火山引擎

Virtio Crypto Akcipher 虚拟化加速和 Crypto Over Fabrics 资源池化方案

皮振伟 <pizhenwei@bytedance.com>

2024-10

CLK 19 湖北 - 武汉

WWW.VOLCENGINE.COM

01 问题背景

非对称加密的负载问题

火山引擎

非对称的密集块计算

- 主流的非对称加密算法 RSA 、 ECDSA 、 DH 、 SM2 、 ECDH 等都使用密集的块运算
- 以 nginx 作为 HTTPS 接入网关为例，约 90% 的 CPU 消耗在加解密阶段
- 密集的计算对 CPU 的 hyper-thread 非常不友好， 2 hyper-thread 几乎退化成一个 core
- 以 3.1G Hz 的 Intel(R) Xeon(R) Platinum 8260 为例，单核 RSA2048 的短链接 QPS 约 1K

硬件加速方案

- 例如 Intel QAT 4xxx 卡，平均单卡 RSA2048 的短链接 QPS 约 100K
- 其他加解密加速卡
- FPGA 方案

虚拟化场景下的非对称加密问题

火山引擎

- 热迁移
- 宿主机硬件绑定
- 资源弹性切分
- CPU& 加解密性能不对称
- 监控 & 告警
- QoS 控制
- Virtio Crypto 于 2016 年底合入 Linux 内核，但是缺少非对称加密支持

