

# 大型企业级云平台跨域资源调度与容灾模型

黄翔<sup>✉</sup>, 冯国聪, 宋云奎, 莫剑峰, 赵磊  
(南方电网数字电网研究院有限公司, 广州 510663)

**摘要:** [目的] 传统大型企业地域分散, 信息化投资不足, 信息安全存在较多隐患。云计算技术将企业的信息化资源进行有效的整合, 建设大型企业级云平台, 成为传统大型企业数字化转型的核心支撑平台, 其中跨域资源调度与容灾是关键技术。[方法] 基于云计算、大数据与传统灾备技术的综合运用, 探索大型跨地域、分层级的企业级云平台中跨域资源调度、备份容灾等云管理需求的实现模型。[结果] 经过实践测试, 初步实现了大型企业云平台多地资源统一管理、IT资源统一使用、智能化完成全局资源均衡与容灾的建设目标。[结论] 推荐的跨域资源调度与容灾模型, 对于大型企业云平台的建设和管理有重要的指导作用。

**关键词:** 大型企业; 企业级云平台; 多数据中心; 资源调度; 备份; 容灾

中图分类号: TM7; TP393.09

文献标志码: A

文章编号: 2095-8676(2020)S2-0011-07

开放科学(资源服务)二维码:



## Research on Cross-domain Resource Scheduling and Disaster Tolerance Model for Large Enterprise Cloud Strategy

HUANG Xiang<sup>✉</sup>, FENG Guocong, SONG Yunkui, MO Jianfeng, ZHAO Lei

(Digital Grid Research Institute, China Southern Power Grid Co., Ltd., Guangzhou 510663, China)

**Abstract:** [Introduction] Traditional large-scale enterprises are geographically dispersed, and information investment is insufficient with high security risks. Cloud computing technology effectively integrates enterprise information resources and builds a large enterprise-level cloud platform. It has become the core supporting platform for the digital transformation of traditional large enterprises. Cross-domain resource scheduling and disaster recovery are key technologies. [Method] Based on the comprehensive application of cloud computing, big data and traditional disaster recovery technology, this paper explored the implementation model of cloud management requirements such as cross-domain resource scheduling and backup disaster recovery in large-scale cross-regional and hierarchical enterprise cloud platforms. [Result] After the practice test, the construction goal of unified management of multi-site resources of large enterprise cloud platform, unified use of IT resources, and intelligent completion of global resource balance and disaster tolerance was initially realized. [Conclusion] The cross-domain resource scheduling and disaster tolerance model recommended in this paper has important guiding role for the construction and management of large enterprise cloud platform.

**Key words:** large-scale enterprise cloud; distributed datacenters; cross-domain resource scheduling; disaster recovery

随着全球经济产业链中配置分工的变化, 我国经济的升级转型正从农业和低端制造业向制造业强国、创新型经济迈进, 并制定了“中国制造2025”的宏大蓝图<sup>[1]</sup>。传统大型企业作为国家经济的主体, 信息化技术是提升企业经营水平、创新企业经

营方式, 攻坚领域核心技术, 提升企业的竞争力的重要助力。

为了解决企业数字化转型、精益化管理等迫切的生存压力, 国内企业非常重视云计算技术的研究和应用, 建设企业私有云平台, 充分发挥资源整合、高效共享、弹性供给、便捷服务的云计算特性, 有效实现公司信息化基础架构的革新, 进一步提升信息化服务创新的能力, 快速响应多元化业务发展的需要。

收稿日期: 2019-10-27 修回日期: 2020-07-13

基金项目: 国家重点研发计划资助“数字电网关键技术”(2020YFB0906000); 国家重点研发计划“基于数字孪生的海量多源异构数据中台建模与融合关键技术研究与应用”(2020YFB0906004)

当前企业私有云案例更多地局限于局部、小范围、一级集中的简单模式。对于地域分布广、资产管理复杂、数字化业务需求庞大的大型企业而言,建立多层级管理、跨域资源调度、具备备份容灾能力的私有云平台,仍是极具挑战性的工作。

