(12) 按照专利合作条约所公布的国际申请

(19) 世界知识产权组织
国 际 局

(43) 国际公布日
2025 年 6 月 19 日 (19.06.2025)

WIPO | PCT

(10) 国际公布号
WO 2025/124178 A1

(51) 国际专利分类号:
H04L 47/125 (2022.01)

(21) 国际申请号: PCT/CN2024/135825

(22) 国际申请日: 2024 年 11 月 29 日 (29.11.2024)

(25) 申请语言: 中文

(26) 公布语言: 中文

(30) 优先权:
202311703070.9 2023年12月12日 (12.12.2023) CN

(71) 申请人:天翼云科技有限公司(**CHINA TELECOM CLOUD TECHNOLOGY CO., LTD.**) [CN/CN]; 中国北京市东城区青龙胡同甲1号、3号2幢2层205-32室 100007 (CN)。

(72) 发明人:刘邦运(**LIU, Bangyun**); 中国北京市东城区青龙胡同甲1号、3号2幢2层205-32室 100007 (CN)。 傅旭明(**FU, Xuming**); 中国北京市东城区青龙胡同甲1号、3号2幢2层205-32室 100007 (CN)。 周玉波(**ZHOU, Yubo**); 中国北京市东城区青龙胡同甲1号、3号2幢2层205-32室 100007 (CN)。 张义飞(**ZHANG, Yifei**); 中国北京市东城区青龙胡同甲1号、3号2幢2层205-32室 100007 (CN)。 吕玉超(**LYU, Yuchao**); 中国北京市东城区青龙胡同甲1号、3号2幢2层205-32室 100007 (CN)。 韩国军(**HAN, Guojun**); 中国北京市东城区青龙胡同甲1号、3号2幢2层205-32室 100007 (CN)。

(74) 代理人: 北京润泽恒知识产权代理有限公司(**BEIJING RUN ZEHENG INTELLECTUAL PROPERTY LAW FIRM**); 中国北京市海淀区中关村南大街甲18号北京国际C座6层606 100081 (CN)。

(81) 指定国(除另有指明，要求每一种可提供的国家保护): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CV, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IQ,

(54) **Title:** METHOD AND APPARATUS FOR SOLVING IPSEC ELEPHANT FLOW ON BASIS OF DPDK

(54) 发明名称: 一种基于DPDK的解决IPSec大象流的方法与装置

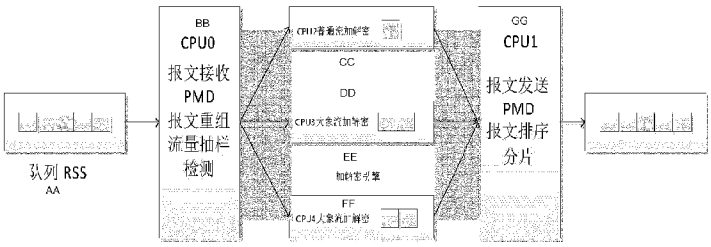

图 1

AA Queue RSS
BB CPU0: message receiving; PMD; message recombination; traffic sampling detection
CC CPU2 common flow encryption/decryption
DD CPU3 elephant flow encryption/decryption
EE Encryption/decryption engine
FF CPU4 elephant flow encryption/decryption
GG CPU1: message sending; PMD; message sorting; shading

(57) **Abstract:** Provided in the present application are a method and apparatus for solving an IPSec elephant flow on the basis of a DPDK. The method comprises: on the basis of a configuration file, initializing a CPU0 and a CPU1 into PMD thread cores, and initializing CPU2-CPU4 into encryption/decryption engine operation cores, creating lock-free rings for communication between the PMD thread cores and the encryption/decryption engine operation cores, and establishing a network card sending queue and a network card receiving queue; the PMD thread cores regularly executing polling operations, acquiring messages, detecting an elephant flow and a common flow, performing load balancing on the elephant flow, and distributing the messages to the encryption/decryption engine operation cores; and encrypting/decrypting the messages in the encryption/decryption engine operation cores, then sending the messages to the CPU1 for sorting, and then sending the messages to the network card sending queue. In this way, the detection of an elephant flow is realized, and multi-core IPSec encryption/decryption is performed on the elephant flow to solve the performance bottleneck of a single-core processor. The present application can be applied to wider application scenarios.

WO 2025/124178 A1

[见续页]

IR, IS, IT, JM, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, MG, MK, MN, MU, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, WS, ZA, ZM, ZW。

(84) 指定国(除另有指明，要求每一种可提供的地区保护): ARIPO (BW, CV, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SC, SD, SL, ST, SZ, TZ, UG, ZM, ZW), 欧亚 (AM, AZ, BY, KG, KZ, RU, TJ, TM), 欧洲 (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, ME, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG)。

本国际公布:

— 包括国际检索报告(条约第21条(3))。

— 在修改权利要求的期限届满之前进行，在收到该修改后将重新公布(细则48.2(h))。

(57) 摘要: 本申请的一种基于DPDK的解决IPSec大象流的方法与装置，通过根据配置文件，根据配置文件，将CPU0和CPU1初始化为PMD线程核，将CPU 2 ~ 4初始化为加解密引擎工作核，创建无锁环用于PMD线程核和加解密引擎工作核间的通信，建立网卡发送队列和网卡接收队列，PMD线程核定期执行轮询工作，获取报文并检测出大象流和普通流，对大象流进行负载均衡，将报文分发到加解密引擎工作核中，对加解密引擎工作核中的报文加解密后发送给CPU1进行排序，然后发送给网卡发送队列，实现了大象流的检测，以及多核IPSec加解密大象流来解决此单核处理器的性能瓶颈，可应用场景更加广泛。

一种基于 DPDK 的解决 IPSec 大象流的方法与装置

技术领域

本申请涉及网络通信技术领域，特别涉及一种基于 DPDK 的解决 IPSec 大象流的方法与装置。

背景技术

随着 SDWAN 的发展，以及企业、政府部门对安全性需求的提高，使用 IPSec 加密传输路径中的数据变成了基本需求，越来越多的组织构建自己的虚拟专用网络，由 IPSec 协议提供隧道安全保障。

