数据中心运维管理方案

第一章 某数据中心基础运维概述

某数据中心的基础运维工作主要包含包括四个部分：基础环境、网络、服务器存储和基础软件。

- 其中第一部分机房基础环境部分，包含机柜位置、空调、消防、安防、弱电、UPS 等最基础的机房环境设施。需要对这些基础环境部分进行运维维护，确保整个机房环境正常稳定。
- 第二部分为网络环境，包括当前数据中心所有的交换机、路由器等设备，以及由这些设备组成的所有网络，需要监控网络运行情况并提出网络风险评估，定期对网络进行优化配置，提高网络运行效率，保证整个网络环境的安全。
- 第三部分服务器和存储部分，包含整个数据中心的小型机、服务器、存储设备、SAN 交换机等设备。这些设备支撑着整个业务系统，是非常重要的基础硬件环境。需要监控这些设备的运行情况，及时处理出现的问题和变更，并基于整个环境提供优化。
- 第四部分为基础软件部分，包括各种操作系统、数据库、中间件、备份软件等等。要求这些软件可以正常工作，并优化配置，为平台和工作站正常服务，当这些软件出现问题时，能发现并提出解决方案；可以协助应用人员解决故障或进行对应的变更、升级等操作。

本方案将基于这几个方面进行设计，确保数据中心正常、高效运行。

第二章 数据中心运维分类

某数据中心运维团队将根据当前数据中心的实际情况和对应的管理制度，通过主动性、预防性维护，执行日常维护作业计划，对告警、性能、运行状态进行检查分析，及时进行数据备份，并定期对备份数据进行恢复性测试验证，对系统运行质量进行分析，并进行维护记录。对监控或维护中发现的问题及时处理，消除隐患，保障平台的稳定运行。我们将基于以下几个方面对运维工作进行描述

2.1 基础环境运维管理

针对基本的机房环境设施，我们的工作内容包含以下这些内容：

1）机房机柜摆放规划和机柜管理；

2）服务器和网络设备摆放规划和日常管理；

3）设备出入机房审批登记管理；

4）内部人员出入机房审批登记管理；

5）外部来宾机房参观审批登记管理；

6）机房电力系统监控、问题及时上报；

7）消防监控系统监控、接收报警短信和联系第三方；

8）空调报警系统监控、接收报警短信和联系第三方；确认空调运行状态良好。清洁机房的空调防尘网。

9）温湿度报警监控、接受报警短信和联系专业第三方；

10）漏水报警系统监控、接受报警短信和联系专业第三方；

11）IC 卡门禁系统日常运维；

12）视频监控系统日常运维；

13）UPS 报警系统监控和联系第三方；

14）机房资产管理系统（CMDB）。

15）机房环境。清理机房的杂物，将机房物品定置。清洁机房门窗、地面。定期清洁电池室的地面；检查机房所有与外界的空洞是否已严密封堵，严密防鼠；检查机房玻璃、地板、天花板、通气口，墙体表面是否正常，外观是否完好，有否出现老化现象。检查机房是否有漏水现象。检查机房墙壁是否有渗水现象。填写巡检记录，有问题及时报告。

16）巡视电池间；检查电池工作状态。

17）确认机房照明良好，出现问题及时报告。

18）视频网络播放系统。定期检查可用性，有问题及时与专业第三方公司联系解决。

19）填写巡检记录。

2.2 网络运维管理

针对数据中心的网络部分，运维内容主要包含以下内容：

1）测试网络接入速度，监控网络访问可用性和访问质量，出现问题第一时间直接联系接入商解决。

2）网络接入商变化时，配合网络接入商对网络变更方案的可行性审查、问题审查。配合网络接入商更替施工。

3）局域网。本地局域网日常管理和维护；VLAN 划分；网络性能优化；故障排除；网络节点周期性检查，发现潜在问题，并解决。

4）无线局域网。负责无线局域网的日常管理和维护；客户端不能正常接入网络的故障排除；网络性能优化；故障排除；网络节点周期性检查，发现潜在问题并解决。

5）远程接入。制定 VPN 使用策略，实施 VPN 用户日常远程接入服务器的管理，以及性能优化和故障排除等。

6）网络病毒查杀和网络安全保护。

7）根据实际项目或安排而产生的其他工作。

2.3 服务器和存储运维管理

2.3.1 服务器运行情况及性能监测

数据中心运维团队将通过综合监控系统实施 7*24 小时平台设备监控，发现告警，并进行处理，解决问题。对系统运行进行实时检查。对监控或维护中发现的问题及时处理，消除隐患，保障平台的稳定运行。并且还提供针对各服务器物理资源的使用情况和操作系统的运行情况、进行实时监控，提供服务器安全监测报告。

主机性能监控的检查列表包括：

- CPU 利用率
- 内存使用情况
- 交换区使用情况
- 磁盘 I/O 情况
- 关键文件系统的状态
- 重要进程的运行情况（例程数量、消耗 CPU、占用内存）
- 操作系统的各类日志文件
- 网络、端口信息

