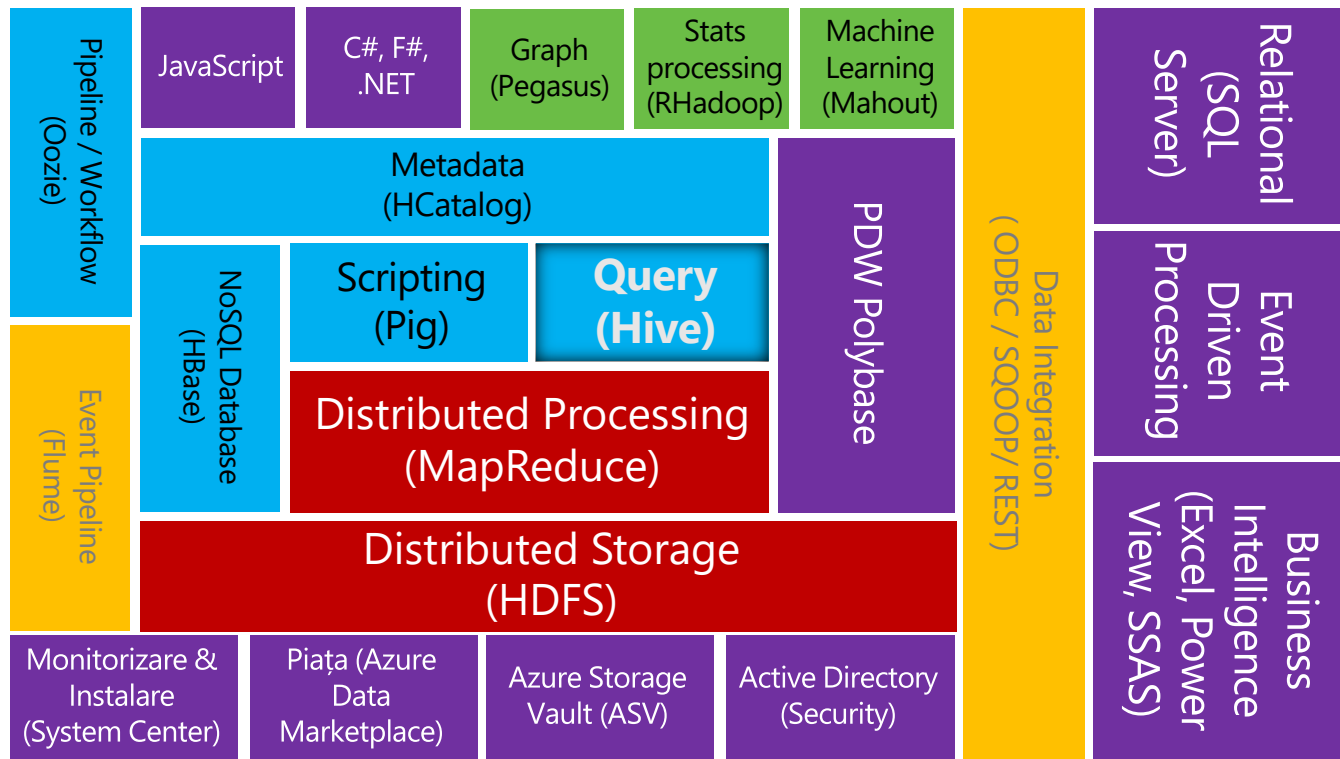
HDInsight

Comunitate Open Source
"Noi Consumăm Cod
Noi Contribuim la
realizarea Codului
Același "Core Code" în
toate distribuțiile"

Partener Microsoft
Important Actor în
Open Source Hadoop
Relație de "Trust" in
Comunitatea Open
Source

Serviciul HDInsight,
Server HDInsight pe
platforma Hortonworks
Funcționalități
adiționale

