CliqueSquare: efficient RDF distributed query processing

OAK TEAM
OAK objectives and collaborations

Efficient Processing of Complex data

- Semantic Web
- Social Networks
- Semi-structured documents
- Data provenance

CliqueSquare
S4
PAX-QUERY
Nautilus

Collaboration with the University of San Diego
Challenges

Semantic web Data is growing

Centralized solutions are no longer efficient
CliqueSquare efficient RDF query processing

Massive parallelism using:
- **Hadoop** open source **MapReduce** implementation

CliqueSquare features:
- *Minimize network transfer* using custom storage scheme
- *Maximize query performance* using local joins and minimal number of jobs (flat plans)
Reducing query cost

Binary plan

CliqueSquare Flat Plan
Next steps

- Support OLAP Queries
- Additional SPARQL Operators
- Integration of analytical tools
Thank you

OAK TEAM WEBSITE
https://team.inria.fr/oak/

CliqueSquare Project page
https://team.inria.fr/oak/projects/cliquesquare/

Ioana Manolescu (Head of the OAK Team)
ioana.manolescu@inria.fr

Benjamin DJAHANDIDEH (Software engineer CliqueSquare Project) 
benjamin.djahandideh@inria.fr