02 Virtio Crypto Akcipher

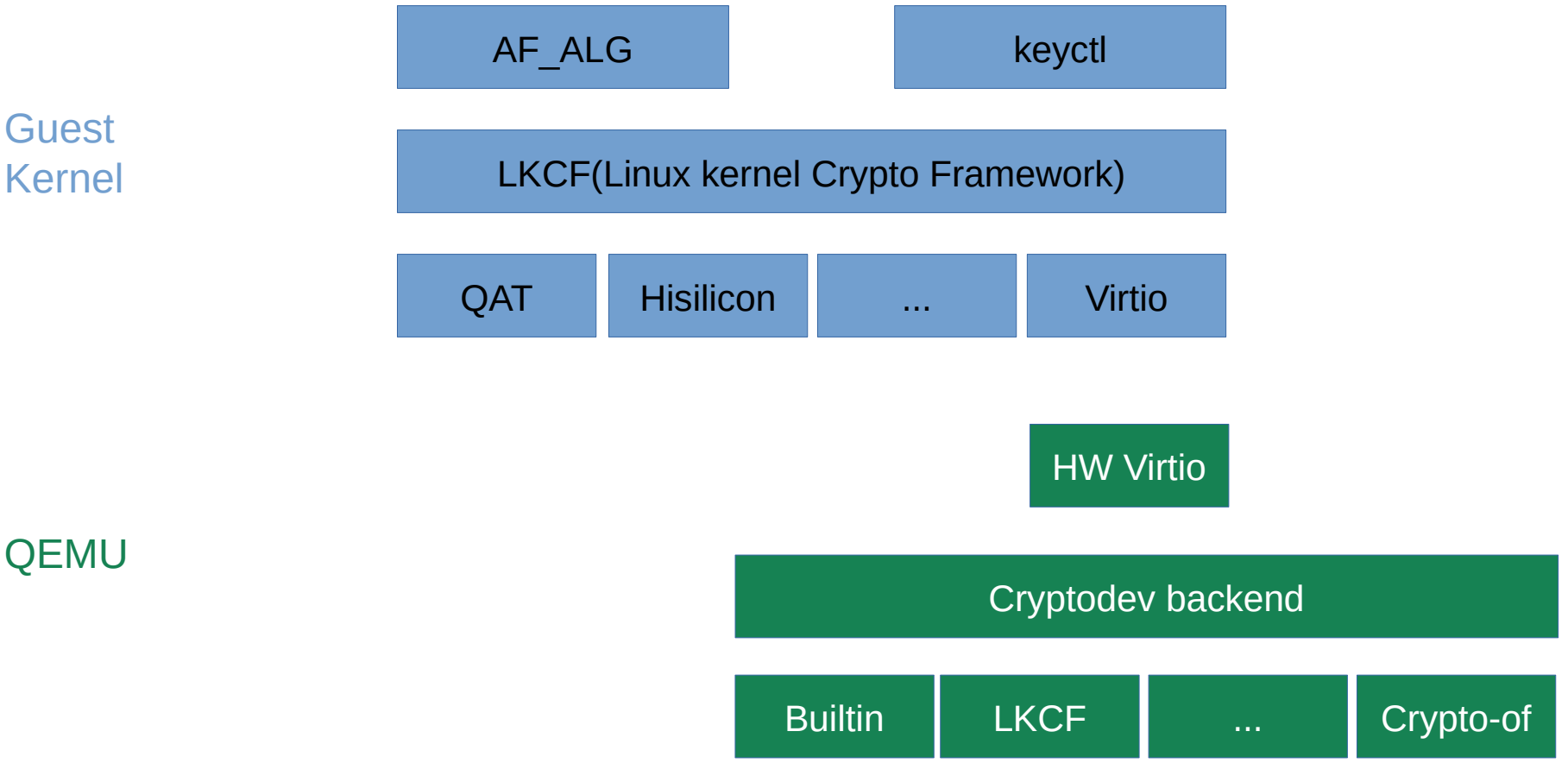
Virtio Crypto Akcipher Spec

火山引擎

- 新增 AKCIPHER service 类型
  - CREATE/DESTROY akcipher session
  - ENCRYPT/DECRYPT/SIGN/VERIFY 操作类型
  - VIRTIO_CRYPTO_KEY_REJECTED
- 新增 RSA/ECDSA 协议支持
- 协议于 2022 年初合入 virtio-spec

https://github.com/oasis-tcs/virtio-spec/commit/abbe8afda8db2089e4498808939311d9a1925cb4

Virtio Crypto



Cryptodev backend

火山引擎

- 新增 AKCIPHER service 后端支持
- 新增 statistics 监控，使用 QMP 命令 query-cryptodev 获取实时监控数据
- 新增 QoS 控制，支持 throttle-bps 和 throttle-ops 同时作用
- 支持 asynchronous 模式

https://github.com/qemu/qemu/blob/master/backends/cryptodev.c

QEMU crypto akcipher

火山引擎

```
int qcrypto_akcipher_encrypt(QCryptoAkCipher *akcipher,
                             const void *in, size_t in_len,
                             void *out, size_t out_len, Error **errp)
{
    const QCryptoAkCipherDriver *drv = akcipher->driver;

    return drv->encrypt(akcipher, in, in_len, out, out_len, errp);
}

...

int qcrypto_akcipher_sign(QCryptoAkCipher *akcipher,
                          const void *in, size_t in_len,
                          void *out, size_t out_len, Error **errp)
{
    const QCryptoAkCipherDriver *drv = akcipher->driver;

    return drv->sign(akcipher, in, in_len, out, out_len, errp);
}

...
```

- 支持 gcrypt
- 支持 nettle

https://github.com/qemu/qemu/blob/master/crypto/akcipher.c

Cryptodev backend builtin

火山引擎

- 新增 QEMU crypto akcipher 后端支持
- builtin 后端驱动并不能提升性能，但适合 debug 和功能展示

https://github.com/qemu/qemu/blob/master/backends/cryptodev-builtin.c
https://github.com/qemu/qemu/blob/master/crypto/akcipher*

Cryptodev backend LKCF

火山引擎

- 使用 keyctl 调用 LKCF, 进一步使用本地 QAT 等硬件加速
- 不支持 AF_ALG 异步模式
- keyctl 是同步函数调用，为了提升并发度，需要使用多线程

https://github.com/qemu/qemu/blob/master/backends/cryptodev-lkcf.c

Libvirt Cryptodev 支持

火山引擎

```
<devices>
  <crypto model='virtio' type='qemu'>
    <backend model='builtin' queues='1'/>
  </crypto>
</devices>
```