运维团队需根据检查列表进行日常检查，并不断地改进日常检查列表，以满足对系统监控的需要。

2.3.2 服务器软硬件兼容性检查

数据中心运维团队在维护系统稳定运行的同时，需主动收集系统关键补丁、软件补丁、硬件微码等信息，在通过数据中心专家评审的前提下，对相关设备进行升级服务，并在升级完成后配合应用方对系统进行测试。升级前后需要和应用方及时做好沟通确认工作，确保不会产生兼容性导致的故障。

2.3.3 磁盘阵列设备管理

运维团队需要对磁盘阵列设备及其相关的部件(如硬盘、控制器等)进行编号，并记录在案，对软件设置中的参数也要进行详细的记录，并在每次变更后及时更新相关的信息。

除此之外，运维团队定期（暂定每半年）对于每个服务器的系统容量监测的审核，并制定相应的容量规划，主要监测文件系统的空间、数据库的空间资源利用情况，分析资源利用趋势，并提供资源情况报表。

文件系统空间管理

- 定期检查文件系统的空间使用情况，根据业务发展需求和新业务的增加，制定合理的空间分配方案，新增、修改或删除空间。
- 对文件系统空间的使用进行监控，发现空间使用不合理或需要清理的协调解决。

数据库空间管理

- 应实时监测数据存储空间的使用情况，根据业务数据的数据量、数据结构以及增长速度，制定合适的数据存储和结构优化策略，动态增加新的空间以存放业务数据；

- 定期检查数据存储空间的使用情况，根据实际情况规划增加新的空间，填写数据库空间新增/修改/删除申请表，经审核后实施，并更新数据库配置状况记录表。

2.3.4 机柜、电源、网线布局管理

运维团队对于新上架安装的设备，需要进行拍照留档，确认各线路位置，并对服务器的电源部分进行编号整理，最终登记在册。

2.3.5 协助第三方维护

对于由专业第三方提供运维的设备，设备出现问题后运维团队需及时通知第三方并告知采购人，视情况严重性，决定是否启动应急预案；配合第三方服务商一起排查和解决问题，实施为了解决故障而进行的系统软硬件的补丁、升级及维护工作。独立处理初级系统故障，与第三方厂商或服务商配合解决高级别系统故障。记录问题、故障的解决办法及解决过程。做出临时的配置变更以排除故障，在必要的时候，提出永久性配置变更建议。

2.4 基础软件运维管理

2.4.1 操作系统

运维团队充分保障服务器操作系统的稳定运行，将提供以下服务内容：

1）系统升级

运维团队在维护系统稳定运行的同时，需主动收集系统关键补丁、软件补丁等信息，在通过数据中心专家评审的前提下，对相关系统进行升级服务，并在升级完成后配合应用方对系统进行测试。升级前后需要和应用方及时做好沟通确认工作，确保不会产生兼容性导致的故障。

2）操作系统稳定性监控定时查看操作系统日志及 IIS 日志，查看 CPU、内存占用率，排除故障。

3）权限与文件管理

服务器应明确责任人及管理帐号持有人，不应出现多人单帐户，单人多帐户的情况，不利于在服务器出现问题后，对服务器进行操作维护、查找问题。

4）定期检查磁盘空间

进行磁盘文件排列的优化和错误扫描，并处理错误；安全地删除系统各路径下存放的临时文件、无用文件、备份文件等等，完全释放磁盘空间。

5）维护系统注册表。

6）系统配置。优化系统配置，关闭无用服务和端口，以最适合系统运行方式，最小化安装等。维护系统配置文档。

7）负责系统用户管理，如增加、删除用户、重置用户密码、管理用户权限等。进行系统用户管理时，记录所有相关的系统变更。

8）对于新安装的服务器，运维团队应负责安装必要的应用软件：如远程监控工具、备份工具、防病毒软件等。

2.4.2 数据库

运维团队将对数据进行日常维护，在数据库性能监控的检查列表包括：

- 资源使用情况
- 运行情况
- 数据库进程状态
- 数据库连接状态
- 数据库进程使用资源

- 数据库的表空间（数据表空间、索引空间、临时表空间等等）使用情况；
- 数据库日志空间
- 回滚段使用情况
- 数据库锁的数量
- 死锁的发生、死锁资源
- 数据库碎片的数量
- 磁盘 I/O
- 数据库运行日志
- 数据库用户登录情况
- 监控结果应做登记管理，如实记录系统日常运行状况及异常情况，填写日常运行情况记录表；
- ……

除此之外，数据库的运维工作还包含一些其他工作，如：

1）数据库备份和恢复

2）做好备份计划，工程师定时完成，因备份占用内存较大，在访问量大的情况下进行。当出现数据问题时，向采购人管理部门通报，说明数据情况，后恢复。

3）访问性能优化及数据库同步

4）服务器管理人员需记录详细的设置；数据库如需要同步，应明确同步时间或实时同步等方式。

5）数据库日志和表空间，定期进行整理，问题解决。

2.4.3 中间件

运维团队针对中间件的运维工作，内容如下：

1) Oracle Weblogic，辅助开发公司进行配置，保留配置文档。模块配置与更新，配合第三方配置.java 及 wls 的版本及更新工作。操作系统模块配置与更新，配合第三方配置操作系统到可用的版本及更新。配合反馈第三方解决服务错误日志中的问题。