由于信息化的深入，越来越多的上云业务、AI 训练使用的大量数据、视频会议或者影视作品的高品质传输，对于 SDWAN IPSec 接入带来了高吞吐的要求，即这些数据在 IPSec 链路中形成了大量、持续传递的大象流。大象流对于带宽形成了一定的要求，对于 IPSec 链路的加解密节点也带来了挑战，即由于加解密比较耗 CPU 资源，大象流的到来可能让一个 CPU 线程无力处理更多的数据，导致传输报文的丢弃。另外，随着国产化信创的发展，更多的企业或部门为了摆脱对因特尔处理器的依赖，选择国产处理器。但就目前的单核处理器性能对比来说，国产处理器的性能还无法比肩因特尔处理器。当前国产服务器采用的普遍做法，是通过叠加 CPU 核数来提升其整体计算能力。使用 DPDK 的 RTC 模型，国产服务器会因为单核性能下降而导致对大象流的处理出现瓶颈问题。

解决 IPSec 大象流的问题，普遍使用卸载功能，通过增强硬件的能力来解决。因特尔公司推出的基于 barefoot 可编程交换芯片，实现更高吞吐；还推出了带有 DLB 功能的 CPU，通过使用 DLB 可以将大象流分发到多个 CPU 上。类似地，英伟达公司推出的新一代 CX7 系列网卡可以使用硬件加解密处理报文，然后使用内层报文的 RSS 功能再送到网卡的各个队列中，从而使得一条大象流分成多条流。这些新出现的硬件解决了大象流以及更高吞吐的问题，只是可能会因为价格成本问题、或者信创领域的特殊要求而无法使用。

发明内容

本申请旨在至少在一定程度上解决相关技术中的技术问题之一。为此，本申请的一个目的在于提出一种基于 DPDK 的解决 IPSec 大象流的方法与装置，本申请基于 DPDK 技术，软件实现 IPSec 大象流的检测，以及多核 IPSec 加解密大象流来解决此单核处理器的性能瓶颈，其可应用场景更加广泛。

本申请的一个方面提供了一种基于 DPDK 的解决 IPSec 大象流的方法，包括：

步骤 S100：根据配置文件，将 CPU0 和 CPU1 初始化为 PMD 线程核，将 CPU 2、CPU 3、CPU 4 初始化为加解密引擎工作核，创建无锁环用于 PMD 线程核和加解密引擎工作核之间的通信，建立网卡发送队列和网卡接收队列；

所述将 CPU0 和 CPU1 初始化为 PMD 线程核是指：CPU0 和 CPU1 都实现为 DPDK 的 PMD 线程核，对网卡发送队列和网卡接收队列执行轮询任务；

所述将 CPU 2、CPU 3、CPU 4 初始化为加解密引擎工作核是指：CPU 2、CPU 3、CPU 4 实现为加解密引擎工作核，实现对报文的加解密处理，其中第一个 CPU 核为 CPU 2，用来执行普通流的报文处理，CPU 3、CPU 4 用来并行执行大象流的报文加解密；

所述创建无锁环用于 PMD 线程核和加解密引擎工作核之间的通信是指：利用 DPDK 的无锁环，在初始化阶段，建立 PMD 线程核即 CPU0 和 CPU1 与加解密引擎工作核之间的通信通道，实现报文核间传递；

所述网卡发送队列用于存储将要被发送到网络的报文；

所述网卡接收队列用于存储从网络中接收到的报文；

步骤 S200：PMD 线程核定期执行轮询工作，获取报文并检测出大象流和普通流，对大象流进行负载均衡，将报文分发到加解密引擎工作核中；

所述 PMD 线程核定期执行轮询工作，获取报文并检测出大象流和

普通流，对大象流进行负载均衡，将报文分发到加解密引擎工作核中的具体步骤为：

步骤 S210：调用 rte_eth_rx_burst()读取网卡接收队列中的报文，按照报文的大小分为分片报文和非分片报文；

所述分片报文是指当报文的大小超过网络链路的最大传输单元时，报文就会被分割成 m 个片段，每个片段就是一个分片报文；

所述非分片报文是指如果报文的大小小于等于网络链路的最大传输单元，整个报文可以在一个报文中传输，而不需要分割分片；

步骤 S220：对分片报文进行重组得到重组报文；

所述对分片报文进行重组是指：将接收到的分片报文存入分片链表，如果尚未收到最后一片分片报文，将已收到的分片报文存入分片链表等待后续分片报文，将按顺序到达的分片报文按照其在原始的报文中的位置信息进行组装得到重组报文；

步骤 S230：对收到的非分片报文进行 sflow 抽样，所述非分片报文包含重组报文和不需要进行分片的报文，统计前 K 个传输流的已知流量数据，如果某个传输流达到大象流的阈值，记录该传输流的五元组信息，并更新大象流的流表数组；

所述对收到的非分片报文进行 sflow 抽样，所述非分片报文包含重组报文和不需要进行分片的报文，统计前 K 个传输流的已知流量数据，如果某个传输流达到大象流的阈值，记录该传输流的五元组信息，并更新大象流的流表数组是指：对于每个非分片报文，执行 sflow 抽样的操作，统计前 K 个传输流的已知流量数据，对每个传输流，判断是否达到大象流的阈值，如果某个传输流达到大象流的阈值，记录该传输流的五元组信息，在流表数组中维护大象流数据，如果传输流是大象流，将其五元组信息添加到流表数组中；

所述大象流是指网络链路中持续进行的占用较大带宽传递数据的过程，也可指此过程中的数据；

步骤 S240：依次用大象流的流表数组对比收到的非分片报文，如果非分片报文属于大象流，则将该非分片报文标记为大象流，并按顺序编

号，将该非分片报文放入加解密引擎工作核 CPU3、CPU 4 的无锁环中；如果是普通流，则将该非分片报文放入加解密引擎工作核 CPU2 的无锁环中；

步骤S300：对加解密引擎工作核中的报文进行加解密后发送给CPU1进行排序，然后发送给网卡发送队列；

所述对加解密引擎工作核中的报文进行加解密后发送给 CPU1 进行排序，然后发送给网卡发送队列的具体方法为：

每个加解密引擎工作核不断轮询自己的无锁环，取出非分片报文；查询IPSec策略判断非分片报文为出方向IPSec报文或入方向IPSec报文；若为出方向 IPSec 报文，根据对该出方向 IPSec 报文进行加密；若为入方向 IPSec 报文，则查询 SA，并进行报文抗重放检查，然后进行解密，更新抗重放窗口；若为普通流，则不进行加解密处理，经过路由查询后，送入负责发送的 CPU1；按照该非分片报文的编号进行排序，然后调用rte_eth_tx_burst()发送给网卡发送队列。

