RACHET: Petascale Distributed Data Analysis Suite

6th SOS Workshop on Distributed Supercomputing
Data Intensive Computing
March 4-6, 2002

N. Samatova, A. Geist, G. Ostroumov
Computer Science and Mathematics Division
Oak Ridge National Laboratory

Scientific Data Sets are Massive

Scientific Data Sets are Complex

Ultimate Goal is Knowledge Extraction

A paradigm shift is needed away from dispersed data collecting repositories towards a conceptually integrated knowledge enabling repository.

Our data infrastructure has to support fundamentally new ways of doing science.

High Performance Knowledge Extraction

Scalable techniques for analysis of dynamically changing, terascale datasets.
- Distributed feature extraction
- Dynamic dimension reduction
- Inference capabilities across multiple sources of information
Need to break the Algorithmic Complexity Bottleneck

Algorithmic Complexity:

<table>
<thead>
<tr>
<th>Data size, n</th>
<th>Algorithm Complexity</th>
<th>n</th>
<th>n\log(n)</th>
<th>n^2</th>
</tr>
</thead>
<tbody>
<tr>
<td>100B</td>
<td>O(n)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10KB</td>
<td>O(\sqrt{n})</td>
<td>10^8 sec.</td>
<td>10^6 sec.</td>
<td>10^4 sec.</td>
</tr>
<tr>
<td>1MB</td>
<td>O(n)</td>
<td>10^8 sec.</td>
<td>10^6 sec.</td>
<td>1 sec.</td>
</tr>
<tr>
<td>100MB</td>
<td>O(n^2)</td>
<td>10^4 sec.</td>
<td>10^3 sec.</td>
<td>3 hrs</td>
</tr>
<tr>
<td>10GB</td>
<td>O(n^2)</td>
<td>10^2 sec.</td>
<td>0.1 sec.</td>
<td>3 yrs.</td>
</tr>
</tbody>
</table>

For illustration chart assumes 10^{-12} sec. calculation time per data point

RACHET: High Performance Framework for Distributed Data Analysis

Strategy
Perform data analysis in a distributed fashion with reasonable data transfer overheads

Key idea
- Compute local analyses using distributed agents
- Merge minimum info into a global analysis via peer-to-peer agents’ collaboration & negotiation

Benefits
- NO need to centralize data
- Linear scalability with data size and with data dimensionality

Parallel Data Analysis

Distributed
- Data distribution is driven by a science application
- Software code is sent to the data
- One time communication
- No assumptions on hardware architecture
- Provide an approximate solution

(Rachet approach)

Centralized
- Data distribution is driven by algorithm performance
- Data is partitioned by a software code
- Excessive data transfers
- Hardware architecture-centric
- Aim for the “exact” computation

Background: Cluster Analysis

Problem Description:

- Given:
  - A data set with N k-dimensional data items (might be distributed across multiple data sites)

- Task:
  - Determine a natural partitioning of the data set into a number of clusters and noise

- Application:
  - Construction of Phylogenetic Trees in Biology
  - File Storage Optimization in DB Management
  - Documents Organization in Search Engines
  - and…. 
Background: Hierarchical Clustering

Distance Matrix

<table>
<thead>
<tr>
<th></th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>0.00</td>
<td>0.25</td>
<td>0.75</td>
<td>0.60</td>
<td>0.40</td>
</tr>
<tr>
<td>B</td>
<td>0.25</td>
<td>0.00</td>
<td>0.20</td>
<td>0.45</td>
<td>0.80</td>
</tr>
<tr>
<td>C</td>
<td>0.75</td>
<td>0.20</td>
<td>0.00</td>
<td>0.60</td>
<td>0.70</td>
</tr>
<tr>
<td>D</td>
<td>0.60</td>
<td>0.45</td>
<td>0.60</td>
<td>0.00</td>
<td>0.50</td>
</tr>
<tr>
<td>E</td>
<td>0.40</td>
<td>0.80</td>
<td>0.70</td>
<td>0.50</td>
<td>0.00</td>
</tr>
</tbody>
</table>

75% 40% 60%

Dendrogram

Spanning Tree with Dissimilarity Measures

Many Cluster Cases Considered

RACHET merges local dendograms to determine global cluster structure of the data

Key step is being able to bound the error of the distributed merge

\[
\frac{1}{N_k N_s} \text{MIN}_n \text{SUM}_{\epsilon_i} \leq \frac{1}{N_k N_s} \text{MAX}_n \text{SUM}_{\epsilon_i} \leq \frac{1}{N_k N_s} \text{MAX}_n \text{SUM}_{\epsilon_i} \]

Transmission\text{\_total} = O(k \cdot N)

Time\text{\_total} = O(S^2) + O(S \cdot N)

Space\text{\_total} = O(S^2) + O(k \cdot N)
RACHET Summary

RACHET is designed for Scientific Data that is:
- Massive
- Distributed
- Dynamic and High-dimensional

Highly Scalable Approach
- Compute local analyses
- Merge info with minimum data transfer
- Visualize global results

Science Applications
- Climate
- Genomics
- Astrophysics
- High Energy Physics