

# 满足可靠性约束的冗余服务池多目标优化策略\*

陈秋红<sup>1</sup>, 景伟娜<sup>1</sup>, 吴长泽<sup>2</sup>

(1. 河南城建学院 计算机科学与工程系, 河南 平顶山 467036; 2. 重庆大学 计算机学院, 重庆 400044)

**摘要:** 针对服务组合中冗余服务池开销较大、服务响应时间较长的问题, 提出一种多目标优化策略, 在满足可靠性约束的前提下, 优化服务池配置、降低冗余服务池开销和缩短服务响应时间。首先分析了在两种不同容错策略下, 满足服务组合特性的服务池多个优化目标; 然后采用基因多目标优化算法进行优化, 全局采用非支配排序策略, 局部采用随机漫步搜索策略。仿真实验通过与现有的研究方法进行对比, 优化效果较好, 表明该方法可用于优化冗余服务池配置。

**关键词:** 可靠性; 冗余服务池; 服务响应时间; 冗余开销

**中图分类号:** TP393      **文献标志码:** A      **文章编号:** 1001-3695(2012)10-3791-03

**doi:**10.3969/j.issn.1001-3695.2012.10.048

## Multi-objective optimization strategy of service pool to satisfy reliability constraint

CHEN Qiu-hong<sup>1</sup>, JING Wei-na<sup>1</sup>, WU Chang-ze<sup>2</sup>

(1. Dept. of Computer Science & Engineering, Henan University of Urban Construction, Pingdingshan Henan 467036, China; 2. School of Computer, Science Chongqing University, Chongqing 400044, China)

**Abstract:** Aimed at the problem that redundant service pool in service composition brings large cost and long response time, this paper presented a multi-objective optimization strategy to reduction the redundant cost and response time under certain reliability. At first, it analyzed the optimization objective of service pool used two kinds of fault tolerant strategies. Then, it used a memetic algorithm based on global nondominated sorting strategy and local random-walk method to solve the optimization problem. At last, simulation shows the strategy is validity and better than the other multi-objective evolutionary algorithms.

**Key words:** reliability; redundant service pool; service response time; redundant cost

在服务计算领域, 服务的松散耦合特性增强了服务的可重用性、灵活性, 但也对保障服务可靠性提出了新的挑战。在现有研究工作中, 保障服务可靠性的主要方案是采用冗余服务池策略<sup>[1,2]</sup>, 即为每个服务绑定一个冗余服务池, 服务池中有若干相同功能的冗余服务, 如果当前服务失效, 则马上采用绑定服务池中的服务进行替换, 以达到提高服务可靠性的目的。这种措施极大地提高了服务可靠性, 但会在服务响应时间及服务开销两方面付出较大的代价。为了有效地缩短服务响应时间和降低服务开销, 合理地配置系统中各个服务的服务池规模显得至关重要。

对于冗余服务池的研究已有较多的研究成果。Liu 等人<sup>[3]</sup>研究了利用绑定冗余服务的方法提高可靠性, 并讨论了绑定不同数量的冗余服务对可靠性提高的影响。文献<sup>[4]</sup>提出了一种根据服务工作流的变化或可靠性需求变更进行动态调整服务池冗余服务的方法; 文献<sup>[5]</sup>评估了采用不同几种基于服务池的容错机制, 并为取得较优的可靠性和服务响应时间, 估算了服务池中的冗余服务数; 文献<sup>[6]</sup>通过探讨在 SOA 系统中存在的“弱点”, 采用优化方法部署冗余资源, 并对冗余资源数进行了量化。现有的这些研究成果均优化了服务池的冗余配置, 但均以可靠性的保障为主要研究目标, 没有探讨在保障可靠性目标的前提下, 如何合理地调整服务

池配置, 以达到降低冗余开销, 缩短服务响应时间, 因而现有方法在保障可靠性目标的前提下, 时间和冗余开销较大。为此, 本文引入多目标决策机制来优化配置服务池, 确定服务池中服务冗余度。

