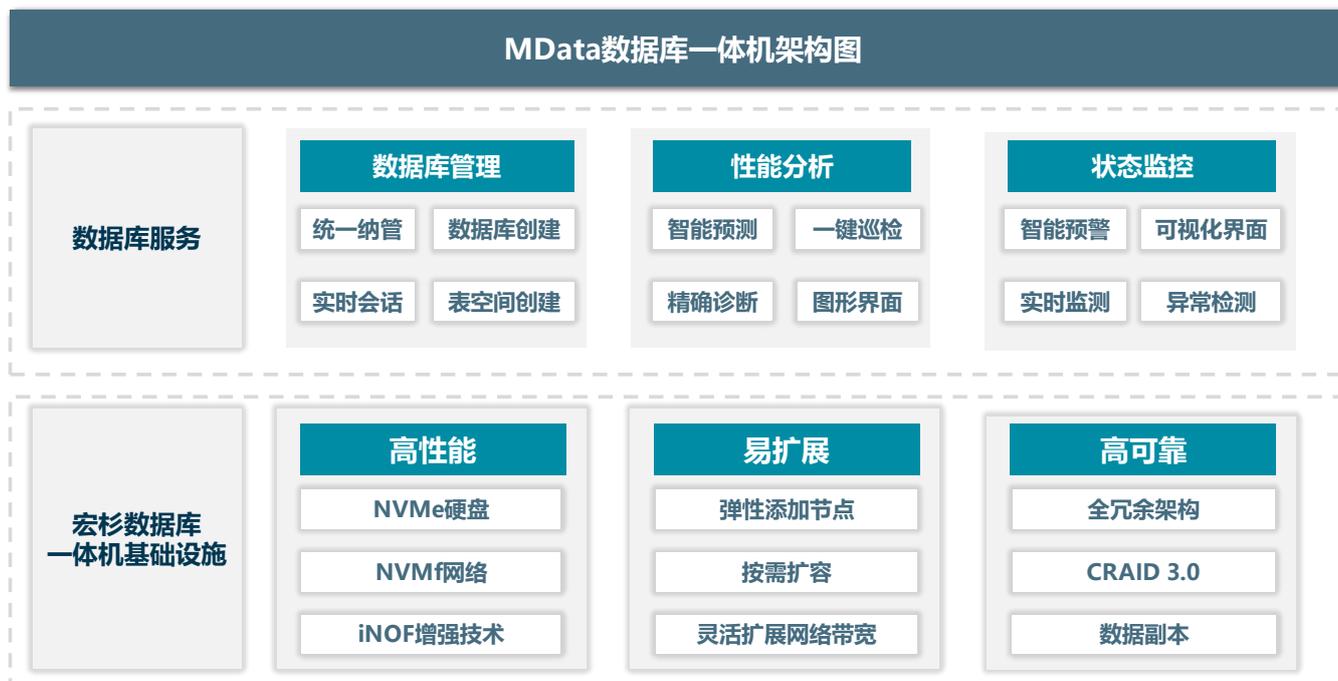


# MacroSAN MData 数据库一体机



## 产品概述

MData 是宏杉科技推出的融合计算、存储、网络和数据库的数据库一体机平台，通过预集成和深度调优，旨在实现一站式搭建核心业务系统，满足 OLTP 和 HTAP 场景下的极致性能需求，为企业核心业务提供高性能、高可靠、高扩展、易运维的数据库支撑。