本文着重讨论大型企业级云平台的顶层架构设计。结合企业的信息化资源现状和投资渠道,大型

企业级云平台应采用多级云架构模式进行建设,由总部主节点和按需建设的多个区域分节点组成,如图1所示。总部主节点建设在总部数据中心,由总部投资兴建,承载一级部署类业务;各分节点建设在各分支的本地数据中心,由分支公司整合利旧资源,遵从总部的云产品体系进行升级改造,承载二级部署类业务。

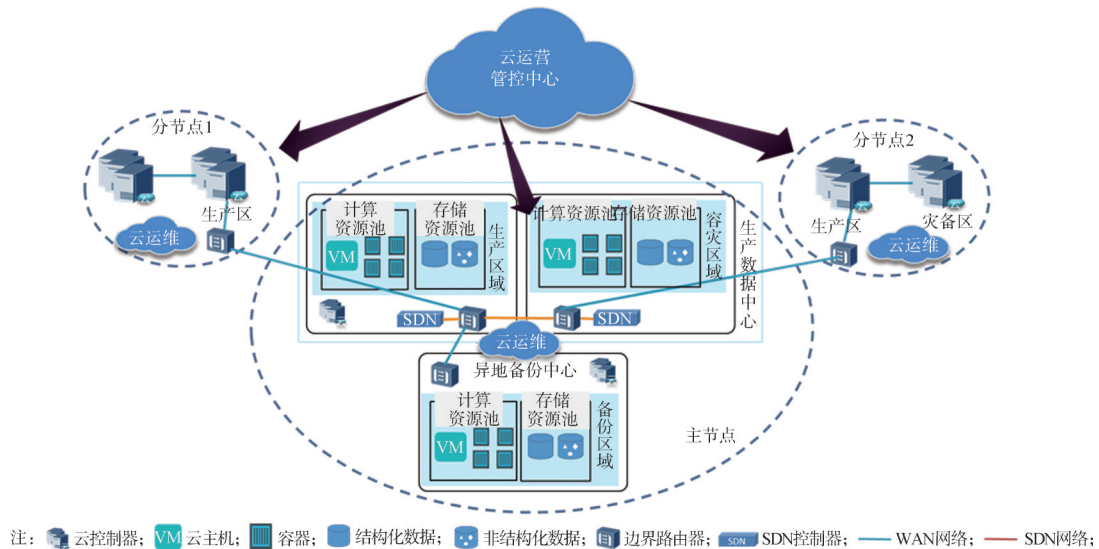


图1 大型企业级云平台顶层架构设计

Fig. 1 Architecture design of enterprise's cloud platform

主节点与分节点之间进行交互,分节点之间不进行交互。主节点与分节点通过云运营管控中心的跨域资源调度和备份容灾体系,形成逻辑上的“大型企业级云平台”。主节点与分节点的云产品体系保持一致,避免多种混合技术带来的集成和调度的困难和工期延误。因此云产品必须具备足够的弹性和兼容性。弹性指云产品需要适应不同的部署资源规模,从研发测试的个位数服务器到上千台的大规模集群;兼容性指云产品要适配主流的服务器、网络产品,同时兼容分节点早期型号的利旧服务器,避免现有投资的浪费。

在规划实施上,大型企业级云平台还面临如下挑战:(1)云平台部署规模的弹性和兼容性;(2)主节点与分节点的数据链路设计;(3)跨域环境的容灾设计。

## 1 云平台总体设计

为了适应云平台规模的巨大差异化,云平台产品必须是分布式架构,并内置集群化、异构兼

容性。

### 1.1 分布式架构

分布式架构<sup>[2]</sup>可以保证各类管理服务可以集中部署,也可以分散部署,这样可以充分利用多台物理服务器的计算能力,实现管理能力的横向扩充。例如,通常云平台都会包含计算、存储、网络、认证、镜像等服务,传统的三层开发模式下这些服务放在一个巨大的进程里,很难进行分节点部署,只能使用中间件组建有限规模的应用集群。而推荐的部署模型应该是基于消息队列的微服务集群,服务之间通过注册的服务地址进行调用,而不是基于IPC(Inter-Process Communication,进程间通信)。原理如图2所示。

