Autonomic Cloud Placement of Mixed Workload: An Adaptive Bin Packing Algorithm

Summary:

Chris Seebeck

Cloud Placement Problem (CPP)

- **Virtual Entity (VE)**
  - Virtual Machines
  - Kubernetes Pods
  - Data Volumes

- **Physical Entity (PE)**
  - CPU, Memory, Disk storage, Nodes, Physical Machines

- **How are VEs assigned to PEs?**
Multi-dimensional bin packing

- Bin packing is the decision of how to place resources
- VEs are scheduled to minimize probability of failure to find placement
- The schedule needs to adapt its placement policy due to resource demands
- Challenge is to keep placement decisions short
Pack Placement Policy

- **Packr**
  - Minimum PE that meets demand of LE request picked regardless of other resources

- **Packdist**
  - Cartesian distance max distance between clusters

- **Accommodate a mix of LE requests by making sure some nodes have large amounts of available resources ready**

- **Pack results in skewed loading of nodes in the cluster**
  - Works against a desired load balanced cluster
Spread Placement Policy

- Selects the assignment for the maximum PE
- Balances load across the cluster
- May not be able to accommodate a LE with a large demand
X Balance Policy

- Devised by the authors
- Linear combination of Pack and Spread
- Selectively packs or spreads
- Use a weight vector of resource type
- If the weight vector is > 0 spread
- If the weight vector is < 0 pack
Optimization Problem 2

• Need to find a PE assignment to an arriving LE Request
  - If no solution is found LE is rejected
  
  • Constraint function
    - Feasibility of allocation
      (a) LE with demand (d) on node (n)
  
  • Objective function
    - Evaluation for placing
      LE with demand (d) on
      node (n) given state (a)

\[
\min_n \phi(n; d, a) \\
\min_n \phi(n; d, a), \quad s.t. \psi(n; d, a) \geq 0, \quad (2)
\]
Adaptive Bin Packing (ABP)

- Adapts to mixed workloads
  - Demand for VEs and usage of PEs
- Equalize variability of demand to variability of allocated resources in the physical cluster
- Predicts demand based on
  - Observed Demand
  - Cluster Loading
ABP Continued

- Biased importance sampling to learn placement decisions
  - Avoids large combinatorial problem

- Solves a variation of optimization problem 2 previously defined

\[
\min_n \left( \gamma^{sys}(n) - \gamma^{dem}(t) \right)^2, \quad s.t. \ \psi(n; d, a) \geq 0, \quad (4)
\]

Optimization Problem 2

\[
\min_n \phi(n; d, a), \quad s.t. \ \psi(n; d, a) \geq 0, \quad (2)
\]
ABP Algorithm

1) Update statistics of relative demand (average vector and covariance matrix) by including demand \( d \) of incoming LE request
   a) \( \nu_r \leftarrow (1 - \alpha) \cdot \nu_r + \alpha \cdot \frac{d_r}{c_r}, \quad r = 1, 2, \ldots, R \)
   b) \( \nu_{r_1,r_2} \leftarrow (1 - \alpha) \cdot \nu_{r_1,r_2} + \alpha \cdot \left( \frac{d_{r_1}}{c_{r_1}} \cdot \frac{d_{r_2}}{c_{r_2}} \right), \quad r_1, r_2 = 1, 2, \ldots, R \)
   c) \( \nu_{r_1,r_2} \leftarrow \nu_{r_1,r_2} - \nu_{r_1} \cdot \nu_{r_2}, \quad r_1, r_2 = 1, 2, \ldots, R \)
   d) Compute \( \mu^{dem}(t), \Sigma^{dem}(t), \) and \( \gamma^{dem}(t) \) using equation 3

2) Collect cluster usage data
   a) Get resource utilization \( \mu_{n,r}, \quad n = 1, 2, \ldots, N, r = 1, 2, \ldots, R \)
   b) Compute \( \mu^{sys}(n), \Sigma^{sys}(n), \) and \( \gamma^{sys}(n) \) using equation 3

3) Construct objectiveFunction given by equation 4

4) \( n, \text{feasible} \leftarrow \text{BSAalgorithm}(\text{objectiveFunction}) \)

5) return \( n \), If feasible, and failure otherwise
**ABP Algorithm Graphically**

*Names of steps created for use in this presentation

**Only represents a description of the process*
Experimental Setup

• **ABP** is compared to existing simple **BP policies** and an optimized policy
  - Pack, Spread, XBalance
• **A simulator, placement policies and Biased Sampling Algorithm (BSA) were written in C**
• **Simulation mimics Kubernetes cluster**
Experimental Setup Cont.

- **Three types of resources**
  - CPU, Memory, GPU
  - GPU is a scarce resource and needs to be packed

- **Placement requests are for single pods**
  - Pod is a Kubernetes resource

- **4,000 requests generated**
  - Generated Poisson Arrivals – a distribution type

- **System evaluated after 60 requests**
Experimental Setup Cont.

- A 32 node cluster containing the previous resources was used
- Load is CPU utilization

<table>
<thead>
<tr>
<th>Type</th>
<th>A</th>
<th>B</th>
<th>C</th>
</tr>
</thead>
<tbody>
<tr>
<td>Demand</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CPU</td>
<td>2</td>
<td>8</td>
<td>16</td>
</tr>
<tr>
<td>Memory</td>
<td>24</td>
<td>32</td>
<td>96</td>
</tr>
<tr>
<td>GPU</td>
<td>0</td>
<td>2</td>
<td>4</td>
</tr>
</tbody>
</table>

Table I
Workload mixed resource demand.
Experimental Setup Cont.

