Cost-efficient deployment of distributed software services

Máté J. Csorba

csorba@item.ntnu.no
Short introduction & contents

Cost-efficient deployment of distributed software services

Cost functions

Bio-inspired decentralized optimization strategies

Cross-validation

- The deployment problem
  - MAPE, notation, target environment
- Intro to CEAS
- How to encode pheromones
- Scenario 1: Collaborating software components
- Scenario 2: Replication management
- Scenario 3: Cloud computing
- Validation using ILP
- Summary & not covered
The deployment problem and its complexity

Focus on:
- load-balancing,
- remote comm.,
- replica management,
- and financial costs
In the context of the MAPE cycle

- REQs 1 – 6:
  1. Efficiency
  2. Robustness
  3. Autonomicity
  4. Adaptivity
  5. Scalability
  6. Generality & Extensibility

Focus of the Thesis

Deployment mapping

Máté J. Csorba
Target environment
Notation example

- **Constants:**
  - \( N = \{n_1, n_2\} \);
  - \( D = \{d_1\} \);
  - \( S = \{S_1\} \);
  - \( C = \{c_i, c_j\} \);
  - \( K = \{k_k\} \);

- **Variables:**
  - \( D_1 = \{d_1\} \);
  - \( H_1 = \{n_1, n_2\} \);
  - \( L_1 = \{30 \Rightarrow n_1, 20 \Rightarrow n_2\} \);
  - \( M_1 = \{c_i \Rightarrow n_1, c_j \Rightarrow n_2\} \);

\( M \): Deployment mapping
Introduction to CEAS

Cross-Entropy Ant System

- Originally: robust, distributed path management in network nodes
  - Schoonderwoerd et al. – ant based routing scheme (1997)
  - Rubinstein – Cross-Entropy for stochastic optimisation (1999)

- 2 types of ants:
  - Explorer => cover up the search space, i.e. do random search
  - Normal => optimize mapping

- 2 phases in an iteration
  - Forward search => ants looking for a deployment mapping
  - Backtracking => ants update the pheromone database
CEAS deployment strategies – how it works

\[ \gamma_r \leftarrow \text{update}(\text{cost}) \]

\[ \text{selget} : h_{n,n,r} \]

\[ F_{\text{w.d.}} \]

\[ \text{Pheromone update} \]

\[ p_{mn,r} = \sum_{k=1}^{r} I(l \in M_{n,r}) \beta \sum_{x=k+1}^{r} I(x \in M_k) H(F(M_k), \gamma_r) \]
CEAS – auxiliary adjustments

- Parallel nests
  - In joining / splitting networks

- Load-sampling
  - To facilitate load-balancing and gather information

- Guided random hop-selection
  - Taboo-listing
  - Cover domains and then nodes

- Binding & release of instances / maintenance
  - Ease convergence
  - Conditional
CEAS – in a class–diagram

- Implemented in Simula/DEMOS
- Peersim (Java) ongoing
### Pheromone encodings

- Tables distributed over the possible hosts

| Encoding     | DB size in a node | Encoding example w/ $|C_i^k| = 4$ |
|--------------|-------------------|---------------------------------|
| bitstring    | $2^{\sqrt[k]{|C_i^k|}}$ | $[0000]b \ldots [1111]b$        |
| per comp.    | $2 \cdot \sqrt[k]{|C_i^k|}$ | $[0/1]; [0/1]; [0/1]; [0/1]$    |
| # replicas   | $|C_i^k| + 1$       | $[0] \ldots [4]$                |

---

![Diagram showing distribution of tables over possible hosts](image-url)
Scenario 1: Collaborating software components

- Deployment of collaborating components – UML 2.0 collaborations
- PAPER A – C
- Target
  - Load-balancing (i.e. min. deviation from average load)
  - Minimization of remote communication
- E.g.
  - $S_1$ : a simple client-server
  - $S_2$ : authorization and authentication server
Scenario 1 (cont’d)

