

集中式数据库上线部署

莫听穿林打叶声，何妨吟啸且徐行。竹杖芒鞋轻胜马，谁怕？一蓑烟雨任平生。料峭春风吹酒醒，微冷，山头斜照却相迎。回首向来萧瑟处，归去，也无风雨也无晴。

苏轼的这首《定风波·莫听穿林打叶声》，是其担任黄州团练副使时，与朋友春日出游过程中，风雨忽至，朋友深感狼狈，词人却毫不在乎，吟咏自若。表达了词人在逆境中始终保持着乐观豁达的人生态度。

在当今风起云涌、英雄辈出的数据库江湖中，能一直不惧风雨砥砺前行非集中式数据库莫属了。自其问世以来，始终高效支撑着商业银行重要系统的稳健运行。尤其是上线实施涉及面很广，需要开发中心和数据中心共同协作才能完成。

很多商业银行在部署数据库时，主要面临这样的难题：数据库安装配置不标准、监控指标不全面、巡检体系不配套等，从而给数据库的运维带来很大压力。产生该问题的根源在于规范缺失或者规范没有得到有效执行。实际上，数据库的部署规范和运维规范是数据库生产运维的基础，只有制定了高质量的规范，并且确保规范能得到有效执行，才能真正确保数据库上线的标准化。

通常商业银行新系统上线基于相关部署规范和运维规范，按照以下步骤进行。

首先，开发中心项目团队提出部署需求，数据中心新系统上线团队收到需求后，开始制订部署方案草案；新系统上线团队接着邀请各个岗位人员对草案进行评审，这些岗位人员包括开发中心项目团队、数据中心数据库岗、备份岗、网络岗、存储岗等人员；草案经多次讨论和修改，最终才能形成正式的部署方案。

