

# 基于业务数据的大型企业 SOA 测试方法

李 鹏<sup>1</sup>, 杨永艳<sup>2</sup>

(1. 中国电子科学研究院系统集成部, 北京 100041; 2. 北京中电普华信息技术有限公司创新研究中心, 北京 100192)

**摘 要:** 对面向服务架构(SOA)测试的困难性和研究现状进行分析, 提出一种基于业务数据的大型企业 SOA 测试方法。在集成环境下, 采用大量历史的真实业务数据作为测试用例, 并编制输入/输出表检测业务流程的正确性。案例分析表明, 该方法在实际项目中具有较强的可操作性, 适用于大型企业的 SOA 测试。

**关键词:** 面向服务架构; 大型企业; 业务数据; 历史数据; 测试用例

## SOA Test Method Based on Business Data in Large Enterprise

LI Peng<sup>1</sup>, YANG Yong-yan<sup>2</sup>

(1. Department of System Integration, China Academy of Electronics and Information Technology, Beijing 100041, China;

2. Innovation Research Center, Beijing China-Power Information Technology Co., Ltd., Beijing 100192, China)

**【Abstract】** This paper analyzes the difficulty and research status of Service Oriented Architecture(SOA) test and proposes an SOA test method based on business data in large enterprise. A large number of real history business data are input into the integrated environment, and a comparison table is used to test the outputs of services and business processes. Case analysis shows that the method is operable in action, and suitable for testing SOA in large enterprise.

**【Key words】** Service Oriented Architecture(SOA); large enterprise; business data; history data; test case

DOI: 10.3969/j.issn.1000-3428.2012.04.020

### 1 概述

面向服务架构(Service Oriented Architecture, SOA)是一种分布式的计算架构模式, 具有天然的松耦合特性, 使服务的可替换性、可复用性成为可能。而且, 通过对服务的编排, 企业可以轻松地开发企业级业务流程, 实现更加复杂的业务功能<sup>[1]</sup>。因此, SOA 被越来越多的大型企业所采用。如何对大型企业中的 SOA 进行测试成为研究热点, 为此, 本文提出一种基于业务数据的大型企业 SOA 测试方法。

### 2 SOA 测试分析

#### 2.1 SOA 测试的困难性

与传统软件系统相比, SOA 的松耦合特性会为测试工作带来一系列的困难, 主要体现在以下 3 个方面:

(1)对于服务的使用方来说, 服务的内部实现机制是不可见的。因此, 服务的演化发展过程也超出了服务使用方可控的范围<sup>[2]</sup>。传统的白盒测试只可能由服务提供方来实施, 而服务的使用方只有在服务运行态下才能获知服务质量。

(2)SOA 的另一个重要特性是服务的动态绑定。动态绑定会导致被测系统在运行时发生变化, 从而导致已有的测试结果失效, 增加了回归测试的工作量。

(3)在通过服务编排实现的业务流程中, 集成后的流程测试是 SOA 测试的最大难题。即便单个服务已经通过了测试, 串联之后的业务流程仍有可能出现问题。另一方面, 流程测试的充分性也很难做到。

#### 2.2 SOA 测试的研究现状

文献[2-3]将 SOA 的服务测试技术的研究分为 3 个阶段:

(1)第 1 阶段: 关注企业内部的测试。主要测试服务的基本功能, 如功能正确性、SOAP 消息和 WSDL 描述等。

(2)第 2 阶段: 关注面向服务的特征, 主要测试服务的发

布、查找和绑定 3 种行为的能力, 以及服务间的异步通信问题和服务质量问题。

(3)第 3 阶段: 关注服务的动态特性和集成的总体测试, 如服务组合测试和版本测试。

从国内外的研究与应用情况来看, 目前关于 SOA 测试的研究内容和研究成果可分为 3 类:

(1)对服务协议、服务状态的测试, 例如文献[2]对 Web 服务测试的理论和进行了综述。

(2)采用自动化生成的测试用例来对 SOA 系统进行测试, 例如, 文献[4]提出了基于测试用例生成器和仿真客户端来测试 SOA 系统, 文献[5]提出了生成基于 XML 语言的测试用例文档的方法。

(3)自动化测试工具的研究, 例如文献[6]提出使用自动化测试引擎对服务进行测试, 文献[7]对 Web 服务注册测试工具进行了研究。

从上述论述可以看出, 目前 SOA 测试的理论研究、应用和测试工具多停留在对服务协议本身或单个服务测试的层面上, 对集成后的业务流程测试则考虑较少。自动化测试工具多采用某种规则生成测试用例, 使用真实数据作为测试用例输入的较少。

