

FRFTA 的性能之所以优于其他方法,从总体上来说,是因为其 3 个主要模块在处理出现的两类故障(临时性故障和永久性故障)时扮演着各自的角色,而且相互之间还起到了互补作用,因此能够具有较好的评估性能.从各个模块的角度来看,更详细的原因如下.

- (1) 在融合网络环境下,可能因为无线信号或者有线线路受到瞬时的干扰,出现瞬时的网络不通的故障,或者是正在调用的服务所在的代理服务器出现了瞬时的系统故障,造成服务暂时不可被调用.此类故障的特点是在很短时间内就能够恢复,对于此类故障宜采取重试调用方法.本文中,服务重试机制采用模糊逻辑的方法得到调整后的重试次数.如果能在调整后的重试次数内成功调用到服务,则继续使用该服务,从而组合服务能够调用最初绑定的原子服务,使其具有较高的组合最优度.这也是在第 2.2 节的各类仿真对比中,FRFTA 方法始终能具有最高的组合最优度的原因.即:服务重试模块起到排除出现的临时性故障的同时,还使组合服务具有较高的组合最优度的作用.如果服务重试失败,证明故障为永久性故障,就需要考虑其他容错机制;
- (2) 如果产生永久性故障的原子服务不具有原子特性,且其存在复制服务时,可以采用服务复制模块提供的服务复制机制进行容错,即:用拥有相同功能属性、不同 QoS 属性的复制服务替代出故障的服务,继续工作流的执行.但这类服务一般有多个,在制定复制方案时需要选择 QoS 属性较好的服务.但不同的 QoS 属性从不同的角度刻画了服务的特性,所以不能只从单个 QoS 属性选择复制服务.本文中的服务复制机制采用多属性决策的方法来综合评价服务,可根据组合服务对可靠性要求的高低选择若干个排在前面的服务作为复制服务.在组合服务的执行阶段,如果事先绑定的原子服务出现永久性故障,就先用贴程度最高的服务替代出故障服务:如果调用成功,则继续工作流的执行;而如果调用不成功,则用贴程度次之的服务替代.依照此规则,直至成功调用到复制服务,即,用服务复制机制排除出现的永久性故障;
- (3) 如果产生的永久性故障的原子服务具有原子特性,或其不存在复制服务,则可以采用服务补偿模块提供的服务补偿机制进行容错,即,回溯到一个可以进行补偿的状态,从此状态开始,为具有原子特性的组合服务片段重新选择原子服务,然后继续执行工作流.在本文中,服务补偿机制通过改进的粒子群算法,在使重新寻找原子服务的时间降低的同时,也使服务组合片段的组合最优度值升高,从而增加整个组合服务的组合最优度.所以,通过服务补偿模块,在排除了需要补偿的永久性故障的同时,也能减少故障排除时间和使组合服务具有较高的组合最优度.

综上所述,FRFTA 独立于组合服务执行逻辑,只在出故障的时候才被触发.且其 3 个模块相互配合,在拥有较高的故障排除率的同时,也能缩短故障排除时间和拥有较高的组合最优度,所以其才能在实验对比中的结果优于其他 4 种方法.

3 相关工作

对于服务组合的容错方法,许多学者提出了相关方法或解决思路,并取得了较好的研究成果.下面将对一些与本文密切相关的研究成果予以分析.

文献[5,11]研究了带局部和全局约束的可靠面向服务系统的最佳容错策略问题,提出了一个系统的和可扩展的框架,通过此框架,可以在组合服务的设计阶段选择最佳容错策略.文献把所研究问题建模为一个优化问题,同时把用户的要求转化为局部和全局约束,并把语义相关任务(task)的最佳容错策略的选择问题也形式化为优化问题的一个约束,并设计出一种启发式算法(FT-HEU)有效地求解优化问题.文献在服务组合的设计阶段,在对各个任务选择原子服务时,同时兼顾选择最优容错策略.但一些实验研究显示,大部分的故障出现在组合服务的执行阶段,所以应重点研究组合服务执行阶段的容错方法.

