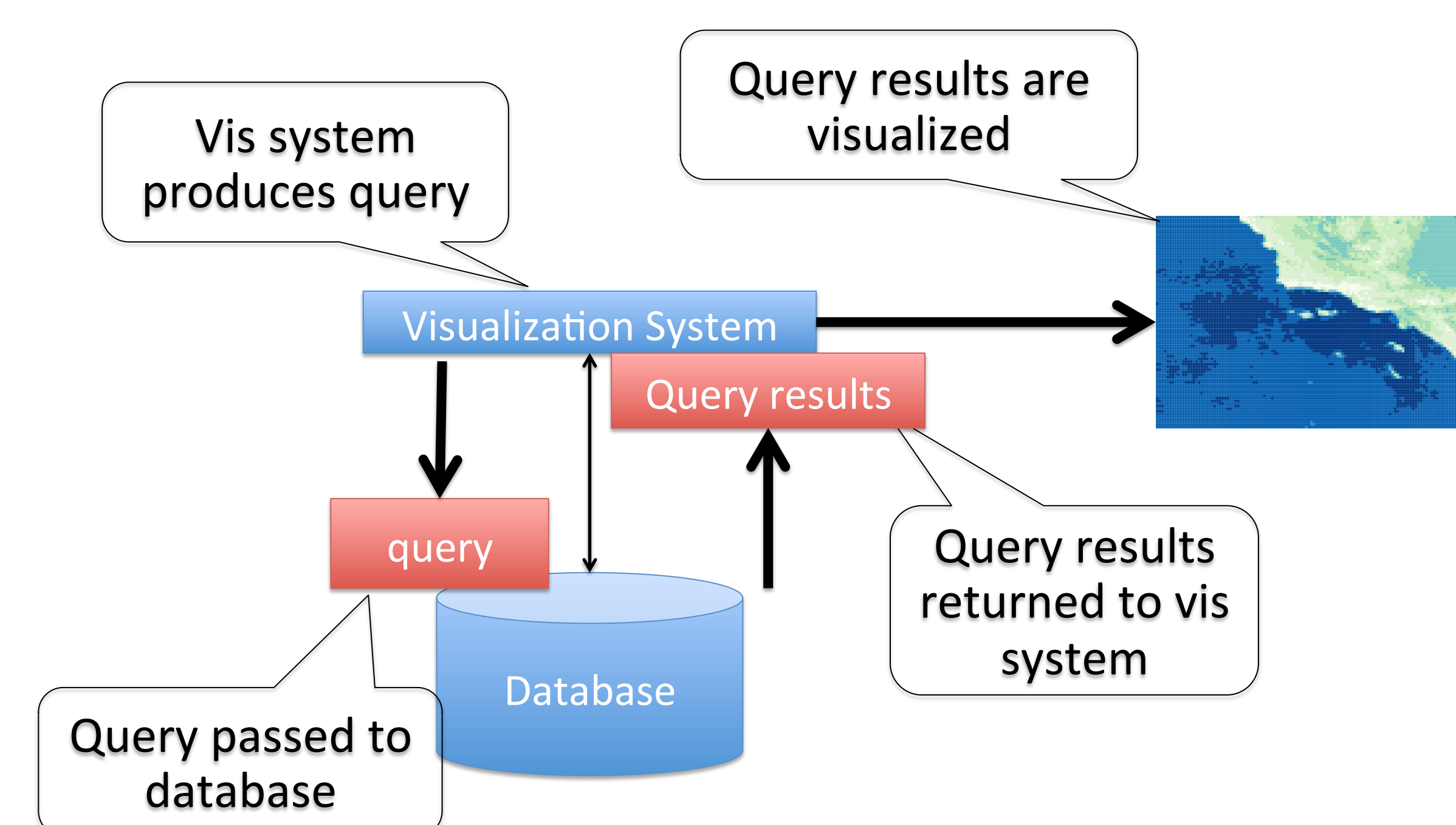


# Dynamic reduction of Query Result Sets for Interactive Visualization

Leilani Battle, Mike Stonebraker, Remco Chang

## Context

- User has a very large data set stored in a database
- User wants a visual overview of the data
- Visualization must be interactive



