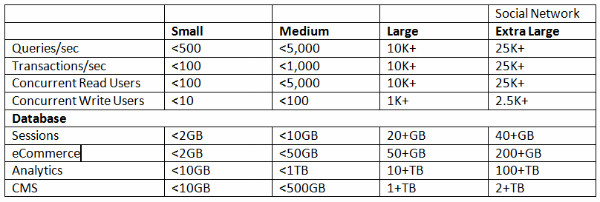
甲骨文发布了《[面向大规模可伸缩网站基础设施的MySQL参考架构](http://www.mysql.com/why-mysql/white-papers/mysql-reference-architectures-for-scalable-web-infrastructure/)》白皮书，针对将MySQL用作数据存储的不同类型和不同规模的网站给出了推荐的拓扑结构。

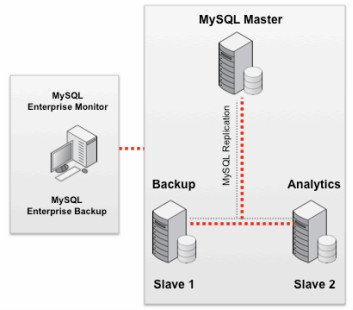
　　根据分别提供4类服务——用户和会话管理、电子商务、分析类应用 (多结构数据)和CMS(元数据)——的网站的规模和可用性要求(如下表所示)，这份白皮书给出了4个参考架构。



　　请注意，这里给出的指导方针只是基本建议，实际应用中需要根据读写模式、负载平衡和所用的缓存机制等因素进行调整。

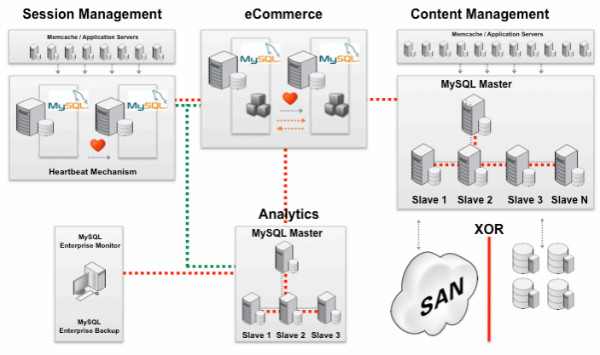
**小型(Small)网站参考架构**

　　这一参考架构可用于上述4类网站的所有小型实现。可以使用MySQL Replication来制作数据的副本以支持备份和分析。



**中型(Medium)网站参考架构**

　　在这种情况下，推荐针对不同类型的活动选择独立的基础设施，考虑每个MySQL服务器最多支持8个应用服务器，如果因伸缩性需求应用服务器数量增加，则添加更多的MySQL从服务器。



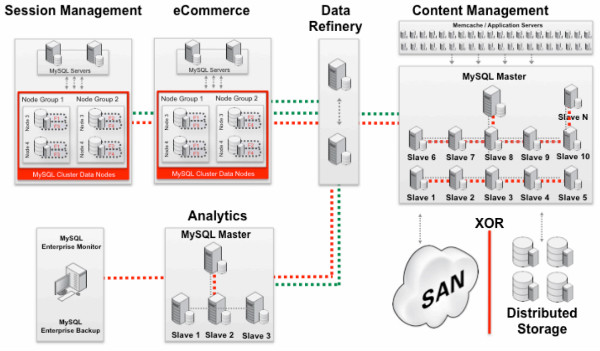
　　为满足会话管理网站和电子商务网站的高可用性要求，可以使用Linux心跳(Heartbeat)和半同步复制。CMS网站通常对读操作的向外扩展有更高要求，假定每个MySQL从服务器最多可以处理3000个并发用户，白皮书建议为每个MySQL主服务器添加20-30个从服务器。CMS系统可将数据保存在一个SAN中，或者保存在连接到该服务器的分布式设备中。

　　会话管理网站和CMS网站推荐使用Memcached，这有助于减轻应用服务器和MySQL服务器的负担。

　　分析类网站的拓扑结构简单一些，1个主服务器加3个从服务器就能解决问题。

**大规模(Large)网站参考架构**

　　针对大规模网站，白皮书推荐使用MySQL Geographic Replication来进行跨数据中心的数据库复制，这种方式支持跨越地理上分离的集群进行异步复制。

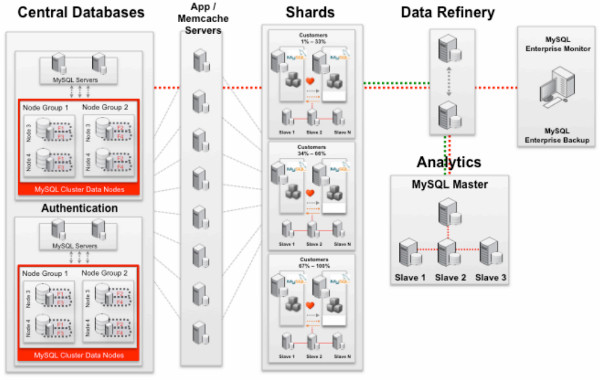


　　会话管理网站和电子商务网站应该使用集群，白皮书声称“4个数据节点，1秒可以支持6000个会话(页面点击)，其中每次页面点击生成8–12个数据库操作”。大规模CMS网站使用的配置与中型网站类似，只是必要时需要多添加一些从服务器。针对分析类应用，这里引入了一个数据提炼(Data Refinery)单元，用于数据的清理和组织。

**超大规模(Extra Large)网站参考架构**

　　针对社交网站，白皮书也给出了相应建议。它声称“网络上流量最大的10个网站有9个部署了MySQL，其中包括Google、Facebook和YouTube”，但是没有说明这些网站用MySQL干什么，不过众所周知的是，LinkedIn成功应用了MySQL。

　　社交网站的拓扑结构利用了中型和大规模网站中实现的概念，包括专用应用服务器、Memcached和数据提炼单元, 但为支持写操作的向外扩展引入了分片(Shard)。MySQL集群被用于用户的认证和查找，当“用于查找的键(key)不止1个”时，直接读写相应的分片。



　　MySQL主服务器和从服务器的推荐规格如下：

* 8–16个x86-64位CPU核心(MySQL 5.5及以上)。
* 4–8个x86 -64位CPU核心(MySQL 5.1及更早版本)。
* 比活动数据多3–10倍的内存。
* Linux、Solaris或Windows操作系统。
* 最少4块磁盘，8–16块磁盘能增加I/O密集型应用的性能。
* 支持电池供电高速缓存的硬件RAID。
* 推荐使用RAID 10。如果负载为读密集型，RAID 5也是合适的。
* 2个网卡和2个供电单元用作冗余。

　　另外，白皮书还有一些针对MySQL集群和数据存储设备的建议，再就是用于监控、备份和集群管理的解决方案。