接下来，数据中心新系统上线团队按照部署方案完成数据库安装配置、监控部署、巡检脚本部署、应急脚本部署、集中备份部署等工作。

最后，部署好的数据库环境将移交给数据中心应用运维保障部门，正式进入实际生产运行阶段。

本章选择最常用的集中式数据库 Db2 和 MySQL，并以某商业银行网络贷款系统作为案例进行深入讲解。

4.1 部署需求

开发中心应用项目团队在提出部署需求时，需要提供待上线系统的逻辑架构图。基于逻辑架构，接下来对功能性需求、非功能性需求、工作负载需求和灾备需求等进行逐项深入分

析，最后形成正式的需求说明书。

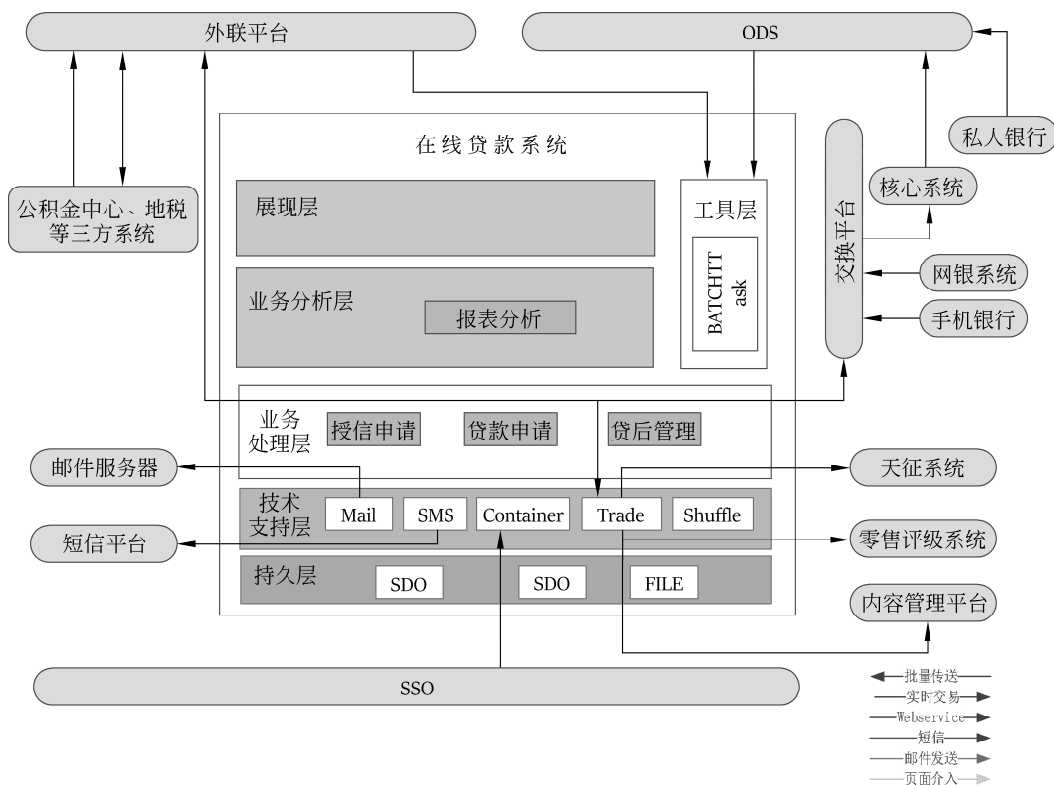
4.1.1 逻辑架构

在待上线系统的逻辑架构图中，需要说明该系统属于哪类业务系统、包括哪些子系统、该系统与外联系统的关系等，具体来说，需要明确以下问题。

- (1) 各子系统功能。
- (2) 本系统和关联系统的关系，特别是要梳理清楚系统之间的数据流向。
- (3) 根据数据流向，梳理清楚对数据库的访问需求。

例如，本章案例网络贷款系统为业务人员提供机构人员维护、贷后管理、客户管理、产品管理、合作方管理等功能，是一个在线贷款受理、贷款审批、放款、回收和贷后管理的系统，对接网银、手机银行，采取全流程“不落地”线上操作模式。根据客户的资产、负债和收入情况，对客户的还款能力进行综合评价，由系统在线自动审批授信额度和贷款利率定价。实现在线实时获取客户的商业信誉信息和经营业绩历史数据，全面了解授信客户的现金流及信用情况。

图 4-1 是某商业银行网络贷款系统的逻辑架构图，整个系统从逻辑上分为持久层、技术支撑层、业务处理层、业务分析层和展现层等各个子系统，其中，持久层负责和数据库交互。



网络贷款系统所涉及的关联系统较多，主要与外联平台、ODS 系统、天征系统、零售评级系统、内容管理平台等系统全面对接。

4.1.2 功能性需求

了解了网络贷款系统的逻辑架构后，接下来需要明确应用对数据库的功能需求。这里重点要考虑 SQL 语句兼容性要求、应用访问数据库编程接口、存储过程开发、工作负载等需求。

1. SQL 语句兼容性要求

美国国家标准化组织（ANSI）是一个建立多种行业标准的组织，SQL 作为关系数据库所使用的标准语言，最初是基于 IBM 的实现在 1986 年被批准的。1987 年，国际标准化组织（ISO）把 ANSI SQL 作为国际标准。这个标准在 1992 年进行了修订，即 SQL-92，1999 年再次修订，即 SQL-99，目前最新版是 SQL-2011。

不同数据库产品所实现的 SQL 标准存在差异，不同应用使用的 SQL 也可能存在兼容性差异，因此在进行功能需求分析时，要明确 SQL 语句兼容性需求。

2. 应用访问数据库编程接口

现在的数据库产品都支持多种编程接口，但在企业级应用中，最常用的为 JDBC 接口和 C/C++接口。在部署中，对客户端使用的 JDBC 驱动版本或者 C/C++驱动版本的选择要重视，确保客户端所使用的驱动版本和数据库服务器版本兼容。

3. 存储过程开发

存储过程是在数据库系统中，一组为了完成特定功能的 SQL 语句集，它编译后存储在数据库服务器端，应用通过指定存储过程名字并提供输入参数来执行它，可以避免客户端和服务端过多的交互。存储过程在执行批处理时具有性能优势，因此使用广泛。但是，由于不同数据库支持的存储过程编写语法不同，所以在将存储过程从一种数据库迁移到另外一种数据库时，迁移的工作量较大，而且存储过程出了问题调试困难。

