



浪潮信息

*KeyarchOS*云峦

KOS 测试报告

量化交易解决方案性能测试报告

*KeyarchOS*云峦

KOS 量化交易解决方案 性能测试报告

KeyarchOS云峦

浪潮电子信息产业股份有限公司

系统软件产品部

2023.9

目录

云峦服务器操作系统 KeyarchOS.....3

1 背景.....5

2 测试环境信息： 6

 2.1 硬件环境：6

 2.2 软件环境：6

 2.3 测试工具：6

 2.4 网路拓扑..... 7

3 调优参数..... 8

 3.1 BIOS 参数:.....8

 3.2 OS 参数:..... 9

 3.2.1 配置网卡中断聚合参数：9

 3.2.2 CPU 亲和性设置10

4 测试结果.....12

 4.1 SWIFT-2200 网卡 KOS 与 Kylin 对比测试结果：12

 4.1.1 TCP 测试结果：12

 4.1.2 UDP 测试结果：13

 4.2 KOS 系统下 SWIFT-2200 网卡与 Solarflare SFC9250 对比测试结果：14

 4.2.1 TCP 测试结果： 14

 4.2.2 UDP 测试结果： 15

云峦服务器操作系统 KeyarchOS

云峦服务器操作系统 KeyarchOS（简称 KOS）是浪潮信息依托十余年的高端主机操作系统研发经验和龙蜥社区最新的成果，打造的面向智算时代的服务器操作系统，支持 x86、ARM 等多种芯片架构和计算场景，性能和稳定性通过了多项权威评测和大规模部署验证。通过软硬协同优化，基于 KOS 的人工智能、虚拟化、云原生等应用在开发效率、运行效率、扩展性等方面都得到了增强。在虚拟化场景，虚拟机迁移加速提升 8 倍；在云原生场景，容器密度提升 10 倍。

KOS 是基于龙蜥操作系统开发的衍生商业版，具备高效软硬协同、稳定可靠、全天候运维、全栈安全可信四大特点，同时依托自身覆盖云、数、智、边等业内最全整机产品线，实现整机系统架构优化，最大程度发挥多元算力价值。

在软硬协同方面，KOS 通过设备管理、设备发现等一系列措施，简化了不同类型设备的管理，目前已与 100+ 服务器整机、350+ 主流 CPU、GPU、FPGA、网卡、存储及网络设备完成兼容认证，实现主流数据库、中间件、虚拟化、容器全覆盖；同时，KOS 也提供了异构算力的编程模型和高效开发库，为上层提供统一的开发接口和运行时支持，屏蔽底层计算的差异，简化上层业务使用各种通用计算设备和专用加速设备。

KOS 通过安全可靠、**RAS** 增强、完善的集群高可用支撑能力，提升整机系统的可靠性。**KOS** 建立了从硬件启动、内核启动、驱动加载、应用执行完整可信链，为用户实现系统运行时持续保护、升级时安全检查、故障或破坏时自动恢复，保障系统安全性。

浪潮信息充分利用整机厂商的硬件优势，在操作系统层面联合固件、硬件进行协同创新，对整机系统 **RAS** 特性进行了增强。其中，**KOS** 具备关键数据冗余机制，可对核心数据内存进行镜像保障业务关键数据可靠，而增强性容错能力则可将引起系统宕机的 **UCE** 进行降级容错处理，大幅降低了系统宕机率。此外，**KOS** 深度挖掘硬件的 **RAS** 能力，支持 **CPU**、内存等核心部件的热替换，能够有效提升系统可维护性，支持内核、应用软件的热升级，能够有效保障用户业务连续性。

1 背景

在证券交易的场景下，从业务需求角度出发，他们更加关心任务响应时间，处理速度。从而在单位时间内能够获得更快的响应速度，更低的响应时延。这样就能更加及时高效的处理数据，带来更大的商业价值

因此在这种用户需求场景下，需要更加快速且高效的一体化系统，从各个方面带来更低的时延。包括但不限于：

- 硬件平台的处理时延
- 操作系统的处理时延
- 应用软件的处理时延
- 网络的传输时延

本文档主要从硬件平台及操作系统的角度给出优化建议。

2 测试环境信息：

2.1 硬件环境：

服务器	CPU 型号	Mem	Net
CS5280H2	Hygon C86 7380 32-core Processor	32G*32	SWIFT-2200 Solarflare SFC9250
CS5280H2	Hygon C86 7380 32-core Processor	32G*30	SWIFT-2200 Solarflare SFC9250