本申请的一个方面提供了一种基于 DPDK 的解决 IPSec 大象流的装置，包括：

初始化创建模块，用于根据配置文件，将 CPU0 和 CPU1 初始化为PMD 线程核，将 CPU 2、CPU 3、CPU 4 初始化为加解密引擎工作核，创建无锁环用于 PMD 线程核和加解密引擎工作核之间的通信，建立网卡发送队列和网卡接收队列；

大象流检测分发模块，用于 PMD 线程核定期执行轮询工作，获取报文并检测出大象流和普通流，对大象流进行负载均衡，将报文分发到加解密引擎工作核中；

加解密发送模块，用于对加解密引擎工作核中的报文进行加解密后发送给 CPU1 进行排序，然后发送给网卡发送队列。

本申请的一个方面提供了一种电子设备，包括存储器、处理器及存储在存储器上并可在处理器上运行的计算机程序，所述处理器执行所述

程序时，以实现一种基于 DPDK 的解决 IPSec 大象流的方法中的步骤。

本申请的一个方面提供了一种可读存储介质，所述可读存储介质存储有计算机程序，所述计算机程序适于处理器进行加载，以执行一种基于 DPDK 的解决 IPSec 大象流的方法中的步骤。

本申请提出的一种基于 DPDK 的解决 IPSec 大象流的方法相对于现有技术，具备以下优点：

当前随着信息数据的巨量增长，信息安全需求的增强，对于 SDWAN IPSec 接入带来了高吞吐的要求，即这些数据在 IPSec 链路中形成了大量、持续传递的大象流。由于加解密非常消耗 CPU 资源，一条 IPSec 大象流很容易达到单核 CPU 线程的性能瓶颈，从而导致报文的丢失。

本申请考虑了这种单核处理 IPSec 大象流面临的瓶颈问题，在 DPDK RTC 模型的基础上，引入流水线模型，通过检测大象流，对大象流进行多核并行处理，从而解决由单核性能瓶颈导致的丢包问题，使得单条 IPSec 大象流的吞吐可以随着 CPU 核数的增加而线性增加。在一些国产信创环境，单核处理性能比较低，但其 CPU 总数较多，本申请能较好地以低成本解决总体性能问题。

附图说明

图 1 为本申请的一种基于 DPDK 的解决 IPSec 大象流的方法的方法流程框图；

图 2 为网卡 RSS 和数据流的示意图；

图 3 为本申请的一种基于 DPDK 的解决 IPSec 大象流的装置的功能模块图；

图 4 为本申请提供的一种电子设备的结构示意图；

图 5 是本申请提供的一种可读存储介质结构示意图。

具体实施方式

为了更好地理解本申请，将参考附图对本申请的各个方面做出更

详细的说明。应理解，这些详细说明只是对本申请的示例性实施方式的描述，而非以任何方式限制本申请的范围。在说明书全文中，相同的附图标号指代相同的元件。表述“和/或”包括相关联的所列项目中的一个或多个的任何和全部组合。

在附图中，为了便于说明，已稍微调整了元素的大小、尺寸和形状。附图仅为示例而并非严格按比例绘制。如在本文中使用的，用语“大致”、“大约”以及类似的用语用作表近似的用语，而不用作表程度的用语，并且旨在说明将由本领域普通技术人员认识到的、测量值或计算值中的固有偏差。另外，在本申请中，各步骤处理描述的先后顺序并不必然表示这些处理在实际操作中出现的顺序，除非有明确其它限定或者能够从上下文推导出的除外。

还应理解的是，诸如“包括”、“包括有”、“具有”、“包含”和/或“包含有”等表述在本说明书中是开放性而非封闭性的表述，其表示存在所陈述的特征、元件和/或部件，但不排除一个或多个其它特征、元件、部件和/或它们的组合的存在。此外，当诸如“...中的至少一个”的表述出现在所列特征的列表之后时，其修饰整列特征，而非仅仅修饰列表中的单独元件。此外，当描述本申请的实施方式时，使用“可”表示“本申请的一个或多个实施方式”。并且，用语“示例性的”旨在指代示例或举例说明。

除非另外限定，否则本文中使用的所有措辞（包括工程术语和科技术语）均具有与本申请所属领域普通技术人员的通常理解相同的含义。还应理解的是，除非本申请中有明确的说明，否则在常用词典中定义的词语应被解释为具有与它们在相关技术的上下文中的含义一致的含义，而不应以理想化或过于形式化的意义解释。

需要说明的是，在不冲突的情况下，本申请中的实施方式及实施方式中的特征可以相互组合。下面将参考附图并结合实施方式来详细说明本申请。

实施例 1

如图 1 所示，为本申请提供的一种基于 DPDK 的解决 IPSec 大象流

的方法的方法流程框图，包括：

步骤S100：根据配置文件，将CPU0和CPU1初始化为PMD线程核，将CPU 2、CPU 3、CPU 4初始化为加解密引擎工作核，创建无锁环用于PMD线程核和加解密引擎工作核之间的通信，建立网卡发送队列和网卡接收队列；

所述将CPU0和CPU1初始化为PMD线程核是指：CPU0和CPU1都实现为DPDK的PMD线程核，对网卡发送队列和网卡接收队列执行轮询任务；

所述将CPU 2、CPU 3、CPU 4初始化为加解密引擎工作核是指：CPU 2、CPU 3、CPU 4实现为加解密引擎工作核，实现对报文的加解密处理，其中第一个CPU核为CPU 2，用来执行普通流的报文处理，CPU 3、CPU 4用来并行执行大象流的报文加解密；