### 1.2 集群化架构

云平台是以分布式架构为基础<sup>[3]</sup>,实现超大规模资源池部署的能力。图2中的最左边传统架构通常会基于商业中间件有中心控制节点的小规模集群来实现有限的服务容量的提升;而右边的大型企业级云平台则采用开源技术无中心控制节点的大规模

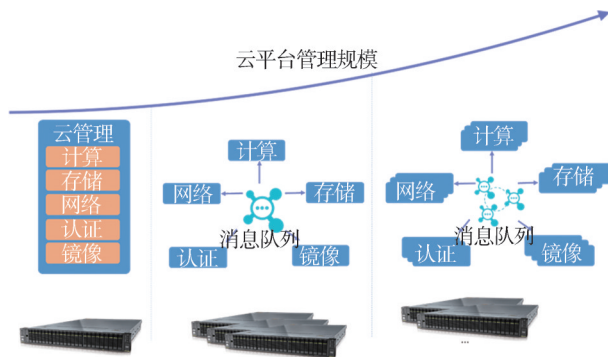


图2 云平台部署规模的弹性

Fig. 2 Elastic deployment of cloud platform

集群来实现无限的服务容量的提升, 具体有两类主流技术路线。

1) 集群软件。对于无状态类服务, 采用负载均衡软件实现大规模服务集群; 对有状态的消息队列、数据库, 采用相应软件本身专用的集群化功能来实现。

2) 容器化服务。基于容器编排, 使用容器编排的默认功能实现各类服务的集群化、数据持久化。

### 1.3 控制执行分离的兼容性架构

为了实现云平台广泛的兼容性, 云平台中各类云平台服务在设计上应采用控制与执行分离的架构设计。控制层制定协议和标准, 并作为云平台服务的对外出口, 较长时间内保持兼容和稳定; 执行层接受控制层的调度, 按照具体的计算资源特性(物理主机、虚拟机、容器)、存储资源特性(分布式存储、集中式存储)、网络资源特性(VLAN网络、VxLAN网络、SDN硬件VxLAN网络)<sup>[4]</sup>, 编写相应的执行层实现。原理如图3所示。

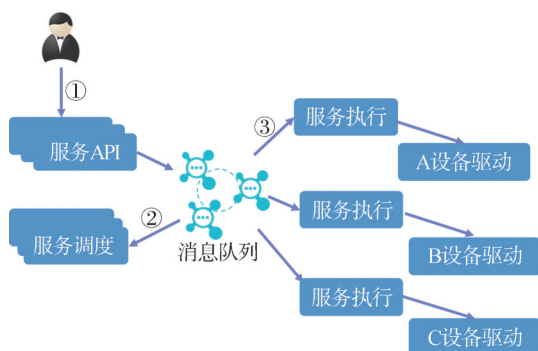


图3 云平台服务的兼容性设计

Fig. 3 The compatibility of cloud service

## 2 主节点与分节点的数据链路设计

### 2.1 技术难点

传统上大型企业的总部与分子公司数据中心之间的通信链路使用传统上用于报表传输的基础数据网, 数据通信的带宽有限、延时很高。因此对于主节点与分节点之间统一“监管控”就必须针对主节点与分节点的数据链路质量进行分层设计。最低要实现可以统一监控, 实现全局资源视图的集中; 其次可以实现统一管理, 实现资源的统一调度; 最好可以实现统一控制, 实现逻辑统一的云数据中心。