2）新软件安装，收集安装光盘、安装合同（可复印学习）、使用说明书、授权书（Liscense）。纸质版文件扫描后入库，电子版文件进入配置库。

2.4.4 备份系统

为保证在系统崩溃或停止运行时能尽快恢复系统，将制定相关的数据备份制度。应针对不同系统制定备份方案，应包括备份方法、频率等。数据备份包括定期和不定期备份。重要数据应每月进行全备份和增量备份；不定期备份应该在数据变更后立即进行，更新前的备份按需要保存一定时间。

2.4.5 应用系统

当前的应用系统及相关的开发工作由第三方公司负责，运维团队主要起配合作用，相关的工作内容如下：

1）当应用出现问题，及时联系第三方解决，并做问题记录。

2）配合第三方进行操作系统、数据库和中间件的系统配置，并做配置记录，在有授权运维的系统中，熟悉应用系统维护方法。

3）配合第三方新应用系统上线，需收集安装文件，源代码，部署文档、运维文档。扫描后，入配置库。与合同库相关联，记录维护期间联系人，原公司质保期。

4）每日上班后、下班前检查可用性，确认无灾难性问题、黑客篡改问题。

5）其他待完成工作，根据实际情况来处理。

第三章 运维工作内容

3.1 日常维护工作

运维团队的值班安排分三班，保持 7x24 小时的人员安排，在任何时间数据中心都由值班人员。运维团队根据数据中心的运维管理制度，通过主动性、预防性维护，执行日常维护作业计划，对告警、性能、运行状态进行检查分析，及时进行数据备份，并定期对备份数据进行恢复性测试验证，对系统运行质量进行分析，并进行维护记录。对监控或维护中发现的问题及时处理，消除隐患，保障平台的稳定运行。

3.2 系统性能监控管理

运维团队通过综合监控系统等实施 7*24a 小时平台设备监控，发现告警，并进行处理，解决问题。使用综合监控系统对系统运行进行实时检查。对监控或维护中发现的问题及时处理，消除隐患，保障平台的稳定运行。

3.3 系统维护管理

故障处理

运维团队负责故障发现、故障分析、故障处理工作，在规定时间内，处理完成故障，同时负责调查故障原因，最后编写详细的《故障报告》，包括故障发生的起止时间、原因、现象、处理过程、处理结果和处理经验。如果故障设备或组件为第三方维保，值班工程师负责和第三方对接，迅速解决问题。

软件和补丁维护

操作系统级别的软件和补丁服务

- 运维团队对于维保设备提供所有软件补丁，提供预警服务，对于软件的维护版本提供补丁，并按稳定性和安全性的要求，提供是否升级的建议，评估风险和制作实施方案。
- 故障经工程师的分析表明它是由一个软件错误所引起的，那么运维团队需提供相应的软件版本和补丁。
- 对于软件版本和补丁的安装，运维团队首先将确认是否可以在对应平台上进行装载。若确认可实施，运维团队则将提供补丁升级服务，升级前要配合相关应用方做好测试。

应急预案及演练

为加强风险管理意识，提高应急预案相关人员的应急处置能力，及时发现应急预案可能存在的问题，确保在紧急情况下，应急预案能够真正发挥作用，需要通过周期性的演习演练来不断检验应急体系应急预案的可靠性、有效性和可操作性。

应急预案的演习演练方式、演习演练频度等内容明确如下：

1、演练分为桌面演练和实战演练两种方式，每次演练都应该有相关技术人员全程参与。

2、定期桌面演练，定期实战演练；

3、每次演练结束之后应进行分析和总结，及时完成应急预案的更新、优化和完善。

协助第三方维护

在服务期内，运维团队将配合第三方或服务商进行系统的升级、替换、新部件（模块）安装等，并在实施完成后确认工作正常。

备份

为保证在系统崩溃或停止运行时能尽快恢复系统，将制定相关的数据备份制度。应针对不同系统制定备份方案，应包括备份方法、频率等。数据备份包括定期和不定期备份。重要数据应每月进行全备份和增量备份；不定期备份应该在数据变更后立即进行，更新前的备份按需要保存一定时间。

系统优化

对于巡检或日常维护过程中发现的系统隐患或系统不是处于满意状态，提供相关系统优化的报告。

对于运行情况跟踪，预防性诊断设备存在的隐患，提供系统优化建议，提供系统规范和流程的建议，提供系统优化概要。

硬件设备统计

运维团队将定期对参保设备进行统计。

质量分析报告

运维团队建立数据中心平台的质量分析报告。每月汇总设备运行质量、系统性能等指标，进行数据中心平台运行质量分析，排除质量隐患，不断提高网络运行质量和服务质量。

运维工程师应每周和每月对于数据中心在网系统运行情况作分析，数据采集、统计和分析系统设备的运行数据，形成系统运行周报和月报。

分析报告，包括优化设备运行的绩效，提高系统稳定性的建议，对于系统扩容和优化投资的建议，提供系统运行情况概要，系统中关键设备的运行情况分析，并能识别和解决潜在问题，做好预警，制定并实施相应的优化措施，并对于系统的扩容和项目投资提供建议报告。