所述创建无锁环用于PMD线程核和加解密引擎工作核之间的通信是指：利用DPDK的无锁环，在初始化阶段，建立PMD线程核即CPU0和CPU1与加解密引擎工作核之间的通信通道，实现报文核间传递；

所述网卡发送队列用于存储将要被发送到网络的报文；

所述网卡接收队列用于存储从网络中接收到的报文；

步骤S200：PMD线程核定期执行轮询工作，获取报文并检测出大象流和普通流，对大象流进行负载均衡，将报文分发到加解密引擎工作核中；

所述PMD线程核定期执行轮询工作，获取报文并检测出大象流和普通流，对大象流进行负载均衡，将报文分发到加解密引擎工作核中的具体步骤为：

步骤S210：调用rte_eth_rx_burst()读取网卡接收队列中的报文，按照报文的大小分为分片报文和非分片报文；

所述分片报文是指当报文的大小超过网络链路的最大传输单元时，报文就会被分割成m个片段，每个片段就是一个分片报文；

所述非分片报文是指如果报文的大小小于等于网络链路的最大传输单元，整个报文可以在一个报文中传输，而不需要分割分片；

步骤 S220：对分片报文进行重组得到重组报文；

所述对分片报文进行重组是指：将接收到的分片报文存入分片链表，如果尚未收到最后一片分片报文，将已收到的分片报文存入分片链表等待后续分片报文，将按顺序到达的分片报文按照其在原始的报文中的位置信息进行组装得到重组报文；

步骤 S230：对收到的非分片报文进行 sflow 抽样，所述非分片报文包含重组报文和不需要进行分片的报文，统计前 K 个传输流的已知流量数据，如果某个传输流达到大象流的阈值，记录该传输流的五元组信息，并更新大象流的流表数组；

所述非分片报文不需要进行重组，因此包含完整的报文数据；

所述对收到的非分片报文进行 sflow 抽样，所述非分片报文包含重组报文和不需要进行分片的报文，统计前 K 个传输流的已知流量数据，如果某个传输流达到大象流的阈值，记录该传输流的五元组信息，并更新大象流的流表数组是指：对于每个非分片报文，执行 sflow 抽样的操作，统计前 K 个传输流的已知流量数据，对每个传输流，判断是否达到大象流的阈值，如果某个传输流达到大象流的阈值，记录该传输流的五元组信息，在流表数组中维护大象流数据，如果传输流是大象流，将其五元组信息添加到流表数组中；

所述大象流是指网络链路中持续进行的占用较大带宽传递数据的过程，也可指此过程中的数据；

相对的，网络链路中进行的占用较小带宽、短时间的数据传输过程，则称为老鼠流或者普通流；

步骤 S240：依次用大象流的流表数组对比收到的非分片报文，如果非分片报文属于大象流，则将该非分片报文标记为大象流，并按顺序编号，将该非分片报文放入加解密引擎工作核 CPU3、CPU 4 的无锁环中；如果是普通流，则将该非分片报文放入加解密引擎工作核 CPU2 的无锁环中；

所述大象流的流表数组使用连续内存，通过以空间换时间、连续内存比特位比较的方式实现高效快速的比对检查，减少执行的 CPU 循环，

降低此处大象流与非分片报文比对检查带来的对网络吞吐性能的影响；

步骤 S300：对加解密引擎工作核中的报文进行加解密后发送给 CPU1 进行排序，然后发送给网卡发送队列；

所述对加解密引擎工作核中的报文进行加解密后发送给 CPU1 进行排序，然后发送给网卡发送队列的具体方法为：

每个加解密引擎工作核不断轮询自己的无锁环，取出非分片报文；查询 IPSec 策略判断非分片报文为出方向 IPSec 报文或入方向 IPSec 报文；若为出方向 IPSec 报文，根据对该出方向 IPSec 报文进行加密；若为入方向 IPSec 报文，则查询 SA，并进行报文抗重放检查，然后进行解密，更新抗重放窗口；若为普通流，则不进行加解密处理，经过路由查询后，送入负责发送的 CPU1；按照该非分片报文的编号进行排序，然后调用 rte_eth_tx_burst()发送给网卡发送队列。

实施例 2

RSS 是当前网卡普遍采用的网络驱动技术，用于缓解单个 CPU 过载而导致的网络吞吐瓶颈，但现在随着 SDWAN IPSec 接入数据的猛增，如图 2 的网卡 RSS 和数据流的示意图，IPSec 大象流的存在再次使得单个 CPU 的处理能力成为瓶颈所在。由于单核 CPU 的处理能力已经因为处理此 IPSec 大象流而达到极限，从服务质量（QoS）的角度来看，大象流和并行到达的普通流量会因此相互影响，并可能均产生不同程度的丢包，进而导致更严重的传输质量的下降。英伟达和因特尔推出的网卡和处理器特定功能，在硬件上解决了这些问题。但在一些场景中，或者基于考虑硬件成本考虑，或者是信创环境要考虑自主可控，不能使用这种新的硬件方案。

本申请主要用于解决 IPSec 大象流面临的性能瓶颈问题，在实现上要尽量做到大象流检测的精准，又要实现 IPSec 加解密的高效，同时不浪费 CPU 资源，基于一般的 SDWAN IPSec 入网实践，给出本申请的最佳实施例：

对应一个队列的加解密工作核可以为 2~4 个，性能较差的低端处理

器可以采用更多的工作核。核间ring的描述符深度可以为512或者1024，在轮询时，一次读取核间传递报文的最大个数为32。

在使用大象流进行抽样检测时，为了使得抽样更加灵活，可将抽样百分率设置为可配置模式。抽样率低，可节省 CPU 资源；抽样率越高，检测准确率越高。默认地，可将抽样率设置为2%。在实现 Top K 算法时，从实际应用以及效率上考虑，以哈希表+堆的方式，可最多支持 20 条大象流的匹配，一般地，可选择最多支持10条大象流进行匹配，以此来节省查找时间。一般地，选择源地址IP、目的地址IP、源端口、目的端口、传输层协议作为一条流的特征，为了灵活高效，可实现配置接口支持手工配置大象流规则，来减少抽样检测带来的资源消耗。

