

用户态协议栈在移动云极速型云盘应用实践

目录

●项目背景

下一代云盘面临的挑战

Gazelle：通用高性能用户态协议栈

●用户态协议栈上云之路

多线程模式：实现相比内核15%的提升

协议栈运维可视化

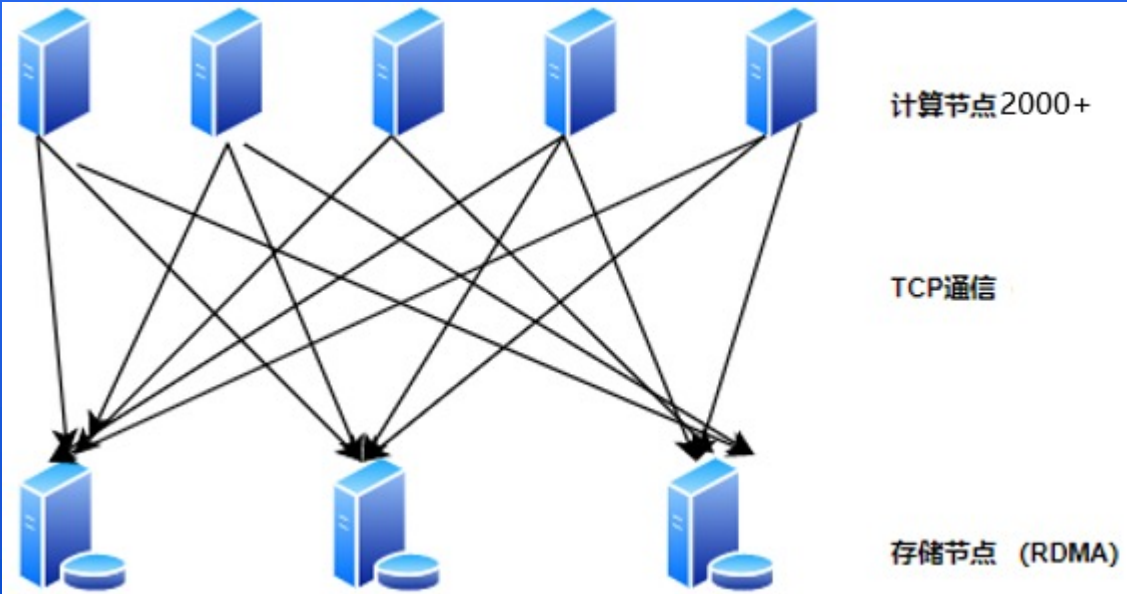
●社区共建

社区共建：深入参与gazelle建设

项目背景

下一代云盘面临的挑战

公有云场景中，计算节点到存储节点网络的有特殊要求：**大规模**（10W+并发）、**兼容性（POSIX API）**、**高性能（IOPS）**，同时还要综合考虑成本支出；以下是主流通信方案的对比：



	组网规模	连接数	性能	成本	应用场景
RDMA IB	数千节点	几百~数万	最高	高	高性能计算
RDMA ROCE V2	数百节点	几千	高	中	高性能计算、存储内部小规模通信
内核态 TCP	数千节点	10万以上	一般	低	低规格存储
用户态 TCP	数千节点	10万以上	高	低	计算节点与存储节点通信

应对方案：

- 移动云下一代云盘将逐渐把通信和存储负载卸载到DPU上。从而提高整体的处理效率。
- 为了解决RDMA大规模组网困难问题，实现降本增效，目前采用用户态TCP+DPU传输方案的极速云盘已在规划中。

