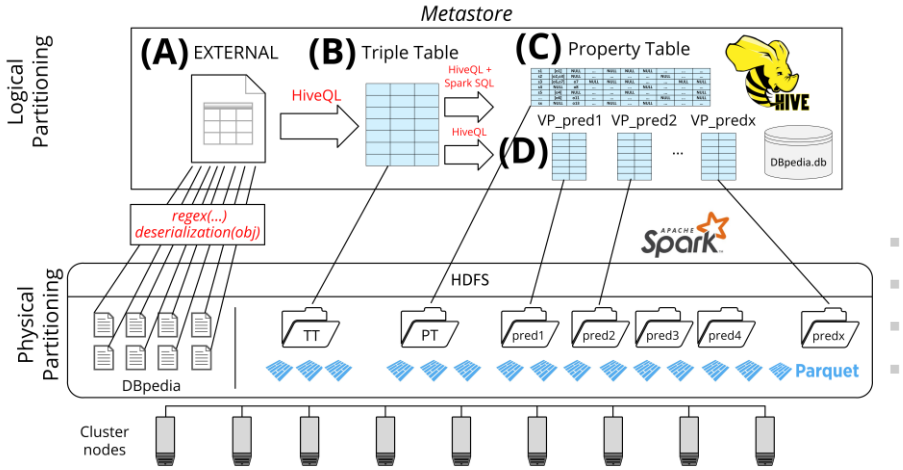
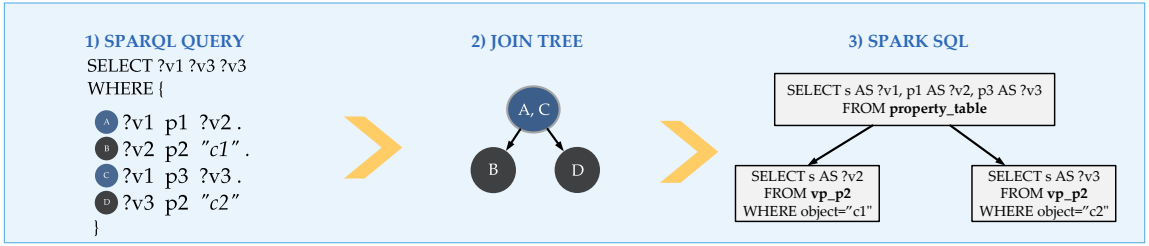
Topic 2: Extending PRoST Features



PRoST

- Based on hadoop-based technologies such as Spark, and Hive
- To evaluate SPARQL queries against large RDF-graphs
- Considers multiple data partitioning strategies
- Query optimizer leverages these different partitioning strategies



Publications (Recommended Literature)

UNI FREIBURG

- Matteo Cossu, Michael Färber, Georg Lausen: PRoST: Distributed Execution of SPARQL Queries Using Mixed Partitioning Strategies. EDBT 2018: 469-472
- Matteo Cossu, Michael Färber, Georg Lausen: PRoST: Distributed Execution of SPARQL Queries Using Mixed Partitioning Strategies. CoRR abs/1802.05898 (2018)
- Victor Anthony Arrascue Ayala, Georg Lausen: A Flexible N-Triples Loader for Hadoop. International Semantic Web Conference (P&D/Industry/BlueSky) 2018

Features (potential topics)

1. Implementing a SPARQL-endpoint
2. Extending query evaluation strategies
3. Comparing PRoST with other systems

Thank you!

UNI FREIBURG

Any questions?