第 34 卷 第 11 期 计 算 机 工 程 2008 年 6 月
Vol.34 No.11 Computer Engineering June 2008
·开发研究与设计技术· 文章编号：1000—3428(2008)11—0252—03 文献标识码：A 中图分类号：N945

基于 *M+N* 结构的 SCP 容灾系统设计

王福阳[1]，李小坚[1]，俞 前[2]
(1. 湖南大学计算机与通信学院，湖南 410082；2. 华为技术有限公司，深圳 518129)

摘 要：将业务接入与业务控制分离，设计并实现基于 *M+N* 结构的业务控制点(SCP)容灾系统。该系统可扩展性和可维护性好、可用性高，且对现网设备和业务影响小。测试及实际应用结果表明，其实现了快速的应用级业务恢复，保证了业务的连续性。可以容忍多个生产 SCP 同时发生故障并提供业务接管。避免数据备份服务的集中运行，增强了容灾 SCP 的业务处理能力。

关键词：移动智能网；业务控制点；容灾；*M+N* 结构；可扩展性；高可用性

Design of SCP Disaster Tolerant System Based on *M+N* Architecture

WANG Fu-yang[1], LI Xiao-jian[1], YU Qian[2]
(1. School of Computer and Communication, Hunan University, Changsha 410082；
2. Huawei Technologies Co.,Ltd., Shenzhen 518129)

【Abstract】 A application layer disaster tolerant system of Service Control Point(SCP) based on "*M+N*" architecture is designed and implemented by separating service access function from service control function. It has good expansibility, high availability, fine maintainability, and small impact on business equipment. The system test and practical application show: the rapid recovery of application level ensures business continuity; multiple points of failure tolerance improves system availability; distribution of data replication enhances the handling capacity of SCP on single equipment.

【Key words】 wireless intelligent network; service control point; disaster tolerant; *M+N* architecture; expansibility; high availability

1 概述

业务控制点(Service Control Point, SCP)是移动智能网系统中集中控制和管理业务的设备系统[1]，简单的本地集群保护无法满足该系统的高可用性要求，需要全面的容灾保护[2]来保障业务运行的连续性和高可靠性。

容灾系统的拓扑结构可分为循环备份、1+1 备份和*N*+1 备份、*M+N*备份[3]。相比之下，*M+N*备份结构独立构建容灾系统，对现网设备和业务影响小、管理维护集中方便，并具有以下优点：(1)可扩展性好，*M+N*结构能灵活配置灾备节点个数(*N*≥1)，以适应容灾系统可用性要求和为系统扩容的需要；(2)可用性高，最大可为*N*台故障生产设备提供业务接管，容忍多点故障；(3)避免负载倾斜，可以选择在负载较轻的备份节点进行业务接管，提高系统的吞吐量；(4)可维护性好，升级和改造容易，当系统维护、升级时，可临时让其他灾备节点接管，提高了业务的连续性。*M+N*容灾结构的建设成本和设备利用率介于 1+1 容灾结构和*N*+1 容灾结构之间，适合为移动智能网中设备数量较多、可用性要求高的SCP设备构建容灾系统。本文设计并实现了基于*M+N*结构的SCP容灾系统。

2 基于 *M+N* 结构的 SCP 容灾系统物理结构

在本地 HA(High Availability)生产系统和远程 DT(Disaster Tolerant)容灾系统的双层次安全体系之上，本文设计基于 *M+N* 结构的 SCP 容灾系统，如图 1 所示。*M* 台 SCP 生产设备均为本地 HA 系统(双机或多节点集群系统)，相互独立，分别通过 DDN/WAN 专线建立至容灾备份设备的远程数据网络连接实现远程数据备份，并通过 7 号信令网络与异地容灾系统建立信令连接。异地容灾系统采取业务接入与业务控制分离，由业务接入点(Service Access Point, SAP)、*N* 节点的 SCP 服务器集群系统和操作维护服务器(Operation And Maintenance Server, OAMS)组成，OAMS 将 SAP 和 SCP 服务器集群系统接入上层网管系统。