加解密引擎工作核在查找使用 SA 时，要考虑核间同步。SA 数据结构的创建删除使用 RCU 机制进行更新。为了实现抗重放窗口检查和更新的高效，PMD 线程核将大象流送入加解密引擎工作核的核间 ring 时，将 N 个连续的报文送入其中一个 ring，再将另外 N 个连续的报文送入下一个 ring。

为了在实现上降低难度，报文需要分片时，可采用先加密后分片的方式进行实现。

实施例 3

如图 3 所示，为本申请提供的一种基于 DPDK 的解决 IPSec 大象流的装置的功能模块图，包括：

初始化创建模块，用于根据配置文件，将 CPU0 和 CPU1 初始化为 PMD 线程核，将 CPU 2、CPU 3、CPU 4 初始化为加解密引擎工作核，创建无锁环用于 PMD 线程核和加解密引擎工作核之间的通信，建立网卡发送队列和网卡接收队列；

大象流检测分发模块，用于 PMD 线程核定期执行轮询工作，获取报文并检测出大象流和普通流，对大象流进行负载均衡，将报文分发到加解密引擎工作核中；

加解密发送模块，用于对加解密引擎工作核中的报文进行加解密后发送给 CPU1 进行排序，然后发送给网卡发送队列。

实施例 4

图 4 是本申请一个实施例提供的电子设备结构示意图。如图 4 所示，根据本申请的又一方面还提供了一种电子设备。该电子设备可包括一个或多个处理器以及一个或多个存储器。其中，存储器中存储有计算机可读代码，计算机可读代码当由一个或多个处理器运行时，可以执行如上所述的一种基于 DPDK 的解决 IPSec 大象流的方法。

根据本申请实施方式的方法或系统也可以借助于图 4 所示的电子设备电子设备的架构来实现。如图 4 所示，电子设备可包括总线、一个或多个 CPU、只读存储器(ROM)、随机存取存储器(RAM)、连接到网络的通信端口、输入/输出组件、硬盘等。电子设备中的存储设备，例如 ROM 或硬盘可存储本申请提供的一种基于 DPDK 的解决 IPSec 大象流的方法。一种基于 DPDK 的解决 IPSec 大象流的方法可例如包括：根据配置文件，将 CPU0 和 CPU1 初始化为 PMD 线程核，将 CPU 2、CPU 3、CPU 4 初始化为加解密引擎工作核，创建无锁环用于 PMD 线程核和加解密引擎工作核之间的通信，建立网卡发送队列和网卡接收队列；PMD 线程核定期执行轮询工作，获取报文并检测出大象流和普通流，对大象流进行负载均衡，将报文分发到加解密引擎工作核中；对加解密引擎工作核中的报文进行加解密后发送给 CPU1 进行排序，然后发送给网卡发送队列。进一步地，电子设备还可包括用户界面。当然，图 4 所示的架构只是示例性的，在实现不同的设备时，根据实际需要，可以省略图 4 示出的电子设备中的一个或多个组件。

实施例 5

图 5 是本申请一个实施例提供的可读存储介质结构示意图。如图 5 所示，是根据本申请一个实施方式的可读存储介质。计算机可读存储介质上存储有计算机可读指令。当计算机可读指令由处理器运行时，可执行参照以上附图描述的根据本申请实施方式的一种基于 DPDK 的解决

IPSec 大象流的方法。存储介质包括但不限于例如易失性存储器和/或非易失性存储器。易失性存储器例如可包括随机存取存储器(RAM)和高速缓冲存储器(cache)等。非易失性存储器例如可包括只读存储器(ROM)、硬盘、闪存等。

另外，根据本申请的实施方式，上文参考流程图描述的过程可以被实现为计算机软件程序。例如，本申请提供了一种非暂时性机器可读存储介质，所述非暂时性机器可读存储介质存储有机器可读指令，所述机器可读指令能够由处理器运行以执行与本申请提供的方法步骤对应的指令，例如：根据配置文件，将 CPU0 和 CPU1 初始化为 PMD 线程核，将 CPU 2、CPU 3、CPU 4 初始化为加解密引擎工作核，创建无锁环用于 PMD 线程核和加解密引擎工作核之间的通信，建立网卡发送队列和网卡接收队列；PMD 线程核定期执行轮询工作，获取报文并检测出大象流和普通流，对大象流进行负载均衡，将报文分发到加解密引擎工作核中；对加解密引擎工作核中的报文进行加解密后发送给 CPU1 进行排序，然后发送给网卡发送队列。在该计算机程序被中央处理单元(CPU)执行时，执行本申请的方法中限定的上述功能。

可能以许多方式来实现本申请的方法和装置、设备。例如，可通过软件、硬件、固件或者软件、硬件、固件的任何组合来实现本申请的方法和装置、设备。用于方法的步骤的上述顺序仅是为了进行说明，本申请的方法的步骤不限于以上具体描述的顺序，除非以其它方式特别说明。此外，在一些实施例中，还可将本申请实施为记录在记录介质中的程序，这些程序包括用于实现根据本申请的方法的机器可读指令。因而，本申请还覆盖存储用于执行根据本申请的方法的程序的记录介质。

另外，本申请的实施方式中提供的上述技术方案中与现有技术中对应技术方案实现原理一致的部分并未详细说明，以免过多赘述。

如上所述的具体实施方式，对本申请的目的、技术方案和有益效果进行了进一步详细说明。应理解的是，以上所述仅为本申请的具体实施方式，并不用于限制本申请。凡在本申请的精神和原则之内，所做的任何修改、等同替换、改进等均应包含在本申请的保护范围之内。

权 利 要 求 书

1. 一种基于 DPDK 的解决 IPSec 大象流的方法，其特征在于，包括：根据配置文件，将 CPU0 和 CPU1 初始化为 PMD 线程核，将 CPU 2、CPU 3、CPU 4 初始化为加解密引擎工作核，创建无锁环用于 PMD 线程核和加解密引擎工作核之间的通信，建立网卡发送队列和网卡接收队列；