3.4 系统配置与支持维护

运维团队的日常工作中，在系统配置和支持方面的工作内容如下：

- 维护系统软硬件配置文档；
- 负责系统用户管理，如增加、删除用户、重置用户密码、管理用户权限等；
- 进行系统用户管理时必须遵循数据中心的账户命名规则及账户密码策略，并文档记录所有相关的系统变更；
- 每月提交系统账户变更月报；
- 配合第三方进行升级、安装系统，及时更新操作系统补丁，进行系统软件备份；
- 根据运维报告及统计报表，每月制定维护作业计划，并提交日常维护报告；

3.5 系统容量管理

运维团队至少每半年进行一次对于每个服务器的系统容量监测的审核，并制定相应的容量规划，主要监测文件系统的空间、数据库的空间资源利用情况，分析资源利用趋势，并提供资源情况月报表。

文件系统空间管理

- 定期检查文件系统的空间使用情况，根据业务发展需求和新业务的增加，制定合理的空间分配方案，新增、修改或删除空间。
- 对文件系统空间的使用进行监控，发现空间使用不合理或需要清理的协调解决。

数据库空间管理

- 应实时监测数据存储空间的使用情况，根据业务数据的数据量、数据结构以及增长速度，制定合适的数据存储和结构优化策略，动态增加新的空间以存放业务数据；
- 定期检查数据存储空间的使用情况，根据实际情况规划增加新的空间，填写数据库空间新增/修改/删除申请表，经审核后实施，并更新数据库配置状况记录表；

3.6 巡检工作

除了依靠数据中心的监控软件，还要求运维团队对服务器、存储、操作系统、数据库、中间件等基础设施进行巡检，并编写巡检报告。通过巡检可以对当前系统的运行状况有一个详细的了解，对巡检中发现的问题可以及时采取预防性措施，降低故障发生的概率，提高系统的可靠性。

巡检工作需要检查以下几个方面：

- 场地环境检查：包括机房的温度、湿度、通风及 UPS 工作状态等的检测；
- 操作系统：检查补丁完整性，记录软件版本，以保证系统发挥最佳性能；
- 外设检查：对网卡或 HBA 卡、磁盘驱动器的读写、磁带机的读写进行检测；
- 网络设备检查：运行环境检查、LED 控制面板、IOS 版本信息、进程状态、内存利用率、接口状态、路由表状态、网络连通性测试；
- 设备清洁：对相关设备进行维护保洁工作，使设备保持良好的运行状态；
- 系统日志检查；
- 文件系统检查、清理；

- 系统配置检查；
- 系统和数据备份检查；
- 系统运行情况分析；
- 系统总体性能评估。

1.机房环境日常检查内容

机房环境服务是为机房设备如小型机、网络设备和存储设备等提供一个安全可靠的物理环境，确保机房设备不会因为环境因素导致不能正常运行或损坏。为了达到此目的，机房环境需具备以下标准：

- 确保机房温度在 24+2℃之间，最大温度变化率不超过 10℃/小时；
- 确保机房湿度在 50+5%之间；
- 确保机房电压在 220V+5%之间，电压频率在 50.5~49.5 之间，瞬间变动电压不超过 220V+/-15%，总谐波不高于 5%；
- 机房电源地线方面确保机房接地线与任何导线完全隔离及绝缘，接地线线径至少为 3.5mm，系统接地电阻在电源插座连线与地线间不大于 2 欧姆，在电源输出座连线与地线间电压小于 1V，在接地线的接地端测的接地电阻不大于 1 欧姆；
- 确保机房为网络设备、空调、视频等提供独立的冗余双电源供应系统，杜绝电源公用现象，确保网络设备电源无隐患；
- 确保机房整洁干净，避免机房在阳光直射之下；
- 确保机房无线电杂波干扰低于 0.5V/米；

2.服务器、存储、操作系统、数据库、中间件巡检及巡检报告内容

针对服务器、存储、操作系统、数据库、中间件等比较重要的组件，数据中心制定了按月巡检的计划，需要按照巡检报告的模板进行检查，巡检报告要涵盖以下内容：

	巡检报告内容
故障检修	✓ 故障现象 ✓ 故障分析、定位 ✓ 故障处理 ✓ 故障原因
巡检保养	✓ 巡检内容 ✓ 巡检发现的问题或隐患 ✓ 隐患产生的后果分析 ✓ 隐患的如何处理

3.7 定期服务报告

系统维护档案，详细记录数据中心相关的设备信息和项目管理信息。在日常运维中，服务报告和技术文档由运维团队的相关人员负责维护和更新。

系统维护档案将分为以下四个部分：

3.7.1 设备配置档案

- 维护设备及软件清单、系统功能、详细配置信息及软件版本和设备 PN 号；
- 设备位置、网络拓扑、设备连接拓扑及各种工程图纸；
- 如果系统发生变更，如实施软件、补丁、微码升级或业务调整，同步更新配置档案；

- 系统双机、备份设置和运行情况。

3.7.2 服务文档

- 技术参数的配置文档；
- 处理故障时的《故障处理报告》；
- 每季度的《季度运维总结》；
- 每次重大故障处理后发布《重大问题分析报告》；
- 共享维护内容及其他技术资源整理知识库；
- 每次巡检时的《巡检报告》；
- 微码更新、性能分析及优化、机房搬迁等服务实施方案、专业服务报告和技术建议等。