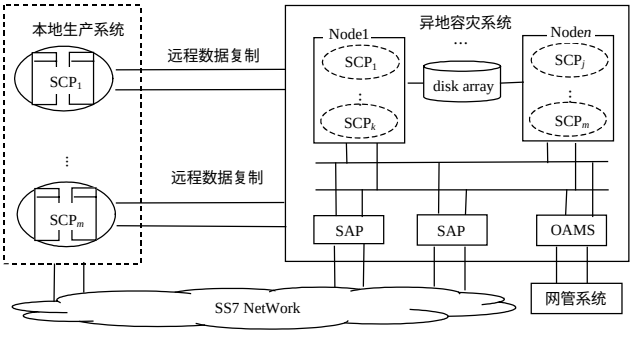

图 1 基于 *M+N* 结构的 SCP 容灾系统物理结构

正常情况下，本地生产SCP对外提供服务，异地DT系统为本地生产SCP提供备份，本地节点故障由本地HA系统优先接管。当本地HA系统接管失败或灾难发生时，借助SS7 网的

基金项目：湖南省自然科学基金资助项目(05FJ3018)；华为技术有限公司科技基金资助项目
作者简介：王福阳(1982－)，男，硕士研究生，主研方向：移动智能网，无线网络与移动计算；李小坚，教授；俞 前，工程师
收稿日期：2007-07-22 **E-mail：**mmyywong@gmail.com

故障检测和信令重定向[4-5]机制，信令转接设备STP将所有上行至故障SCP的信令消息转发至异地DT系统前端的业务接入点SAP。SAP接收信令消息，启动SCP服务器集群系统中目的容灾SCP的业务接管，创建至SCP应用服务器的连接通道，把SSP上行的业务请求消息透传给容灾SCP处理，并把容灾SCP下发的响应消息传给SSP，从而保证业务的连续性。其他与SCP连接设备如SMP、计费采集设备等，通过IP地址切换可以重新连接到对应容灾SCP。

3 基于 *M+N* 结构的 SCP 容灾系统软件结构

基于 *M+N* 结构的 SCP 容灾软件系统由容灾集群子系统、数据复制子系统、容灾业务子系统和容灾管理子系统组成。其中，容灾业务子系统和容灾管理子系统是 SCP 应用级容灾系统的关键部分。

3.1 容灾集群子系统

容灾集群子系统由 *N* 节点集群系统和存储域网络 SAN 构成。*N* 节点集群系统中通过 SAN 共享备份数据，集群管理软件通过节点间心跳消息通知，可以检测节点故障或人为的服务切出。

3.2 数据复制子系统

数据复制子系统完成生产设备和容灾设备间的数据备份和反向同步，是实现应用级容灾的基础。其实现取决于远程数据复制技术[3]的选择，当前成熟的产品主要有VERITAS VVR和Informix IDS-ER等。

3.3 容灾业务子系统

基于业务接入与业务控制分离，容灾业务子系统分为容灾业务接入服务端(DTSAS)和容灾业务控制服务端(DTSCS)，为容灾 SCP 提供基于 Standby, Active 和 Reserved 状态的业务接入和业务控制服务，如图 2 所示。处于 Standby 状态时，运行生产 SCP 至容灾 SCP 的数据复制服务，容灾 SCP 不对外提供业务接入；处于 Active 状态时，停止生产 SCP 至容灾 SCP 的数据复制服务，容灾 SCP 接管业务；处于 Reserved 状态，生产设备和网络故障恢复正常，运行容灾 SCP 至生产 SCP 的反向数据复制，容灾 SCP 提供受控服务，限制新呼入业务请求。其中，Reserved 状态为灾难恢复至服务回迁的过渡状态，提供受控接入并运行反向数据同步以实现快速的灾难恢复并倒回，同时避免在业务运营时期直接中断容灾 SCP 应用服务运行，人为造成当前已经接管的呼叫全部丢失及由此带来的计费损失。

