Interactive Data Exploration (IDE)

- Exploratory analysis: ad-hoc & repetitive
  - Questions are not well defined
  - “Interesting” can be complex

- Human-in-the-loop operation
  - Online answering

Constraint Programming Approach

- Decision variables describe the object of search
  - Left-most corner, side lengths

- Constraints describe interesting properties
  - Area 10,000 mi²
  - Average brightness more than 0.8

What about DBMSs?

- No native support for exploratory constructs
  - Inefficient handling of arbitrary constraints
  - No power set traversal

- No support for efficient interactivity
  - Unable to steer the search to the results

OR-Tools + SciDB: constrained search inside data arrays

CP Solver
- Searches in memory
- Works on a synopsis
- Extensive pruning

Validator
- Verifies candidates
- Accesses the disk
- Separate thread

Dynamic Balancing
- Solvers exchange sub-trees
- Validators request chunks
- Validators forward candidates

Preliminary single-instance results

<table>
<thead>
<tr>
<th>Method</th>
<th>Results</th>
<th>First result</th>
</tr>
</thead>
<tbody>
<tr>
<td>SciDB (3h)</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>Searchlight (1 h)</td>
<td>50</td>
<td>18 sec</td>
</tr>
</tbody>
</table>

Time-limited search (32GB, 256Mb memory)

<table>
<thead>
<tr>
<th>Method</th>
<th>Total time</th>
<th>First result</th>
</tr>
</thead>
<tbody>
<tr>
<td>SciDB</td>
<td>17 min</td>
<td>17 min</td>
</tr>
<tr>
<td>Searchlight</td>
<td>12 sec</td>
<td>1 sec</td>
</tr>
</tbody>
</table>

Complete search (32GB, 256Mb memory)