宏杉 MData 数据库一体机架构

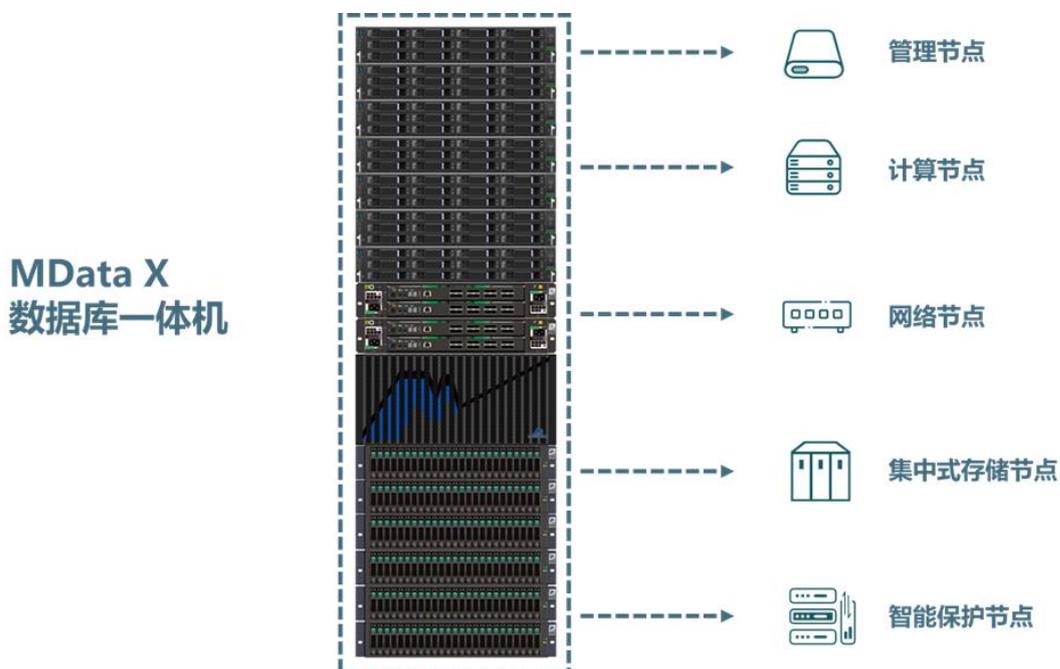
根据产品定位及形态不同，MData 分为两大系列产品：

(1) MData X 系列：定位于高性能集中式数据库一体机，基于高性能服务器、高带宽无损以太网交换机、低延迟集中式存储构建，具有极致性能、稳定可靠的特点；

(2) MData O 系列：定位于高扩展分布式数据库一体机，基于高性能服务器、高带宽无损以太网交换机或 IB 交换机、高扩展分布式存储构建，具有架构灵活、并发访问的特点。

