解决方案简介

阿里云第八代企业级 ECS 实例采用第五代英特尔® 至强® 可扩展处理器 为企业云服务提供更优安全防护

阿里云

“阿里云从成立的第一天起，安全可信就是第一属性。阿里云一直致力于通过各种方式使阿里云对客户更透明，且用户数据的保护是其中不可分割的重要部分。除了严密的安全规约外，我们也在通过各种硬件安全技术来实现用户'可验证'的数据保护机制。第五代英特尔® 至强® 可扩展处理器在强劲的算力之外，其提供的英特尔® TDX 技术也有力地支持我们为客户提供了更便捷和更多样化的机密计算服务。”

—刘煜堃
阿里云高级安全专家
阿里云安全团队

帮助用户保护敏感的数据和工作负载，并保持从云到边缘的合规是我们一直致力于实现的目标。第五代英特尔® 至强® 可扩展处理器提供硬件级安全功能和可信服务，能够为用户的应用与数据筑牢安全屏障。我们将围绕云安全，与阿里云进行持续深度合作，拓展机密计算的应用实践，打造灵更活和友好的安全云计算环境。

李亚东
英特尔中国至强客户解决解决方案集团总经理

挑战

云服务的安全性正在得到越来越多的关注，传统的安全保护方案通常会面临如下挑战：

- 传统的数据安全解决方案更多的聚焦于保护静态和传输中的数据，但难以保护使用中的数据，带来了一定的安全隐患。
- 单一采用应用程序级可信边界，不仅会增加客户将全部应用程序、工作负载部署到安全云环境的难度，同时逐一对应用程序、工作负载开展改造，也会带来巨大的工作量。
- 在深度学习模型训练、推理等应用中，需要涉及到大量数据的流通，这个过程可能导致数据泄露、数据不当利用、数据篡改等风险。

解决方案概述

作为新一代的数字基础设施，云计算服务正逐渐成为各行业实施数字化转型，实现业务再突破的重要引擎之一。而随着数据价值的不断突显与相关政策的持续出台，云服务的安全性也正受到越来越多客户的关注。作为全球领先的云服务提供商与先进机密计算服务的领跑者，阿里云正与英特尔展开合作，将全新第五代英特尔® 至强® 可扩展处理器引入其最新的企业级 ECS 实例 g8i 中，以应对更为多样化的云服务模式需求。

英特尔全新发布的第五代英特尔® 至强® 可扩展处理器在为阿里云第八代企业级 ECS 实例提供强劲的算力支持之外，新处理器内置的英特尔® Trust Domain Extension（英特尔® TDX）技术，也为阿里云向客户提供面向虚拟化实例的机密计算新方案

提供了坚实的技术保障，助力客户在不改变现有应用程序的情况下，为其基础设施即服务（Infrastructure as a Service，IaaS）和平台即服务（Platform as a Service，PaaS）应用分别构建基于硬件设备的可信执行环境（Trusted Execution Environment，TEE），如机密虚拟机或机密容器。同时，英特尔® TDX 技术使用便捷，客户能在阿里云环境中大规模部署并实现实时迁移，拥有更灵活和友好的保密云计算环境。

阿里云通过英特尔® TDX 为客户提供机密计算安全保护

云环境中的数据按其所处状态，可分为三类，即 Data in Transit（传输状态）、Data at Rest（存储状态）以及 Data in Use（使用状态）。对于前两种状态，云服务提供商正借助安全访问、数据加密以及各类加密传输协议等技术来为客户打造更为安全可信的云端环境。而对于使用状态中的数据，机密计算是实现其有效保护的良策。机密计算为客户敏感数据提供了基于硬件的TEE环境，通过隔离保护的方式来防止未经授权的入侵者访问或修改处理中的数据，从而成为了目前云服务中常见的、面向应用运行时的数据安全技术方案。

随着更多企业级业务系统与云服务相融合，客户的大多数应用程序或工作负载都是以虚拟机或容器的方式部署到云环境中，并需要获得更大的可信边界。例如在金融、医疗等数据敏感性行业，客户往往希望其整个云环境，包括虚拟机、容器和应用程序中的数据都能处于机密计算环境的保护下。