文献[16]引入一种外部 Web 服务,被称为故障避免服务(fault avoidance service,简称 FAS),其可以作为原子服务被直接包含进组合服务中.在组合服务的执行过程中,FAS 周期性地探测组合服务绑定的未被执行的原子服务的可用性.如果发现某个原子服务不可用,则在组合服务调用该服务前,FAS 用实现相同功能不同 QoS 的可

用原子服务替代不可用原子服务.该方法虽然避免了组合服务而去调用一个不可用的原子服务,但其不区分是临时性故障还是永久性故障.即使原子服务出现的是临时性故障,也要更新原子服务.如果在发现一个服务不可用且是临时性故障时,那么只要在组合服务真正调用该服务前能够从故障中恢复,就没有必要将其更新为其他服务,从而可以减少不必要的更新开销,且能使组合服务调用到最佳原子服务.

文献[12,14,36]采用复制策略(replication strategy),也称为冗余服务策略,来保证组合服务的可靠执行.服务复制是实现高可靠服务的一种有效方法.通过同时使用多个拥有相同功能属性的服务来实现同一个任务,在一个服务出现调用失败时,其他服务可以继续提供所需功能.文献[14]认为:在一个业务流程中,协调服务(coordination service)的不可用会直接影响所有合作伙伴(partners)的可用性,所以应该为协调服务提供更高的可用性.该文献在广域网环境下,采用服务复制机制,设计和实现了一种基础设施.该基础设施无缝地提高了协调服务的可用性.但该文献仅为协调服务提供高可用性,而为了保证融合网络环境下的组合服务的可靠执行,必须提高所有原子服务的可用性.文献[12]提出和实现了一种分布式复制策略评价和选择框架,基于此框架,可以对各种复制方案通过理论公式和实际实验进行系统的比较,方便用户找到最优复制方案.但此框架在参与构建复制策略的原子服务较少时选择最优复制方案计算量较少,但随着参与原子服务的数量的增加,计算量会呈指数级增长,会在很大程度上影响选择速度.文献[36]在文献[12]的基础上采用有向无环图(directed acyclic graph,简称 DAG)建模复制方案,通过 DAG 能够探索到所有可能的复制方案,然后采用改进遗传算法(genetic algorithm)寻找最优解.在参与复制策略的原子服务较多的情况下,执行速度也较快.但上述 3 个文献都是基于出故障的服务存在替代服务的假设,然而在一些情况下也许不存在冗余服务,如果这样,就不能采用复制策略进行容错.

文献[13]认为:虽然服务组合语言,即业务流程执行语言 BPEL 为从错误(error)处进行回滚(rollback)操作提供了补偿机制(compensation mechanism),但此类补偿存在很多问题,比如其在补偿后不能保证组合服务的功能属性.在文献[13]中提出了一种基于遗传算法的自动化算法,该方法通过计算可以得到在补偿后能够保证组合服务功能属性的恢复计划(recovery plan).给定一个拥有很大状态空间的组合服务,提出的方法不需要在组合服务的全状态空间中探索最优恢复计划,而是在部分状态空间寻找近似最优恢复计划,因此,该方法能够有效地选择恢复计划.此外,恢复计划是根据服务的 QoS 选择得到的,一个最优 QoS 的恢复计划能够保证组合服务有效地从故障的状态恢复过来.文献[15]提出了 EXTRA 机制(exception handing+transaction)来提高组合服务的可靠性,该机制是结合了异常处理(exception handing)和事务技术(transaction techniques)的混合容错机制.EXTRA 采用 8 类高层异常处理机制来修复发生在组合服务执行过程中的故障,同时也定义了一个基于服务状态转移的终止协议(service-transfer-based termination protocol,简称 STTP),如果产生了一个不可修复的故障,STTP 可以使组合服务终止在一个一致的状态.但上述两个文献提出的方法时间复杂度均偏高,如果将其应用到组合服务的实际容错中,会使组合服务的总执行时间变长,降低了用户体验.与上述研究不同,在本文中,组合服务执行逻辑与容错逻辑彼此独立.只有在执行逻辑中监测到故障的情况下,才会触发容错逻辑排除出现的故障.

4 结论与未来工作展望

本文提出了一种融合网络环境下快速可靠的服务组合容错方法.该方法在组合服务的执行阶段不断收集组合服务的执行状态:如果没有异常,容错逻辑就不阻碍组合服务的正常执行;而当发现某一原子服务出现故障时,则根据故障的类型,采用不同的机制进行容错.

- 如果为临时性故障,就采用服务重试机制进行容错,即,通过重复调用的方式探测故障原子服务的可用性;
- 如果出现故障的原子服务在重试持续时间内还是不能被成功地调用,本文就认为该原子服务出现了永久性故障,如果其不具有原子性,就采用服务复制机制进行容错;如果出故障的原子服务具有原子性,就采用服务补偿机制进行容错.