Gazelle: 通用高性能用户态协议栈



高性能

- 零拷贝
- 无锁线程协议栈
- 灵活scale-out, 自适应调度, 亲和管理

通用性

- POSIX API兼容
- 支持RTW/RTC网络模型
- 兼容内核协议栈

即插即用

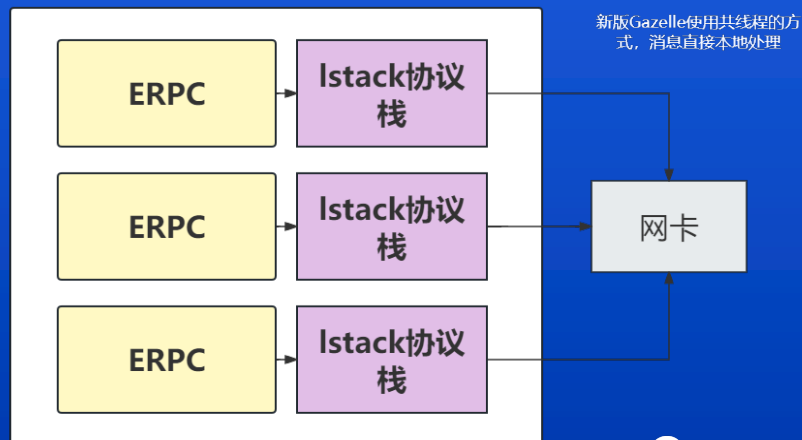
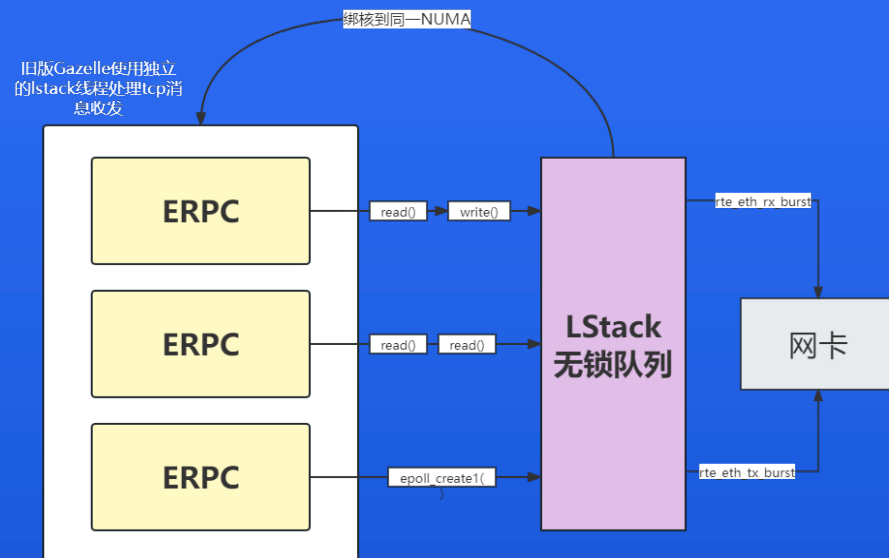
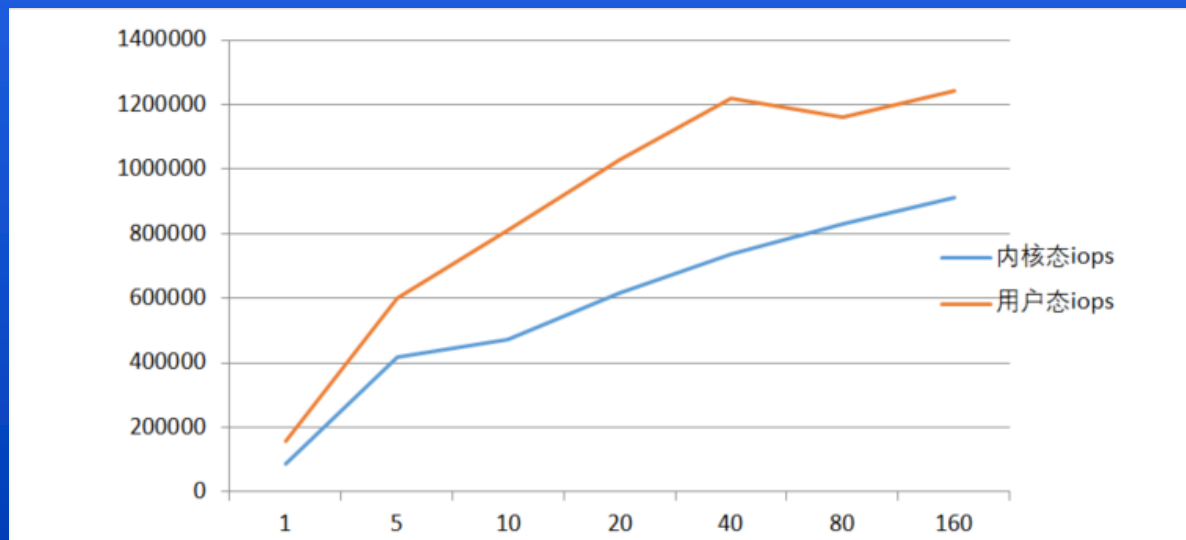
- LD_PROLOAD应用免配套