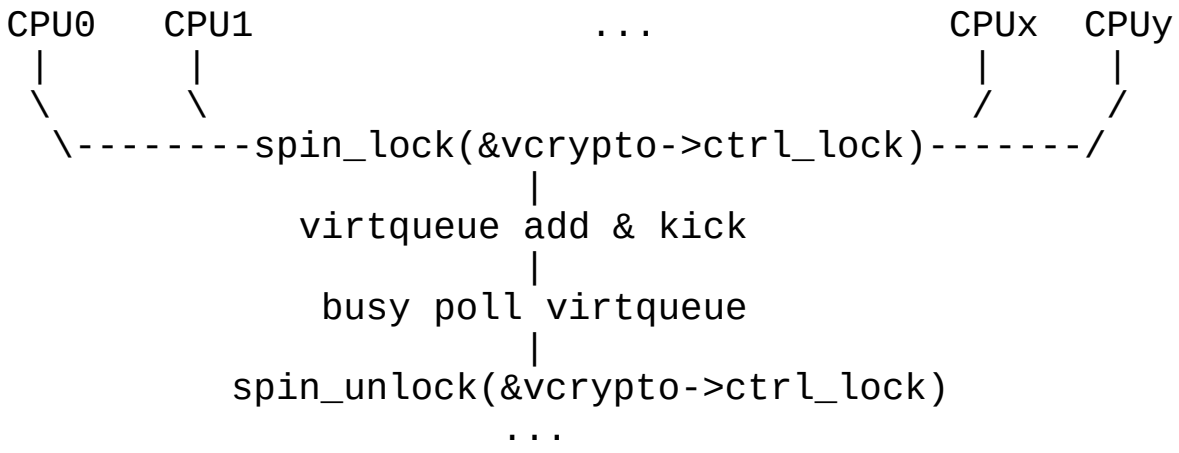
Linux kernel 支持 virtio crypto akcipher

火山引擎

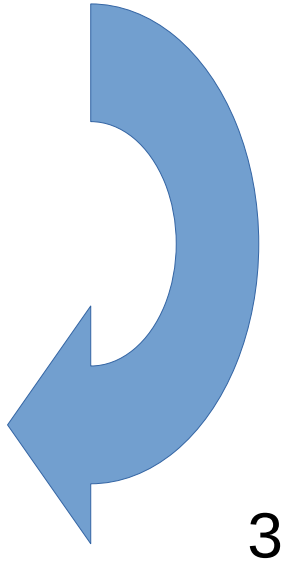
```
        {
                .algonum = VIRTIO_CRYPTO_AKCIPHER_RSA,
                .service = VIRTIO_CRYPTO_SERVICE_AKCIPHER,
                .algo.base = {
                        .encrypt = virtio_crypto_rsa_encrypt,
                        .decrypt = virtio_crypto_rsa_decrypt,
                        .base = {
                                .cra_name = "rsa",
                                .cra_driver_name = "virtio-crypto-rsa",
                                .cra_priority = 150,
                                .cra_module = THIS_MODULE,
                                .cra_ctxsize = sizeof(struct virtio_crypto_akcipher_ctx),
                        },
                },
                .algo.op = {
                        .do_one_request = virtio_crypto_rsa_do_req,
                },
        },
        {
                .algonum = VIRTIO_CRYPTO_AKCIPHER_RSA,
                .service = VIRTIO_CRYPTO_SERVICE_AKCIPHER,
                .algo.base = {
                        .base = {
                                .cra_name = "pkcs1pad(rsa,sha1)",
                                .cra_driver_name = "virtio-pkcs1-rsa-with-sha1",
                                .cra_priority = 150,
                        },
                },
        },
```