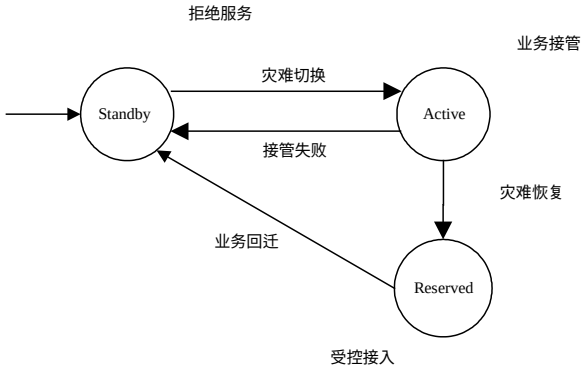

图 2 容灾 SCP 业务接入控制状态

3.3.1 容灾业务接入服务端

DTSAS 在 SAP 上运行，接收转发至容灾 SCP 的信令消息，根据各容灾 SCP 当前受控状态对信令进行相应接入控制。

(1)自动灾难切换

容灾 SCP 初始工作在 Standby 状态。为避免信令链路间断造成误切换，DTMS 接收信令消息并解析其目的地址和业务消息类型，累计 SSP 上行至目的 SCP 的业务请求次数，当满足灾难切换触发条件——在设定时间内接收到指定次数时，认为对应生产 SCP 发生故障，向 DTMS 报告自动灾难切换的业务接入事件。当 DTMS 回复业务接管请求时，DTSAS 切换对应容灾 SCP 至 Active 接入控制状态，并创建到目的 SCP 业务控制服务器的转发连接。

(2)信令接入控制

DTSAS根据当前容灾SCP的控制状态，对上行至目的SCP的信令进行相应接入控制：当目的容灾SCP处于Standby状态时，将所有接收到的信令消息均作为异常信令来处理，回复TC-U-ABORT[6]结束对话；处于Active状态时，DTSAS通过至目的容灾SCP的转发连接，将信令消息透传至容灾业务控制服务端进行业务接管；处于Reserved状态时，若接入新发起业务请求则回复TC-U-ABORT，否则直接转发。

容灾业务接入服务端结构如图 3 所示。

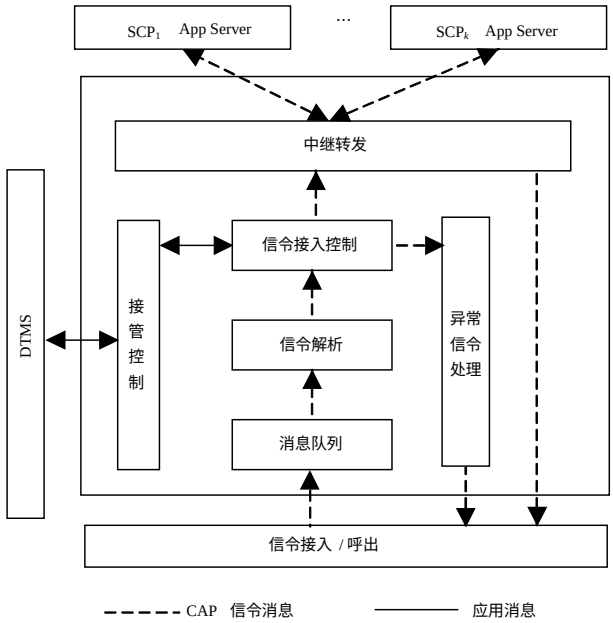

图 3 容灾业务接入服务端

3.3.2 容灾业务控制服务端

容灾业务控制服务端包含 *M* 个容灾 SCP，它们与生产 SCP 的业务应用子系统一一对应且相互独立，借助集群子系统构建成资源组，分别占用不同的卷组、文件系统、虚拟 IP 地址等资源，尽可能均匀地部署在 *N* 个节点。初始均处于备份状态，在容灾管理子系统控制下，完成数据库服务器和 SCP 应用服务器的启动以及业务的加载，从而实现容灾 SCP 的业务接管。

3.4 容灾管理子系统

容灾管理子系统提供业务接入管理和业务控制管理，实现与业务接入设备和业务控制设备的互操作。容灾管理子系统采用 CORBA/TMN 集成体系结构，包括容灾管理客户端(DTMC)、容灾管理服务端(DTMS)和容灾管理代理端(DTMA)。其中，DTMC 通过 IIOP(因特网 ORB 互操作协议)接入 DTMS，DTMS 通过网元协议 MML(人机交互语言)与 DTSAS 和 DTMA 通信，如图 4 所示。