### 3 基于业务数据的大型企业 SOA 测试方法

#### 3.1 基本思想

大型企业中的 SOA 系统更应该注重集成测试和业务流程测试。笔者认为, 基于历史真实业务数据的测试方法是在大型企业 SOA 系统中行之有效的测试方法, 理由如下:

**作者简介:** 李 鹏(1982—), 男, 工程师、硕士, 主研方向: 数据库技术, SOA 架构分析; 杨永艳, 工程师

**收稿日期:** 2011-08-02 E-mail: duizhang123@163.com

(1)业务的本质就是数据。大多数业务操作,实际上就是对业务数据的修改,因此,基于历史真实业务数据的测试就成为测试服务功能正确性的直接入口。

(2)业务流程的分支本质上是由数据不同造成的,因此,保证足够的真实数据测试样本,就能最大限度地接近全覆盖测试。

(3)真实数据可能出现的问题要远远多于测试人员主观想象的范围。因此,有必要使用真实数据进行测试。

(4)在大型企业中,需要做纵向的数据一致性(即一定时间区间内的数据总和)校验。使用历史真实的数据进行测试,可以利用历史报表对数据进行比对。

(5)大型企业的信息化建设起步相对较早,一般都已经建设了信息系统,因此,具有一定的业务数据储备量,可以支持基于数据的测试。

### 3.2 案例分析

在基于数据的测试方法中,最重要的工作是准备测试数据和测试用例,为了能较清晰地阐述基于业务数据的测试步骤,本文采用案例分析的方式加以描述。

案例说明:某大型企业 A 选择了 SOA 作为其信息化建设架构,企业系统架构如图 1 所示。

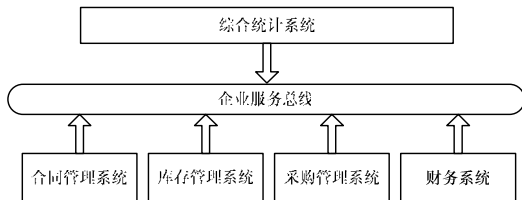


图 1 企业 A 的 SOA 系统架构

综合统计系统是企业的报表处理及数据分析系统,所有关键的业务数据都会准实时地同步到综合统计系统中,以便为企业管理者提供决策支持服务。企业 A 开发了一个大宗订单处理流程,如图 2 所示。

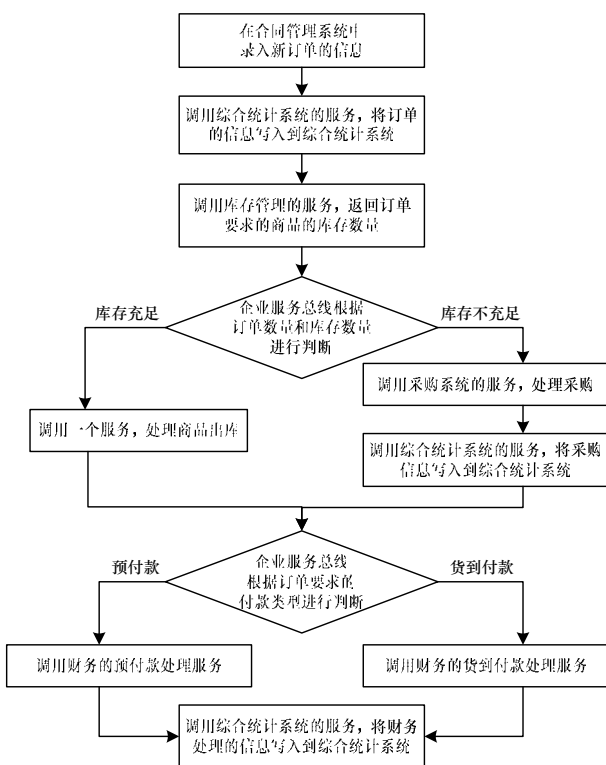


图 2 企业 A 的大宗订单处理流程

该流程是一个进销类企业中常见的业务流程,涉及到合同管理系统、库存管理系统、采购管理系统、财务系统等。该业务流程的起点是合同录入,后续环节根据起点环节输入的数据自动流转,非常适合使用基于历史真实数据的测试方法进行测试。本案例中的一个测试用例如表 1 所示。