Arhitectura HDInsight / Hadoop



Legend

Red = Core Hadoop

Blue = Data processing

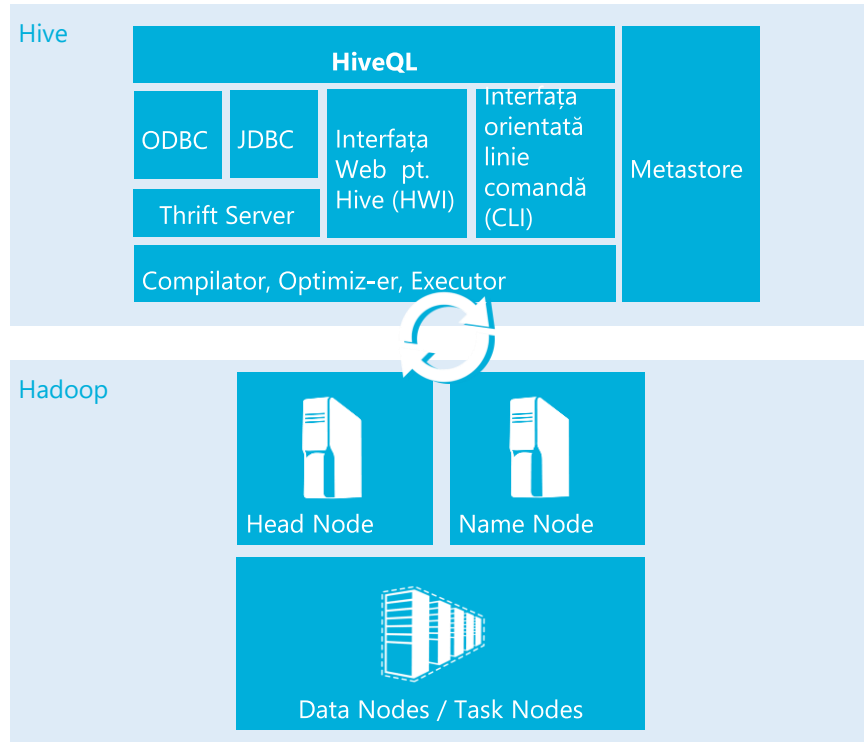
Purple = Microsoft integration points and value adds

