Replicated Client-Server Execution to Overcome Unpredictability in Mobile Environment

Bing You and Hao Chu

Intelligent Space Lab
National Taiwan University
Outline

- Problem
- Related Work
- Replicated Client-Server Model
- Experiments
- Implementation
- Conclusion
Problem

- Dynamic factors in mobile environment that affect response time
  - Wireless network bandwidth
  - Server loads
  - Usage patterns

- The optimal client thickness (application partition) depends on these dynamic factors.
Related work

- Adaptation methods
  - Chroma from CMU
  - Replets from NTT DoCoMo USA Lab
  - Agilos from UIUC
  - Based on a closed control loop to runtime repartition the application between client and server.
    - Resource monitor
    - Resource prediction
    - Application reconfiguration
Related work (cont')
Related work (cont’)

- Major limitation of adaptation methods
  - Require predictable resources
- What happen if resources are unpredictable?
  - Frequent application reconfigurations
  - Each reconfiguration incurs computing overheads.
  - Incorrect reconfiguration
    - Poor response time
Replicated Client-Server Model

- How to get good response time under unpredictable resources without app reconfiguration?
  - Replicated client-server model
The best of thin and thick client models

Replicated client-server model (cont')
Replicated Client-Server Model vs. Adaptation Methods

- Use replication to solve unpredictability.
- No need to reconfigure under changing resource conditions.
- Give good response time under unpredictable resources.
Experiment Setup

- Using a component-based sample J2EE application to show the impact of changing resource conditions and usage patterns on application response time.
Impact of usage patterns on response time

- two possible usage patterns when a user views his/her shopping cart
  - No sign-on: views his/her shopping cart prior to sign-on.
  - After sign-on: views his/her shopping cart after sign-on.
- Different usage patterns can change the optimal application partition.

Response time

<table>
<thead>
<tr>
<th>Usage Patterns</th>
<th>Thick Client Partition</th>
<th>Thin Client Partition</th>
</tr>
</thead>
<tbody>
<tr>
<td>No sign-on</td>
<td>2.489 s</td>
<td>9.016 s</td>
</tr>
<tr>
<td>After sign-on</td>
<td>45.065 s</td>
<td>11.241 s</td>
</tr>
</tbody>
</table>
Impact of network bandwidth on response time

- Handoff between WLAN and GPRS network
- Varying the network bandwidth can change the optimal application partition.

### Response time

<table>
<thead>
<tr>
<th>Network Bandwidth</th>
<th>Thick Client Partition</th>
<th>Thin Client Partition</th>
</tr>
</thead>
<tbody>
<tr>
<td>GPRS</td>
<td>2.489 s</td>
<td>9.016 s</td>
</tr>
<tr>
<td>WLAN</td>
<td>2.987 s</td>
<td>0.614 s</td>
</tr>
</tbody>
</table>
Implementation and evaluation

- A preliminary implementation on HP iPAQ
  - request dispatcher, computation coordinator, database access interceptor and application-level session state synchronization

- show the system overhead in replicated execution.

### Response time

<table>
<thead>
<tr>
<th>Usage Patterns</th>
<th>Thick Client Partition</th>
<th>Thin Client Partition</th>
<th>Replicated Execution (Overhead)</th>
</tr>
</thead>
<tbody>
<tr>
<td>No sign-on</td>
<td>3.1 s</td>
<td>13.1 s</td>
<td>5.5 s (2.4 s)</td>
</tr>
<tr>
<td>After sign-on</td>
<td>17.3 s</td>
<td>15.0 s</td>
<td>18.6 s (3.6 s)</td>
</tr>
</tbody>
</table>
Conclusion

- Unpredictability in mobile environment
  - The adaptive system may frequently reconfigure.

- Replicated client-server model.
  - Give good response time (at a cost of replicated execution overhead)
  - Work well under resource unpredictability

- Future work
  - Refine our implementation
  - Code download mechanism