作为亚太地区最早部署机密计算的云服务提供商，阿里云一直以来都在开展机密计算技术推广。在其上一代 ECS 云实例（包括 g7t、c7t 和 r7t 安全增强型实例）中，就已经通过引入英特尔® SGX，以“飞地”的形式在内存中为客户应用程序开辟出可信的TEE环境，更好地保证了重要代码和数据的机密性与完整性。

单一使用英特尔® SGX 提供的应用程序级可信边界，不仅会增加客户将全部应用程序、工作负载部署到安全云环境的难度，同时逐一对应用程序、工作负载开展改造，也会带来巨大的工作量。此时，阿里云等云服务提供商就需要寻求一种具有弹性可信边界、且易于将应用程序部署在其中的分级机密计算新方案。

为了解决这一问题，阿里云与英特尔一起，基于其强大的云计算技术积累，在全新的第八代企业级 ECS 实例 g8i 中引入了第五代英特尔® 至强® 可扩展处理器。新一代处理器所内置的英特尔® TDX 技术与 ECS g8i 实例搭载的可信平台模块（Trusted Platform Module，TPM）安全芯片相配合，可为大型互联网、新金融、医疗保健、知识产权等业务场景提供更高安全等级的数据保护能力和云上可信运行环境，进一步帮助客户实现数据可用不可见的愿景。

第五代英特尔® 至强® 可扩展处理器拥有更可靠的性能，更出色的能效。它在运行各种工作负载时均可实现显著的每瓦性能增益，在 AI、数据中心、网络和科学计算的性能和总体拥有成本（TCO）方面亦有更出色的表现。在安全性方面，第五代英特尔® 至强® 可扩展处理器提供了应用隔离、虚拟机层面的隔离、独立认证、内存保护、全内存加密等能力，并支持用户从目前市场上广泛部署的数据中心机密计算方案中进行选择。

图 1. 第五代英特尔® 至强® 可扩展处理器带来多层面的提升

借助英特尔® 虚拟机扩展（Intel® Virtual Machine Extension，英特尔® VMX）技术与英特尔® 多密钥全内存加密（Intel® Multi-Key Total Memory Encryption，英特尔® MK-TME）技术，英特尔® TDX 为云实例提供了一种被称为“信任域（Trust Domain，TD）”的全新虚拟访客环境。TD 可与其它 TD、实例，以及底层系统软件、管理软件实现相互隔离。而这些安全策略的实施，是由运行在安全仲裁模式（Secure-Arbitration Mode，SEAM）下的 TDX 安全服务模块来完成。

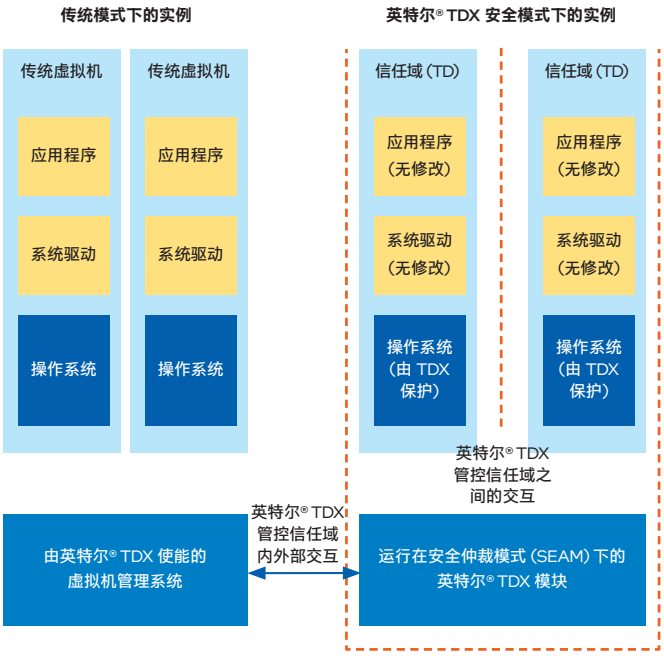

图 2. 英特尔® TDX 技术架构

这一架构中，英特尔® TDX 借助英特尔® MK-TME 技术为 TD 提供了数据机密性和完整性。英特尔® MK-TME 技术支持使用多种密钥对内存进行加密：

- 一方面，其提供的私密密钥，可用于对专用内存（放置 TD 的机密数据）进行加密；
- 另一方面，其提供的共享密钥则用于对共享内存进行加密，用于与 TD 外部的代理进行通信，以执行 I/O 操作，如网络访问、存储服务、调用管理程序服务等。