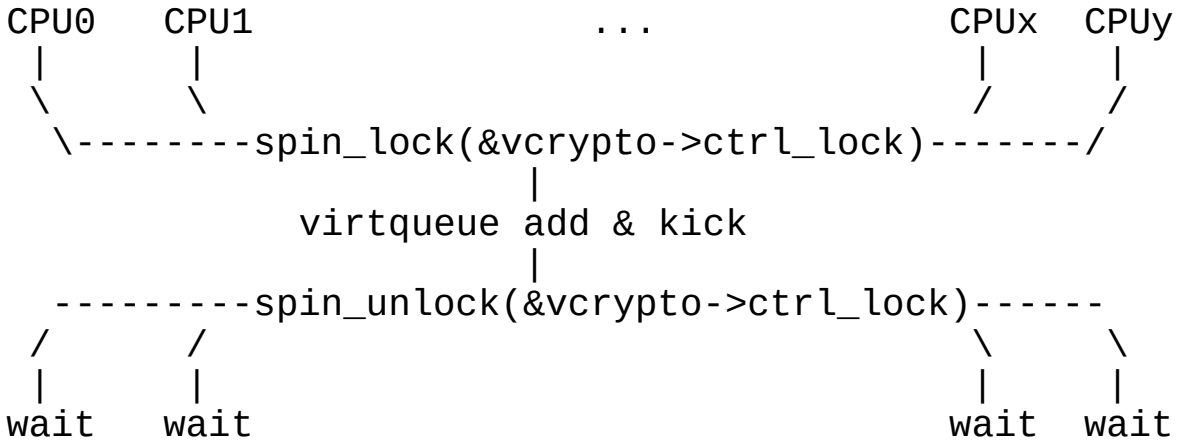
https://github.com/torvalds/linux/blob/master/drivers/crypto/virtio/virtio_crypto_akcipher_algs.c

Linux kernel 优化 virtio crypto control queue



800% CPU, ~40K/s QPS





300% CPU, ~200K/s QPS

https://github.com/torvalds/linux/commit/977231e8d45657871a86fe3c7bed94921d04e447

03 Crypto Over Fabrics

Crypto Over Fabrics

火山引擎

- Crypto-Manager： 负责资源管理和调度
- Crypto-Server：提供 crypto 服务
- Crypto-client：使用 crypto 服务

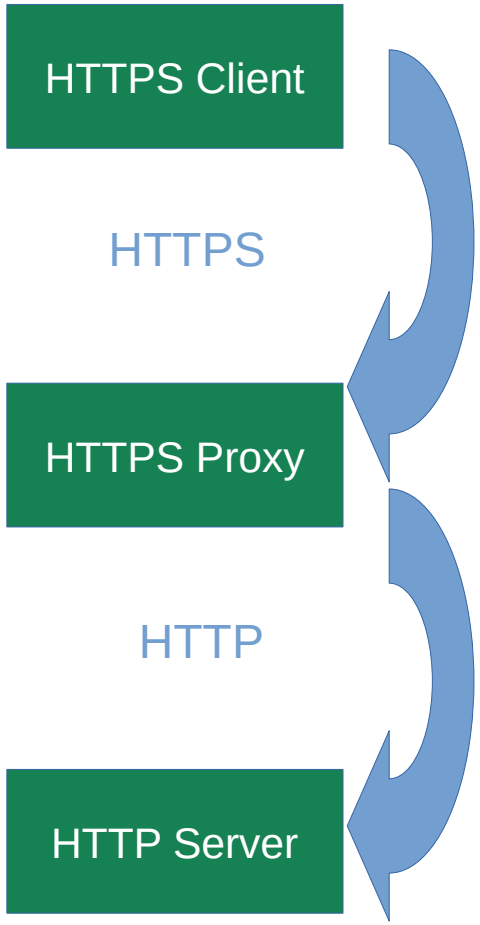
ServerA 向 Manager 注册，提供 X QPS 服务；管控节点向 Manager 申请 Y QPS 的 quota， Manager 返回 ServerA 的服务地址； QEMU 使用 Crypto-of 的 client 连接到 Crypto-Server 服务，对虚拟机提供 akcipher 加速服务。

04 性能表现

nginx&lkcf-engine

```
worker_processes 80;
events {
        use epoll;
        worker_connections 8192;
        multi_accept on;
        accept_mutex on;
}
ssl_engine lkcf-engine;

http {
        server {
                listen 443 ssl reuseport backlog=131072 so_keepalive=off;
                keepalive_timeout 0s;
                ssl_verify_client off;
                ssl_session_tickets off;
                ssl_ciphers AES128-SHA;
                ssl_prefer_server_ciphers on;
                ssl_certificate server.crt;
                ssl_certificate_key server.key;
                location / {
                        root /root/ng-benchmark;
                        index index.html;
                }
                location /proxy {
                        proxy_pass http://localhost/bg/;
                }
        }
}
```