### 1 服务池模型分析

根据文献<sup>[5]</sup>, 在服务池模型中共有两种容错策略, 即顺序策略(sequential)和并行策略(parallel)。对于顺序策略, 如图 1(a)所示, 当故障发生时, 服务池中的冗余服务按备选顺序进行替换; 对于并行策略, 如图 1(b)所示, 服务池中的冗余备选服务并行执行, 以提高服务可靠性。

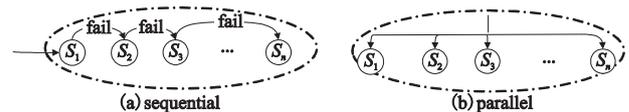


图1 服务池模型容错策略

服务池的可靠性在这两种不同容错策略下是相等的, 用  $R_p$  表示服务池的可靠性,  $n$  表示服务池中冗余服务数,  $R_m$  表示服务池中冗余服务的平均可靠性, 服务池的可靠性为

$$R_p = 1 - \prod_{i=1}^n (1 - R_i) = 1 - (1 - R_m)^n \quad (1)$$

服务池的冗余开销和响应时间, 在两种不同容错策略模式

**收稿日期:** 2012-02-20; **修回日期:** 2012-03-22      **基金项目:** 国家自然科学基金资助项目(90818028); 河南省重点科技攻关项目(102102210020)

**作者简介:** 陈秋红(1979-), 女, 河南平顶山人, 讲师, 硕士研究生, 主要研究方向为网络安全、容错计算等(46979213@qq.com); 景伟娜(1977-), 女, 河南襄县人, 讲师, 硕士, 主要研究方向为信息安全; 吴长泽(1980-), 男, 硕士, 博士, 主要研究方向为服务计算、容错计算。

下是不同的。对于顺序策略,冗余开销和响应时间均较大;而对于并行策略,响应时间不会有太大变化,但冗余开销会很大。用  $T_m$  和  $C_m$  分别表示服务池中冗余服务的平均响应时间和平均开销。服务池的响应时间  $T_p$  和开销  $C_p$  分别为

$$T_p = \left\{ \begin{array}{l} \sum_{i=1}^n iT_i R_i (1 - R_i)^{i-1} = \sum_{i=1}^n iT_m R_m (1 - R_m)^{i-1} \quad \text{sequential} \\ \sum_{i=1}^n N_i T_i / N = T_m \quad \text{parallel} \end{array} \right\}$$

其中:  $N_i$  表示第  $i$  次调用的服务数,  $N$  表示调用服务总数。

$$C_p = \left\{ \begin{array}{l} \sum_{i=1}^n iC_i R_i (1 - R_i)^{i-1} = \sum_{i=1}^n iC_m R_m (1 - R_m)^{i-1} \quad \text{sequential} \\ \sum_{i=1}^n C_i = nC_m \quad \text{parallel} \end{array} \right\}$$

在服务应用过程中,每个服务除了单独提供服务外,还可进行服务组合,以提供功能更多的服务。在服务组合中,每个服务环节均需绑定冗余服务池以提高服务组合的可靠性。由于服务组合具有分支、循环等较复杂的情况,需根据服务组合特性对服务组合中服务池模型的可靠性、响应时间以及开销进行进一步分析。

根据基于架构的可靠性模型和层次化方法<sup>[7]</sup>,用  $V_j$  表示对服务  $j$  访问的平均次数,则可得服务组合的可靠性为

$$R_c = \prod_{j=1}^m R_{pj}^{V_j} = \prod_{j=1}^m (1 - (1 - R_{mj})^{n_j})^{V_j}$$

服务组合中所有服务池总开销  $C_c$  为

$$C_c = \sum_{j=1}^m V_j C_{pj} = \left\{ \begin{array}{l} \sum_{j=1}^m V_j \left( \frac{C_{mj}}{R_{mj}} \right) (1 - (1 - R_{mj})^{n_j} (R_{mj} n_j + 1)) \quad \text{sequential} \\ \sum_{j=1}^m V_j n_j C_{mj} \quad \text{parallel} \end{array} \right\}$$