• Three phased experiment
• Phase 1
  – Only type A pods workloads
• Phase 2
  – Mixed A and B pods workloads
• Phase 3
  – All types workloads
Average Utilization

- Collected over entire experiment
  - Rejection probability: failure to place from lack of resources
  - Average: load on the system
  - StDev: std deviation of utilization across cluster

<table>
<thead>
<tr>
<th>%</th>
<th>Reject</th>
<th>PACK</th>
<th>SPREAD</th>
<th>XBAL</th>
<th>ABP</th>
</tr>
</thead>
<tbody>
<tr>
<td>Average</td>
<td>CPU</td>
<td>0.39</td>
<td>0.37</td>
<td>0.39</td>
<td>0.39</td>
</tr>
<tr>
<td></td>
<td>Memory</td>
<td>0.36</td>
<td>0.34</td>
<td>0.36</td>
<td>0.36</td>
</tr>
<tr>
<td></td>
<td>GPU</td>
<td>0.50</td>
<td>0.46</td>
<td>0.50</td>
<td>0.50</td>
</tr>
<tr>
<td>StDev</td>
<td>CPU</td>
<td>0.29</td>
<td>0.18</td>
<td>0.14</td>
<td>0.22</td>
</tr>
<tr>
<td></td>
<td>Memory</td>
<td>0.31</td>
<td>0.23</td>
<td>0.12</td>
<td>0.22</td>
</tr>
<tr>
<td></td>
<td>GPU</td>
<td>0.35</td>
<td>0.21</td>
<td>0.35</td>
<td>0.32</td>
</tr>
</tbody>
</table>

Table II: Results summary.
GPU Demand

- GPU demand of the rejected pods
- GPU node capacity is 4
  - Pod with GPU demand of 4 takes whole node

<table>
<thead>
<tr>
<th>GPU demand</th>
<th>PACK</th>
<th>SPREAD</th>
<th>XBAL</th>
<th>ABP</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>0</td>
<td>0</td>
<td>5</td>
<td>2</td>
</tr>
<tr>
<td>3</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>4</td>
<td>18</td>
<td>186</td>
<td>12</td>
<td>11</td>
</tr>
<tr>
<td>Total</td>
<td>18</td>
<td>186</td>
<td>17</td>
<td>13</td>
</tr>
</tbody>
</table>

Table III

Relation between rejections and GPU demand.
Final Results

- Variation of workload
  - Analytic
  - Observed
- Both sets were close in value
### Final Results

#### Table VI
**AVERAGE RESOURCE USAGE $\mu^{\text{sys}}$.**

<table>
<thead>
<tr>
<th></th>
<th>CPU</th>
<th>Memory</th>
<th>GPU</th>
</tr>
</thead>
<tbody>
<tr>
<td>PACK</td>
<td>0.459</td>
<td>0.396</td>
<td>0.703</td>
</tr>
<tr>
<td>SPREAD</td>
<td>0.465</td>
<td>0.404</td>
<td>0.703</td>
</tr>
<tr>
<td>XBAL</td>
<td>0.457</td>
<td>0.393</td>
<td>0.703</td>
</tr>
<tr>
<td>ABP</td>
<td>0.455</td>
<td>0.390</td>
<td>0.703</td>
</tr>
</tbody>
</table>

#### Table VII
**COVARIANCE MATRICES RESOURCE USAGE $\Sigma^{\text{sys}}$.**

<table>
<thead>
<tr>
<th></th>
<th>CPU</th>
<th>Memory</th>
<th>GPU</th>
</tr>
</thead>
<tbody>
<tr>
<td>PACK</td>
<td>0.035</td>
<td>0.030</td>
<td>0.055</td>
</tr>
<tr>
<td>Memory</td>
<td>0.030</td>
<td>0.043</td>
<td>0.013</td>
</tr>
<tr>
<td>GPU</td>
<td>0.055</td>
<td>0.013</td>
<td>0.175</td>
</tr>
<tr>
<td>SPREAD</td>
<td>0.036</td>
<td>0.038</td>
<td>0.051</td>
</tr>
<tr>
<td>Memory</td>
<td>0.038</td>
<td>0.044</td>
<td>0.047</td>
</tr>
<tr>
<td>GPU</td>
<td>0.051</td>
<td>0.047</td>
<td>0.095</td>
</tr>
<tr>
<td>XBAL</td>
<td>0.033</td>
<td>0.029</td>
<td>0.061</td>
</tr>
<tr>
<td>Memory</td>
<td>0.029</td>
<td>0.033</td>
<td>0.039</td>
</tr>
<tr>
<td>GPU</td>
<td>0.061</td>
<td>0.039</td>
<td>0.143</td>
</tr>
<tr>
<td>ABP</td>
<td>0.090</td>
<td>0.087</td>
<td>0.122</td>
</tr>
<tr>
<td>Memory</td>
<td>0.087</td>
<td>0.094</td>
<td>0.106</td>
</tr>
<tr>
<td>GPU</td>
<td>0.122</td>
<td>0.106</td>
<td>0.191</td>
</tr>
</tbody>
</table>

#### Table VIII
**COEFFICIENT OF VARIATION $\gamma$.**

<table>
<thead>
<tr>
<th></th>
<th>Analytic</th>
<th>Observed</th>
<th>PACK</th>
<th>SPREAD</th>
<th>XBAL</th>
<th>ABP</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0.761</td>
<td>0.679</td>
<td>0.456</td>
<td>0.422</td>
<td>0.457</td>
<td>0.636</td>
</tr>
</tbody>
</table>

$\gamma$