PMD 线程核定期执行轮询工作，获取报文并检测出大象流和普通流，对大象流进行负载均衡，将报文分发到加解密引擎工作核中；

对加解密引擎工作核中的报文进行加解密后发送给 CPU1 进行排序，然后发送给网卡发送队列。

2. 如权利要求 1 所述的一种基于 DPDK 的解决 IPSec 大象流的方法，其特征在于，所述将 CPU0 和 CPU1 初始化为 PMD 线程核是指：CPU0 和 CPU1 都实现为 DPDK 的 PMD 线程核，对网卡发送队列和网卡接收队列执行轮询任务。

3. 如权利要求 2 所述的一种基于 DPDK 的解决 IPSec 大象流的方法，其特征在于，所述将 CPU 2、CPU 3、CPU 4 初始化为加解密引擎工作核是指：CPU 2、CPU 3、CPU 4 实现为加解密引擎工作核，实现对报文的加解密处理，其中第一个 CPU 核为 CPU 2，用来执行普通流的报文处理，CPU 3、CPU 4 用来并行执行大象流的报文加解密。

4. 如权利要求 3 所述的一种基于 DPDK 的解决 IPSec 大象流的方法，其特征在于，所述创建无锁环用于 PMD 线程核和加解密引擎工作核之间的通信是指：利用 DPDK 的无锁环，在初始化阶段，建立 PMD 线程核即 CPU0 和 CPU1 与加解密引擎工作核之间的通信通道，实现报文核间传递。

5. 如权利要求 4 所述的一种基于 DPDK 的解决 IPSec 大象流的方法，其特征在于，所述 PMD 线程核定期执行轮询工作，获取报文并检测出大象流和普通流，对大象流进行负载均衡，将报文分发到加解密引擎工作核中的具体步骤为：

调用 rte_eth_rx_burst()读取网卡接收队列中的报文，按照报文的大小分为分片报文和非分片报文；

对分片报文进行重组得到重组报文；

对收到的非分片报文进行 sflow 抽样，所述非分片报文包含重组报文和不需要进行分片的报文，统计前 K 个传输流的已知流量数据，如果某个传输流达到大象流的阈值，记录该传输流的五元组信息，并更新大象流的流表数组；

依次用大象流的流表数组对比收到的非分片报文，如果非分片报文属于大象流，则将该非分片报文标记为大象流，并按顺序编号，将该非分片报文放入加解密引擎工作核 CPU3、CPU 4 的无锁环中；

如果是普通流，则将该非分片报文放入加解密引擎工作核 CPU2 的无锁环中。

6. 如权利要求 5 所述的一种基于 DPDK 的解决 IPSec 大象流的方法，其特征在于，所述对收到的非分片报文进行 sflow 抽样，所述非分片报文包含重组报文和不需要进行分配的报文，统计前 K 个传输流的已知流量数据，如果某个传输流达到大象流的阈值，记录该传输流的五元组信息，并更新大象流的流表数组是指：对于每个非分片报文，执行 sflow 抽样的操作，统计前 K 个传输流的已知流量数据，对每个传输流，判断是否达到大象流的阈值，如果某个传输流达到大象流的阈值，记录该传输流的五元组信息，在流表数组中维护大象流数据，如果传输流是大象流，将其五元组信息添加到流表数组中。

7. 如权利要求 6 所述的一种基于 DPDK 的解决 IPSec 大象流的方法，其特征在于，所述对加解密引擎工作核中的报文进行加解密后发送给 CPU1 进行排序，然后发送给网卡发送队列的具体方法为：

每个加解密引擎工作核不断轮询自己的无锁环，取出非分片报文；

查询 IPSec 策略判断非分片报文为出方向 IPSec 报文或入方向 IPSec 报文；

若为出方向 IPSec 报文，根据对该出方向 IPSec 报文进行加密；

若为入方向 IPSec 报文，则查询 SA，并进行报文抗重放检查，然后进行解密，更新抗重放窗口；

若为普通流，则不进行加解密处理，经过路由查询后，送入负责发送的 CPU1；

按照该非分片报文的编号进行排序，然后调用 rte_eth_tx_burst()发送给网卡发送队列。

8. 一种基于 DPDK 的解决 IPSec 大象流的装置，其特征在于，包括：

初始化创建模块，用于根据配置文件，将 CPU0 和 CPU1 初始化为 PMD 线程核，将 CPU 2、CPU 3、CPU 4 初始化为加解密引擎工作核，创建无锁环用于 PMD 线程核和加解密引擎工作核之间的通信，建立网卡发送队列和网卡接收队列；

大象流检测分发模块，用于 PMD 线程核定期执行轮询工作，获取报文并检测出大象流和普通流，对大象流进行负载均衡，将报文分发到加解密引擎工作核中；

加解密发送模块，用于对加解密引擎工作核中的报文进行加解密后发送给 CPU1 进行排序，然后发送给网卡发送队列。

9. 一种电子设备，其特征在于，包括存储器、处理器及存储在存储器上并可在处理器上运行的计算机程序，所述处理器执行所述程序时，以实现如权利要求 1-7 任一所述的一种基于 DPDK 的解决 IPSec 大象流的方法中的步骤。

10. 一种可读存储介质，其特征在于，所述可读存储介质存储有计算机程序，所述计算机程序适于处理器进行加载，以执行如权利要求 1-7 任一项所述的一种基于 DPDK 的解决 IPSec 大象流的方法中的步骤。

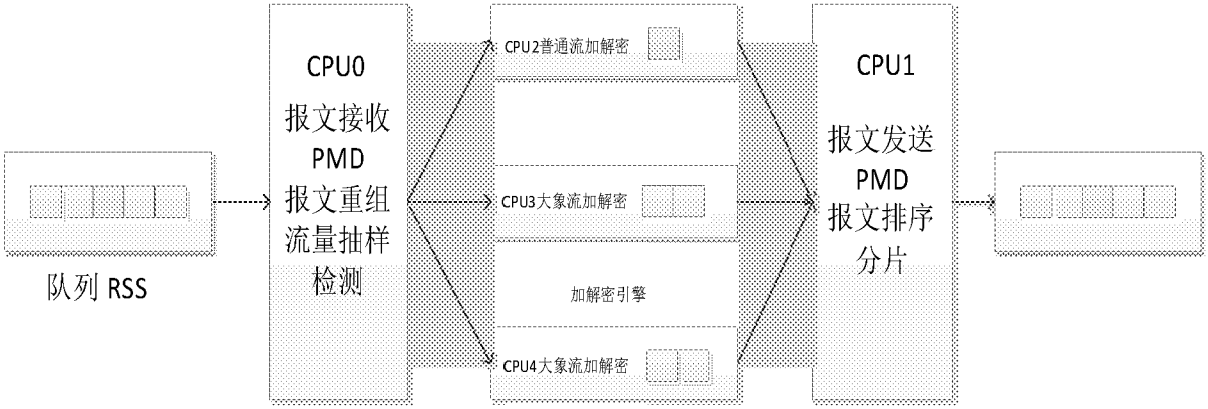
CPU0
报文接收
PMD
报文重组
流量抽样
检测
队列 RSS
CPU2普通流加解密
CPU3大象流加解密
加解密引擎
CPU4大象流加解密
CPU1
报文发送
PMD
报文排序
分片