### 2.2 设计模型

为了减轻主节点与分节点之间在受限数据链路上的实时数据同步的通信压力, 企业云的主节点与分节点必须采用统一的监控接口和标准, 以降低系统间耦合的兼容性风险。对于主节点与分节点之间统一“监管控”就必须针对主节点与分节点的数据链路质量进行分层设计, 监控、管理、和控制对于数据链路的需求递增。

#### 2.2.1 统一监控

建设以分布式消息队列为核心的异步数据同步机制, 在有限的带宽上实现最大的数据转发量。在数据持久化层, 引入时序数据库(TSDB, Time Series and Spatial-Temporal Database)<sup>[5]</sup>。时间序列数据库主要用于指处理带时间标签(按照时间的顺序变化, 即时间序列化)的数据, 带时间标签的数据也称为时间序列数据, 使得时序大数据可以高效存储和快速处理海量时序大数据, 采用特殊数据存储方式, 极大提高了时间相关数据的处理能力, 相对于关系型数据库它的存储空间减半, 查询速度极大的提高。在数据展现层, 引入轻量级可定制商业智能分析(BI, Business Intelligence)系统<sup>[6]</sup>, 可方便地从时间序列数据库中进行数据的高速加工和多样可视化展现。

由于统一监控的数据是异步且可以动态调整数据上报周期, 可以基于大型企业的总部与分子公司数据中心之间的传统通信链路完成。

#### 2.2.2 统一管理

为了在跨地区的高质量链路上实现大型企业级云平台的统一管理, 通常会采用区域(Region)的设计概念。一个区域包含多个可用域, 一个可用域由多台服务器和网络设备、存储设备等组成。区域

用于机房之间计算节点的隔离。每个区域内的计算、存储和网络资源互相独立，共享认证管理等基础服务。一个区域通常部署一套云平台，通过云平台实现对该区域下所有可用域资源的统一管理和调度。可用域用于将不同机柜、不同规格的计算节点进行隔离，形成逻辑资源池。原理如图4所示。

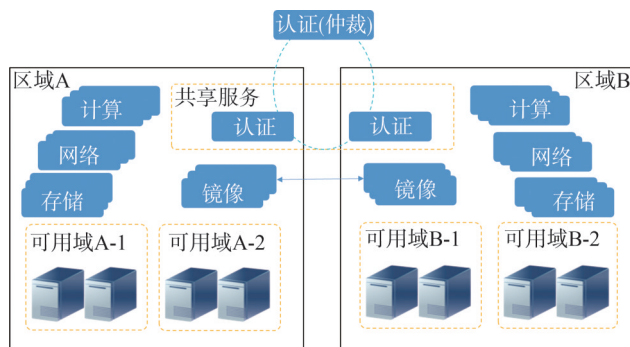


图4 跨区域统一认证

Fig. 4 Uniform authentication for multiregion

统一管理的数据延时需要专用链路来保障，可以是远程打通的VPN网络，可以是三层互通，也可以是二层直通，通常建议延时不超过20 ms。为了保障数据镜像同步的实时性，可根据业务镜像的同步数据量进行带宽规划。

### 2.2.3 统一控制

使用SDN软件定义网络技术来实现对云透明的Overlay网络，是构建跨域云平台的必要技术。网络硬件厂商的软件定义网络技术使用VxLAN网络通道来实现网络服务中的基础网络服务（端口、网络），并使用防火墙、路由器等网络硬件实现网络扩展服务（路由器、防火墙等）。网络硬件厂商的软件定义网络技术要求采购全新的网络设备，成本高，并可以跨越物理数据中心实现跨域大二层网络模型，适用于大规模的网络环境。原理如图5所示。

## 3 跨域环境的容灾设计

### 3.1 技术难点

利用跨域云平台环境，可实现业务的跨域部署和容灾。(1) 基于专用链路，实现业务应用网络的跨数据中心的大二层软件定义网络，部署逻辑上统一的业务应用集群；(2) 基于镜像同步机制，实现云内业务应用的跨数据中心的快速恢复；(3) 利用云平台中的备份、容灾等技术，可实现业务应用在

