

Our take on BigData 100

Stephen Brobst, Yogesh Dandawate,
Boris Glavic, Mod Marathe, Tilmann
Rabl, Rekha Singhal

Workload Classification

- 2 dimensions
 - Queries: SQL vs. NoSQL
 - Data: Structured vs. Unstructured

| Data\Processing | SQL | NoSQL |
|-------------------|--------|---------------|
| Relational Data | TPC | Clustering |
| Unstructured Data | Weblog | Path Analysis |

Metrics and Properties

- Throughput
- Responsiveness
- Robustness
- Efficiency
- Scalability
- Fault Tolerance

Concerns

- How to enforce late binding
- Approximate queries
- Verify the results