图 1

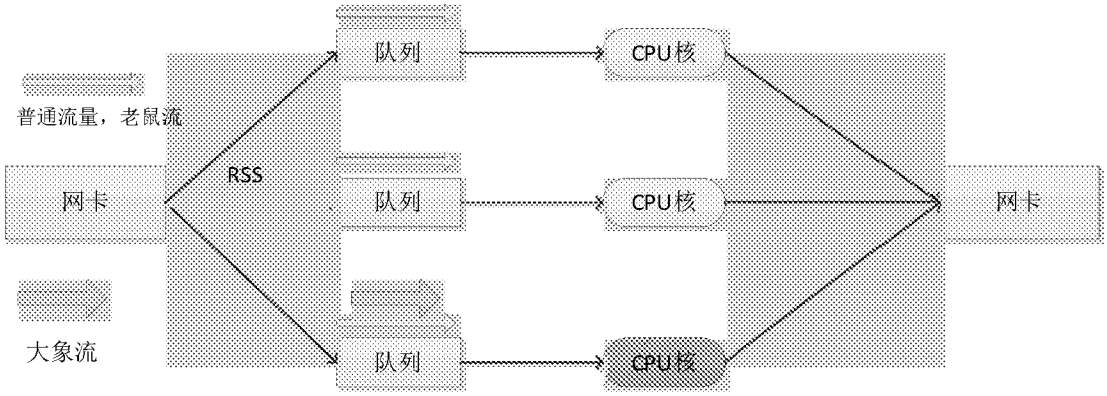
普通流量，老鼠流
网卡
RSS
大象流
队列
队列
队列
CPU核
CPU核
CPU核
网卡

图 2

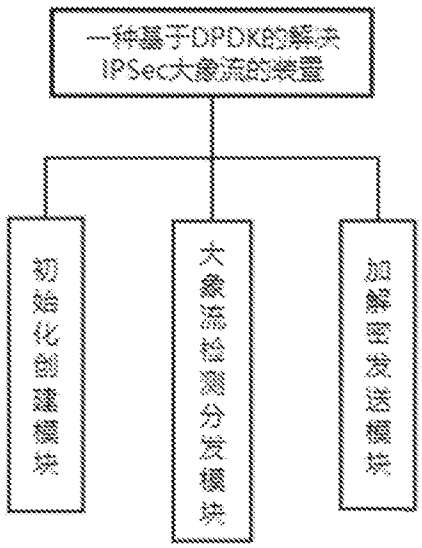
一种基于DPDK的解决
IPSec大象流的装置
初始化配置模块
大象流检测分发模块
加解密处理模块

图 3

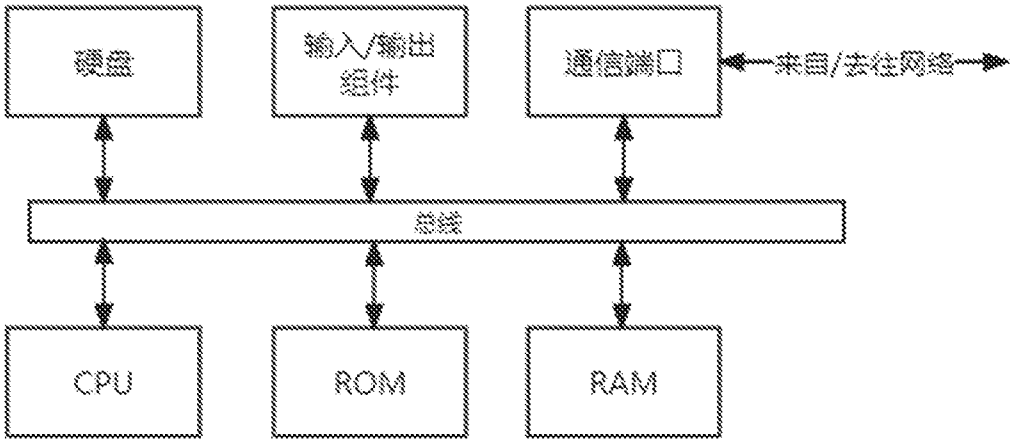

图 4

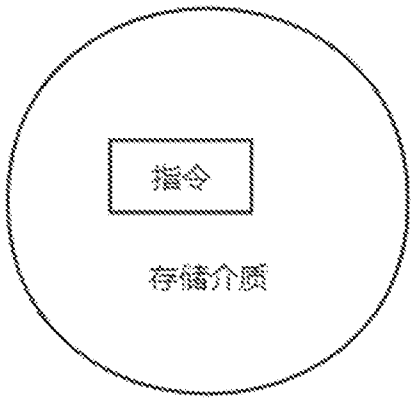

图 5

INTERNATIONAL SEARCH REPORT

International application No.
PCT/CN2024/135825

A. CLASSIFICATION OF SUBJECT MATTER

H04L 47/125(2022.01)i

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

IPC: H04L H04W

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

CNTXT, CNKI, DWPI, ENTXT, 3GPP: CPU, 核, 多核, 并行, DPDK, IPSec, 加密, 解密, 大象流, 老鼠流, 普通流, 无锁环, 队列, 网卡, 负载均衡, core? , multi w core? , parallel, encryption, decryption, elephant w stream? , mouse stream? , common w stream? , lock w free w ring, lockless w ring, queue, network w card, load w balancing