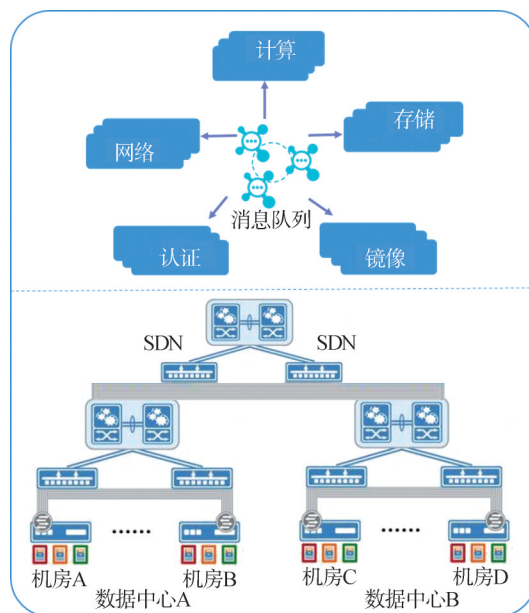


图5 基于SDN硬件网络实现跨域云的统一管理

Fig. 5 SDN based uniform multiregion management

多个数据中心的备份和按需恢复，以及基于业务应用的数据卷的实时复制的应用容灾。

### 3.2 设计模型

为了实现业务的高可用，利用跨域云平台环境实现业务的跨域部署和容灾。

#### 3.2.1 统一监控

在实现基于大二层网络的跨域云平台的基础上，业务应用可以轻松实现跨域部署。同一个业务应用中的云主机使用二层IP互通，即使物理上部署在不同的云数据中心。原理如图6所示。

#### 3.2.2 镜像同步容灾

为了保障多个云上部署的业务的基础镜像的一致性，主节点与分节点的镜像服务需要建立协同机制，支持多租户的镜像权限同步和控制，保障租户镜像的安全性和私密性。云平台镜像可通过云管理平台在多个节点中同步维护云平台镜像。云管理平台通过侦听不同云平台中的镜像管理操作，可即时触发镜像的同步操作，同步后的镜像可保持ID、名称、附加属性一致。云平台镜像支持私有属性，租户可以保护自己的应用镜像不被其它租户访问。在多个云平台之间同步镜像时，需要在同步镜像时不随意扩大镜像的可见范围，避免数据安全风险。

为了在容灾场景中快速恢复应用，可基于同步的云平台镜像恢复业务，实现云内业务应用的跨数

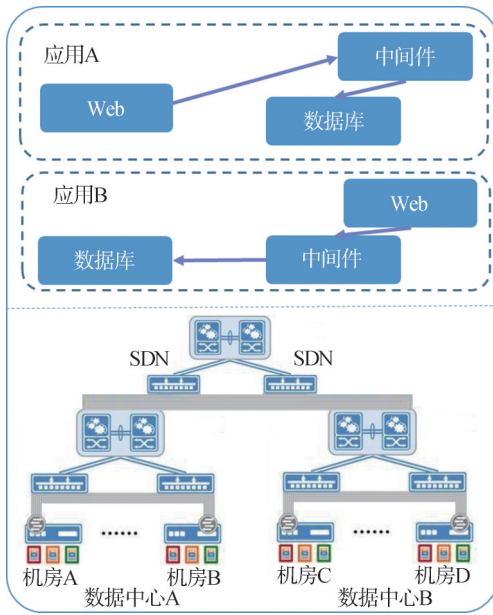


图6 基于SDN硬件网络实现跨域业务应用部署  
Fig. 6 SDN based multiregion application deployment

据中心的快速容灾。应用在节点1中完成部署后,云管理平台自动将部署后的镜像同步到节点2。在灾难恢复场景中,可以在节点2中基于云平台镜像恢复虚拟机,减少了云平台镜像的复制时间,显著缩短了应用容灾的恢复时间。原理如图7所示。