Yellow = Data Movement

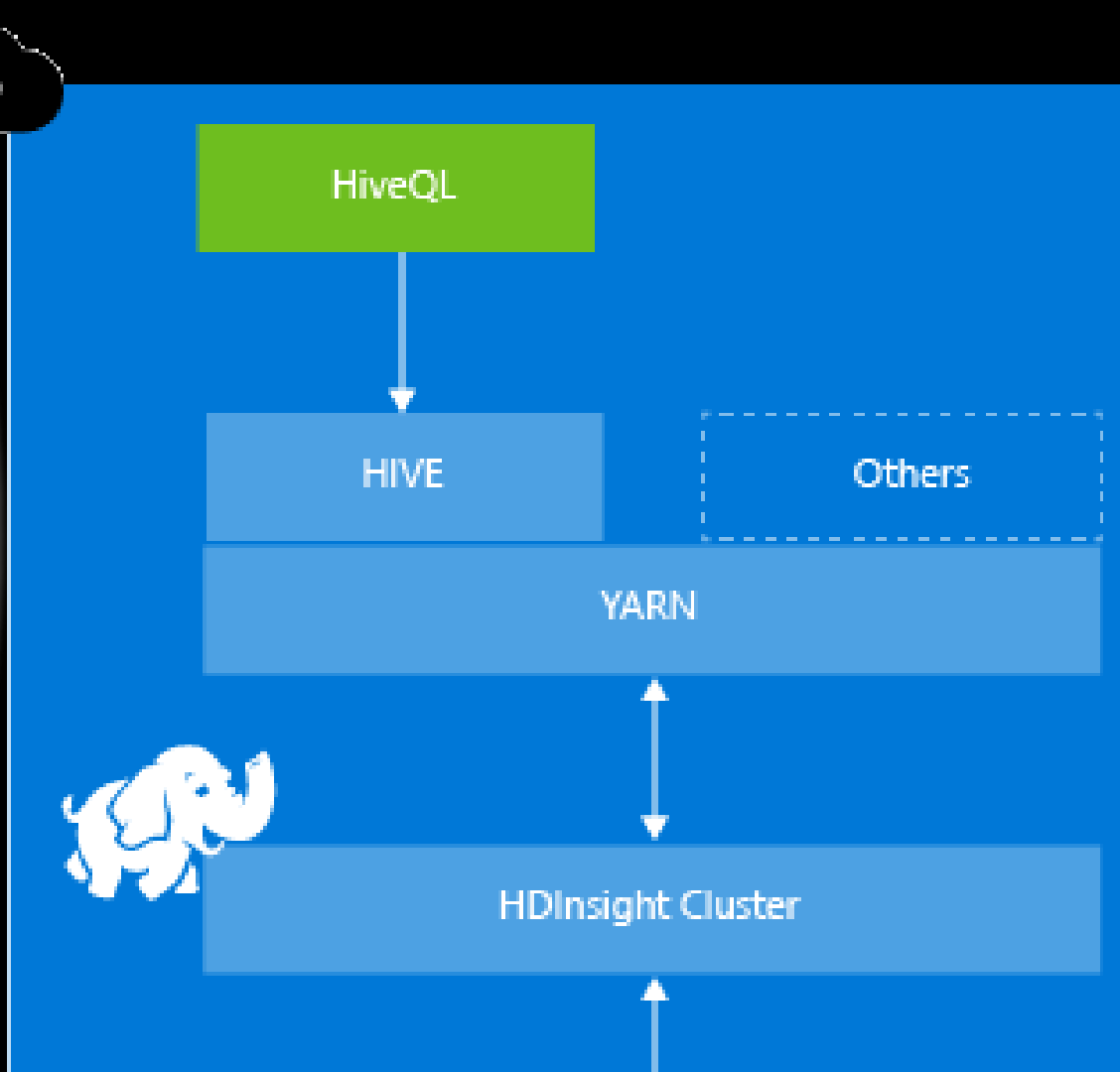
Green = Packages

White = Coming Soon

Arhitectura HIVE



Hadoop 2.0 Arhitectura Hive



Hive for Hadoop

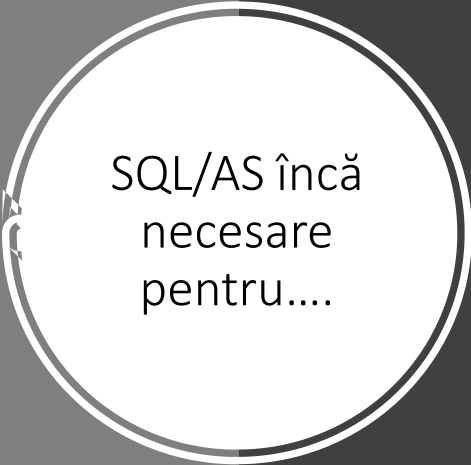
- Facilitează lucrul cu instrumentele BI prin ODBC
- Structură fără model relațional complet.
- Utilizarea funcționalităților HiveQL
- Acces simplificat la datele Hadoop

Caracteristici Hive

- Orientat batch
- Orientat Data Warehouse
- Seturi de date
- Genereaza/executa MapReduce
- Indexare limitată, fără statistici(încarca statisticile asociate tabelelor în Derby), fara cache
- Optimizare in programare
- Numai adaugare

Hive pentru SQL Pro

- Asocierea proiectelor cu tehnologia adecvată
- Noi oportunități de business și tehnologice
- Căutare, arhivare, prototip, pre-agregare,
- algoritmi rafinați, etc



SQL/AS încă
necesare
pentru....