2.2 软件环境：

操作系统	驱动
KOS 5.8	SWIFT-2200 驱动 SFC9250 驱动
Kylin V10 SP2	SWIFT-2200 驱动 SFC9250 驱动

2.3 测试工具：

sfnt-pingpong:

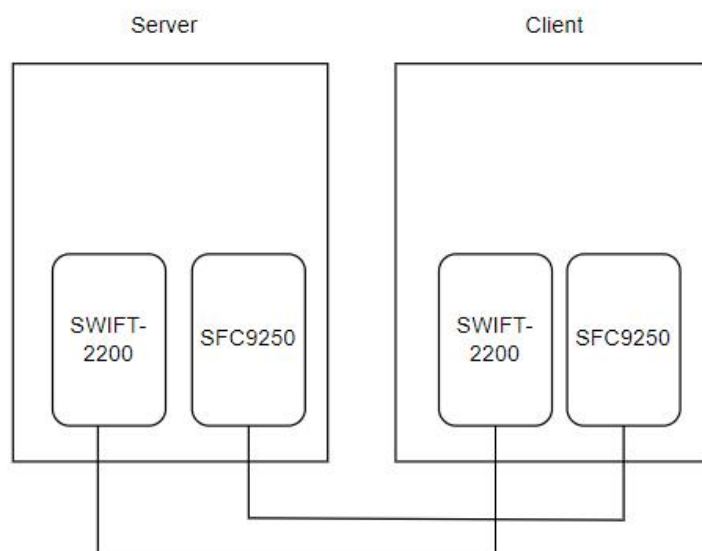
sfnttest 是 Solarflare 开发的一套用于测量网络性能的工具。具备 sfnt-stream、sfnt-pingpong 等子组件。它与传统的

工具不同之处在于它提供的一些附加选项，这些选项允许观测特定的应用程序行为。

sfnettest is a suite of tools for measuring network performance developed by Solarflare. It differs from existing tools in some of the options it provides, which allow probing of particular application behaviours.

2.4 网路拓扑

在本次测试场景中，使用两台完全相同配置的服务器 (CS5280H2)。因此在硬件环境下在处理器、内存、硬盘、BIOS 等硬件方面具有相似的性能水平。两台服务器均各自搭载了两个网卡，每个同品牌的网卡之间采用了光纤直连的连接方式进行通信。



3 调优参数

3.1 BIOS 参数:

在处理网卡延迟时，BIOS 参数可以对性能和响应时间产生显著影响。常见的 BIOS 设置包括 PCI/PCIe 配置，电源管理选项，中断设置，虚拟化支持和高级芯片组设置。PCI/PCIe 配置决定了网卡的带宽和错误处理，正确设置可以减少延迟。电源管理选项可影响网卡进入节能状态的频率，在延迟敏感场景中可能需要禁用以保持性能。中断设置包括中断分配和轮询参数，优化它们有助于减少中断相关的延迟。虚拟化支持参数对虚拟机性能和延迟有影响，可根据需要启用。高级芯片组设置可能包括 DMA 设置，优化可降低数据传输延迟。在降低网卡延迟时，根据具体需求精心配置这些 BIOS 参数至关重要，以实现最佳性能和响应时间。

参数	建议设置	说明
Hyper Threading Technology	Disabled	处理器超线程
CPU P-State Control	P0	CPU 处理器依据运算量负载轻重
CPU C-State Control	Disabled	CPU 节能省电模式
Channel interleaving	Enabled	Channel 交织模式，使能后能够提升 ddr 访问性能
Core Performance Boost	Enabled	内核超频设置
Determinism Slider	Performance	选择电源功率确定性模式
Bank Interleaving	Enabled	Bank 交织模式，使能后能够提升 ddr 访问性能
SR-IOV Support	Disabled	PCIE 虚拟化特性支持

3.2 OS 参数:

3.2.1 配置网卡中断聚合参数:

网卡的中断聚合，也称为中断共享或中断汇聚，是一项常用的技术，旨在减少系统中断的数量，提高系统性能和效率。其作用如下：

1.降低中断处理开销： 在传统设置中，每个网络接口通常都有自己的中断，当有数据包到达时，会触发中断，这会增加 CPU 的负担。中断聚合将多个网络接口的中断合并成一个，减少了中断处理的开销。