4. 工作负载要求

系统对数据库的负载可分为事务型（On-Line Transaction Processing, OLTP）、分析型（On-Line Analysis Processing, OLAP）和混合负载。

如表 4-1 所示，OLTP 负载要求非常快的处理查询和事务，例如，股票交易系统、银行核心系统是典型的 OLTP 系统；OLAP 负载要求处理复杂查询以及报表等，例如，数据仓库系统、决策支撑系统是典型的 OLAP 系统；混合负载是 OLTP 负载和 OLAP 负载的组合，例如，商业银行的理财平台，既有基金产品的申购和赎回这样的 OLTP 负载，也有按照某区域对理财用户进行统计分析的 OLAP 负载。

表 4-1 工作负载种类

特点	OLTP 负载	OLAP 负载	HTAP 混合负载
负载特点	事务处理为主	查询、分析、报表和跑批为主	混合事务处理、查询、分析、报表和批处理等负载
SQL 语句种类	语句简单，不会使用数据库内置的 olap 函数	语句复杂，会使用数据库内置的 olap 函数	既有简单的语句，也有使用 olap 函数的复杂语句
多表关联	不超过 3 张表	超过 3 张表以上	既有不超过 3 张表的关联，也有超过 3 张表的关联
子查询	很少	较多	较多
返回结果集大小	小（10 万条记录以下）	大（10 万条记录以上）	大（10 万条记录以上）

续表

特点	OLTP 负载	OLAP 负载	HTAP 混合负载
索引	少量	较多	较多
事务吞吐量	高	低	较高
分区技术	较少	较多	较多

就网络贷款系统来说,经过和开发中心项目组确认,该系统工作负载为混合负载,既有日间联机交易,也有日终批处理。

4.1.3 非功能性需求

系统的非功能性需求,主要从性能、高可用、可扩展性三个方面分别考虑,具体来说,包括以下内容。

(1) 从性能指标来看,主要从以下三方面来衡量。

① 响应时间:应用向数据库服务器发起请求,服务器完成处理并返回结果给应用,中间总共消耗的时间。它反映了数据库服务器的处理效率。

② 事务吞吐量:通常用每分钟处理的事务数(TPM)或者每秒钟处理的事务数(TPS)来计算,这个指标反映了数据库的事务处理能力。

③ 资源利用率:数据库服务器在处理事务或者响应查询的过程中,CPU、I/O 以及磁盘等系统资源的使用情况,这个指标反映了系统资源是否被有效利用。

就网络贷款系统来说,根据开发中心项目组和业务部门的沟通结果来看,其日间联机和日终批处理的性能要求,具体如下。

① 日间联机性能要求:最大在线连接数 1000 人,最大并发数 200,简单查询应在 20 ms 内响应,复杂查询则应在 1s 内响应。

② 日终批处理性能要求:日终批处理不应影响日间,应在晚上业务清闲时完成,日终批处理时间应小于 2 h,最长不超过 4 h。

(2) 根据系统的高可用要求,有以下分类。

① 不间断运行(7×24 h):这种系统要求全天候运行,系统负载在白天和晚上会有变化,但都不允许停机。

② 平时 24 h 运行(5×24 h):这种系统在周一~周五的白天和晚上都要运行,周末可以停机维护。

③ 平时 8 h 运行(5×8 h):这种系统在周一~周五的白天运行,平时晚上和周末都可以停机维护。

④ 8 h 运行(7×8 h):这种系统要求在周一~周日的白天运行,平时晚上和周末晚上可以停机维护。

就网络贷款来说,根据开发中心项目组和业务部门的沟通结果来看,确认该系统工作时间为 7×24 h 不间断运行,系统需要达到 99.99%可用性,即每年最大停机时间为 52.6 min。关于系统可用性和每年最大停机时间对应关系,如表 4-2 所示。

(3) 可扩展性。

数据库提供了垂直(Scale Up)和水平(Scale Out)两种扩展方式。如图 4-2 所示,垂直扩展方式中,通过在单个服务器内增加 CPU、内存以及存储来提升处理能力;在水平扩展方