用户态协议栈上云之路

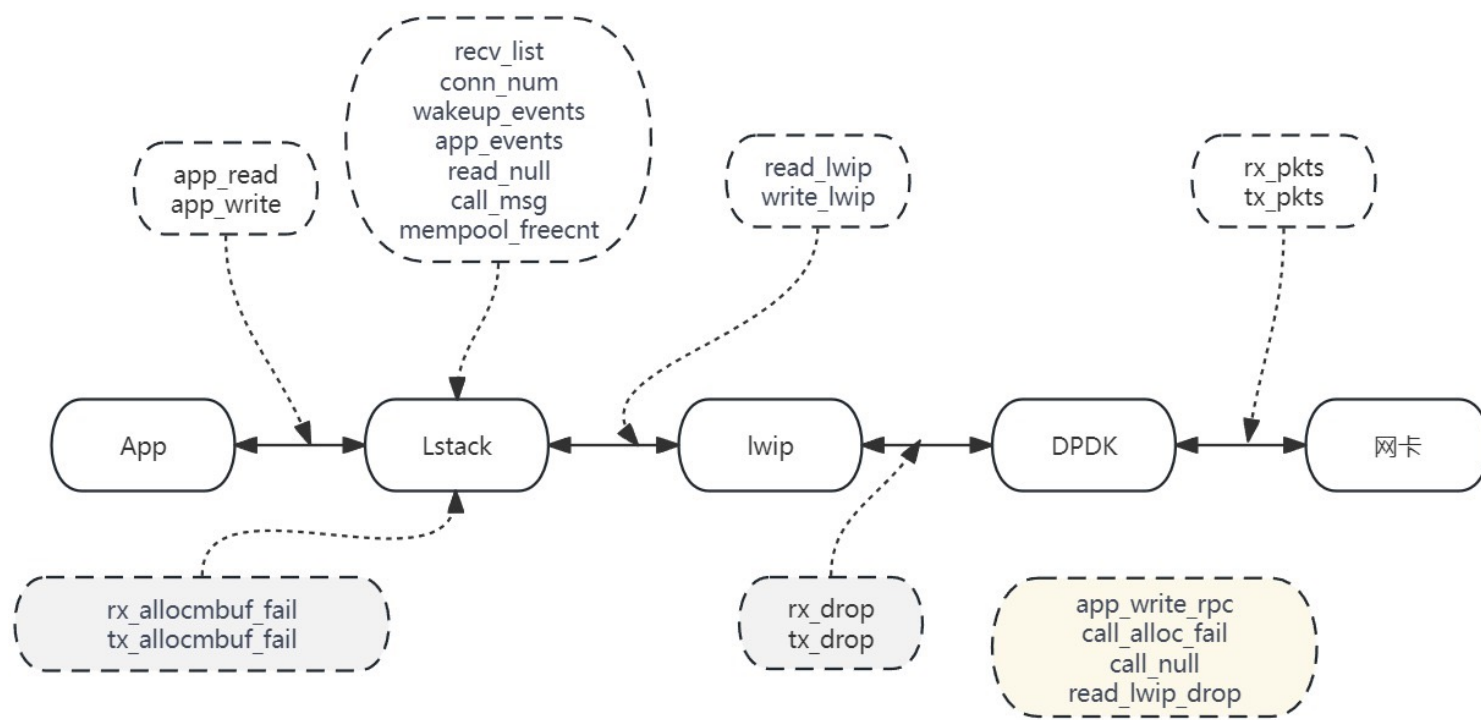
共线程模式：实现相比内核15%提升

共线程模型优势：

- 免去了线程间耦合，减少了伪共享等进程间通信开销。
- 避免频繁地在多个任务之间切换，减少因切换任务而产生的额外开销，从而提高系统的运行效率。
- 不依赖于传统的加锁机制来保证数据的一致性和并发控制，这样可以避免因锁竞争带来的性能损耗，提高系统的并发处理能力。



协议栈运维可视化



协议栈视角的运维可视化

- 网卡指标可视化
- 分层信息可视化

连接粒度运维可视化

- 收发包计数、链接状态、option属性等

社区共建

社区共建：深入参与gazelle建设

在过去一年中国移动云能力中心投入**10+人**参与Gazelle社区合作，提交多个核心功能；

1、支持ipv6协议

- a) 增加客户端api对ipv6协议的支持；
- b) 优化gazelle对ip协议的兼容性，支持ipv4、ipv6双栈模式；
- c) 支持ICMPv6协议；
- d) 支持TSO分段；
- e) 增加对ipv6连接的监控输出；

2、支持bond6模式

- a) 优化参数配置；
- b) 增加bond6模式配置；

3、支持VLAN通信

- a) 增加lstack和LwIP对vlan头的封装解析处理；
- b) 支持网卡的硬件卸载（VLAN_STRIP、VLAN_INSERT）；
- c) 增加ltran对vlan头的解析处理；

4、dfx相关

- a) 增加ipv6相关的gazellectl运维能力支持
- b) 增加gazellectl -s中tcp乱序统计量

master

zhengjiebing

开始日期

结束日期

搜索

排序方式: 按提交时间

2023-12-07 (1)

Z

dfx: gazellectl -c support ipv6

zhengjiebing 提交于 8天前

61bf272

浏览文件

2023-12-05 (1)

Z

add struct gz_addr_t

zhengjiebing 提交于 10天前

7d721f7

浏览文件

2023-11-28 (2)

Z

support vlan offload

zhengjiebing 提交于 16天前

a449111

浏览文件

Z

add vlan_id in netif

zhengjiebing 提交于 17天前

fcdf8a0

浏览文件

2023-11-23 (2)

Z

ltran support vlan

zhengjiebing 提交于 21天前

0b7daa1

浏览文件

Z

cfg: rm ipv6_enable

zhengjiebing 提交于 22天前

d0c8c25

浏览文件

2023-11-18 (2)

Z

enable ipv6

zhengjiebing 提交于 26天前

480ef35

浏览文件

Z

ipv6

zhengjiebing 提交于 27天前

894c59a

浏览文件

2023-11-16 (1)

THANKS

THANKS

THANKS