

虚拟化技术在气象业务应用中的效益分析

■ 李楠 聂元丁

在气象业务高速发展、建设过程中，系统重复建设、IT资源浪费、冗余等问题越发凸显。为解决上述问题，通过国家气象信息中心气象资源池项目的建设，采用虚拟化技术，建立了高效、集约、可动态调度的计算存储资源池，极大提高了物理服务器和存储系统的硬件利用率，解决了之前物理服务器和存储资源利用率低，系统部署慢等问题，节约了IT系统基础设施的投资成本，达到集约高效的目的。

虚拟化技术起源于20世纪60年代末，当时计算机系统较为复杂、昂贵，一个学术科研机构一般只有一个或几个大型计算机系统，而且早期操作系统存在只能单任务处理的问题，为方便多个用户/应用共享计算机资源，发展出了可以将硬件资源进行逻辑划分的虚拟化技术。经过信息技术数十年的发展和演进，虚拟化技术也得到了极大的完善和进步，但究其本质仍是一种资源管理技术，用于将实体资源进行抽象，打破实体结构间不可分割的障碍，供用户更灵活的使用资源。随着21世纪云计算的兴起，虚拟化作为其中一项核心技术，在各个领域均呈现爆发式增长。气象行业由于具有资料量庞大，业务系统复杂等特点，在扩展性、易用性、功能性等方面对基础平台提出较高要求。在传统基础设施平台难以满足气象信息业务高速发展的带来的技术需求时，虚拟化技术的应用，可作为有力的技术驱动力，支撑业务发展。

1 虚拟化技术在气象业务的应用

1.1 虚拟化技术概述

虚拟化是指将计算机、存储等硬件的各种实体资源予以抽象、转换，形成可供多个上层系统、应用共享使用的状态的过程。根据对象的不同，可分为平台虚拟化、应用程序虚拟化、内存虚拟化、存储虚拟化、网络虚拟化等。当前应用最为广泛的是平台虚拟化，其旨在通过技术手段将硬件平台与终端用户使用的操作系统分开。平台虚拟化根据实现方式的不同，又可以分为完全虚拟化和部分虚拟化。完全虚拟化在硬件和虚拟机操作系统之间增加软件层，将两者进行逻辑分离，其优势在于虚拟机所使用的操作系统无需任何改动，主要产品有VMware-ESXi、KVM、Xen-HVM等。部分虚拟化利用虚拟化软件实现对底层硬件的共享访问，针对部分特权指

令，用户虚拟机可以与硬件进行直接交互，其优势在于虚拟机运算执行效率高，接近物理机效能，主要产品有Hyper-V、Xen-PV等。

当前，虚拟化在多个技术领域均有具体实现，如计算资源相关的服务器虚拟化、桌面云、弹性计算服务，以及其他资源相关的SDN、虚拟交换机、虚拟防火墙、分布式存储等都是虚拟化技术应用的产物。

1.2 技术应用背景

气象业务系统的发展历程与业界信息技术的发展是高度一致的，未应用虚拟化技术之前，在IT基础设施方面，主要使用的有高性能计算机、小型机以及x86服务器三种硬件架构。三种架构面向的业务功能及运行模式不同，在技术更新换代过程中的替代性也有所差异。高性能计算机主要用于复杂的数值模式计算，需要消耗海量的计算资源。由于高性能计算机系统的高可靠性和针对特定复杂问题的高度优化特性，使其短时间内无法被其他架构取代。小型机主要用于对性能和可靠性要求较高的业务系统，如交易系统、核心数据库等，此种平台优势在于系统稳定性高，单机性能强，但在业务量爆发膨胀的情况下，难以实现横向扩展，业务规模受到单机能力限制。x86服务器主要用于网站、数据传输、软件开发等轻量业务系统使用，此类业务对资源需求量不大，如使用独立的服务器，会造成较为严重的系统资源浪费，同时硬件设备采购周期较长，不利于业务快速上线。为解决或缓解上述软硬件在业务使用中存在的问题，气象信息系统的建设过程中引入虚拟化技术，并基于此技术建设IT基础设施云平台。