作为一种全新的机密计算技术，英特尔® TDX 使TEE环境的可信边界获得了有效扩展，从而让不同类型下的云服务，无论是 IaaS 或是 PaaS 中的云工作负载都能通过英特尔® TDX 整体纳入机密计算的数据保护之下。一般地，客户可以选择运行两种常用的机密计算方案，机密虚拟机（TD VM）和机密容器（TD CC）。机密虚拟机是运行在 TD 中的虚拟机实例，而机密容器是将机密计算与云原生容器集成，以保护 Kubernetes 上运行的敏感数据和应用程序。

阿里云第八代企业级 ECS 实例 g8i 可为客户提供机密虚拟机和机密容器两种使用模式：

1. **机密虚拟机：** 作为全球首个基于英特尔® TDX 的公共云实例服务，客户可以按需订购机密计算环境，并通过英特尔® TDX 提供的“Lift-and-Shift（直接迁移）”方法，将传统应用程序升级为机密计算应用程序。此外，客户还可以通过利用远程认证功能（如实例内的磁盘加密等）来构建更加安全的解决方案。同时，双方还合作将英特尔® TDX 引入了阿里云创建的开源 Linux 发行版 OpenAnolis 中；
2. **机密容器：** 由阿里云新实例提供的机密容器构建的基于虚拟化实例的TEE环境，能将特定容器组（例如 POD）与其他容器组及底层管理程序实现隔离以更好地保证数据安全性。同时双方也正推动基于 CoCo（云原生计算基金会“ CNCF ”的一个机密容器沙盒项目）的解决方案成为 OpenAnolis 的一部分。

无论哪种方式，客户都可以在云实例中轻松地搭建起自己的应用程序和数据，并受到可信赖的安全保护，使应用程序与数据都与外部环境隔离，以防止未经授权的访问。

此外，与同类技术相比，英特尔® TDX 具备较低的性能损耗，有助于用户在保证数据安全性的同时，降低对于负载性能的影响，这在 AI 推理等数据保护需求规模庞大、性能密集型的应用中，有着重要意义。

采用英特尔® TDX技术的 BigDL LLM 隐私保护方案

阿里云推出了采用英特尔® TDX 技术的 BigDL LLM 隐私保护方案，可以在英特尔® TDX 技术的加持下实现对分布式节点或 AI 管道的保护，从而让客户在不牺牲数据隐私的前提下将更多的数据运用到 AI 应用中，有效挖掘数据价值，为客户构建更为高效的隐私保护机器学习方案，助力大模型的广泛应用。