服务组合的服务响应时间  $T_c$  为所有非并行分支服务响应时间的总和,即  $T_c = \sum \max \{ T_{\text{sub-composition}} \}$ 。服务组合的服务响应时间为

$$T_c = \sum_{j \in Z} V_j T_{pj} = \left\{ \begin{array}{l} \sum_{j \in Z} V_j \left( \frac{T_{mj}}{R_{mj}} \right) (1 - (1 - R_{mj})^{n_j} (R_{mj} n_j + 1)) \quad \text{sequential} \\ \sum_{j \in Z} V_j T_{mj} \quad \text{parallel} \end{array} \right\}$$

服务池多目标的优化是要在满足可靠性约束的基础上,使得开销和响应时间最小,则可将多目标优化归约为下式。

顺序策略下,优化目标为

$$\left\{ \begin{array}{l} \min C_c = \sum_{j=1}^m V_j \left( \frac{C_{mj}}{R_{mj}} \right) (1 - (1 - R_{mj})^{n_j} (R_{mj} n_j + 1)) \\ \min T_c = \sum_{j \in Z} V_j \left( \frac{T_{mj}}{R_{mj}} \right) (1 - (1 - R_{mj})^{n_j} (R_{mj} n_j + 1)) \\ \text{s. t. } R_c = \prod_{j=1}^m (1 - (1 - R_{mj})^{n_j})^{V_j} \geq R_0 \end{array} \right.$$

平行策略下,优化目标为

$$\left\{ \begin{array}{l} \min C_c = \sum_{j=1}^m V_j n_j C_{mj} \\ \min T_c = \sum_{j \in Z} V_j T_{mj} \\ \text{s. t. } R_c = \prod_{j=1}^m (1 - (1 - R_{mj})^{n_j})^{V_j} \geq R_0 \end{array} \right.$$

## 2 多目标基因优化算法

### 1) 多目标基因算法

为了求解第 1 章提出的优化目标式,采用多目标基因算法进行优化。全局搜索过程采用 NSGA II<sup>[8]</sup> 中的非受支配排序

策略,并用随机漫步邻居搜索方法来改善局部最优解。其伪代码如下所示。

```
Set: Qt: the t generation offspring population,
    Pt: the t generation parent population,
    N: individuals, t: generation counter,
    tmax: the maximum generation number,
    Ft: nondominated fronts, I(dt): crowding distance
```

```
Set Q0 = ∅, t = 0.
```

```
Initialize P0 = { p1, p2, ..., pN }.
```

```
While t < tmax
```

```
Use genetic operator to generate Qt ← Pt,
```

```
Rt = Pt ∪ Qt,
```

```
Calculate objective values and constraints values of Rt,
```

```
Calculate Ft and I(dt) of Rt.
```

```
Sort Rt in descending order and select the first individual.
```

```
Qt = Local_search_procedure( local best individual x, t)
```

```
Calculate Ft and I(dt) of Rt.
```

```
Rt = Pt ∪ Qt,
```

```
Calculate Ft and I(dt) of Rt.
```

```
Sort Rt again in descending order and fill Pt+1 with first N individuals.
```

```
t = t + 1.
```

```
Output Pt
```

### 2) 遗传算子

遗传算子采用单点交叉算子与位变异算子相结合的方式。位变异算子设定为一个随机数,其取值范围为 1 到预先设定的最大变异值。

### 3) 局部搜索策略

在得到局部最优解后,对邻居采用随机漫步迭代搜索。整个过程分为三步:a)从局部最优个体中随机选择一个个体,将它的服务池中冗余度加 1 或减 1;b)判断新生产的个体是否为非受支配解;c)如果是非受支配解,则将其归为可接收解,然后重新搜索该新个体的邻居。其伪代码如下所示。

```
Set: x: local best individual, T: iteration times,
```

```
    i: random index, M: predefined max number,
```

```
Local_search_procedure(x, t)
```

```
While t > T
```

```
    + t, xt = x,
```

```
    if (xt[i] > 1) xt[i] = xt[i] - 1, generate neighbors.
```

```
    if xt constraint_dominate x, Q = Q ∪ xt.
```

```
    Local_search_procedure(xt, t).
```

```
    if (xt[i] < M)
```

```
        xt[i] = xt[i] + 1, generate neighbours.
```

```
        if xt constraint_dominate x, Q = Q ∪ xt.
```

```
    Qt = Qt ∪ Q,
```

```
Output Qt
```