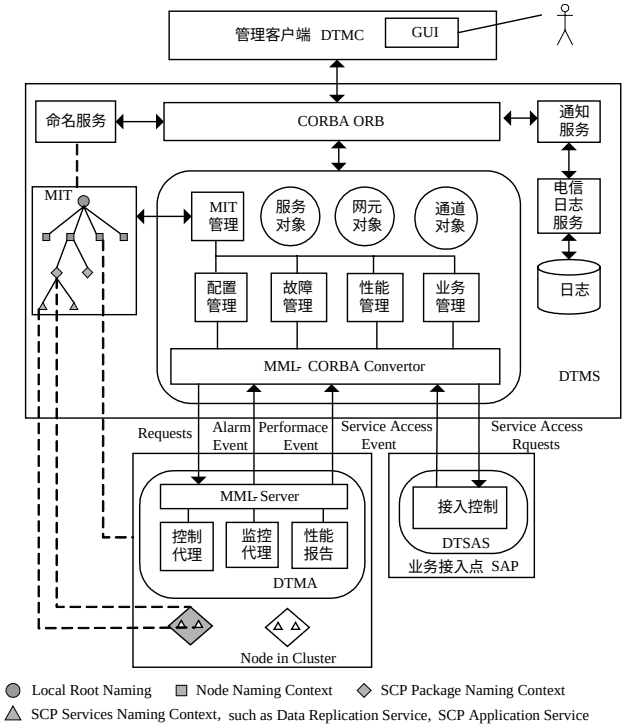

图 4　容灾管理子系统

3.4.1　容灾管理客户端

基于开放的 CORBA 管理接入，容灾管理客户端 DTMC 集成在网管系统中或为独立的客户端，为管理操作员提供 GUI 接口。

3.4.2　容灾管理服务端

容灾管理服务端 DTMS 在 OAMS 上运行，主要包含以下子模块：

(1)协议转换器子模块，完成协议网关的功能，提供 MML 命令和 CORBA 对象操作之间的转换。

(2)MIT 管理子模块，负责生成、查询、修改和删除 MIT 管理信息实例树。其中，MIT 以命名服务中的命名图存在，表示容灾系统中的网元节点、节点上启用的容灾 SCP 资源组和资源组中运行的应用服务，维护信息模型中的包容关系。

(3)配置管理子模块，处理 DTMA 发送的事件，包括管理对象创建、对象删除、属性值改变和状态改变，据此修改 MIT；为所有事件创建结构化通知，通过通知/日志通道上报至 DTMC，并通过电信日志服务记入日志。

(4)故障管理子模块，处理 DTMA 和业务管理子模块产生的告警，维护告警列表，通过通知服务同步告警消息至 DTMC，并记入日志。

(5)性能管理子模块，负责建立性能测量任务和收集性能数据。当接受 DTMC 和业务子系统的访问时，若本地数据是最新则直接获取，否则发送性能查询命令至 DTMA 以获取网元节点的当前性能数据。

(6)业务管理子模块，接收并处理 DTMC 操作请求以及 DTSAS 自动灾难切换的业务接入事件，将业务控制请求和业务接入请求通过协议转换器子模块，分别发送至 DTMA 和 DTSAS，实现对容灾 SCP 业务接入和业务控制的管理，为容灾 SCP 提供灾难切换、灾难恢复和业务回迁。业务管理子模块采取节点资源独占式的业务接管控制，即单个节点最多只允许 1 个容灾 SCP 处于非 Standby 状态。当接收容灾 SCP 的灾难切换请求时，业务管理子模块访问 MIT 和性能管理子模块，选择当前未被独占、启用 SCP 资源组数量最少且性能参数最优的节点，然后发送业务接管控制请求至 DTMA 以在选定节点上运行容灾 SCP 的业务接管，成功后将发送业务接入请求至 DTSAS，允许对应容灾 SCP 的业务接入。若节点选择失败或接管失败则拒绝灾难切换请求，并产生告警消息至故障管理子模块。