https://github.com/bytedance/openssl-lkcf-engine

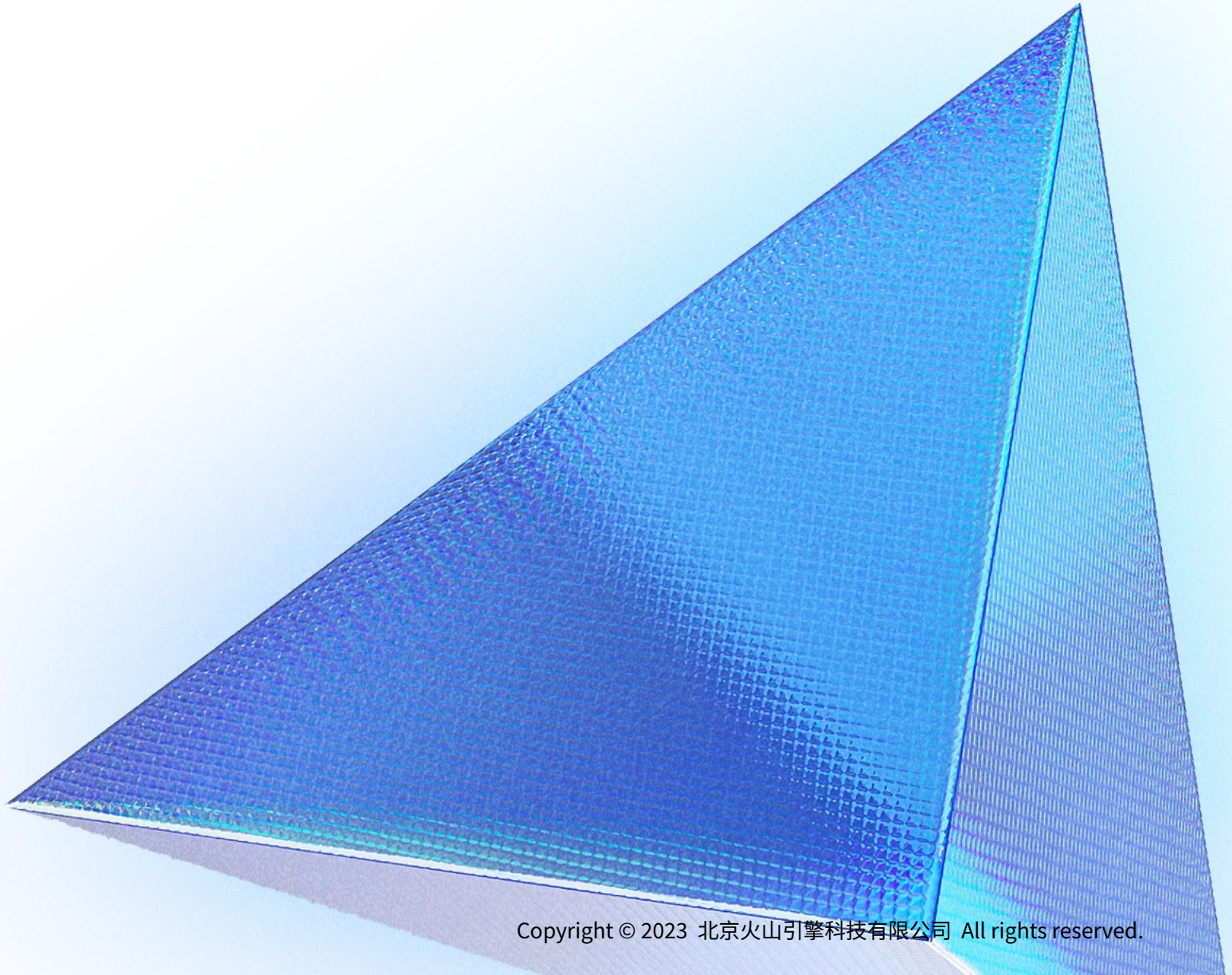
2C4HT 的性能表现

火山引擎

	访问 proxy	通过 proxy 访问后端
CPU	2207	2082
Virtio crypto + crypto-of 卸载	7902	6314

火山引擎

THANKS



皮振伟 <pizhenwei@bytedance.com>

2024-10

CLK 19 湖北 - 武汉

WWW.VOLCENGINE.COM