## 3 仿真实验

### 1) 实验设置

实验采用文献[9]中的服务组合实例,如图 2 所示。根据文献[9],可得到服务组合的可靠性、开销、响应时间,分别为  $R_c = R_1 R_2 R_3^{3.03} R_4 R_5 R_6^{0.7} R_7^{0.2} R_8^{0.1} R_9 R_{10} R_{11} R_{12}$ ,  $C_c = C_1 + C_2 + 3.03C_3 + C_4 + C_5 + 0.7C_6 + 0.2C_7 + 0.1C_8 + C_9 + C_{10} + C_{11} + C_{12}$  and  $T_c = T_1 + T_2 + \max \{ T_5, 3.03T_3 + T_{11}, T_4 + 0.7T_6 + 0.2T_7$

+ 0.1T<sub>8</sub> + T<sub>9</sub> + T<sub>10</sub> + T<sub>12</sub>。对于单个服务的可靠性、开销和响应时间,进行模拟产生样本数据。模拟数据保障单个服务的可靠性大于 85%,开销小于 100,服务响应时间小于 500 ms。生成的模拟数据样本如表 1 所示。

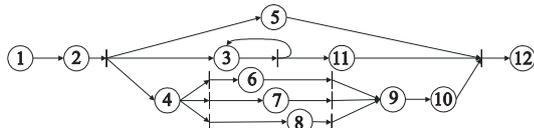


图2 服务组合实例

表 1 服务组合数据样本

	6	S2	S3	S4	S5	S6	S7	S8	S9	S10	S11	S12
R	97.4	85.5	88	92.2	90.7	89.2	87.5	88.3	89.4	95.5	92.5	94.1
C	36.7	60	52.5	10	6.7	70	26.7	13.3	13.3	24.4	5	18
T	45	71.75	117	105.2	108.2	125.2	110.3	125.2	105.4	129	114	96

2) 实验结果

根据以上实验设置,将本文算法与 NSGA II<sup>[8]</sup> 和 PHC-NSGA-II<sup>[10]</sup> 进行对比。NSGA II 是广泛采用的多目标优化算法,PHC-NSGA-II 是一种在 NSGA II 基础上构建的新的多目标优化算法。

采用表 1 所示样本数据,将可靠性约束值从 0.9 ~ 0.99 进行调整,采用三种不同算法每次实验运行 10 次,将所得结果进行对比。其实验结果如图 3 所示。在顺序容错策略下,本文算法接近或优于现有的两种算法。可靠性约束值越大,则在减小开销方面,本文算法更具优势。可靠性约束值越小,则本文算法在缩短响应时间方面更具有优势。当可靠性目标趋近于 95% 时,三种算法的效果差别较小。

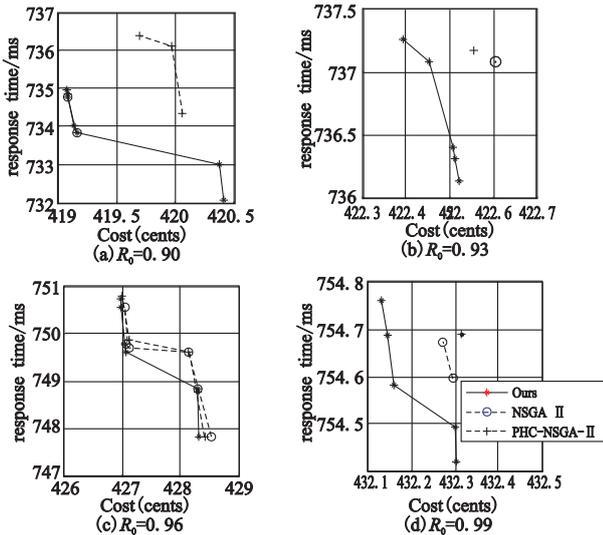


图3 顺序容错策略下对比结果

如图 4 所示,在平行容错模式下,服务响应时间基本无变化。可靠性约束目标值越大,则本文算法与现有两种算法比较,在降低开销方面的优势越大。

根据实验结果,本文提出的多目标优化算法在保障可靠性的前提下,与现有的两种多目标优化算法<sup>[8,10]</sup> 相比,更能有效地进行多目标优化,缩短响应时间,降低冗余开销。