2.提高系统性能： 通过减少中断数量，系统的 CPU 资源可以更专注于应用程序处理，而不是处理大量的网络中断。这可以提高整个系统的性能和响应速度。

3.减少中断竞争： 中断共享可以避免多个网络接口之间的中断竞争，因为它们现在共享一个中断。这可以降低系统中的竞争情况，提高了稳定性。

4.简化管理： 管理多个网络接口的中断可能会变得复杂。中断聚合简化了这一管理过程，因为只需要关注一个聚合中断，而不是多个独立的中断。

因此，网卡的中断聚合有助于降低系统中断的开销，提高性能，减少竞争，同时简化管理。这对于高性能网络应用和需要处理大量网络流量的服务器非常有益。

低时延模式下的中断参数配置为 0，使用配置命令：

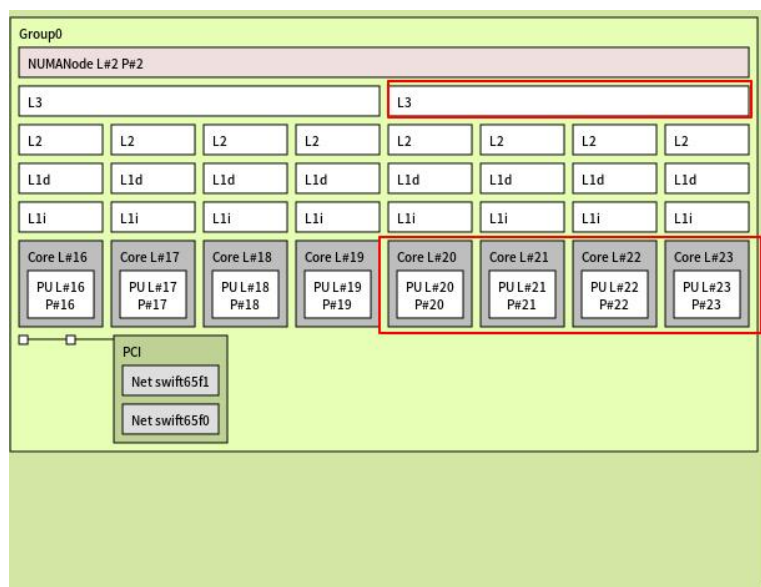
```
ethtool -C swift1f0 rx-usecs 0
```

3.2.2 CPU 亲和性设置

CPU 亲和性是多核处理器上任务分配的关键概念，通过将特定任务与特定核心绑定，提高性能和资源利用效率。它

功能强大，首先，可显著提升性能，减少任务之间资源竞争，适用于计算密集型应用。其次，有助于精确控制资源分配，确保关键任务获得足够处理能力，对实时和多线程应用至关重要。此外，减少了资源浪费，避免了频繁上下文切换，提高系统资源利用。降低延迟也是其功能之一，适用于需要快速响应的应用，如游戏和实时媒体处理。最后，保持负载平衡，防止核心过度负荷。总而言之，CPU 亲和性是优化多核处理器性能的重要工具，适用于大数据处理、科学计算和实时响应应用。

具体到本次测试用例，将测试进程绑定在与网卡中断所在 CPU 核心共享 L3cache 的 CPU 核心上，如图，网卡中断所在 cpu21,因此将进程绑定在 cpu20 cpu22 cpu23 上即可。

