Just-in-time Data Virtualization
Manos Karpathiotakis, Anastasia Ailamaki

Meaningful analysis relies on data variety
- Diverse Data
- Ad-hoc Queries

Can we execute fast queries over arbitrary data?

Proteus: A shape-shifting query engine

Query Types
- OLAP
- Navigate
- (Un)nest

• Express queries in unifying algebra
• Generate a new engine per query

Query Engine

Proteus vs. State of the art
- TPC-H data; caching off
SELECT COUNT(*) FROM lineitem
WHERE val1 <[X] AND val2<[Y] AND val3<[Z]

Symantec Spam E-mail Analysis

Flexible and fast by specializing & adapting