3.7.3 服务总结

运维团队根据自身的工作内容，在每季度需要对自己的工作进行汇总，并生成《季度运维总结》。

报告中的具体内容包括：

- 故障处理及备件更换情况汇总；
- 设备状况分析及评价；
- 人员出勤情况，工作量，或资源使用情况，包括第三方供应商服务情况；
- 重大事件和变更情况；
- 配置管理相关信息；
- 趋势信息；
- 下一步工作计划；

3.8 运行维护优化评估

（1） 建立基于数据中心的基础运维服务管理框架体系及运维团队，根据网络的现状提出整体安全规划，包括日常维护计划、安全风险控制计划、应急响应计划等

（2） 提供风险评估、灾难恢复、应急响应、安全培训服务并提供报告

（3） 安全检测

每季度定期对服务范围内的对网络设备、服务器操作系统、数据库系统、应用软件系统的安全策略和安全配置进行检查和测试，从中获得相关的信息、发现系统面临的威胁以及存在的安全性。

（4） 安全评估。

每季度对服务范围内的整体网络系统进行全面、统一的系统性的安全风险评估，识别和控制网络中的关键资产及可能会产生的安全风险，并对所发现的问题提供优化、改进建议。并根据评估的结果为关键资产建立应急响应预案以及细微调整其后安全维护服务所要监控的内容。

（5） 策略优化

根据安全评估的结果每半年对系统策略及网络系统进行优化设计，制定调整系统策略优化、网络拓扑优化、安全域规划与配置、IP 规划、VLAN 优化等策略，并根据实际情况调整与实施。

（6） 应急预案与演练

根据数据中心的现状，模拟实际灾难发生场景，提供各种应急预案，经过采购人讨论，协助采购人实施演练。

（7） 培训

运维服务期内，安排以运维管理、安全为主题的培训，数量为 4~5 人次，按要求制定相应的培训计划。

（8） 资料收集存档

参与机房运维涉及的专业第三方机构合同的起草、谈判，与采购人一起对第三方机构进行管理。整理收集涉及到的第三方合同，中间文档、过程记录，备查，按照采购人规定进行提交。

3.9 应急保障措施和组织

3.9.1 应急响应系统

运维团队在处理紧急情况和重大事项时，会启用应急指挥系统：

接口人：应用系统下，各个相关方的固定接口人，一般为项目经理

运维团队：事故发生期间提供直接的技术咨询、指导服务，负责直接处理故障。

二线专家：严重事件由承保的第三方服务商或原厂商的二线专家最快速度到达现场处理事故。

3.9.2 应急响应过程

应急响应过程划分为四个主要阶段：应急准备、监测与预警、应急处置措施和总结改进。

a）应急准备阶段的工作包括：组建应急响应组织，确定应急响应制度，系统性识别运行维护服务对象及运行维护活动中可能出现的风险，定义应急事件级别，制定预案，开展培训和演练；

b）监测与预警阶段的工作包括：进行日常监测，及时发现应急事件并有效预警，进行核实和评估，以规定的策略和程序启动预案，并保持对应急事件的跟踪；

c）应急处置阶段的工作包括：采取必要的应急调度手段，基于预案开展故障排查与诊断，对故障进行有效、快速的处理与系统恢复，及时通报应急事件，提供持续性服务保障，进行结果评价，关闭事件；

d）总结改进阶段的工作包括：对应急事件发生原因、处理过程和结果进行总结分析，持续改进应急工作，完善信息系统。

3.9.3 制定应急保障预案及演练

为了应对业务系统可能出现的紧急故障，运维团队将定期模拟故障演练服务。运维团队有一套整体的应急方案，以确保数据中心在系统发生突发事件或灾难情况下能够迅速恢复 IT 服务，从而保证系统业务的持续运行。根据普遍认可的最佳实践指导原则，IT 应急和 IT 灾难恢复的定义应该是：

“计算机系统灾难是指任何造成计算机系统不能处理业务的时间超过了可容忍程度的事故。应急方案是指计算机系统灾难发生后，按照既定的应急恢复方案在一定时间内恢复系统运行和业务处理的过程。”

为了应对生产系统可能出现的紧急故障（重大、严重故障），数据中心将从事前预防和事后处理两个方面制定紧急故障应处理预案。

（A）事前预防：

- 应急涉及到多个层面的配合，每方都需要指定专人负责在紧急故障发生时及时沟通
- 数据中心专家支持团队进行系统风险评估，提出系统整改建议，制定紧急故障应急处理预案
- 进行一定次数的实际演练，包括后备系统切换测试、备份数据还原测试
- 对流程进行持续性跟踪，系统出现变更后，重新评估流程的有效性

(B) 事后处理:

- 响应时间：由工程师立即做出响应
- 故障修复：由经验丰富的专家支持团队提供专人支持，包括搭建测试环境、远程和现场故障诊断和排除；同时启动紧急故障处理流程，按既定程序做应急处理