- Actualizări, OLTP, ACID
- Submulțimi, indexari/agregări, optimizări, caching
- Aplicații, date, structuri, infrastructură deja existentă
- Fiecare interogare trebuie să fie rapidă

HIVE Fundamente

Creare Tabelă

Tabela ne-partitionată

```
CREATE EXTERNAL TABLE tabelaUnPart (type string COMMENT 'tip de tabela')
```

```
ROW FORMAT DELIMITED
```

```
FIELDS TERMINATED BY '\t'
```

```
STORED AS TEXTFILE
```

```
LOCATION '/user/demo/ckTabela';
```

Partiționată

```
CREATE EXTERNAL TABLE tabelaPart (type string COMMENT 'type of sectiune tabela')
```

```
COMMENT 'SQL Sectiune tabela'
```

```
PARTITIONED BY (an string)
```

```
ROW FORMAT DELIMITED
```

```
FIELDS TERMINATED BY '\t'
```

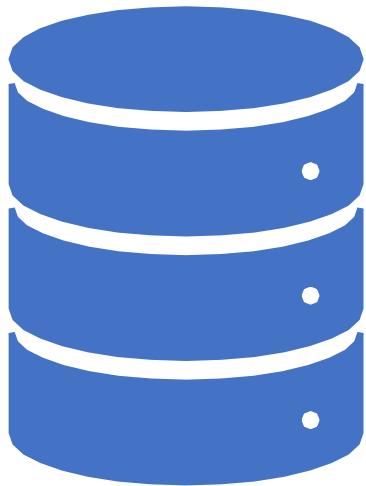
```
STORED AS TEXTFILE;
```

```
ALTER TABLE tabelaPart ADD PARTITION (Year = '1')
```

```
LOCATION '/user/demo/cktabla1';
```

```
ALTER TABLE tabelaPart ADD PARTITION (Year = '2')
```

```
LOCATION '/user/demo/ckTabela2';
```



Detalii Hive Table

- DATA TYPES
- EXTERNAL / INTERNAL
- PARTITIONED BY | CLUSTERED BY | SKEWED BY
- Terminators
- ROW FORMAT DELIMITED | SERDE
- STORED AS
- Fields/Collection Items/Map Keys
- TERMINATED BY
- LOCATION

MetaDate

- Metadatele sunt încărcate în baze de date
- MetaStore cum ar fi:
 - Derby
 - SQL Azure
 - SQL Server
- Vizualizarea Schemei
- SHOW TABLES 'ta.*';
- DESCRIBE tabelaunpart;
- DESCRIBE tabelaunpart.type;
- DESCRIBE EXTENDED tabelaunpart;
- DESCRIBE FORMATTED tabelaunpart;
- SHOW FUNCTIONS "x.*";
- SHOW FORMATTED INDEXES ON tabelaunpart;

Tipuri de Date

Primitive

- Numerice: Int, SmallInt, TinyInt, BigInt, Float, Double
- Character: String
- Speciale: Binary, Timestamp

Colectii

- STRUCT<Oras:String, Judet:String> | Struct ('Cluj', țCluj')
- ARRAY <String> | Array ('Cluj', 'Cluj')
- MAP <String, String> | Map ('City', 'Cluj', 'State', 'Cluj')
- UNIONTYPE <BigInt, String, Float>

Proprietăți

- Lungime stringului nu conteaza
- NULL depinde de SerDe

Stocare – Externă și Internă

```
CREATE EXTERNAL TABLE tabelaUnPart(...)  
  LOCATION '/user/demo/tabela';  
  LOCATION 'hdfs:///user/demo/tabela';  
  LOCATION 'asv://user/demo/tabela';
```

- Utilizăm EXTERNAL când datele au fost utilizate în afara Hive.
- Este necesar ca datele sa persiste și după DROP TABLE
 - Utilizăm o locație indicată de utilizator cum ar fi ASV(Azure Storage Volume)
 - Hive should not own data and control settings, dirs, etc.
- Use INTERNAL when you want Hive to manage the data and storage for short term usage
- Creating table based on existing table (AS SELECT)