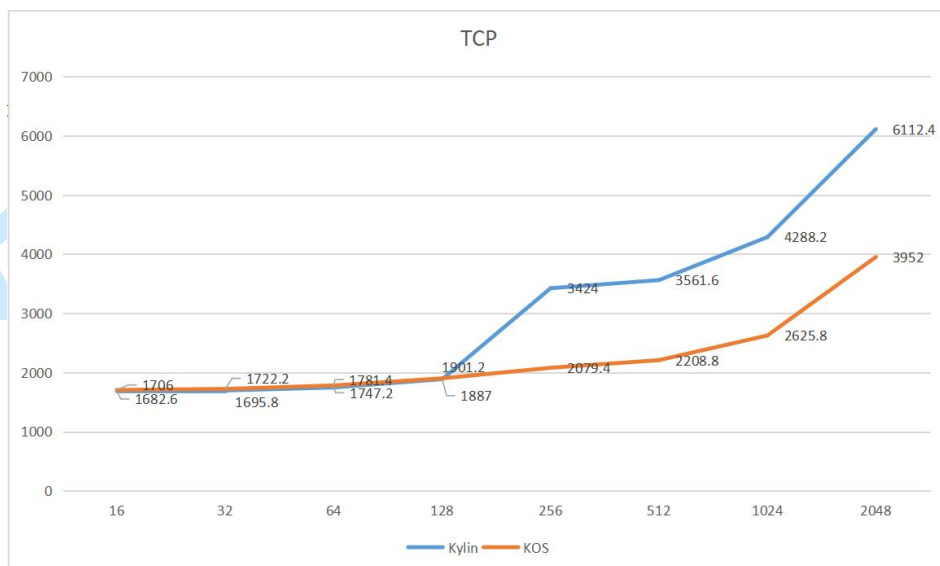


4 测试结果

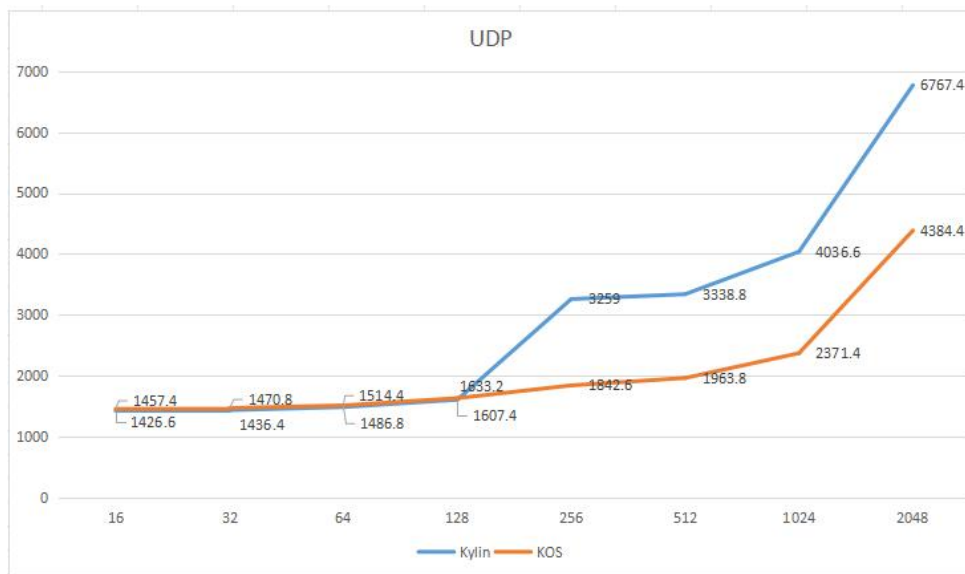
4.1 SWIFT-2200 网卡 KOS 与 Kylin 对比测试结果：

在相同的硬件环境下，KOS 与 Kylin V10 采用相同的操作系统软件配置（包括内核参数以及软件包）。使用相同的测试工具与测试参数进行横向对比测试。

4.1.1 TCP 测试结果：



4.1.2 UDP 测试结果：



测试结论一：

本次测试分别在 KOS 5.8 和 Kylin 操作系统 V10 操作系统基础上，基于 Solarflare 研发的 SFNT 测试工具对中科驭数的网卡延迟性进行了测试。本次测试工具具体使用的是 sfnt-pingpong 命令，测试相同硬件场景、配置参数下的 KOS 与 Kylin 的延迟性能（数值越低，性能越好）。测试最终结论为：

1) 基于 KOS 5.8 和 KylinV10 在相同的硬件环境与软件参数环境下，大包（包大小>128K）KOS 的延迟性能比 Kylin 高 40%-50%。

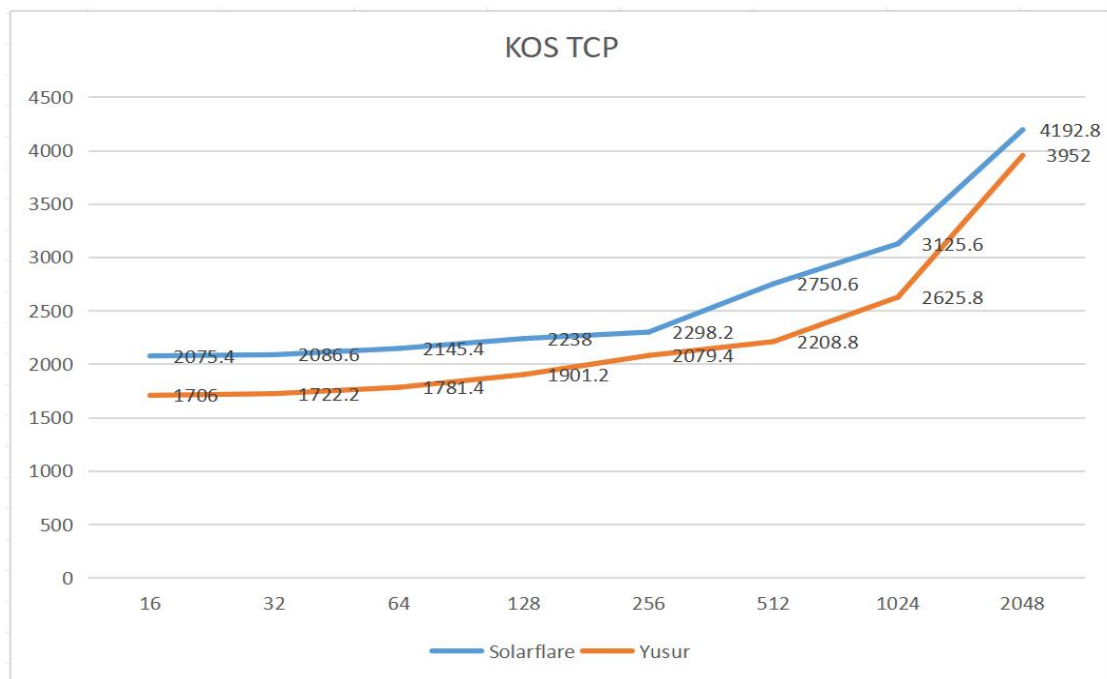
2) 基于 KOS 5.8 和 KylinV10 在相同的硬件环境与软件参数环境下，小包（包大小≤128K）的延迟性能基本一致。