- Deployment cost function

\[ F_1(M) = \sum_{n=1}^{\left| N \right|} |\hat{I}_n - T| + F_K(M) \]

with global load-balance estimate

\[ T = \frac{\sum_{i=1}^{\left| C \right|} f_c^{(e)}_{c_i}}{\left| N \right|} \]

and communication cost

\[ F_K(M) = \sum_{j=1}^{\left| K \right|} I(q_0(M, k_j, 1) \neq q_0(M, k_j, 2)) \cdot f_{k_j}^{(c)} \]
Scenario 1 (cont’d)

- Example run from PAPER A

2^7 = max DB size

117 = optimal cost

exploration
Scenario 1 (cont’d)

- Deployment costs for multiple services
  - Eliminating global knowledge (~$T$)
  - Multiplicative (~better convergence)

- Deployment cost function

\[
F_2(M, H, L) = \left[ \sum_{\forall n \in H} C_0(n) \right] \cdot (1 + \omega \cdot F_K(M))
\]

where

\[
C_0(n) = \left( \sum_{i=0}^{\hat{i}_n} \frac{1}{\left[ \sum_{\forall n \in H} \hat{i}_n \right] + 1 - i} \right)^2
\]
Scenario 1 (cont’d)

- Examples with 3 services from PAPER C

Node error & repair

New node

Cost

Iterations

www.ntnu.no
Máté J. Csorba
PhD Thesis defense, 12.05.2011
Scenario 2 : Replication management

- Deployment of replicas
- PAPER D – E
- Target
  - Load-balancing
  - Node- & domain-disjointness
- E.g.
  - $S_3$: a simple DB replica
  - $S_4$: 3-way replication

![Diagram showing replication management]

$S_3$: 15
  - $c_{11}$
  - $c_{12}$

$S_4$: 10
  - $c_{21}$
  - $c_{22}$
  - $c_{23}$

Nodes: $n_1$, $n_2$, $n_3$
- $n_1$ connected to $n_2$ by $d_1$
- $n_3$ connected to $n_2$ by $d_2$

#S = 2
#N = 3
Scenario 2 (cont’d)

- Deployment cost function

\[ F_3(|D|, M, H, L) = \frac{1}{|D|} \cdot \sum_{\forall n \in H} C_1(M, -, n, 0) \cdot \sum_{\forall n \in H} C_1(-, L, n, 1) \]

\(1/W\) \quad \ast \quad \text{Domain-disjointness} \quad \text{Node-disjointness} \quad \text{Load-balancing} \quad X^2 \quad \ast \quad Y^2

where

\[ C_1(M, L, n, x) = \left( \sum_{i=0}^{\vartheta(n,x)} \frac{1}{\left[ \sum_{\forall n \in H} \vartheta(n,x) \right] + 1 - i} \right)^2 \]

and

\[ \vartheta(n,x) = (1-x) \cdot |m_n| \cdot w + x \cdot \hat{l}_n \quad \text{for} \quad x \in \{0, 1\} \]
Scenario 2 (cont’d)

- Example from PAPER E
  - 65 replicas, 11 nodes in 5 domains
Scenario 2 (cont’d)

- Example from PAPER E

- Costs of service $S_{10}$
- 2000 explorer iterations
- Splitting of $d_i$ after 4000 iterations
- Merging the regions after 5000 iterations
Scenario 3: Cloud computing

- Deployment of VM replicas in a cloud-like environment
- PAPER F
- Target
  - Load-balancing
  - Node- & domain-disjointness
  - Minimize financial costs

- 5 private (costs 0), 2 public clouds (costs 1 and 10)
- 18 clusters, 130 nodes (cluster sizes of 10 or 5 nodes)
- 5 x 25 services, replication level 5, => 625 VMs
Scenario 3 (cont’d)

- Deployment cost function

\[ F_4(|D|, M, H, L, z) = F_3(|D|, M, H, L) \cdot g(M, z) \]

- where the weighting function is defined as

\[ g(M, z) = \begin{cases} 
  g^{(l)}(M, z) = 1 + z \cdot F_F(M) \\
  g^{(e)}(M, z) = 2 - e^{-(z \cdot F_F(M))^2} 
\end{cases} \]

- and the sum of financial costs is

\[ F_F(M) = \sum_{n_i \in q_1(M, C)} f_{n_i}^{(f)} \]
Scenario 3 (cont’d)