表 1 服务流程测试用例

步骤	涉及系统	数据变化
步骤 1	合同管理系统	增加如下数据: 订单号=“D001” 商品=“商品 a” 数量=“30” 金额=“30 000” 日期=“2006-1-30” 付款方式=“预付款”
步骤 2	综合统计系统	增加该订单数据
步骤 3	库存管理系统	返回商品 a 的库存数量: 50
步骤 4	企业服务总线	根据商品 a 的库存数量做判断
步骤 5	库存管理系统	商品 a 的数量减少 30
步骤 6	企业服务总线	根据订单付款方式做判断
步骤 7	财务系统	增加该订单的预付款记录, 且金额=“30 000”
步骤 8	综合统计系统	增加订单的预付款记录, 且金额=“30 000”

操作步骤如下:

(1)明确业务场景。场景是对用户行为的描述,大多数业务可以由若干个场景来描述。基于场景组织测试用例,可以提高测试用例的覆盖程度,也是常规测试中的必要步骤。

在理想情况下,应该根据被测试对象(单个服务或是业务流程)的业务特征枚举出所有的场景。但在复杂的业务流程中,业务场景的数量有时会出现组合爆炸的情况。此时,应按照出现概率对业务场景进行排序,选出主要的业务场景进行优先处理。

(2)准备测试数据。针对上一个步骤分析出的主要业务场景列表,可以从历史真实数据中为每种场景找出数据样本。数据样本从旧系统的数据库中根据条件查询可得。需要注意的是,由于企业中的新旧系统交替,历史数据的数据结构以及质量可能不完全满足新系统输入的需要,需要一些 SQL 语句对数据的结构进行转换,并进行去空格、去乱码等数据清洗工作。样本的选择还应该具有普遍的统计性。如果历史库中有 2006 年、2007 年、2008 年这 3 年的数据,则应该从每一年的数据中选择一些。避免集中在某个时间段中选取样本,以提高缺陷暴露的可能性。

对于不属于主要业务场景的小概率业务场景,可采用时间覆盖的方式,即从历史数据库中选择一段连续时间内的数据作为输入样本(例如可选择连续 3 个月的数据),以提高测试的覆盖度。

另外,在流程集成测试和压力测试的过程中,还需要做纵向数据一致性校验。在本案例中,可以选择某一个历史统计月份的所有订单数据作为输入,将综合统计系统生成的订单月报表和订单财务报表与历史已有报表进行比对。

(3)编制输入/输出表。本步骤的主要工作是根据每组测试数据样本,编制输入/输出对比表,以便能够校验测试用例的输出结果是否正确。在通过服务编排实现的业务流程中,需要对每个服务执行后的数据结果进行检查。

编制输入/输出表的工作可以交给客户,或者在客户监督下完成。具体操作是:针对每组业务输入数据,按照业务逻辑进行转换,得出其应该输出的结果。在实际测试过程中可

以借助 Excel 宏等编程工具,来减少数据计算中的工作量。

(4)编写自动化脚本。自动化脚本主要有3个功能:1)根据输入数据列表,模拟客户端发起服务请求;2)自动获取服务输出以及测试用例中的输入/输出列表,并自动对比测试结果;3)如果发现不一致,自动报告问题。

(5)运行测试用例。上述步骤都准备完成之后,就可以使用自动化测试脚本完成测试了。如果系统的状态发生了改变,使用自动化测试脚本进行回归测试即可。

### 3.3 测试方法评价

基于历史真实业务数据的测试方法具有如下优点:

(1)易操作。在复杂业务环境的测试过程中,由于业务数据的内在关联性,编制模拟测试数据往往具有很高的难度。本测试方法的步骤具有较强的可操作性,适合于在大型企业的 SOA 系统实施项目中应用。

(2)测试覆盖度高。使用模拟数据的测试方法中,由于测试人员对于业务过程的主观认识的局限性,制造的数据往往千篇一律,很难覆盖到所有方面;而使用真实业务数据的测试则使测试过程最大限度地接近系统真实运行环境,测试覆盖度高。

(3)缺陷暴露率高。在使用模拟数据的测试方法中,由于模拟数据本身的业务性差,因此测试人员很难得知服务处理业务数据的过程是否正确。而在使用真实数据的测试方法中,由于输入/输出对比表中的输出都是已有而合理的,因此可以反向暴露出更多的服务处理逻辑缺陷。

但本文测试方法的局限性也较为明显,例如:

(1)需要对历史的真实业务数据进行提取和处理,对测试人员的业务能力要求较高。

(2)没有解决 SOA 测试中缺陷定位的难题。

(3)不适用于需要人工参与的业务流程测试,例如审批类型的业务流程。为了对人机交互类型的服务实现自动测试,还需要额外开发模拟人工处理的功能。

## 4 结束语

本文在分析现有 SOA 测试方法的基础上,提出一种面向业务数据的大型企业 SOA 测试方法,在企业级 SOA 系统开发中具有较高的可操作性。由于本文方法需要历史真实数据的支撑,因此更适用于大型企业的信息系统升级改造过程。下一步将对自动化测试工具、自动化数据比对工具、测试缺陷自动定位等问题进行研究。