式中，通过增加服务器节点来提升处理能力。

表 4-2 系统可用性和每年最大停机时间

系统可用性	每年最大停机时间
99.9%	8.76 h
99.99%	52.6 min
99.999%	5.26 min

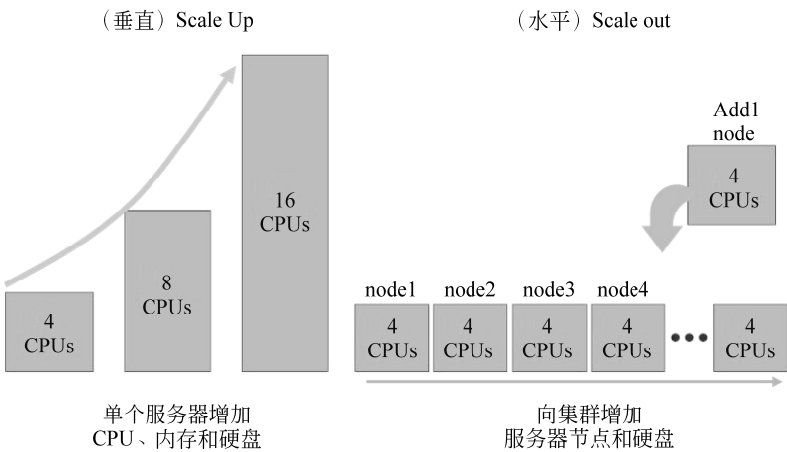


图 4-2 数据库两种扩展方式

就网络贷款系统将来的性能需求来说，由于采用集中式架构，所以只能考虑纵向扩展，即通过在当前节点增加物理资源的方式进行硬件扩容。

4.1.4 容灾需求

根据银监会《商业银行数据中心监管指引》，总资产规模一千亿元人民币以上且跨省设立分支机构的法人商业银行及省级农村信用联合社应设立异地灾备中心，重要信息系统灾难恢复能力应达到《信息安全技术—信息系统灾难恢复规范》中定义的灾难恢复等级第 5 级（含）以上；其他法人商业银行应设立同城模式灾备中心并实现数据异地备份，重要信息系统灾难恢复能力应达到《信息安全技术—信息系统灾难恢复规范》中定义的灾难恢复等级第 4 级（含）以上。

重要信息系统是指银行业金融机构支撑关键业务，其信息安全和系统服务安全关系公民、法人和组织的权益或社会秩序和公共利益，甚至影响国家安全的信息系统。主要包括面向客户、涉及账务处理且时效性要求较高的业务处理类、渠道类和涉及客户风险管理等业务的管理类信息系统，以及支撑系统运行的机房和网络等基础设施。

根据第 1 章相关内容，读者了解到商业银行通常采用“两地三中心”的容灾架构，即生产中心机房、同城灾备中心机房和异地灾备中心机房。这些机房需经过第三方检测，按照 GB50174 A 级机房标准建设，供电采用双电源双回路，低压采用 2N 冗余，高压及制冷采用 N+1 冗余，使用柴油发电机作为应急电源并有持续供油能力，满足银行业务连续性对基础设施高可用的要求。

不同的商业银行，会对其管理的应用系统进行评级，根据不同的评级，应用系统的重要

性评级划分为以下三类。

第一类，对应银监会重要信息系统，主要是指关键应用系统，具体包括面向客户、涉及账务处理且时效性要求极高，其服务中断将对商业银行或客户利益造成严重损害的应用系统。这类系统需搭建同城灾备和异地灾备环境。

第二类，主要是指重要应用系统，具体包括面向客户、涉及账务处理且时效性要求较高，其服务中断将对我行或客户利益造成较大损害的应用系统。这类系统搭建同城灾备环境。

第三类，主要是指一般应用系统，具体包括分析决策类和协同办公类系统、不涉及账务处理的应用系统，以及非直接面向客户的应用系统。这类系统无须搭建同城或者异地灾备环境。

就网络贷款系统来说，属于第一类重要系统，需要搭建同城灾备和异地灾备环境。根据中国人民银行《银行业重要信息系统灾难恢复管理规范》，不同的系统采用不同的灾备等级，如表 4-3 所示。

