Exercise 11
Released: June 25, 2019 · Discussion: July 1, 2019

1 Fragmentation

1. Describe the principle of horizontal and vertical fragmentation. Discuss for which kind of queries which fragmentation technique might be beneficial?

2. The lecture discussed three correctness rules for fragmentations. Name them and discuss why they are beneficial for a distributed DBMS.

3. How can you calculate a horizontal or vertical fragmentation? How can the original relation be reconstructed from the partitions?

4. Given the following relations Projects and Employees:

<table>
<thead>
<tr>
<th>PID</th>
<th>Title</th>
<th>Office</th>
<th>Budget</th>
<th>eid</th>
<th>Name</th>
<th>Proj.</th>
<th>Salary</th>
<th>Department</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Aquarius</td>
<td>London</td>
<td>16000</td>
<td>628</td>
<td>J. Smith</td>
<td>1</td>
<td>58000</td>
<td>Research</td>
</tr>
<tr>
<td>2</td>
<td>Eridanus</td>
<td>Paris</td>
<td>21000</td>
<td>262</td>
<td>D. Miller</td>
<td>4</td>
<td>184000</td>
<td>Research</td>
</tr>
<tr>
<td>3</td>
<td>Centaurus</td>
<td>Paris</td>
<td>17000</td>
<td>381</td>
<td>P. Hanks</td>
<td>1</td>
<td>52000</td>
<td>Marketing</td>
</tr>
<tr>
<td>4</td>
<td>Andromeda</td>
<td>Rome</td>
<td>29000</td>
<td>725</td>
<td>D. Clark</td>
<td>3</td>
<td>55000</td>
<td>Development</td>
</tr>
<tr>
<td>5</td>
<td>Pegasus</td>
<td>London</td>
<td>23000</td>
<td>395</td>
<td>P. Jones</td>
<td>4</td>
<td>143000</td>
<td>Development</td>
</tr>
</tbody>
</table>

5. Divide Projects horizontally on the attribute Office and derive a horizontal fragmentation for Employees.

6. Now divide Employees horizontally on the attribute Department and derive the fragmentation for Projects.

7. Discuss how the correctness criteria are affected by the choice of the fragmentation scheme in the previous two assignments.
2 Allocation

1. What is an allocation? Explain the connection between allocation and fragmentation.
2. For which criteria can an allocation be optimized?
3. Discuss the advantages and disadvantages of replication!
4. Explain the heuristics for calculating an allocation that were presented in the lecture.

3 Distributed Transactions

1. Describe the function of the Two-Phase Commit Protocol.
2. How can serializability be realized in a distributed database?