## 产品架构

### ❖ MData X 系列架构

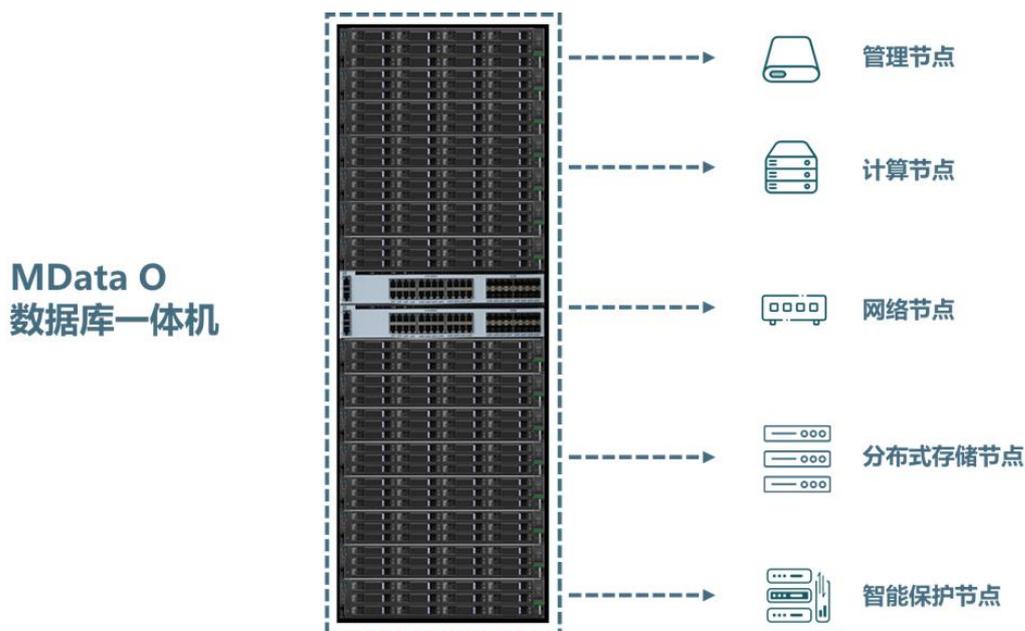


- **管理节点：**采用 1 台服务器，部署一体机管理软件，基于 web 界面提供对数据库一体机平台的系统监控和统一管理。管理软件还提供了告警功能，可通知管理员或运维人员在一体机出现问题时采取必要的措施。
- **计算节点：**采用 2 节点及以上宏杉高性能服务器，采用高能效冗余电源，避免单点故障产生的宕机，优化的散热系统和监控系统保障了系统运行过程中的稳定可靠，减少风险，较大程度降低业务中止带来的损失。
- **网络节点：**采用 RoCE 无损交换机构建计算、存储互连网络，利用 RDMA 的内核旁路和零拷贝特性提供的低延迟特性，从而减少了 CPU 占用，减少了内存带宽瓶颈，构建高带宽、低延迟、零丢包的网路传输链路。
- **集中式存储节点：**采用至少一套高性能 MS 集中式存储作为用户业务系统数据存储节点，通过冗余交换机与前端服务器相

连，实现前端各业务系统分散数据的集中存储，实现数据的统一管理和保护，简化数据管理并带来数据管理的高效性。

- **智能数据保护节点：**采用宏杉数据备份解决方案，部署宏杉科技高可用系列备份一体机，采用了业内领先的双控制器架构，实现硬件模块全冗余、硬件故障自动切换、组件无线缆连接、在线灵活扩展等特性。

## ❖ MData O 系列架构



- **管理节点：**采用 1 台服务器，部署一体机管理软件，基于 web 界面提供对数据库一体机平台的系统监控和统一管理。管理软件还提供了告警功能，可通知管理员或运维人员在一体机出现问题时采取必要的措施。
- **计算节点：**采用 2 节点及以上宏杉高性能服务器，采用高效冗余电源，避免单点故障产生的宕机，优化的散热系统和监控系统保障了系统运行过程中的稳定可靠，减少风险，较大程度降低业务中止带来的损失。
- **网络节点：**采用 RoCE 无损交换机或 IB 交换机构建计算存储互连网络，利用 RDMA 的内核旁路和零拷贝特性提供的低延迟特性，从而减少了 CPU 占用，减少了内存带宽瓶颈，构建高带宽、低延迟的网络传输链路。
- **分布式存储节点：**采用一套 3 节点及以上宏杉万象分布式存储，基于分布式架构，实现高效访问、智能负载均衡、性能随节点扩容线性增长的特性。
- **智能数据保护节点：**采用 1 台服务器，作为独立的备份节点，对一体机平台上的数据库做集中备份和统一管理，实现 CDP 备份，灾备保护等功能。

## 产品特性

### ❖ 架构高可用

- **计算高可用：**通过部署多个计算节点组成集群实现计算高可用。组建高可用集群技术是实现数据库高可用的一种常见方法，利用至少两台或多台数据库服务器节点组合在一起以共同处理数据请求，并在故障情况下提供无缝切换。数据库集群的高可用性设计旨在最小化服务中断，确保数据库系统在故障或意外情况下继续提供可靠的服务。这对于支持关键业务应用程序和大规模数据存储的系统非常重要。不同的数据库软件和云平台提供了各种实现高可用性的工具和解决方案，MData 数据库机对其具有良好的兼容性，用户可以根据特定需求进行选择 and 配置。
- **网络高可用：**采用冗余链路的配置实现内部网络的互通，每个计算节点都通过 2 条链路分别连接到 2 台 RoCE 或 IB 交换机上，实现了数据访问路径的冗余，保证在单链路出现故障时，业务不会中断。
- **存储高可用：**MData X 系列支持 RAID 5、RAID 6、RAID X-3 多种 RAID 方式，允许在同一个硬盘组中任意三块硬盘同时发生完全物理故障，数据不丢失，业务不中断。在更换三块新硬盘后，支持并行重建。MData O 系列数据库一体机可支持存储层设置多副本（两副本或三副本）数据存储模式，保证数据的冗余性、高可用性。