表 4-3 系统分级和灾备等级对照

系统分级	定 义	灾备等级 (国标)
第一类	(1) 短时间中断将对国家、外部机构和社会生产产生重大影响的系统； (2) 短时间中断将严重影响单位关键业务功能并造成重大经济损失的系统； (3) 单位和用户对系统短时间中断不能容忍的系统	5 级
第二类	(1) 短时间中断将影响单位部分关键业务功能并造成较大经济损失的系统； (2) 单位和用户对系统短时间中断具有一定容忍度的系统	3 级
第三类	(1) 短时间中断将影响单位非关键业务功能并造成一定经济损失的系统； (2) 业务功能容许一段时间中断的系统	2 级

4.1.5 某银行网络贷款系统部署需求结论

经过上述分析，最终完成了某商业银行网络贷款系统的部署需求：该系统为业务交易类；属于第一类重要系统；要求系统服务时间为 7×24 h，达到 99.99%可用性；供客户、业务管理员和系统管理员使用；功能需求上，要求支持 SQL-99 标准、混合负载；容灾需求上要求同时支持同城和异地灾备，因此采用国标 5 级灾备等级，如表 4-4 所示。

表 4-4 网络贷款系统部署需求表

系统名称	网络贷款系统
系统分类	业务交易类（√） 运营支撑类（×） 管理分析类（×） 办公管理类（×） 开发支撑类（×） 运维支撑类（×）
重要性等级	第一类（√） 第二类（×） 第三类（×）

续表

系统名称	网络贷款系统
系统服务时间	7×24（√）
	5×24（×）
	5×8（×）
	7×8（×）
用户类型	客户、业务管理员、系统管理员
功能性需求	支持 SQL-99（√）
	通过 JDBC 接口访问数据库（√）
	存储过程（×）
	工作负载：混合负载（√）
非功能性需求	日间联机性能要求： 最大在线连接数 1000 人，最大并发数 200，简单查询应在 20ms 内响应，复杂查询则应在 1s 内响应。 日终批处理性能要求： 日终批处理不应影响日间，应在晚上业务清闲时完成，日终批处理时间应小于 2h，最长不超过 4h
	可用性要求： 99.9%（×） 99.99%（√） 99.999%（×）
	扩展性要求： 纵向扩展（√） 横向扩展（×）
容灾需求	同城灾备（√）
	异地灾备（√）

4.2 数据库部署环境规划

完成了部署需求，接下来就是针对需求进行部署环境规划。从硬件方面，需要考虑计算资源、存储资源和网络资源；从软件方面，需要真正做好数据库本身规划；从容灾方面，需要规划好同城和异地规划。

4.2.1 计算资源规划

在为数据库服务器规划计算资源时，商业银行需要根据实际情况综合考虑，选择最合适的方案。从实践上来说，建议形成低、中、高不同配置的套餐，供新系统部署时使用，下面具体分析。

1. 选择小型机还是 x86 服务器

当前，很多商业银行在一些重要系统上还在继续使用 IBM 或者 HP 小型机，其主要优势是小型机稳定、故障率低。但随着技术潮流的推动，x86 服务器的性能和可靠性提升很快，

很多关键系统开始逐渐迁移到 x86 平台上。如果没有特殊要求, 建议大胆采用 x86 服务器作为数据库的运行环境。

2. 物理机和虚拟机

从实践上考虑, 对于 OLAP 负载和混合负载, 由于对 I/O 吞吐能力和 I/O 响应时间具有较高要求, 建议使用物理机; 对于 OLTP 负载, 如果对 I/O 吞吐能力和 I/O 响应时间没有特殊要求, 可以选用虚拟机。

3. 规划 CPU 内核数和内存

根据最佳实践, 需要选择好具有足够内核数的 CPU, 随后根据内核数配备相应的物理内存。通常 CPU 内核数和内存最佳比例为 1:8, 例如, 4 核的 CPU 对应的内存配置为 32GB。

4.2.2 存储资源规划

关于数据库所使用的存储资源, 首先是进行资源选择, 选择本地盘还是集中式存储, 随后是进行文件系统规划。

1. 本地固态硬盘还是 SAN 存储

随着存储技术的进步, SAN 存储所使用的磁盘开始从 SAS 盘逐渐过渡到固态硬盘。当前, 很多 x86 服务器所使用的内置本地盘, 也开始逐渐过渡到固态硬盘, I/O 性能可以和使用固态硬盘的存储媲美。那么, 对数据库服务器来说, 如何选择呢?