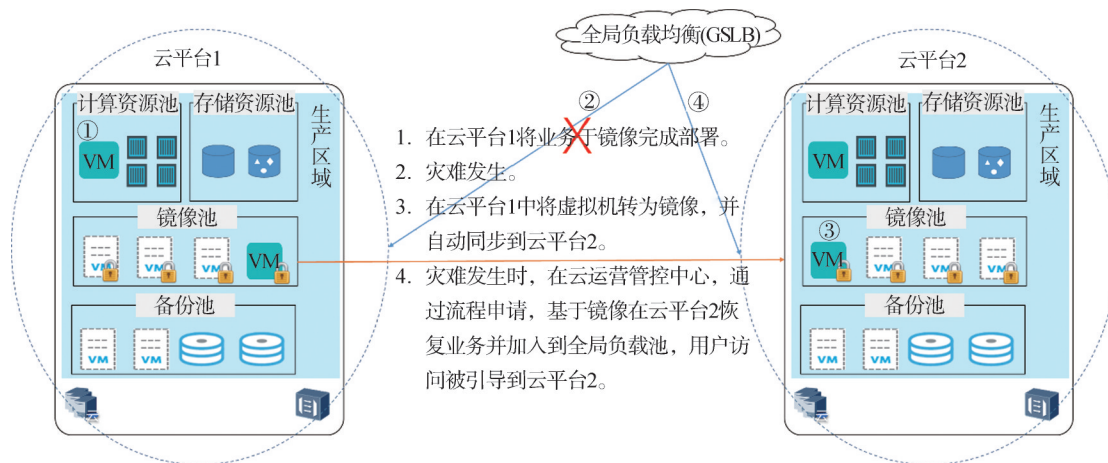


图7 基于云平台镜像同步的容灾模型  
Fig. 7 Image based disaster recovery model

存储数据的容灾根据存储技术不同而异,集中式存储支持实时数据异地复制,分布式存储支持异步数据异地复制。由于延时的敏感性,存储数据的容灾通常要求延时低于5 ms,物理距离小于100 km。

### 3.2.3 云平台容灾

利用云平台中的备份、容灾等技术,可实现业务应用在多个数据中心的备份和按需恢复,以及基于业务应用的数据卷的实时复制的应用容灾。

#### 1)应用数据的备份与恢复

应用数据备份遵从传统应用备份的方案,恢复时需要手工在云平台中进行恢复。云主机未损坏的情况下可以直接覆盖数据;云主机损坏的情况下需要先恢复云主机,再恢复应用数据。

#### 2)云平台数据的备份与恢复

云平台数据备份可实现业务应用在多个数据中心的备份和按需恢复,可以直接将备份数据恢复为云主机。

备份网络通常要单独建设LAN-Free专用跨数据中心网络;容灾网络需要高速低延时网络。<sup>[7]</sup>

#### 3)云平台数据的容灾

云平台数据的容灾基于存储的双活可实现业务的连续性,并在业务灾难情况下自动启动灾难恢复<sup>[8-9]</sup>。云平台数据的容灾在存储数据容灾的基础上,实现了云平台中云主机数据的同步,并由云平台主导数据的同步方向和云主机的灾难恢复。原理如图8所示。

## 3 大型企业级云平台的效益分析

长期以来,传统大型企业在信息化基础设施建设方面存在明显短板,如数据中心分散(总部和每个分支独立建设,没有统一的建设和运行标准)、机房空间有限(部分只能租用公有云,存在信息安