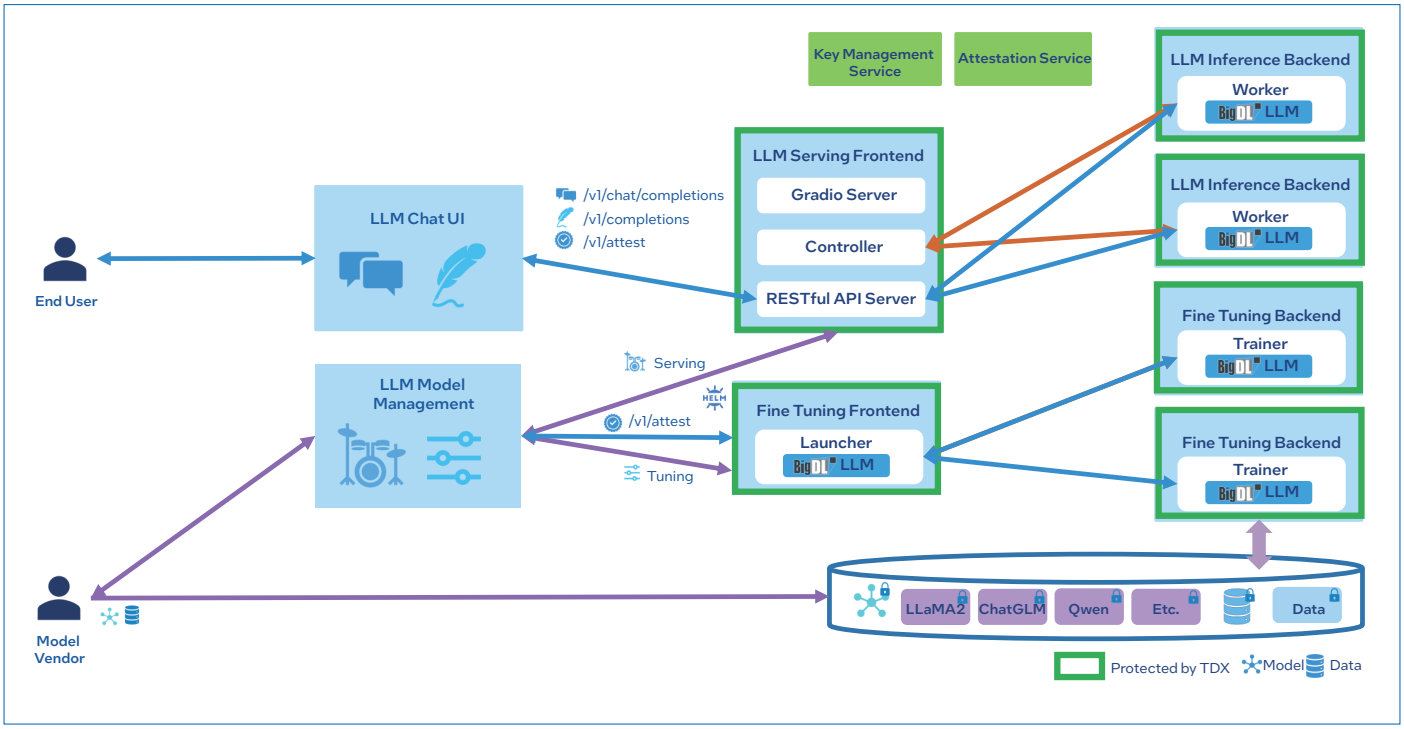

图 3. 采用英特尔® TDX 的 BigDL LLM 服务和调优架构

除了安全性之外，采用第五代英特尔® 至强® 可扩展处理器的阿里云 ECS g8i 实例还在性能、能效等方面表现出了强大的优势，助力加速云上的各种负载。以大模型推理为例，在通义千问大模型（Qwen-7 B）中，即使启用英特尔® TDX，第五代英特尔® 至强® 可扩展处理器（TDX）的性能相较于第三代英特尔® 至强® 可扩展处理器，依然可实现 3.64 倍的提升，相比第四代处理器也有 1.18 倍的提升[1]。

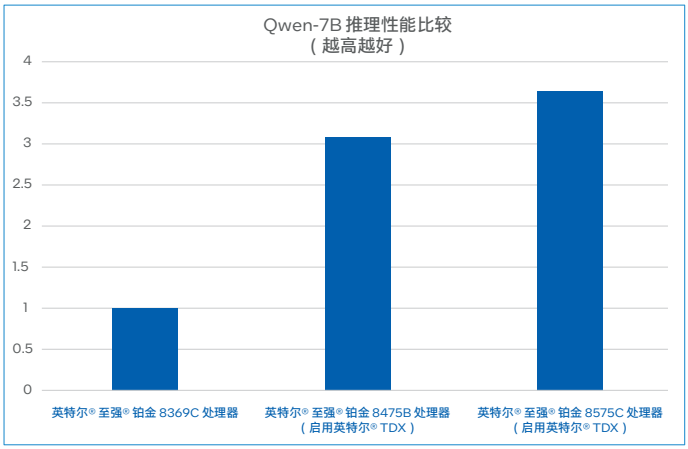

图 4. Qwen-7B 推理性能比较

[1] 数据援引自阿里云截止 2023 年 11 月的内部测试结果。测试配置 – 配置 1：英特尔® 至强® 铂金 8369C 处理器；配置 2：英特尔® 至强® 铂金 8475B 处理器，启用英特尔® TDX；配置3：英特尔® 至强® 铂金 8575C 处理器，启用英特尔® TDX。英特尔并不控制或审计第三方数据。请您审查该内容，咨询其他来源，并确认提及数据是否准确。

展望

随着云服务安全关注度的不断提升，阿里云与英特尔都深刻地意识到，促进机密计算的技术发展和普及，应用和生态也是非常关键的一环。因此，双方已在多个领域携手开展基于英特尔® TDX 技术的云服务安全实践，双方还将在更多领域进行安全领域的前沿探索，为大模型等应用提供可信赖的安全能力的支撑。