### ❖ 数据高可靠

- **全冗余架构：**数据从服务器写入硬盘，会经过主机通道、控制器、缓存、硬盘通道、硬盘柜、硬盘等组件，在整个数据的通道上，采用了模块化、全冗余的架构设计，任意组件发生故障时，能够实现快速的故障隔离和组件更换，确保业务连续。
- **A-A 架构：**MData X 系列存储节点采用控制器采用双 Active-Active 架构，使得工作负载均衡分配，可共享访问缓存、前/后端 IO 命令以及硬盘空间，提升存储系统访问性能。
- **缓存掉电保护：**采用缓存降落技术，当存储突然掉电后，通过存储自带的电池，将缓存数据下刷到硬盘中进行永久保存，能够保证缓存中的数据不丢失。
- **RAID 机制：**MData X 系列采用 RAID 作为数据管理机制，支持宏杉 CRAID 技术。CRAID 技术是宏杉科技针对传统 RAID 的缺陷，在传统 RAID 技术之上的革新。CRAID1.0 技术提升了故障硬盘的重建效率，CRAID2.0 技术允许 RAID 组中任意坏三块盘，数据不丢失，而 CRAID3.0 技术则提升了数据读写性能和缩短重建时间。RAID 与 CRAID 可共存，支持专用热备、全局热备、空闲硬盘热备，热备盘与热备空间可共存。

- **多副本机制:** MData O 系列数据库一体机采用数据多副本机制保障数据可靠性。提供支持两种冗余级别，两副本和三副本。数据被复制到多个节点上，这些节点可以位于不同的物理位置或数据中心，每个副本都是相互独立的，可以独立地读取和写入数据，当一个节点发生故障或不可用时，其他副本可以继续提供数据服务，从而保证系统的可用性。通过设置不同的冗余级别，可以容忍不同数量的节点发生故障，确保业务不中断、数据不丢失，两副本模式可允许一个存储节点发生故障，三副本模式可允许两个存储节点发生故障。

## ❖ 访问高性能

- **RDMA:** RDMA (Remote Direct Memory Access) 是一种在网络上的两台计算机的主内存之间交换信息的方式，而无需通过两台计算机的处理器 (CPU)、缓存或操作系统。由于 RDMA 绕过操作系统，因此它通常是通过网络传输数据的最快且成本最低的机制。
- **NVMe-oF:** NVMe-oF (NVMe over Fabrics) 是 NVMe 协议的扩展，NVMe-oF 是提供封装了 NVMe 命令的远程网络协议，可通过 PCIe 以外的网络互连实现基于 NVMe 的通信，NVMe-oF 通过 Fabric 网络来承载 NVMe 协议，在传输网络上具备超低延时、超高性能的特点，克服了 PCIe 总线的局限性，可以将“外部”存储直接或通过交换机连接到服务器，使得主机通过 NVMe 协议访问远端 SSD 成为可能，同时仍将 NVMe 用作基本通信机制。不同 MData 数据库一体机系列支持不同 NVMe-oF 连接方式：MData X 系列数据库一体机产品除了支持基本的 FC、iSCSI 访问方式之外，还支持 25Gb/100Gb NVMe over RoCE 及 16G/32G NVMe over FC 访问，保障端到端的 NVMe，为用户核心业务赋能，提供高性能、低延时的数据库基础设施；MData O 系列数据库一体机产品除了支持基本的 FC、iSCSI 访问方式之外，还支持 25Gb/100Gb NVMe over RoCE 及 100Gb IB。
- **iNOF:** MData X 系列数据库一体机能够提供基于 iNOF (Intelligent Lossless NVMe Over Fabric) 的端到端 NVMe over RoCE 存储解决方案，实现以太网存储的零丢包，高性能，秒切换，易配置。为数据中心核心业务提供高性能、低延时的端到端 NVMe 存储平台。