## Problems Visualizing Big Data

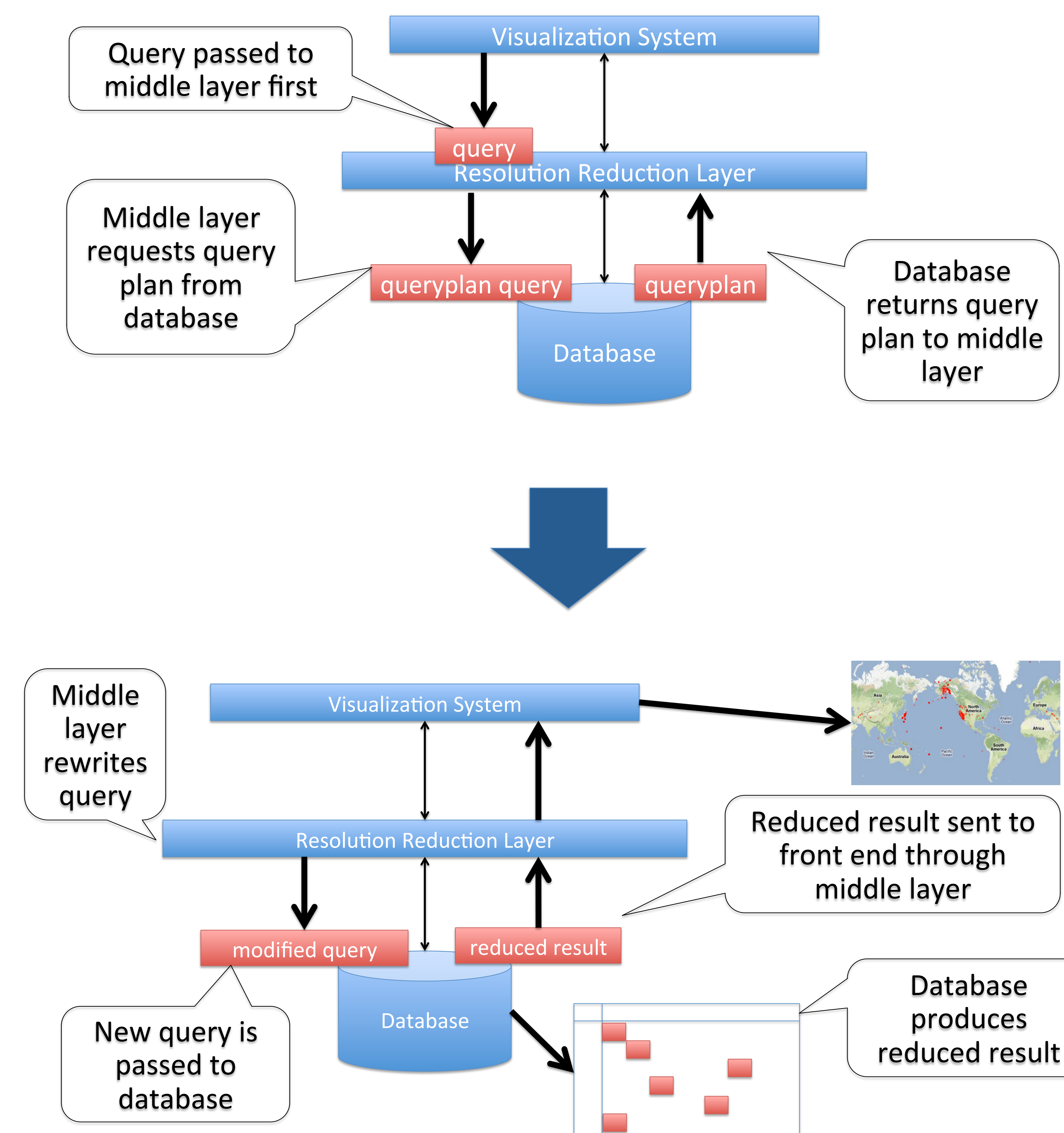
- Performance
  - Vis systems don't scale well for big data
  - Or are turning into databases
- Over-plotting
  - Makes visualizations unreadable
  - Waste of time/resources



## Resolution Reduction

- When a query will return too much data, reduce the results
  - Aggregate, sample, filter, etc.
- Use query plan info to estimate size of results

## Architecture



## ScalaR

- Scalable vis system for data exploration
  - Web front-end
  - Uses SciDB ([www.scidb.org](http://www.scidb.org))
- Visualizes query results
- Performs Resolution Reduction

## Plotting Satellite Imagery Data

- NDVI calculations over NASA satellite imagery data
- 27 GB sparse array in SciDB