3.4.3　容灾管理代理端

容灾管理代理 DTMA 在每台容灾集群节点上运行，主要包括 MML-Server 子模块、控制代理子模块、监控代理子模块和性能报告子模块。其中，MML-Server 子模块负责 DTMS 与 DTMA 间命令行消息解析、分发和命令封装。控制代理子模块接收 DTMS 的 Request 消息，根据图 2 所示容灾 SCP 的状态行为，控制容灾 SCP 资源组在节点间的切换、双向数据复制服务以及容灾 SCP 业务接管服务的启动与停止。监控代理模块负责配置管理对象的事件上报，监控业务应用服务和数据复制服务的运行，接收集群管理软件的失效异常事件消息通知，触发告警消息上报。性能报告子模块负责实施性能测量任务，周期性地收集本节点当前性能数据(CPU 使用率、内存使用率、磁盘使用情况和网络 I/O 吞吐占用率等信息)，接受 DTMS 查询并返回最新的收集数据。

4　性能测试与分析

构建 4+2 SCP 容灾系统，其中，Node1 为 SCP1 和 SCP3 提供备份；Node2 为 SCP2 和 SCP4 提供备份。具体测试环境如下：容灾设备(Node1, Node2)为 2×HP RP7410，操作系统为 HP-UX 11.11，集群管理软件为 Mc/ServiceGuard A.11.09，数据复制软件为 VERITAS VVR/VxVm3.1，磁盘阵列为 VA7110。在容灾系统侧进行了 3 组呼叫对比测试：(1)Node1 无备份单独运行 SCP1，作为基准测试；(2)在 Node1 和 Node2 上分别运行容灾 SCP1 和容灾 SCP2，且为 SCP3 和 SCP4 备份；(3)Node1 运行容灾 SCP1，且为其他 3 台 SCP 备份。采用 MGTS 仿真模拟 SSP，发起呼叫 *CAPS* 数均为 150，通话时长为 60 s，测试时间为 10 min，测试结果如图 5 所示。

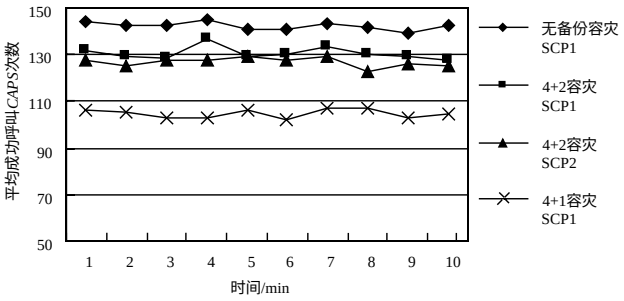

图 5　4+2 结构 SCP 容灾系统性能测试

测试结果表明，对比无备份基准测试，4＋2 SCP 容灾系统提供容灾备份，容灾 SCP 平均最大呼叫 *CAPS* 下降了 8%～10.3%，而 4＋1 结构提供容灾备份，容灾 SCP 平均最大呼叫 *CAPS* 下降了 26.4%。由此可见，*M*+*N* 容灾系统可同时为 *N* 个故障 SCP 提供业务接管，且相比 *N*+1 结构，*M*+*N* 容灾系统中数据备份服务分布在 *N* 个节点，减少了对单个容灾设备处理能力的占用，提升了容灾 SCP 的业务处理吞吐量。

5　结束语

本文采用业务接入与业务控制分离的双层结构，设计并实现了基于 *M*+*N* 容灾结构的 SCP 容灾系统，提出基于 Standby, Active 和 Reserved 状态的业务接入控制，保证了业务的连续性，避免了直接倒回而造成计费损失。本文采用基于 CORBA/ TMN 集成体系结构的容灾管理子系统，为快速、方便地进行容灾系统业务接入与业务控制管理提供了开放的网管系统接入。

系统测试表明，相比 *N*+1 结构，*M*+*N* 结构 SCP 系统最

（下转第 282 页）