本文采用改进的 PSO 算法,快速地为组合服务片段寻找到最优原子服务,保证组合服务的继续正确执行.大

量的实验结果验证了本文所提方法的有效性。

虽然通过实验对比,我们的方法比其他方法更具有一定的优势,但仍有改进空间,例如:对于服务重试机制,我们假定采用了固定重试频率,只对重试次数进行了研究.而实际应用中,可能会根据组合服务特点选择不同的重试频率,所以需要进一步研究不同重试频率下对重试次数的控制.另外,如何在实际的融合网络环境中验证并部署 FRFTA,也是我们未来的研究工作。

致谢 在此,我们对提供 QWS 数据集的加拿大圭尔夫大学的 Eyhab Al-Masri 和 Qusay H.Mahmoud 博士以及提供 WS-DREAM 的香港中文大学的郑子彬博士和 Michael R. Lyu 教授表示感谢。

References:

- [1] Zheng ZB, Zhang Y, Lyu MR. Distributed QoS evaluation for real-world Web services. In: Proc. of the 17th Int'l Conf. on Web Services (ICWS 2010). Miami: IEEE, 2010. 83–90. [doi: 10.1109/ICWS.2010.10]
- [2] Hamadi R, Benatallah B. A Petri net-based model for Web service composition. In: Proc. of the 14th Australasian Database Conf. (ADC 2003), Vol.17. Darlinghurst: Australian Computer Society, 2003. 191–200. <http://dl.acm.org/citation.cfm?id=820121>
- [3] Bultan T, Fu X, Hull R. Conversation specification: A new approach to design and analysis of e-service composition. In: Proc. of the 12th Int'l Conf. on World Wide Web (WWW 2003). New York: ACM Press, 2003. 403–410. [doi: 10.1145/775152.775210]
- [4] Lucchi R, Mazzara M. A pi-calculus based semantics for WS-BPEL. The Journal of Logic and Algebraic Programming, 2007,70(1): 96–118. [doi: 10.1016/j.jlap.2006.05.007]
- [5] Zheng ZB, Lyu MR. Selecting an optimal fault tolerance strategy for reliable service-oriented systems with local and global constraints. IEEE Trans. on Computers, 2015,64(1):219–232. [doi: 10.1109/TC.2013.189]
- [6] Yu WD. A software fault prevention approach in coding and root cause analysis. Bell Labs Technical Journal, 1998,3(2):3–21. [doi: 10.1002/bltj.2101]
- [7] Littlewood B. Stochastic reliability-growth: A model for fault-removal in computer-programs and hardware-designs. IEEE Trans. on Reliability, 1981,30(4):313–320. [doi: 10.1109/TR.1981.5221099]
- [8] Randell B. System structure for software fault tolerance. In: Proc. of the Int'l Conf. on Reliable Software. New York: ACM Press, 1975. 437–449. [doi: 10.1145/800027.808467]
- [9] Vergara S, Acciani G, Amoroso V. Inferential statistics for monitoring and fault forecasting of PV plants. In: Proc. of the IEEE Int'l Symp. on Industrial Electronics (ISIE 2008). Cambridge: IEEE, 2008. 2414–2419. [doi: 10.1109/ISIE.2008.4677264]
- [10] Mansour HE, Dillon T. Dependability and rollback recovery for composite Web services. IEEE Trans. on Services Computing, 2011,4(4):328–339. [doi: 10.1109/TSC.2010.16]
- [11] Zheng ZB, Lyu MR. A QoS-aware fault tolerant middleware for dependable service composition. In: Proc. of the 39th Int'l Conf. on the Dependable Systems & Networks (DSN 2009). Lisbon: IEEE, 2009. 239–248. [doi: 10.1109/DSN.2009.5270332]
- [12] Zheng ZB, Lyu MR. A distributed replication strategy evaluation and selection framework for fault tolerant Web services. In: Proc. of the 15th Int'l Conf. on Web Services (ICWS 2008). Beijing: IEEE, 2008. 145–152. [doi: 10.1109/ICWS.2008.42]
- [13] Tan TH, Chen M, André É, Sun J, Liu Y, Dong JS. Automated runtime recovery for QoS-based service composition. In: Proc. of the 23rd Int'l Conf. on World Wide Web (WWW 2014). New York: ACM Press, 2014. 563–574. [doi: 10.1145/2566486.2568048]
- [14] Salas J, Perez-Sorrosal F, Patiño-Martínez M, Jiménez-Peris R. WS-Replication: A framework for highly available Web services. In: Proc. of the 15th Int'l Conf. on World Wide Web (WWW 2006). New York: ACM Press, 2006. 357–366. [doi: 10.1145/1135777.1135831]
- [15] Liu A, Li Q, Huang L, Xiao M. Facts: A framework for fault-tolerant composition of transactional Web services. IEEE Trans. on Services Computing, 2010,3(1):46–59. [doi: 10.1109/TSC.2009.28]
- [16] Gülcü K, Sözer H, Aktemur B, Ercan AÖ. Fault masking as a service. Software: Practice and Experience, 2014,44(7):835–854. [doi: 10.1002/spe.2255]
- [17] Fan XQ, Jiang CJ, Wang JL, Pang SC. Random-QoS-Aware reliable Web service composition. Ruan Jian Xue Bao/Journal of Software, 2009,20(3):546–556 (in Chinese with English abstract). <http://www.jos.org.cn/1000-9825/3339.htm> [doi: 10.3724/SP.J.1001.2009.03339]