3) TCP 与 UDP 测试结果表现不存在差异。KOS 在大包的性能均高于 Kylin40%以上。

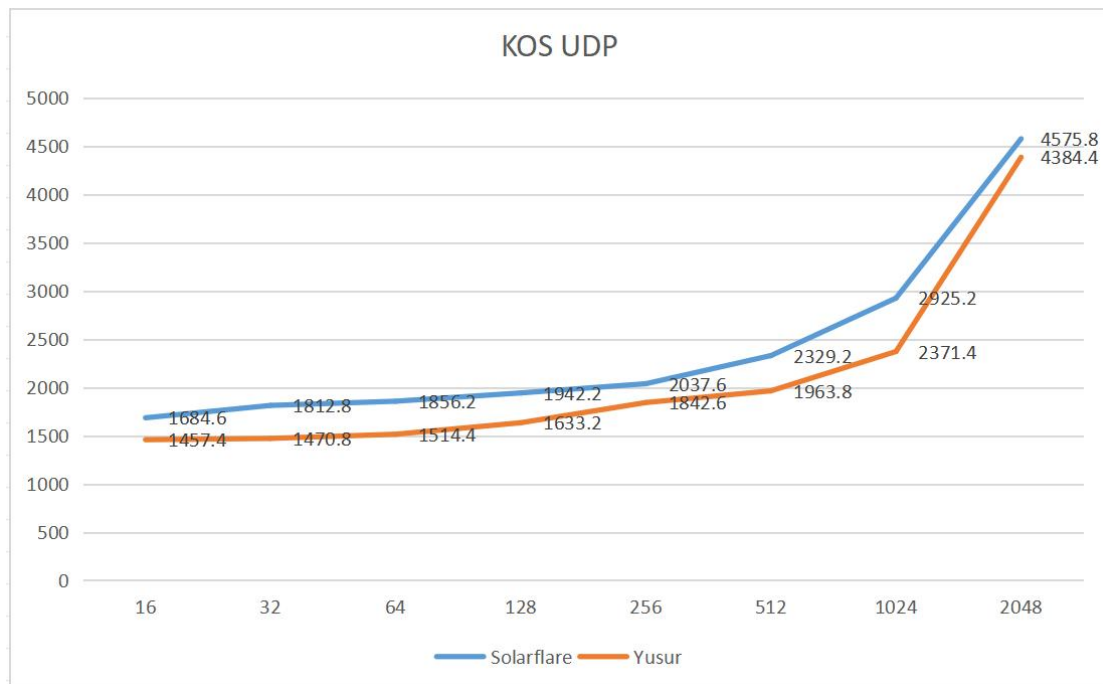
4.2 KOS 系统下 SWIFT-2200 网卡与 Solarflare SFC9250 对比测试结果：

在相同的软件环境下，对中科驭数的网卡与同等级的 Solarflare 网卡进行横向对比测试。除网卡外，其他的硬件配置保持一致，例如 BIOS 配置，操作系统参数，测试参数等完全一致。

4.2.1 TCP 测试结果：



4.2.2 UDP 测试结果:



测试结论二:

本次测试在 KOS 5.8 的软件环境以及除网卡外其他的硬件配置完全一致的场景下，基于 Solarflare 研发的 SFNT 测试工具对网卡的延迟性进行了测试。测试工具使用的是 sfnt-pingpong，测试相同操作系统、BIOS、BMC 等其他硬件环境下的中科驭数与 Solarflare 网卡延迟性能（数值越低，性能越好）。测试最终结论为：

1) 中科驭数的网卡延迟性能相对 Solarflare 的性能整体高 10%以上。