

# Cypher for Apache Spark

Max Kießling

# CAPS - The Spark SQL for graphs

## Spark SQL

```
spark
  .read.parquet("hdfs://...")
  .createTempView("persons")

spark.sql(
  """
  |SELECT * from persons"
  """).stripMargin).show()
```

id	name	age
0	Alice	25
1	Bob	23

(2 rows)

## Cypher for Apache Spark

```
val graph = GraphSources.fs("hdfs://...")
  .parquet
  .graph("myGraph")

graph.cypher(
  """
  |MATCH (p:Person)-[:KNOWS]->(p2:Person)
  |RETURN p1, p2
  """).stripMargin).show
```

n	n:Person	n_name	m	m:Person	m_name
0	true	Alice	1	true	Bob

(1 row)

# **CAPS** - Advanced Cypher Features

## **Graph Catalog**

Manages available graph source and graph views

## **Multiple Graph Queries**

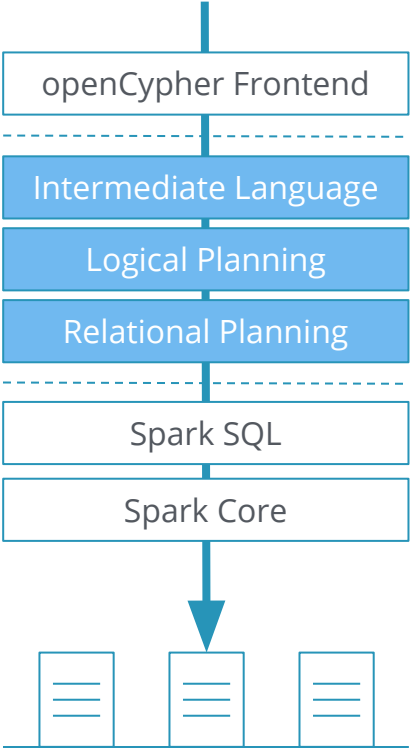
Query and produce multiple graphs in a single Cypher Query

## **Graph Data Source - Neo4j, SQL, Filesystem**

Read and Write graphs to different data sources including Neo4j, RDBMS and Filesystem

# CAPS - Architecture

```
MATCH (n:Person)-[:LOVES]->(s:System)  
WHERE n.name = 'Alice'  
RETURN n.name, s.name
```





**(:Thank)-[:]->(:You)**

[github.com/openCypher/cypher-for-apache-spark](https://github.com/openCypher/cypher-for-apache-spark)