Storage – Partition , și Bucket

- CREATE EXTERNAL TABLE tabelaPart (...)
- PARTITIONED BY (an string)
- CLUSTERED BY (type) into 256 BUCKETS
- **Partitia**
 - Directoare pentru fiecare combinație distinctă de valori partitionate
 - Cheia de partitionare nu poate fi definită în tabela.
 - Permite eliminarea de partiții
 - Util în căutările pe nivele
 - Poate conduce la o slabă performanță dacă partiția nu este referită în query
- **Bucket-uri**
 - Splitarea datelor bazată pe hash-ul unei coloane
 - Un fișier HDFS per bucket în cadrul partiției
 - Performanța poate fi îmbunătățită pentru agregări și joinuri
 - Exemplificare
 - set hive.enforce.bucketing = true;

Storage – Formatul Fisierului

```
CREATE EXTERNAL TABLE tabelaPart (...)  
ROW FORMAT DELIMITED  
FIELDS TERMINATED by '\001'  
STORED AS TEXTFILE, FISIERERC, FISIERE  
de SECVENTE, AVRO(structure complexe  
de date cu MapReduce)
```

- Format
 - În general se sterg headerele înaintea încărcării fișierelor
 - TEXTFILE este comun, util cind datele sunt alfanumerice si partajate
 - Format de stocare extensibil.
 - Extensibil pe disk/reprezentare in-memory



Încărcare—
SerDes

```
CREATE EXTERNAL TABLE ClientSerDeUtil(...)  
  ROW FORMAT SERDE  
  'com.cloudera.hive.serde.JSONSerDe'  
  LOCATION ....
```

- SerDes(Serializare/Deserializare)
 - Crează la nivel de utilizator Java Serializare/Deserializare
 - Include input/output parsing, optimizare
 - In mod uzual suprascrive CREATE TABLE DDL
 - SerDes uzuale: CSV, XML, JSON
 - bibliotecă: org.apache.hadoop.hive.serde2

Storage —HDFS si ASV

- ASV://containername@account.blob.core.windows.net/user/demo/
- HDFS:///user/demo/
- Format de stocare
 - HDFS este Hadoop Distributed File System
 - ASV este Azure Storage Vault utilizând un API în top HDFS
 - ASV permite reutilizarea peste clustere si cu alte aplicatii
 - ASV disponibilizeaza date rapid către clustere HDInsight

CREARE INDEX

- `CREATE INDEX tabelaPart_idx`
 - `ON TABLE tabelaPart (type)`
 - `AS 'org.apache.hadoop.hive.ql.index.compact.CompactIndexHandler'`
 - `WITH DEFERRED REBUILD`
 - `IN TABLE tabelaPart_index;`
 - `ALTER INDEX tabelaPart_idx ON tabelaPart REBUILD;`
-
- Chei
 - Fara chei
 - Datele asociate indexului reprezinta o alta tabela
 - Se impune REBUILD pentru a include noile date
 - `SHOW FORMATTED INDEXES` on ckTable;
 - Indexarea poate fi utila la
 - Creare unui număr mare de partiții mici
 - `GROUP BY`
-



Creare View-uri

- `CREATE VIEW tabelaUnAn (type)`
- `AS SELECT type`
- `FROM tabelaPart`
- `WHERE year = 1`
- `ORDER BY type;`
- Secvență Simplă de Cod
 - `SELECT * FROM tabelaUnAn;`
 - `DESCRIBE FORMATTED tabelaUnAn;`
- Puncte cheie
 - Nu este materializata
 - Poate avea `ORDER BY` sau `LIMIT`

Interrogari

Interogări

- `SELECT c.state_fips, c.county_fips, c.population`
- `FROM census c`
- `WHERE c.median_household_income > 100000`
- `GROUP BY c.state_fips, c.county_fips, c.population`
- `ORDER BY county_fips`
- `LIMIT 100;`
- Puncte cheie
 - Cash minimal, statistici, sau optimizări
 - În general citește întregul volum de date pentru fiecare interogare
- Performanță
 - Ordinea coloanelor, tablelele pot face diferența în zona performanței
 - Eliminarea partițiilor filtrarea rangului

Sortare

- ORDER BY
 - Sortul final realizat cu un reducer poate ridica probleme.
- SORT BY
 - Sortarea în cadrul unui reducer poate aduce performante mai bune
- DISTRIBUTE BY
 - Determina modul de interconectare a datelor implica distribuția reduce-rilor.
- SORT BY + DISTRIBUTE BY = CLUSTER BY
 - Poate copia comportamentul ORDER BY, cu performante mai bune în context de distribuire.