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
PX	CN 117857459 A (TIANYI CLOUD TECHNOLOGY CO., LTD.) 09 April 2024 (2024-04-09) claims 1-10	1-10
Y	CN 115967751 A (BEIJING TOPSEC NETWORK SECURITY TECHNOLOGY CO., LTD. et al.) 14 April 2023 (2023-04-14) description, paragraphs 6-93 and 109-118	1-10
Y	CN 111669374 A (CHENGDU DBAPPSECURITY CO., LTD.) 15 September 2020 (2020-09-15) description, paragraphs 6-17	1-10
A	CN 117177298 A (XIAMEN MEIYA PICO INFORMATION CO., LTD.) 05 December 2023 (2023-12-05) entire document	1-10
A	US 2021075730 A1 (INTEL CORP.) 11 March 2021 (2021-03-11) entire document	1-10

☐ Further documents are listed in the continuation of Box C. ☑ See patent family annex.

* Special categories of cited documents:
"A" document defining the general state of the art which is not considered to be of particular relevance
"D" document cited by the applicant in the international application
"E" earlier application or patent but published on or after the international filing date
"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
"O" document referring to an oral disclosure, use, exhibition or other means
"P" document published prior to the international filing date but later than the priority date claimed
"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
28 April 2025	**29 April 2025**

Name and mailing address of the ISA/CN	Authorized officer
China National Intellectual Property Administration (ISA/CN) **China No. 6, Xitucheng Road, Jimenqiao, Haidian District, Beijing 100088**	Telephone No.

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.
PCT/CN2024/135825

Patent document cited in search report	Publication date (day/month/year)	Patent family member(s)	Publication date (day/month/year)
CN 117857459 A	09 April 2024	None	
CN 115967751 A	14 April 2023	None	
CN 111669374 A	15 September 2020	None	
CN 117177298 A	05 December 2023	None	
US 2021075730 A1	11 March 2021	US 11575607 B2	07 February 2023
		US 2022286399 A1	08 September 2022
		US 2024267334 A1	08 August 2024
		US 2023231809 A1	20 July 2023
		EP 4028883 A1	20 July 2022
		EP 4028883 A4	08 March 2023
		EP 4028883 B1	04 December 2024
		WO 2021050951 A1	18 March 2021

国际检索报告

国际申请号
PCT/CN2024/135825

A. 主题的分类

H04L 47/125(2022.01)i

按照国际专利分类(IPC)或者同时按照国家分类和IPC两种分类

B. 检索领域

检索的最低限度文献(标明分类系统和分类号)

IPC: H04L H04W

包含在检索领域中的除最低限度文献以外的检索文献

在国际检索时查阅的电子数据库(数据库的名称，和使用的检索词（如使用））

CNTXT，CNKI，DWPI，ENTXT，3GPP: CPU，核，多核，并行，DPDK，IPSec，加密，解密，大象流，老鼠流，普通流，无锁环，队列，网卡，负载均衡，core？，multi w core？，parallel，encryption，decryption，elephant w stream？，mouse stream？，common w stream？，lock w free w ring，lockless w ring，queue，network w card,load w balancing

C. 相关文件

类 型*	引用文件，必要时，指明相关段落	相关的权利要求
PX	CN 117857459 A (天翼云科技有限公司) 2024年4月9日 (2024 - 04 - 09) 权利要求1-10	1-10
Y	CN 115967751 A (北京天融信网络安全技术有限公司等) 2023年4月14日 (2023 - 04 - 14) 说明书第6-93、109-118段	1-10
Y	CN 111669374 A (成都安恒信息技术有限公司) 2020年9月15日 (2020 - 09 - 15) 说明书第6-17段	1-10
A	CN 117177298 A (厦门市美亚柏科信息股份有限公司) 2023年12月5日 (2023 - 12 - 05) 全文	1-10
A	US 2021075730 A1 (INTEL CORP) 2021年3月11日 (2021 - 03 - 11) 全文	1-10

☐ 其余文件在C栏的续页中列出。

☑ 见同族专利附件。

* 引用文件的具体类型：

“A” 认为不特别相关的表示了现有技术一般状态的文件

“D” 申请人在国际申请中引证的文件

“E” 在国际申请日的当天或之后公布的在先申请或专利

“L” 可能对优先权要求构成怀疑的文件，或为确定另一篇引用文件的公布日而引用的或者因其他特殊理由而引用的文件（如具体说明的）

“O” 涉及口头公开、使用、展览或其他方式公开的文件

“P” 公布日先于国际申请日但迟于所要求的优先权日的文件

“T” 在申请日或优先权日之后公布，与申请不相抵触，但为了理解发明之理论或原理的在后文件

“X” 特别相关的文件，单独考虑该文件，认定要求保护的发明不是新颖的或不具有创造性

“Y” 特别相关的文件，当该文件与另一篇或者多篇该类文件结合并且这种结合对于本领域技术人员为显而易见时，要求保护的发明不具有创造性

“&” 同族专利的文件

国际检索实际完成的日期	国际检索报告邮寄日期
2025年4月28日	2025年4月29日
ISA/CN的名称和邮寄地址 中国国家知识产权局 中国北京市海淀区蓟门桥西土城路6号 100088	受权官员 亓晓旭 电话号码 (+86) 010-62089384

国际检索报告
关于同族专利的信息

国际申请号
PCT/CN2024/135825

检索报告引用的专利文件			公布日 (年/月/日)	同族专利			公布日 (年/月/日)
CN	117857459	A	2024年4月9日	无			
CN	115967751	A	2023年4月14日	无			
CN	111669374	A	2020年9月15日	无			
CN	117177298	A	2023年12月5日	无			
US	2021075730	A1	2021年3月11日	US	11575607	B2	2023年2月7日
				US	2022286399	A1	2022年9月8日
				US	2024267334	A1	2024年8月8日
				US	2023231809	A1	2023年7月20日
				EP	4028883	A1	2022年7月20日
				EP	4028883	A4	2023年3月8日
				EP	4028883	B1	2024年12月4日
				WO	2021050951	A1	2021年3月18日