1.3 技术应用情况

2012年开始，信息中心开始试点使用服务器虚拟化技术。采购批量x86架构服务器，通过安装虚拟化

软件，将CPU和内存资源进行池化，通过SAN网络或以太网与存储资源进行连接，构成服务器虚拟化资源池（图1）。提供中国气象局机关及直属单位单位使

用，当前资源池平台包括38台x86服务器，5套本地存储，3套网络存储，在局域网、互联网两个网络区域提供服务。

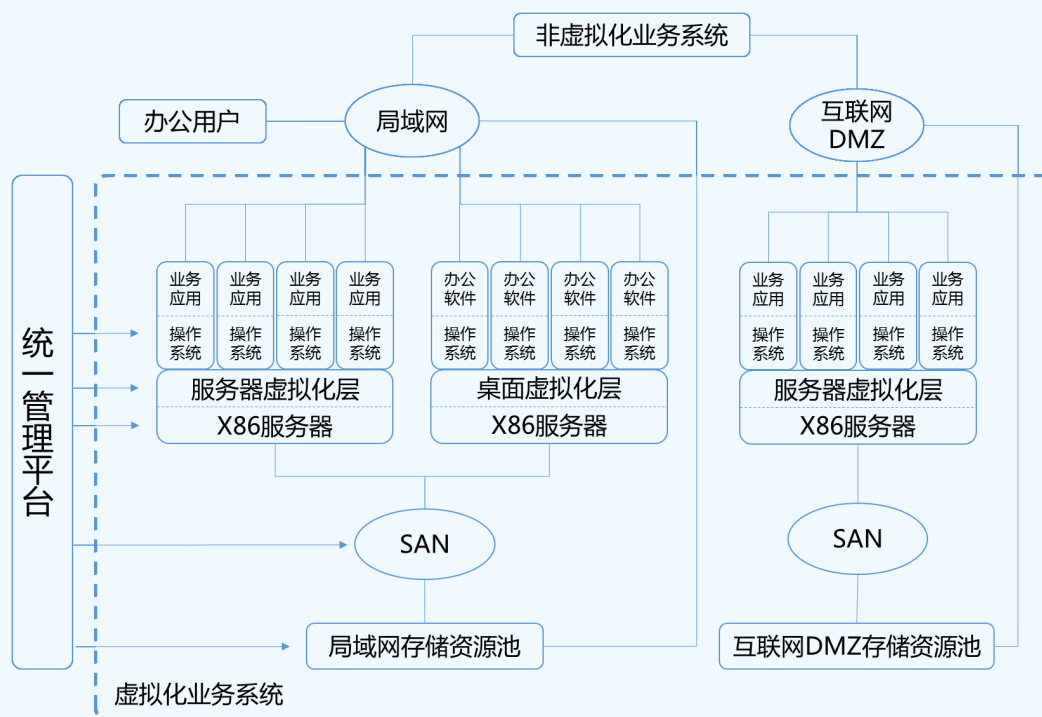


图1 资源池平台架构

其中，业务资源池已为中国气象局机关及国家气象中心、国家气候中心等八个直属单位分配资源，共承载包括CIPAS2.0，强天气预报系统，国家级气象业务内网，中国气象数据网，军事保障平台等231个业务系统共532台虚拟机运行。资源池中CPU处于超分配状态，集约化比（物理服务器数量:虚拟服务器数量）达到1:14，集约化效益显著。业务资源池资源分配及使用情况如表1所示。

表1 业务资源池资源分配及使用情况

物理机数量	虚拟机数量	集约比	资源	总量	分配量	使用率
38	532	1:14	CPU(核)	2740	3093	12.99%
			内存(TB)	12.1	6.74	40.05%
			本地存储(TB)	698	454	65.04%
			网络存储(TB)	2468	997	21.34%

2014年，信息中心小规模试点桌面云技术，采用4台刀片服务器和1套磁盘阵列搭建桌面云平台，已为61个用户创建桌面虚拟机，部分用户工作环境已切换为云桌面办公，工作时间在线用户30人以上。

2015年开始，在虚拟化资源池运维中发现，本地磁盘阵列存储扩充规模有限、难以支持大数据量存储需求的情况越发凸显，因此，引入分布式存储技术。

将分布式存储通过以太网，连接至局域网，为资源池及其他业务提供网络存储服务。目前分布式存储系统在局域网和互联网区域均已部署，由于分布式架构对横向扩充具有较强的支持能力，经过多次扩充后，总量已达到2.4 PB，已分配996.8 TB，使用率为21%，为46个业务提供大容量存储服务。

2 虚拟化在气象业务中的效益分析

2.1 服务器虚拟化效益

采用服务器虚拟化技术带来多方面效益，从场地环境方面进行考量，38台4U服务器需4个机柜，如果所有虚拟机全部采用物理服务器承载，根据1.3节中统计的虚拟机分配数量，则一共需要532个独立物理服务器，如果假定每个服务器为4U，需要至少54个机柜，如果以主流的2U服务器计算，也需要至少26个机柜，因此采用虚拟化技术节省了80%以上的场地空间。能耗方面，平均每台x86服务器功率约为400 W，38台物理服务器相比532台物理服务器，一年可节省电量约173万度。除场地环境外，虚拟化技术也极大缩短了业务部署上线周期，将持续数月的设备招标、采购、部署流程缩短到数十分钟。此外，还可提高IT