- [18] Wu GQ, Wei J, Huang T. A dynamic QoS assessment approach for internetware with uncertainty reasoning. *Ruan Jian Xue Bao/ Journal of Software*, 2008,19(5):1173–1185 (in Chinese with English abstract). <http://www.jos.org.cn/1000-9825/20080509.htm> [doi: 10.3724/SP.J.1001.2008.01173]
- [19] Liu XZ, Huang G, Mei H. Consumer-Centric service aggregation: Method and its supporting framework. *Ruan Jian Xue Bao/ Journal of Software*, 2007,18(8):1883–1895 (in Chinese with English abstract). <http://www.jos.org.cn/1000-9825/20070804.htm> [doi: 10.1360/jos181883]
- [20] Wang SG, Sun QB, Zou H, Yang FC. Particle swarm optimization with skyline operator for fast cloud-based Web service composition. *Mobile Networks and Applications*, 2013,18(1):116–121. [doi: 10.1007/s11036-012-0373-3]
- [21] Wang SG, Zheng ZB, Sun QB, Zou H, Yang FC. Cloud model for service selection. In: *Proc. of the Conf. on Computer Communications Workshops (INFOCOM WORKSHOPS 2011)*. Shanghai: IEEE, 2011. 666–671. [doi: 10.1109/INFCOMW.2011.5928896]
- [22] Ma Y, Wang SG, Sun QB, Yang FC. Web service quality metric algorithm employing objective and subjective weight. *Ruan Jian Xue Bao/ Journal of Software*, 2014,25(11):2473–2485 (in Chinese with English abstract). <http://www.jos.org.cn/1000-9825/4508.htm> [doi: 10.13328/j.cnki.jos.004508]
- [23] Tao F, Zhao D, Hu Y, Zhou Z. Resource service composition and its optimal-selection based on particle swarm optimization in manufacturing grid system. *IEEE Trans. on Industrial Informatics*, 2008,4(4):315–327. [doi: 10.1109/TII.2008.2009533]
- [24] Wang LX. *Adaptive Fuzzy Systems and Control: Design and Stability Analysis*. Upper Saddle River: Prentice-Hall, 1994.
- [25] Lee CC. Fuzzy logic in control systems: Fuzzy logic controller. Part II. *IEEE Trans. on Systems, Man and Cybernetics*, 1990, 20(2):419–435. [doi: 10.1109/21.52552]
- [26] Yu H, Liu Z, Li YJ. Key nodes in complex networks identified by multi-attribute decision-making method. *Acta Physica Sinica*, 2013,62(2):20204–020204 (in Chinese with English abstract). [doi: 10.7498/aps.62.020204]
- [27] Saaty TL. Decision making with the analytic hierarchy process. *Int'l Journal of Services Sciences*, 2008,1(1):83–98. [doi: 10.1504/IJSSci.2008.01759]
- [28] Wang SG, Sun QB, Yang FC. Web service dynamic selection by the decomposition of global QoS constraints. *Ruan Jian Xue Bao/ Journal of Software*, 2011,22(7):1426–1439 (in Chinese with English abstract). <http://www.jos.org.cn/1000-9825/3842.htm> [doi: 10.3724/SP.J.1001.2011.03842]
- [29] Alrifai M, Risse T. Combining global optimization with local selection for efficient QoS-aware service composition. In: *Proc. of the 18th Int'l Conf. on World Wide Web (WWW 2009)*. New York: ACM Press, 2009. 881–890. [doi: 10.1145/1526709.1526828]
- [30] Zeng L, Benatallah B, Ngu AH, Dumas M, Kalagnanam J, Chang H. QoS-Aware middleware for Web services composition. *IEEE Trans. on Software Engineering*, 2004,30(5):311–327. [doi: 10.1109/TSE.2004.11]
- [31] Zeng L, Benatallah B, Dumas M, Kalagnanam J, Sheng QZ. Quality driven Web services composition. In: *Proc. of the 12th Int'l Conf. on World Wide Web (WWW 2003)*. New York: ACM Press, 2003. 411–421. [doi: 10.1145/775152.775211]
- [32] Shi Y, Eberhart R. A modified particle swarm optimizer. In: *Proc. of the Int'l Conf. on Evolutionary Computation (CEC'98)*. Anchorage: IEEE, 1998. 69–73. [doi: 10.1109/ICEC.1998.699146]
- [33] Van Den Bergh. An analysis of particle swarm optimizers [Ph.D. Thesis]. Pretoria: University of Pretoria, 2006.
- [34] Al-Masri E, Mahmoud QH. Investigating Web services on the World Wide Web. In: *Proc. of the 17th Int'l Conf. on World Wide Web (WWW 2008)*. New York: ACM Press, 2008. 795–804. [doi: 10.1145/1367497.1367605]
- [35] Zhang YL, Zheng ZB, Lyu MR. Exploring latent features for memory-based QoS prediction in cloud computing. In: *Proc. of the 30th IEEE Symp. on Reliable Distributed Systems (SRDS 2011)*. Madrid: IEEE, 2011. 1–10. [doi: 10.1109/SRDS.2011.10]
- [36] Liu A, Li Q, Huang L. Quality driven Web services replication using directed acyclic graph coding. In: *Proc. of the 12th Int'l Conf. on Web Information System Engineering (WISE 2011)*. Berlin: Springer-Verlag, 2011. 322–329. [doi: 10.1007/978-3-642-24434-6_28]