4 结束语

本文在分析两种不同容错模式下的服务池多目标优化模型基础上,采用基因多目标优化算法进行优化,在满足可靠性约束的前提下,合理配置服务组合中绑定的服务池冗余度,以利于有效降低冗余服务的开销和缩短服务响应时间。实验结

果表明,本文算法与现有的 NSGA II<sup>[8]</sup> 和 PHC-NSGA-II<sup>[10]</sup> 相比较,优化效果更好。

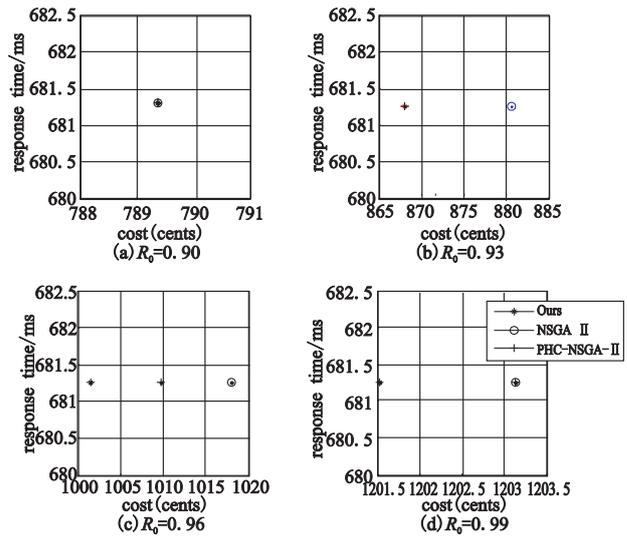


图4 平行容错策略下对比结果

参考文献:

- [1] WANG Li-jun, BAI Xiao-ying, ZHOU Li-zhu, et al. A hierarchical reliability model of service-based software system [C]//Proc of the 33rd IEEE International Computer Software and Applications Conference. Washington DC: IEEE Computer Society, 2009: 199-208.
- [2] HUANG Gang, ZHOU Li, LIU Xuan-zhe, et al. Performance aware service pool in dependable service oriented architecture [J]. Journal of Computer Science and Technology, 2006, 21(4): 565-573.
- [3] LIU L, RUSSELL D, WEBSTER D. Delivering sustainable capability on evolutionary service-oriented architecture [C]//Proc of the 12th IEEE International Symposium on Object Component Service-Oriented Real-Time Distributed Computing. Washington DC: IEEE Computer Society, 2009: 12-19.
- [4] LIU L, WEBSTER D, XU J. Enabling dynamic workflow for disaster monitoring and relief through service-oriented sensor networks [C]// Proc of the 5th International ICST Conference on Communications and Networking. Washington DC: IEEE Computer Society, 2010: 1-7.
- [5] ZHENG Z B, LYU M R. WS-DREAM: a distributed reliability assessment mechanism for Web services [C]//Proc of IEEE International Conference on Dependable Systems and Networks with FTCS and DCC. Washington DC: IEEE Computer Society, 2008: 392-397.
- [6] LUO J, LI Y, PERSHING J A. A methodology for analyzing availability weak points in SOA deployment frameworks [J]. IEEE Trans on Network and Service Management, 2009, 6(1): 31-44.
- [7] TRIVEDI K S. Probability and statistics with reliability, queuing and computer science applications [M]. 2nd ed. New York: John Wiley, 2002.
- [8] DEB K, PRATAP A, AGARWAL S. A fast and elitist multiobjective genetic algorithm: NSGA- II [J]. IEEE Trans on Evolutionary Computation, 2002, 6(2): 182-197.
- [9] XIA Yun-ni, WANG Han-ping, HUANG Yu, et al. A stochastic model for workflow QoS evaluation [J]. Scientific Programming, 2006, 14(3-4): 251-265.
- [10] BECHIKH S, BELGAMI N, SAID L B, et al. PHC-NSGA- II: a novel multi-objective memetic algorithm for continuous optimization [C]//Proc of the 20th IEEE International Conference on Tools with Artificial Intelligence. Washington DC: IEEE Computer Society, 2008: 180-189.