基础设施可靠性，不再因物理设备故障导致业务系统停止服务，通过备份系统对虚拟服务器进行定期整机备份，在操作系统崩溃或软件故障时也可快速恢复系统。另外，出现资源不足情况后，可快速实现虚拟硬件（CPU、内存、存储等）扩充，避免因硬件能力限制导致业务发展。

通过对基础设施资源池中477台部署代理程序（可提取资源使用率）的虚拟机进行分析统计，虚拟

机CPU平均使用率为0.22%，本地存储平均使用率为0.28%。如无虚拟化环境，所有虚拟机将使用物理机承载，因此虚拟机使用率可类比为传统环境物理机使用率。结合资源池计算资源使用率12.99%和存储使用率65.04%进行分析，虚拟化技术的应用，使计算资源使用率提高59倍以上，存储资源使用率提高232倍以上，体现巨大集约化效益（表2）。

表2 服务器虚拟化效益分析

对比内容	传统服务器环境	服务器虚拟化环境
操作系统/物理服务器	1:1	1:14（单位物理机承载业务系统数量多）
基础平台就绪时间	数月	30 min内
场地空间及空调电源 （按照每个操作系统计算）	约400 W	<50 W
计算资源使用率	平均小于0.5%	大于10%
存储资源使用率	平均小于1%	大于60%
平台可靠性	物理服务器故障会导致其上业务停止	集群内单个服务器故障不影响虚拟机使用
数据安全性	仅可备份数据、文件，恢复后需重新安装操作系统及部署应用	可进行整机备份及恢复，无需重新部署业务系统
资源扩充	扩充资源需采购硬件、停机实施。扩充周期长，会对业务造成影响	可在线扩充虚拟机的CPU、内存、存储等虚拟硬件资源，无需停机
系统管理	部分对业务服务器的操作（如服务器开机、安装调试操作系统等）需要到机房现场进行	所有对业务虚拟机的操作可远程执行，操作便捷，时效高

2.2 桌面虚拟化效益

桌面虚拟化带来的效益（表3）是多方面的，首先，方便为新员工、临时用户、测试环境等快速部署桌面环境，资源使用紧张时，可在线扩充，资源使用完后可及时回收，避免造成办公电脑闲置问题，增加整体硬件利用率。同时，平台已为所有桌面虚拟机提供整机备份服务，方便故障恢复，不会因硬件故障导致数据丢失。另外由于桌面操作系统运行于服务器中，平台可支持通过笔记本电脑、平板电脑、手机等设备远程连接，可实现移动办公。

表3 桌面虚拟化效益分析

对比内容	传统办公环境	桌面虚拟化环境
办公环境就绪时间	数天	30 min内
移动办公	无法实现	可使用笔记本、平板电脑、手机通过软件连接虚拟桌面
资源扩充	需要额外采购硬件，存在兼容性风险，扩充周期长	可随时在线扩充，即时生效
办公环境	占用空间多、噪音大、发热量高	占用空间少、噪音小、发热量低
数据安全	硬盘为单点故障，依靠用户个人技术能力进行备份	后台每日整机备份，可以整机恢复系统

2.3 分布式存储效益

分布式存储的应用，为海量数据的在线存储提供可能。传统磁盘阵列存储整机性能受限于存储控制器性能限制，一般存储总量小于10 PB，而且存储系统总带宽等于控制器带宽。分布式存储基于分布式架构，可以扩充节点对存储系统进行扩容，由于其线性

扩充特性，单存储系统可以扩充至100 PB以上，并且每个存储节点均可与外部系统进行流量交互，使整个存储系统的吞吐量可以随存储节点数量增加呈线性增长（表4）。从容量和吞吐量两方面，满足大数据访问需求。

表4 分布式存储效益分析

对比内容	传统磁盘阵列	分布式存储
挂载使用方式	需连接专用SAN网络使用	接入以太网，网络可达即可使用
最大容量	小于10 PB	大于100 PB
吞吐量	受限于磁盘阵列控制器性能	吞吐量随节点数量线性增加
可靠性	磁盘阵列控制器为单点故障点	无管理节点、控制器，不存在单点故障