附中文参考文献:

- [17] 范小芹,蒋昌俊,王俊丽,庞善臣.随机 QoS 感知的可靠 Web 服务组合. *软件学报*,2009,20(3):546–556. <http://www.jos.org.cn/1000-9825/3339.htm> [doi: 10.3724/SP.J.1001.2009.03339]
- [18] 吴国全,魏峻,黄涛.基于非确定性推理的网构软件服务质量动态评估方法. *软件学报*,2008,19(5):1173–1185. <http://www.jos.org.cn/1000-9825/20080509.htm> [doi: 10.3724/SP.J.1001.2008.01173]
- [19] 刘让哲,黄罡,梅宏.用户驱动的服务聚合方法及其支撑框架. *软件学报*,2007,18(8):1883–1895. <http://www.jos.org.cn/1000-9825/20070804.htm> [doi: 10.1360/jos181883]

- [22] 马友,王尚广,孙其博,杨放春.一种综合考虑主客观权重的 Web 服务 QoS 度量算法.软件学报,2014,25(11):2473-2485. <http://www.jos.org.cn/1000-9825/4508.htm> [doi: 10.13328/j.cnki.jos.004508]
- [26] 于会,刘尊,李勇军.基于多属性决策的复杂网络节点重要性综合评价方法.物理学报,2013,62(2):20204-020204. [doi: 10.7498/aps.62.020204]
- [28] 王尚广,孙其博,杨放春.基于全局 QoS 约束分解的 Web 服务动态选择.软件学报,2011,22(7):1426-1439. <http://www.jos.org.cn/1000-9825/3842.htm> [doi: 10.3724/SP.J.1001.2011.03842]



张俊娜(1979-),女,河南周口人,博士生,副教授,主要研究领域为服务计算.



孙其博(1975-),男,博士,副教授,CCF 专业会员,主要研究领域为网络服务与网络智能化,物联网应用技术.



王尚广(1982-),男,博士,副教授,博士生导师,CCF 高级会员,主要研究领域为服务计算,移动云计算,车联网,网络安全.



杨放春(1957-),男,博士,教授,博士生导师,CCF 杰出会员,主要研究领域为融合通信软件,网络安全,网络智能化.