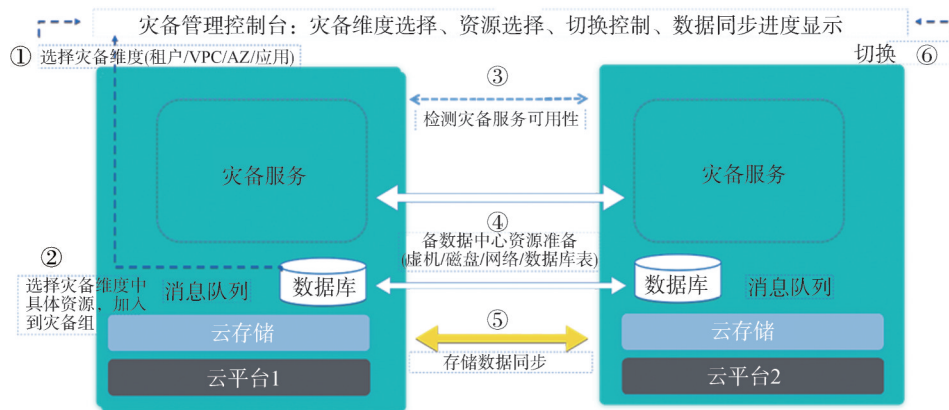


图8 基于云平台数据的容灾模型

Fig. 8 Disaster recovery model for cloud data

全及数据风险，难以融合公司本地化IT资源）、机房环境不达标（基建投入不足，部分采用办公空调，恒温恒湿不达标）、安全防护低（受限于信息化支出和本地信息化团队的技能，安全水平参差不齐）等，信息系统的安全稳定运行存在较多隐患，数字化转型的新业务开展困难（数据不流通、属地定制化开发部署、软硬件绑定导致成本高企），制约业务发展。

对于大型企业级云平台的生产效益，本文认为可以从提升业务服务水平、业务创新、成本控制三大角度去考虑。首先，大型企业级云平台整合了大型企业内部分散的资源，做到了应用部署的全局化视角，可以立足本地，在全企业内部实现业务的多活、容灾和数据备份，对于物联网等新形态业务也可以做到就近部署和访问，大大提升了业务的服务水平。例如：云平台镜像的容灾恢复方式比传统备份方式少了数据恢复的过程，将RTO从1h缩短到10min，并且容灾的跨度从局部地区到了全国。其次，业务创新对于信息化资源的同构的依赖性，产生了对大型企业级云平台的强需求。业务创新对于需求的变化敏感，如果每个分支机构的修改化需求都需要定制化软件和差异化硬件来实现，对于业务创新来说是巨大的桎梏。大型企业级云平台实现了各地差异化的硬件的标准化、服务化，通过配置化来实现创新业务的快捷部署。例如：以微服务架构开发创新业务，以DevOps模式根据需求滚动更新版本，以配置管理适应各分支的业务定制，可以简化创新应用的落地。最后，大型企业级云平台从以下方面加强了信息化系统的成本控制：（1）大规模

云降低了云的建设成本；（2）云运维的自动化程度高，降低了云的运维成本；（3）业务应用的部署在云的各个分支节点趋同，甚至可以从主节点统一下发，降低了应用的部署成本；（4）云业务的弹性伸缩、故障自愈，降低了业务的运营成本；（5）全局统一的“监管控”，增强了企业数据信息的透明度。

#### 4 结论

本文分析了建设大规模、跨地域的全局统一的企业云关键技术的前瞻性研究。并以此为基础提出了大型企业云跨域资源调度与容灾模型的顶层设计、总体架构、技术选型。重点说明了三项企业云跨域资源调度与容灾的关键技术，包括云平台部署规模的弹性和兼容性、主节点与分节点的数据链路设计、跨域环境的容灾设计等。最后，验证了企业云跨域资源调度与容灾模型的应用场景，对于大型企业云平台的建设和管理有重要的指导作用。在模型验证中需要综合云计算、大数据与传统灾备软件，技术实现较为复杂，用户体验一致性不高，完整性有待进一步增强。

#### 参考文献：

- [1] 中华人民共和国工业和信息化部.《中国制造2025》“1+X”规划体系[J]. 机械工业标准化与质量,2017(3):4-4. Ministry of Industry and Information Technology. Made in China 2025 “1+X” planning system [J]. Machinery Industry Standardization & Quality, 2017(3):4-4.
- [2] 欧伟新,陈国忠.云计算及其分布式架构[J]. 电信快报, 2012(1):30-33. OU W X, CHEN G Z. Distributed architecture of cloud computing [J]. Telecom Express, 2012(1):30-33.