2.4 效益总结

通过以上对五年来虚拟化技术在气象业务中应用的情况分析，无论是在服务器、桌面、存储等方面的使用，均带来多重效益，主要体现在资源就绪时间，资源灵活性和系统可靠性三个方面。在现今业务应用开发、迭代速度加快，数据可靠性地位提高的信息技术背景下，虚拟化技术应用带来的效益为气象业务应用的快速发展提供了强有力支持。

3 虚拟化技术在气象业务的发展前景

当前虚拟化技术的应用仅仅解决了硬件资源不集约、硬件使用率低等问题，烟囱式的业务设计模式并未得到有效解决。仅是将物理服务器部署形式变为虚拟服务器部署，在操作系统和基础软件层依然有大量

的资源冗余、重复建设现象，并未将气象业务人员彻底从系统运维工作中解放，仍需进行大量基础维护工作。

为了解决以上遗存问题，除已经应用的服务器虚拟化和存储虚拟化技术外，还可引入当前技术成熟度较高，且应用较为广泛的还有容器虚拟化及网络虚拟化技术。通过容器虚拟化，可以使不同应用系统复用操作系统资源，方便业务人员对于应用系统的调试和部署，极大减少部署业务的时间成本和人力成本。网络虚拟化的部署则可以按需对业务进行网络安全策略发放，对不同业务进行安全隔离。未来对于容器及网络虚拟化的应用，将降低业务开发、部署及网络运维方面的工作难度，释放相关业务人员运维压力。

通过服务器和存储虚拟化技术的应用，已在气象

业务应用中取得了巨大效益，解决众多业务对平台资源的需求问题，节省大量机房场地能源成本，实现部分传统平台不具备的功能。未来还将在系统建设中应用其他虚拟化及云计算相关技术，解决应用级使用问题，满足气象业务的发展诉求。

深入阅读

沈文海, 2012. 从云计算看气象部门未来的信息化趋势. 气象科技进展, 2(2): 49-56.

王彬, 2014. 高性能计算技术在气象部门的应用. 计算机工程与设计, 35(4): 1476-1479.

中国信息通信研究院, 2016. 云计算白皮书(2016年).

(作者单位: 国家气象信息中心)

(上接207页)

核汇交资料或气象大数据在流程、质量或效率上的指标。分析主场景以桑基图展示各汇交阶段，基于时序的数据包中各成分在汇交过程中的转化情况。

面向数据的主题分析是在基本维度的基础上，以气象大数据维度的元数据为主，可考核各数据源维度上的汇交资料转化为气象大数据的流程、质量或效率。分析主场景以树图展示各类气象资料通过元数据规范化识别并归入气象大数据目录的吸纳情况。

跨行业、跨部门、跨系统的主题分析是在基本维度基础上，以数据源维度中行业、数据生产者、业务系统三种情况为主，可考核汇交资料维和气象大数据维度上同一度、差异度和对立度等指标，以解析相关行业、部门和系统的数据质量。分析主场景以旭日图展示各数据的细分溯源、数据具体构成以及存储情况。

面向服务的主题分析是在基本维度基础上，对汇交日志数据进行时序显著性、空间显著性、关联关系等深度分析，通过构建数据资源业务价值评估指标，进行相应的服务价值评估。分析主场景以时间序列、地图等为背景，采用点、线、面的不同标识来呈现其显著性和关联性。

4 展望

现有数据可视化技术为可视化分析提供了丰富的表现形式，但在气象大数据资源汇交这类过程复杂性的可视化分析领域中，主要挑战来自于业务对象关系、业务观测指标、数据描述维度、应用分析场景的复杂性，但却是决定可视化内容的关键。本文结合气象大数据资源汇交的业务过程，汇交资料与气象大数据间汇交的内在规范性联系，构建了气象大数据资源汇交的数据模型、监控指标、分析设计框架和分析场景，解决了气象大数据资源汇交可视化多维分析的一些基础性问题，但可视化分析效果还需结合可视化技术，在构建集成的、实时的、交互式的气象大数据资源汇交全景工具中改进。

参考文献

- [1] 霍亮, 朝乐门. 可视化方法及其在信息分析中的应用. 情报理论与实践, 2017, 40(4): 111-116.
- [2] 任利敬, 赵正旭, 徐骞. 三维地图式可视化大数据. 测绘与空间地理信息, 2017, 40(2): 63-65.
- [3] 杨彦波, 刘滨, 祁明月. 信息可视化研究综述. 河北科技大学学报, 2014, 35(1): 91-102.
- [4] 洪文学, 王金甲. 可视化和可视化分析学. 燕山大学学报, 2010, 34(2): 95-99, 105.
- [5] 赵克勤, 集对分析及其初步应用[M]. 浙江科学技术出版社, 2000.