- Combination of cost function terms in $F_4$

$F_3(\theta) \star \frac{1}{W} \star X^2 \star Y^2 \star 1-\exp^{-Z^2}$
Scenario 3 (cont’d)

- Results of 100 simulation runs each setting
  - No cluster weights
  - Linear weighting
  - Exponential weighting
Validation using ILP

- PAPER G
- ILP (AMPL/CPLEX)

| Example | $|C|$ (bound) | $|K|$ | $|N|$ | Optimum cost | Sim. avg. | Sim. stdev. |
|---------|-------------|------|------|-------------|----------|------------|
| Example 1 | 10(3) | 14 | 3 | 117 | 117.21 | 2.1 |
| Example 2 | 15(5) | 21 | 4 | 180 | 193 | 9.39 |
| Example 3 | 15(5) | 28 | 6 | 274 | 277.9 | 5.981 |

![Graph with nodes and edges representing examples and their costs]
Validation using ILP (cont’d)

- ILP built with objective:
  \[
  \min \left( \sum_{j=1}^{|N|} \Delta_j + \sum_{i=1}^{|K|} f_i \cdot \text{col}_i \right) \sim F_1()
  \]

- and corresponding constraints:
  \[
  \sum_{j=1}^{|N|} m_{i,j} = 1; \quad \forall c_i
  \]
  \[
  m_{i,j} \geq b_{i,j}; \quad \forall c_i, \forall n_j
  \]
  \[
  \sum_{i=1}^{|C|} e_i \cdot m_{i,j} - T \leq \Delta_j; \quad \forall n_j
  \]
  \[
  T - \sum_{i=1}^{|C|} e_i \cdot m_{i,j} \leq \Delta_j; \quad \forall n_j
  \]
  \[
  m_{i,j} + m_{k,j} \leq (2 - \text{col}_l); \quad k_l = (c_i, c_k), \forall c_i, c_k, \forall n_j
  \]
  \[
  m_{i,j1} + m_{k,j2} \leq 1 + \text{col}_l; \quad k_l = (c_i, c_k), \forall c_i, c_k, \forall n_{j1}, n_{j2}
  \]

- with the variables:
  \[
  T, \quad b_{i,j}, \quad m_{i,j}, \quad \text{col}_i, \quad \Delta_j
  \]

- Cross-validation of exact solutions vs. heuristics
  - Global-view vs. decentralized

<table>
<thead>
<tr>
<th>Example 1</th>
<th>Example 2</th>
<th>Example 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Simplex iterations</td>
<td>86</td>
<td>495</td>
</tr>
<tr>
<td>Branch and cut nodes</td>
<td>0</td>
<td>5</td>
</tr>
</tbody>
</table>
Summary

- **CEAS**
  - Heuristic optimization
  - Decentralized algorithm, ant-like agents

- **Application domains**
  - Deployment of collaborating sw components
  - Deployment of replicas
  - VM instance placement in private and public clusters

1) Cost functions
2) Heuristic algorithms
3) Modeling scenarios
4) Cross-validation
Publications

- Category I.  – Validation of mappings
- Category II.  – Load-balancing and communication costs
- Category III. – Load-balancing and replication costs
- Category IV. – Cluster costs

PAPER A
Cost Efficient Deployment of Collaborating Components

PAPER B
Adaptable Model-based Component Deployment Guided by Artificial Ants

PAPER C
Component Deployment Using Parallel Ant-nests

PAPER D
Foraging for better deployment of replicated service components

PAPER E
Laying Pheromone Trails for Balanced and Dependable Component Mappings

PAPER F
Ant system for service deployment in private and public clouds

PAPER G
Swarm Intelligence Heuristics for Component Deployment

2008 2009 2010
Not covered, not implemented, not yet given up

- Translation of the simulations to PeerSim
- Additional ILP models
- The power saving dimension
- Migration costs
- Scaling and larger problem sizes
- Incremental scaling
- Coordination and designated nest selection
- Better and broader service models
- Introducing clients
- Costs derived from service models
- Deployment diagrams
- Feedback to functional design
- Experiments using a middleware platform, e.g. DARM
Thank You for your attention!

* and many thanks to Poul E. Heegaard, Peter Herrmann, and Hein Meling