Join-uri

- Tipuri de Join în HIVE
 - Equality
 - OUTER - LEFT, RIGHT, FULL
 - LEFT SEMI
- Nu permite lucrul cu join-uri de tip:
 - Non-Equality
 - IN/EXISTS sub-queries (rewrite as LEFT SEMI JOIN)
- Caracteristici
 - Job-uri MapReduce multiple cu exceptia cazului in care avem aceleasi coloane in join in toate tabelele
 - Plaseaza cele mai mari tabele ultimele pentru economie de memorie
 - Join-urile sunt solutionate de la stînga la dreapta in ordinea interogării.
 - JOIN ON este complet evaluată inaintea pornirii lui WHERE.

HIVE EXPLICAȚII

- `EXPLAIN SELECT * FROM tabelaPart;`
- `EXPLAIN SELECT * FROM tabelaPart WHERE year > 1;`
- `EXPLAIN EXTENDED SELECT * FROM tabelaPart;`
- Caracteristici
- Nu executa interogarea
- Prezintă parsing-ul
- Lista stage-urilor, fișierelor temp, dependențelor, modes, output operators, etc.
 - ABSTRACT SYNTAX TREE:
 - `(TOK_QUERY (TOK_FROM (TOK_TABREF (TOK_TABNAME tabelaPart))) (TOK_INSERT (TOK_DESTINATION (TOK_DIR TOK_TMP_FIL E)) (TOK_SELECT (TOK_SELEXPR TOK_ALLCOLREF))))`
- STAGE DEPENDENCIES:
 - Stage-0 is a root stage
- STAGE PLANS:
 - Stage: Stage-0
 - Fetch Operator
 - limit: -1

Configurare

Configurare HIVE

- `cd %hive_home%\bin`
- `<install-dir>` curent: `C:\Hadoop\hadoop-1.1.0-SNAPSHOT`
- Hive configurare implicita `<install-dir>/conf/hive-default.xml`
- Configurare `<install-dir>/conf/hive-site.xml`
- Hive configurare director `HIVE_CONF_DIR` variabile de mediu
- Log4j configurare `<install-dir>/conf/hive-log4j.properties`
- Log: `c:\Hadoop\hive-0.9.0\logs\hive.log`

Utilizare HIVE

- BI în Big Data
- “Polenizare încrucișată” a capabilităților SQL!
- Corelatii-incrucisate cu Hadoop, join-uri, filtre,etc
- Permite încărcarea rezultaelor intermediare pentru accelerarea/simplificarea interogărilor
- Procesare orientată pe procese batch
- O abordare insight E2E poate fi mult mai rapida
- Proiecte potrivite la tehnologii potrivite

Pasul Următor

- De citit
 - http://sqlblog.com/blogs/lara_rubbelke/archive/2012/09/10/big-data-learning-resources.aspx
 - [Programming Hive Book](#)
 - <http://blogs.msdn.com/cindygross>
- Descarca Microsoft HDInsight Server <http://microsoft.com/bigdata> (On-Prem CTP)
- Gindeste cum poate Big Data fi utilizat in viata de zi cu zi.
- Incearca sa gindesti proiectul cit mai critic

Referințe pentru Big Data

Hadoop: The Definitive Guide by Tom White

SQL Server Sqoop <http://bit.ly/rulsjX>

Hive <http://hive.apache.org>

Excel2Hadoop via Hive ODBC <http://tinyurl.com/7c4qjjj>

Hadoop pe Azure Videos <http://tinyurl.com/6munnx2>

Microsoft Big Data <http://microsoft.com/bigdata>