如表 4-5 所示为本地固态硬盘和存储的对比, 对追求高性能并能容忍一定故障率的场景, 建议采用本地固态硬盘; 对追求高可用和存储级容灾的场景, 建议采用 SAN 存储。

表 4-5 本地固态硬盘和存储优劣对比

优劣对比	本地固态硬盘	SAN 存储
优势	性能高于存储	(1) 存储一般有多个控制器、多块磁盘, 这样的冗余设计, 相较本地盘, 存储的可用性更高; (2) 便于数据迁移; (3) 有丰富的附加功能, 例如存储复制实现容灾
劣势	(1) 故障率高于存储; (2) 扩容比存储复杂, 需要热插拔磁盘; (3) 本地固态硬盘所采用的 RAID 卡存在单点, 如果 RAID 卡损坏将影响整个 RAID 组	(1) 存储资源相较本地盘成本高, 在应用数据持久化上需要多套系统共用一台存储设备, 有可能存在一套系统的热点 IO 影响其他系统的问题; (2) 相较本地固态硬盘, 存储需要经过 SAN, 从性能上不如本地磁盘

2. 规划文件系统

完成了存储资源选择后, 接下来就是规划数据库文件系统了。不论是 Db2 还是 MySQL, 数据和日志需使用不同的文件系统, 其中, 日志包括事务日志和归档日志。接下来, 分别以 Db2 和 MySQL 为例加以说明。

如表 4-6 所示, 根据 Db2 相关技术规范, 某商业银行规划的 Db2 数据库服务器文件系统, 使用了 LVM 进行逻辑卷管理。总计两个卷组, 一个是 db2datavg, 在该卷组上创建了两个逻辑卷 lvdb2data 和 lvdb2log, 大小分别为 300GB 和 30GB, 对应的文件系统为 /db2data 和 /db2log, 分别用于存放数据文件和事务日志; 另外一个卷组是 arclogvg, 在该卷组上只创建了一个逻辑卷 lvarclog, 大小为 60GB, 对应的文件系统为 /db2arclog, 用于存放归档日志。这些文件系统为 Db2 实例用户和组所有, 实例用户为 db2inst1, 组为 db2iadm1。

表 4-6 某商业银行 Db2 数据库服务器文件系统

卷组名称	逻辑卷名称	逻辑卷大小	文件系统	属主/组	备注
db2datavg	lvdb2data	300 GB	/db2data	db2inst1: db2iadm1	数据文件
	lvdb2log	30GB	/db2log	db2inst1: db2iadm1	事务日志
arclogvg	lvarclog	60GB	/db2arclog	db2inst1: db2iadm1	归档日志

如表 4-7 所示，根据 MySQL 相关技术规范，某商业银行规划的 MySQL 数据库服务器文件系统，也使用了 LVM 进行逻辑卷管理。总计两个卷组，一个是 mydatavg，在该卷组上创建了一个逻辑卷 lvmydata，大小为 150GB，对应的文件系统为 /mysqldata，用于存放数据文件；另外一个卷组是 binlogvg，在该卷组上创建了一个逻辑卷 lvbinlog，大小为 100GB，对应的文件系统为 /mysqllog，用于存放二进制日志。这些文件系统为 MySQL 实例用户和组所有，本例实例用户为 myinst1，组为 myiadm1。

表 4-7 某商业银行 MySQL 数据库服务器文件系统

卷组名称	逻辑卷名称	逻辑卷大小	文件系统	属主/组	备注
mydatavg	lvmydata	150 GB	/mysqldata	myinst1: myiadm1	数据文件
binlogvg	lvbinlog	100GB	/mysqllog	myinst1: myiadm1	二进制日志

4.2.3 网络资源规划

规划了文件系统，随后为数据库服务器申请网络资源，首先需要规划好 IP 地址，随后开通网络访问控制。

1. 规划 IP 地址

需要规划好数据库服务器主机对应的 IP 标识、IP 地址和 IP 类型。其中，IP 标识和服务器的主机名对应；IP 地址对应具体网段；IP 类型包括服务 IP 和管理 IP，服务 IP 是可浮动的，用于配置双机 HA 使用，管理 IP 对应该服务器的物理 IP 地址。