序号	内容	服务规范	补充说明
1	服务范围	为生产系统或其它关键业务系统制定紧急故障应急处理预案，并对预案进行持续性改进	
2	服务时间	紧急故障预案制定：双方协商	
		紧急故障处理：全年 7×24 小时	
3	服务方式	远程或现场	
4	服务发起	由数据中心负责人提出服务请求	
5	紧急故障应急处理流程涵盖范围	软件介质、安装文档、系统配置文档完备，并由双方专人保管，随时可以查阅	
		双机配置正确，处于自动切换状态	
		数据备份系统可靠运行，数据已得到安全备份，并有完善的数据恢复流程文档	
		接受服务请求，在无法处理的情况下立即转交专家支持团队	
		制定远程专家故障诊断和修复流程，专家支持工程师远程尽快排除系统故障	
6	实际演练	根据要求，组织一定次数的实际演练	
		演练内容包括：	
		服务流程演练：定期执行流程运转演练，保障流程的顺畅运行	
		容灾切换演练：定期执行双机热备切换演练，测试双机冗余的可用性	
		业务替代演练：定期进行冷备机启动业务替代演练，备份系统启动演练，测试灾难情况下冷备系统的可用性	
		备份演练：定期在测试机上演练灾难情况下的备份恢复，测试灾难情况下备份恢复的可用性	

应急演练：

应急演练计划至少每季度一次进行测试和演练，以保证：

- 计划内容能够反映当前的状况；
- 计划的有效性和可操作性；
- 应急演练人员熟悉应急恢复流程。

所有测试和演练的结果应当依据事先确定好的标准，来判断测试和演练是否成功。如：多长时间恢复服务，会出现多少问题，及问题的严重性等。在测试完成后应记录下结果，并根据需要对应急恢复计划进行修订。针对演练或测试过程中出现的问题和失败应该进行说明并体现在相应的改进计划中。

3.10 IT 运维服务工具

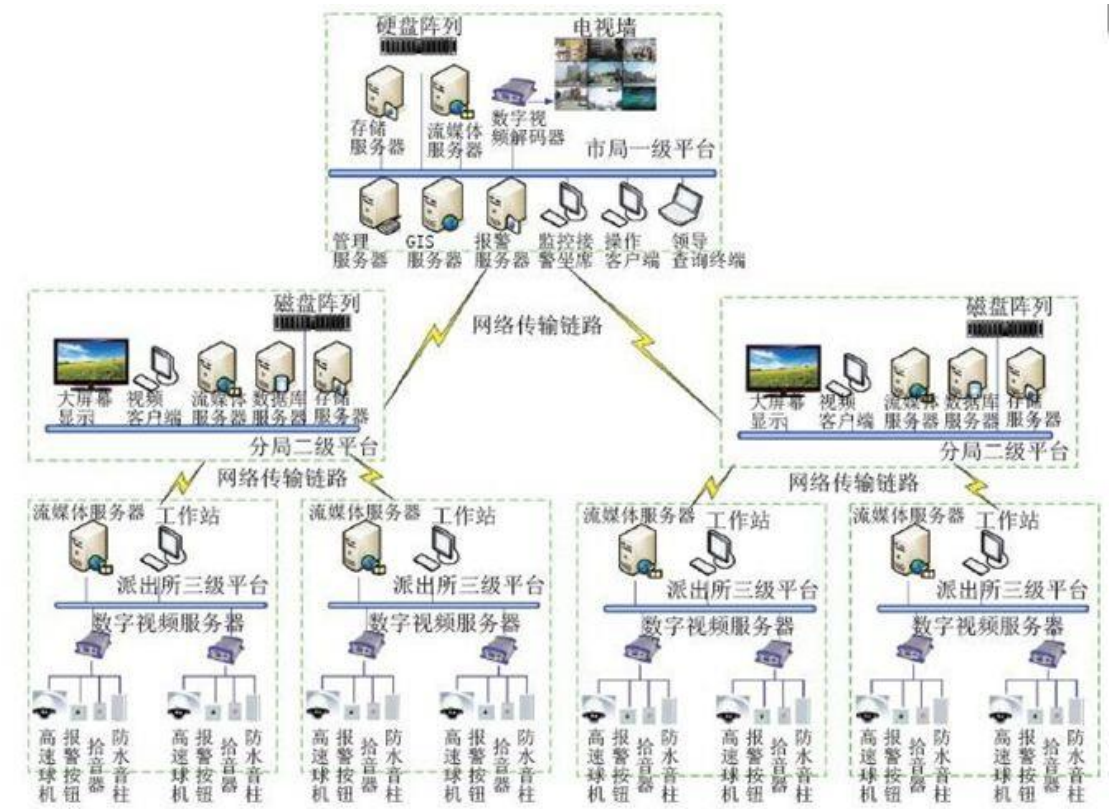
3.10.1 运维监控平台

运维服务事件管理系统是支撑运维管理组织中各运维角色按照规定的运维事件流程开展运维活动的信息化系统。一方面，该系统要支持运维服务提供者对运维服务事件管理对象进行管理，以实现运维服务的能力；另一方面，要支持运维服务提供者按照商定的服务级别协议方便地向运维服务使用者提供运维服务；同时，要支持运维服务管理者对整个运维服务事件的考核、监督和评估。运维服务事件管理工具是构成运行管理体系不可缺少的元素，从被动管理向主动管理转化的重要部分，为整个运行管理体系的高效实施奠定了基础。