- [3] 刘颖. 基于分布式系统的微服务架构演进 [J]. 通讯世界, 2018(7):97-98.  
LIU Y. Evolution of micro-service architecture based on distributed system [J]. World of Communication, 2018 (7) : 97-98.
- [4] 蒋穗,梁勇,诸旻. 云计算数据中心SDN引入研究 [J]. 中国新通信,2016,18(9):18-19.  
JIANG S, LIANG Y, ZHU W. Study on the SDN in cloud computing centers [J]. China New Communications, 2016, 18 (9):18-19.
- [5] 黄红远,李高明,欧阳为年. 一种新型时序数据库在大型配电网自动化系统数据中心中的应用 [J]. 中国电业(技术版), 2014(7):35-37.  
HUANG H Y, LI G M, OUYANG W N. Application of a new time series database in data center of large distribution network automation system [J]. China Electricity Industry (Technical Edition), 2014(7):35-37.
- [6] 张静妙,高双喜. 商业智能(BI)市场的应用和发展趋势分析 [J]. 集团经济研究,2007(22):69-70.  
ZHANG J M, GAO S X. Trend analysis of business intelligence (BI) market [J]. Group Economic Research, 2007(22) : 69-70.
- [7] 王剑鸿. 基于SAN的LANFree集中备份和恢复系统 [J]. 电力信息化,2008,6(12):36-40.  
WANG J H. SAN-based LAN free centralized backup and recovery system [J]. Power Information, 2008, 6(12):36-40.
- [8] 张崧,熊祥鸿,刘建清. 数据容灾技术比较及在电力系统数据中心建设中的应用 [J]. 华东电力,2012,40(12):2249-2252.  
ZHANG Y, XIONG X H, LIU J Q. Comparison of data disaster recovery technology in the construction of power system data center [J]. East China Electric Power, 2012, 40 (12) : 2249-2252.
- [9] 陈亮,陈远斌,姚尧. 高校“双活”架构的虚拟化数据中心 [J]. 科学技术创新,2017(21):150-151.  
CHEN L, CHEN Y B, YAO Y. Virtualized data center of

"double live" architecture in universities [J]. Science and Technology Innovation, 2017(21):150-151.

#### 作者简介:



黄翔

**黄翔** (通信作者)

1982-, 男, 湖南常德人, 南方电网数字电网研究院有限公司高级工程师, 中国科学院软件研究所博士, 主要从事云计算、大数据、人工智能等方面研究与工程工作 (e-mail) huangxiang@csg.cn。

#### 冯国聪

1981-, 男, 广东恩平人, 南方电网数字电网研究院有限公司副总监, 中山大学大学硕士, 主要从事电网数字化、云计算、物联网、大数据、人工智能等方面研究与工程工作 (e-mail) fenggc@csg.cn。

#### 宋云奎

1979-, 男, 吉林敦化人, 南方电网数字电网研究院有限公司工程师, 中国科学院软件研究所硕士, 主要从事云计算、服务计算和大数据等方面研究与工程工作 (e-mail) songyk@csg.cn。

#### 莫剑峰

1989-, 男, 广东江门人, 南方电网数字电网研究院有限公司高级工程师, 暨南大学学士, 主要从事云计算、大数据、人工智能等方面研究与工程工作 (e-mail) mojfc@csg.cn。

#### 赵磊

1979-, 男, 河南焦作人, 南方电网数字电网研究院有限公司网络工程师, 北京邮电大学在职硕士研究生, 主要从事云计算、物联网等方面研究与工程工作 (e-mail) zhao-lei@csg.cn。

(责任编辑 李辉)