如表 4-8 所示，网络贷款系统数据库部署在两台服务器上，主机名分别为 fbmsdb1 和 fbmsdb2，fbmsdb1 和 fbmsdb2 具有管理 IP，采用双机互备部署，通过服务 IP（标识为 fbmsdb_svc）对外提供服务，相关 IP 地址和 IP 类型如表 4-8 所示。

表 4-8 网络贷款系统数据库服务器规划

卷组名称	逻辑卷名称	逻辑卷大小	文件系统
fbmsdb_svc	××.××.××.××	服务 IP	fbmsdb_svc
fbmsdb1	××.××.××.××	管理 IP	fbmsdb1
fbmsdb2	××.××.××.××	管理 IP	fbmsdb2

2. 开通网络访问

规划好 IP 地址后，接下来需要考虑访问控制。数据库服务器和外部系统之间通过网络交互，如果交互的网络上存在防火墙，这时就要填写网络访问申请，以确保系统之间可以互联互通。在网络访问申请中需要说明源系统和目标系统相关信息，包括源地址/目标地址、源端口/目标端口、连接方式等。

经过调研，网络贷款系统数据库服务器和 ODS 系统、天征系统存在访问关系，并且和这

些系统之间的网络上存在防火墙,因此需要按照表 4-9 开通访问关系,具体内容见表 4-9 所示。

表 4-9 网络贷款系统数据库服务器网络访问关系

源系统主机名	源主机 IP 地址	源端口	目标系统名称	目标主机 IP 地址	目标端口	连接方式
fbmsdb_svc	×.×.×.×	Any	ODS 系统	×.×.×.×	×××× ×××× ××××	长连接
fbmsdb_svc	×.×.×.×	Any	天征系统	×.×.×.×	×××× ×××× ××××	长连接

4.2.4 数据库选型

这里的数据库选型是狭义的概念,是指在上线时要根据部署需求,完成数据库种类选择和版本选择以指导数据库部署。下面具体探讨这两个问题。

1. 数据库种类选择

在选择数据库时,不同的商业银行规划标准不同,但出发点都是根据所在商业银行的数据库技术路线规划进行。

从商业银行集中式数据库的使用现状来看,目前使用最广的还是 Db2 或者 Oracle 数据库;但开源数据库 MySQL 社区版或者 PostgreSQL 数据库也开始逐渐增多,如表 4-10 所示,具体选择原则建议如下。

(1) 针对 OLTP 负载:选择范围较广,可以从 Db2、Oracle、MySQL 或 PostgreSQL 这四种数据库中进行选择。从自主掌控的角度考虑,建议优先选择开源的 MySQL 或 PostgreSQL;否则建议选择 Db2 或者 Oracle;对自主掌控有要求,但由于应用中使用了大量的存储过程、多表关联查询这些功能性需求,建议继续沿用 Db2 或者 Oracle。

(2) 针对 OLAP 负载:选择范围也较广,如果要从数据库层面解决,建议选择 Db2 或者 Oracle 数据库,这两种商业数据库都支持 OLAP 负载;或者利用大数据平台解决,建议选择基于 Hadoop 的大数据平台。

(3) 针对混合负载:如果要从数据库层面解决,建议选择 Db2 或者 Oracle;否则,建议选择基于 Hadoop 的大数据平台。

表 4-10 工作负载实现方式

负载种类	通过数据库实现	其他方式实现
OLTP 负载	Db2、Oracle、MySQL、PostgreSQL	N/A
OLAP 负载	Db2、Oracle	基于 Hadoop 的大数据平台
HTAP 混合负载	Db2、Oracle	基于 Hadoop 的大数据平台

2. 数据库版本选择

选择了合适的数据库类型后,接下来要选择合适的数据库版本。根据商业银行对数据库基础软件的版本管理办法,通常会以季度为单位对所使用的数据库软件版本进行评估,随后给出推荐版本,在新建系统中优先使用新版本,对存量系统制订版本升级计划。例如,某商