监控拓扑

当前数据中心采用了某运维监控平台，对数据中心设备进行监测。用户通过客户端登录运维监控平台，查看所有被监控设备的运行情况。当前监控平台支持

机房环境、网络设备、存储设备、服务器设备、系统和数据库等组件的监控，支持故障预警等服务。



主机监控

为确保数据中心服务器高速、稳定运转，运维监控平台从多个方面对主机服务器的硬件设备及操作系统进行监控管理和性能管理。它通过采集服务器的 CPU、内存、硬盘、网卡等硬件的关键运行参数，以及软件和应用程序的进程、服务、端口等的运行状况，对系统日志进行分类扫描查询。通过数据采集和分析，运维监控平台能够及时对影响用户服务器运行性能的故障事件发送报警，并采取相应的故障处理措施，保证服务器的正常安全运行。

Windows 服务器监控

运维监控平台对服务器的监控支持 Agent 代理、SNMP 和 WMI 非代理三大方式，方便不同用户对服务器全面监控的需求。运维监控平台服务器主要监测指标如下

Ping	包成功率(%)
	数据往返时间(ms)
	状态值(200 表示成功 300 表示出错)
CPU	CPU 使用率(%)
	列速度/次数每秒(次数)
	处理器数量 (个数)
虚拟内存	内存使用率（%）
	可用内存量（M）
	总内存空间（M）
物理内存	内存使用率 （%）
	内存可用大小（M）
	总内存量（M）
Windows 日志监测	检查行数
	过滤行数

Top-5CPU 使用	CPU 使用率（%）
	cpuTop1-Name()
	cpuTop1-使用率（%）
	cpuTop2-Name()
	cpuTop2-使用率（%）
	cpuTop3-Name()
	cpuTop3-使用率(%)
	cpuTop4-Name()
	cpuTop4-使用率（%）
	cpuTop5-Name()
	cpuTop5-使用率（%）
Top-5 虚拟内存使用	内存总使用量（M）
	MemTop1-进程名（）
	MemTop1-内存使用（M）
	MemTop2-进程名（）
	MemTop2-内存使用（M）
	MemTop3-进程名（）
	MemTop3-内存使用（M）
	MemTop4-进程名（）
	MemTop4-内存使用（M）
	MemTop5-进程名（）

Top5-物理内存使用	内存总是用量（M）
	MemTop1-进程名（）
	MemTop1-物理内存使用（M）
	MemTop2-进程名（）
	MemTop2-物理内存使用（M）
	MemTop3-进程名（）
	MemTop3-物理内存使用（M）
	MemTop4-进程名（）
	MemTop4-物理内存使用（M）
	MemTop5-进程名（）

	MemTop5-物理内存使用（M）
磁盘指标	Disk 使用率(%)
	磁盘总量(MB)
	剩余空间(MB)
监测 Windows 账户是否被修改	账户个数（）
	系统账户（）
	匹配状态（）
Windows 进程	进程总数(个)
Windows 服务	服务总数（个）
网卡流量	接收流量(Kbit/s)
	接受流量百分比（%）
	发送流量(Kbit/s)
	发送流量百分比（%）
	接收包数（packets/s）
	发送包数（packets/s）
端口	平均响应时间(ms)
CPU 指标	CPU 使用率(%)

Linux 服务器监控

运维监控平台对 Linux 服务器的监控支持 Agent 代理、SNMP 和 SSH、Telnet 非代理三大方式，方便不同用户对服务器全面监控的需求。运维监控平台 Linux 服务器主要监测指标如下

Ping	服务成功率（%）
	平均响应时间（%）
CPU 指标	CPU 使用率(%)
磁盘指标	Disk 使用率(%)
	剩余空间(MB)
交换分区	内存使用率（%）
	可用内存量（M）
	总内存量（M）
内存指标	Memory 使用率(%)
	剩余空间(MB)
	错误页/秒(页/秒)
	内存总量(MB)
监测 Linux 事件日志	检查的总行数(行)
	匹配行数(行)
Linux 进程监测	运行实例个数(个)
端口	平均响应时间(ms)

Top-5CPU 使用	CPU 使用率（%）
	cpuTop1-Name()
	cpuTop1-使用率（%）
	cpuTop2-Name()
	cpuTop2-使用率（%）
	cpuTop3-Name()
	cpuTop3-使用率(%)
	cpuTop4-Name()
	cpuTop4-使用率（%）
	cpuTop5-Name()
	cpuTop5-使用率（%）

Top-5 虚拟内存使用（）	内存总使用量（M）
	MemTop1-进程名（）
	MemTop1-内存使用（M）
	MemTop2-进程名（）
	MemTop2-内存使用（M）
	MemTop3-进程名（）
	MemTop3-内存使用（M）
	MemTop4-进程名（）
	MemTop4-内存使用（M）
	MemTop5-进程名（）
	MemTop5-内存使用（M）

Top5-物理内存使用	内存总是用量（M）
	MemTop1-进程名（）
	MemTop1-物理内存使用（M）
	MemTop2-进程名（）
	MemTop2-物理内存使用（M）
	MemTop3-进程名（）
	MemTop3-物理内存使用（M）
	MemTop4-进程名（）
	MemTop4-物理内存使用（M）
	MemTop5-进程名（）
	MemTop5-物理内存使用（M）

网络设备监控

运维监控平台可以从各个方面对数据中心的网络设备进行监测和管理，内容包括网络设备的可用性、设备性能、流量管理等等。运维监控平台的网络设备管理系统支持的网络设备，包括各种类型的交换机、路由器、防火墙、VoIP 网关设备和其他启用了 SNMP 协议的网络设备。