## ❖ 空间高利用

- **重复数据删除技术:** MData X 系列数据库一体机采用了宏杉自研重复数据删除技术 (Deduplication)。重复数据删除是一种数据缩减技术，可对存储容量进行有效优化。它通过删除数据集中重复的数据，只保留其中一份，从而消除冗余数据。这种技术可以很大程度上减少对物理存储空间的需求，减少传输过程中的网络带宽，有效节约设备采购与维护成本。同时它也是一种绿色存储技术，能有效降低能耗。

- **在线压缩技术：**MData X系列采用了宏杉在线数据压缩，通过存储系统内置的数据压缩功能模块，在数据写入的第一时间，就对数据进行压缩处理，而不是等数据写入完成再进行压缩。存储在对数据进行压缩处理时，前端业务系统是无感知的。宏杉科技存储通过强大的硬件资源以及优化的压缩算法，将数据压缩给前端业务系统的影响降低到最小，最大限度地保证业务的畅通访问。
- **精简配置：**MData数据库一体机提供精简配置功能。可以创建容量大于存储池物理容量的逻辑卷以满足复杂业务场景的容量规划需求，相比直接分配物理存储资源，“按需分配”可显著提高存储空间利用率，节省成本。精简配置过程为先分配虚拟卷，之后存储空间按需扩容。可以实现卷动态扩容，按策略自动扩容无需停机，达成空间最大化利用，降低运维难度的目标。

## ❖ 管理高效率

- **可视化界面：**一体机管理软件为用户提供了可视化的监控管理界面。用户可以根据需要将各种监控项的图表、数据和信息集成在一个页面上，以一目了然地查看系统状态。
- **监控功能：**基于web界面，为运维人员提供全面的系统监控以及网络状态监测，可以实时浏览到一体机平台的各种关键参数，例如CPU使用率、内存使用情况、网络流量和硬盘容量，以确保系统正常运行。同时，一体机通过SNMP（简单网络管理协议），ping，端口监视等方法提供对远程服务器/网络状态的监测，用户能在远程就轻松得知一体机目前的运行状态
- **告警功能：**一体机管理软件还提供了告警功能，一体机运行出现异常状况时，可以通过多种方式（例如电子邮件、短信、钉钉等）自动通知管理员或运维人员。管理员或运维人员根据告警信息对一体机平台进行处理，保障一体机正常可靠地运行。

## 典型配置

### ❖ MData X 系列典型配置

产品型号	通用型	国产型	
<b>计算节点</b>			
<b>CPU</b>	2×Intel Gold 处理器	2× 国产 X86 处理器	
<b>内存</b>	256GB	256 GB	
<b>系统盘</b>	2×480GB SSD	2×480GB SSD	
<b>业务网卡</b>	4×10GE 以太网卡/2×25GE 以太网卡*1		
<b>内部互联网卡</b>	2×100GE 以太网卡*2		
<b>网络节点</b>			
<b>网络类型</b>	无损 RoCE 网络		
<b>端口速率</b>	100GE		
<b>存储节点</b>			
<b>CPU</b>	Intel 至强可扩展处理器	国产 X86 处理器	
<b>内存 (每双控)</b>	128GB-1.5TB	256GB-4TB	128GB-1TB
<b>硬盘容量</b>	20TB~100TB NVMe		20TB~100TB NVMe

## ❖ MData O 系列典型配置

产品型号	通用型	国产型
<b>计算节点</b>		
<b>CPU</b>	2×Intel Gold 处理器	2× 国产 X86 处理器
<b>内存</b>	512 GB	512 GB
<b>系统盘</b>	2×1.92TB SSD	2×1.92TB SSD
<b>业务网卡</b>	4×10GE 以太网卡/2×25GE 以太网卡*1	
<b>内部互联网卡</b>	2×100GE 以太网卡*2	
<b>网络节点</b>		
<b>网络类型</b>	无损 RoCE 网络/IB 网络	
<b>端口速率</b>	100GE	
<b>存储节点</b>		
<b>存储节点数量</b>	3 节点及以上	
<b>CPU</b>	Intel 至强可扩展处理器	国产 X86 处理器
<b>网络</b>	100GE 以太网卡	
<b>硬盘容量</b>	20TB~100TB NVMe	