### 参考文献

- [1] 刁显峰,支兴超.基于 SOA 的业务流程通信方法[J].计算机工程,2010,36(14):67-69.
- [2] 杨利利,李必信.Web 服务测试问题综述[J].计算机科学,2008,35(19):258-265.
- [3] De B. Web Services: Challenges and Solutions[EB/OL]. [2003-01-01]. http://www.wipro.com.
- [4] 姜国彬.测试 SOA 系统的方法与模型研究[D].上海:复旦大学,2006.
- [5] 李长青.基于 SOA 的软件自动化辅助测试方法研究[D].重庆:西南大学,2008.
- [6] 李长青,张为群.基于 SOA 的异构软件自动化测试方法研究[J].计算机科学,2007,34(12):278-282.
- [7] 吴 蕾.SOA 框架下面向 Web 服务的注错测试工具研究[D].合肥:合肥工业大学,2006.

编辑 顾姣健

(上接第 59 页)

$C_0 = C_0 \cup \{a_3\}$ ,  $R = pos_{C_0}(d) = \{U_4\}$ ,  $U' = U' - R = \{U_1, U_7\}$ , 转 Step3。

$SIG_{C_0}(a_4) = 0$ ,  $SIG_{C_0}(a_5) = 2$ ,  $SIG_{C_0}(a_6) = 0$ ,  $C_0 = C_0 \cup \{a_5\}$ 。

(3)  $R = pos_{C_0}(d) = \{U_1, U_7\}$ ,  $U' = U' - R = \emptyset$ 。

Step4  $U' = U_0$ , 计算  $U'' = U' / C_0 = U'$ 。

Step5  $SIG_{C_0 - \{a_1\}}(a_1) = 3$ ,  $SIG_{C_0 - \{a_2\}}(a_2) = 0$ ,  $C_0 = C_0 - \{a_2\}$ 。

Step6 输出约简为  $\{a_1, a_3, a_5\}$ 。

文献[4]的算法为 Step1~Step3,得到的约简为  $\{a_1, a_2, a_3, a_5\}$ 。与文献[4]算法相比,本文算法由于多了 Step4、Step5,结果少了  $a_2$ ,可见文献[4]的算法存在冗余属性,计算量上增加了 2 次相对支持度的计算。

文献[2-3, 5, 7]的算法由于在原论域上计算,每次计算属性重要度时须比较的元素为 12 个,而本文算法的平均比较元素为 4.25 个。文献[2, 5, 7]还需要计算核,计算量也很大,文献[2-3,7]算法中则含冗余属性。文献[5]虽然也得出了  $C_0 = \{a_1, a_3, a_5\}$  的结果,但其需要计算核,需检验所有的属性是否冗余,比本文的计算要复杂得多。

## 6 结束语

本文提出一种基于论域压缩的启发式属性约简算法,对文献[4]算法中时间复杂度的一个错误进行修正,并加入二次

约简过程以改善其冗余属性。因此,本文算法保持了文献[4]算法快速计算的优点,避免了其约简不完备的缺点。如何将这种约简方法推广到其他类型的粗糙集中是下一步的研究方向。

### 参考文献

- [1] Wong S K M, Zisrko W. On Optional Decision Rules in Decision Tables[J]. Bulletin of Polish Academy of Sciences, 1985, 33(6): 663-676.
- [2] 苗夺谦,胡桂荣.知识约简的一种启发式算法[J].计算机研究与发展,1999,36(6):681-684.
- [3] 徐章艳,刘作鹏,杨炳儒,等.一个复杂度为  $\max(O(|C||U|), O(|C^2||U/C|))$  的快速属性约简算法[J].计算机学报,2006,29(3):391-399.
- [4] 廖毅强,桂现才.信息系统属性约简的快速算法[J].计算机工程与设计,2008,29(18):4804-4806.
- [5] 施化吉,秦 川,陈海军,等.基于粗糙集的启发式属性约简算法[J].计算机工程与设计,2008,29(19):5014-5015.
- [6] 张文修.粗糙集理论与方法[M].北京:科学出版社,2001.
- [7] 高学东,丁 军.一种新的信息系统属性约简算法[J].系统工程理论与实践,2007,27(1):131-136.

编辑 顾姣健