运维监控平台监测对象主要包括网络设备（路由器、交换机、防火墙）的状态，如端口，路由器 CPU 负载等，支持 Cisco、华为、港湾、Juniper 等各主流厂家的路由器、交换机，支持 Netscreen、Cisco、天融信等主流厂商的防火墙等网络安全设备。

- 网络设备监控

CPU	cpu 使用率（5 秒）（%）
内存	内存使用率（%）
Ping	服务成功率（%）
	平均响应时间（ms）
设备运行时间	运行时间（）
接口信息	接收流量（kbit/s）
	发送流量（kbit/s）
	每秒发送数据包（个/s）
	每秒接收数据包（个/s）
	发送丢包率（%）
	接收丢包率（%）
	发送错误率（%）
	接收错误率（%）
	带宽使用率（%）
	广播包（包/秒）

- 安全设备监控

不同类型设备，所监控的内容会有不同。

CPU 状态	CPU 利用率(%)
Memory 状态	分配内存数(bytes)
	剩余内存数(bytes)
Ping	服务成功率(%)
	平均响应时间(ms)
	状态值(200 表示成功 300 表示出错)
Session	活动会话数（）
	分配会话数（）
	失败会话数（）
Firewall	攻击包个数（sync attack）
	攻击包个数（tear drop attack）（）
	攻击包个数（Source Route Option Attack）（）
	攻击包个数(Ping of Death Attack)()
	攻击包个数(Address spoofing Attack)()
	攻击包个数(Land attack)
	攻击包个数(Icmp flood Attack)
	攻击包个数(Udp flood Attack)
	攻击包个数（weired netbios attack）()
	攻击包个数(Port Scan Attempt Attack)
	攻击包个数(Address Sweep Attempt Attack)

应用监控

运维监控平台的应用监测模块可以全面智能的监测用户各种与应用相关的服务。运维监控平台对各种数据库、中间件和 WEB 从应用可用性、系统资源占用和性能指标三个方面提供全面的监测管理策略，确保应用的运行正常。

- Oracle 监控

数据库用户连接数	当前用户连接数
数据库性能	游标数
	Session 数
	每秒交易数
	数据库锁数
	死锁数
	缓冲池命中率
	库 cache 命中率
	数据库响应时间
表空间	使用率
	已用量
	剩余率
	剩余量
	总大小
进程监控	进程的内存利用

- MS-SQL 监控

Buffer 状况	Buffer Cache 命中率
	每秒 Lazywrites 数
	每秒数据库页读取数
	数据库页写入数
	每秒刷新到磁盘的页数
	缓冲区池停留秒数
内存统计	使用内存总量
	连接内存
	内存授权进程数
	优化内存数
	SQL 缓存内存数
用户统计	用户连接数
	每秒登录数
	每秒注销数
Cache 统计	缓存命中率
	缓存使用页数
	缓存对象数
	每秒缓存使用数
请求统计	每秒请求批数
	每秒自动参数尝试数
	每秒的 SQL 编译数
	每秒 SQL 重新编译数

监测器参数设置

监测平台中所有监测器，都可以设置重试次数、超时等。

- 监测器间隔：5 秒 至 指定小时，如每 10 秒监测一次，或每 5 小时监测一次；
- 监测器工作计划：可以设置 7X24 或 5X8 工作时间；
- 错误后重试：任意重试次数，但建议不超过 99；
- 错误频率：监测器发生错误后，调整监测器的监测间隔，如 CPU 监测器原监测间隔为 10 分钟一次，发生错误后，监测间隔调整为 1 分钟一次；
- 故障处理记录：针对监测器，记录故障处理的内容；
- 阀值设置：最多可以设置 8 个阀值检测条件，每个阀值检测条件之间可以用：并、或的关系。

拓扑管理

自动发现能够自动识别设备类型，包括各种服务器类型、路由器、交换机、等等，以及它们之间的关系，并且自动将它们存储到公用对象库中对应的类中。

- 故障告警管理

  运维监控平台故障管理系统是管理数据中心的设备、网络和业务所出现的故障；帮助网管人员采集、统计和分析来自网络各方面的报警信息和故障信息，准确预警、定位和解决网络中的故障。

- 故障告警方式

  运维监控平台提供短信息、语音、声音、远程声音、邮件、脚本等多种方式及时发出警报。可以及时通企业的网管人员发现、定位和处理故障，

让系统的管理从被动变为主动，可有效地预防故障发生，也可在故障发生时快速进行定位，及时处理好故障。

- 报警控制台

  关于对警报和故障的管理，运维监控平台主要通过报警控制台来进行。运维监控平台报警控制台包括四个方面：配置文件及接口数据、故障事件搜集、故障事件过滤、告警呈现。

用户权限设计

运维监控平台支持精细的用户分级管理功能，用户按照权限分为超级管理员和一般管理员两类：超级管理员具备全部管理功能，可以为一般管理员配置不同的用户名、密码和权限；一般管理员具备部分管理功能（例如只读）。对一般管理员的功能限制主要从两方面来进行，一方面是管理对象权限设置，另一方面是管理功能权限设置，对于一般管理员的管理对象权限设置可以精确到对任意管理对象和管理对象权限的自由组合。