(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 10 月 2 日 (02.10.2025)

WIPO | PCT

(10) 国际公布号
WO 2025/199904 A1

(51) 国际专利分类号:
H04L 12/18 (2006.01)

(21) 国际申请号: PCT/CN2024/084592

(22) 国际申请日: 2024 年 3 月 28 日 (28.03.2024)

(25) 申请语言: 中文

(26) 公布语言: 中文

(71) 申请人: 新华三技术有限公司 (**NEW H3C TECHNOLOGIES CO., LTD.**) [CN/CN]; 中国浙江省杭州市滨江区长河路466号 310052 (CN)。

(72) 发明人:陈京生(**CHEN, Jingsheng**); 中国北京市朝阳区广顺南大街8号院1号楼利星行中心A座640室 100102 (CN)。 雷菊艳(**LEI, Juyan**); 中国北京市朝阳区广顺南大街8号院1号楼利星行中心A座640室 100102 (CN)。

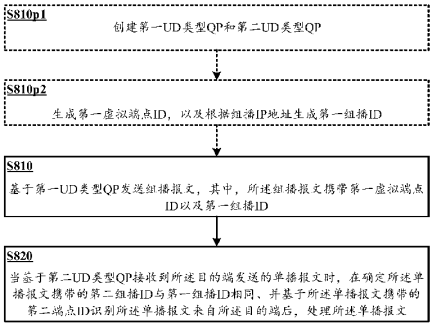
(74) 代理人: 北京博思佳知识产权代理有限公司 (**BEIJING BESTIPR INTELLECTUAL PROPERTY LAW CORPORATION**); 中国北京市海淀区上地三街9号嘉华大厦B座409 100085 (CN)。

(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, 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,

WO 2025/199904 A1

(54) **Title:** RDMA-BASED MULTICAST METHOD AND NODE DEVICE FOR DISTRIBUTED COLLECTIVE COMMUNICATION

(54) 发明名称: 基于RDMA的组播方法及分布式集合通信的节点设备



| 图 8|

S810 p1 Create a first UD-type QP and a second UD-type QP

S810 p2 Generate a first virtual endpoint ID, and generate a first multicast ID on the basis of a multicast IP address

S810 Send a multicast packet on the basis of the first UD-type QP, wherein the multicast packet carries the first virtual endpoint ID and the first multicast ID

S820 When a unicast packet sent by each destination end is received on the basis of the second UD-type QP, process the unicast packet after it is determined that a second multicast ID carried in the unicast packet is the same as the first multicast ID and after it is identified that the unicast packet comes from the destination end on the basis of a second endpoint ID carried in the unicast packet

(57) **Abstract:** Disclosed are an RDMA-based multicast method and a corresponding node device, which are applicable to distributed collective communication involving a source end and N destination ends. On the basis of an example of the method, a source end sends a multicast packet on the basis of a first UD-type QP, wherein the multicast packet carries a first virtual endpoint ID and a first multicast ID; and each destination end can send a unicast packet to the source end on the basis of a second UD-type QP, wherein the unicast packet carries a second virtual endpoint ID and a second multicast ID. The first multicast ID of the source end and the second multicast

[见续页]

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))。

IDs of the N destination ends are all the same, and the first virtual endpoint ID of the source end is different from the second virtual endpoint IDs of the N destination ends. In this way, the multicast packet sent by the source end can be forwarded to all destination ends by means of a multicast function of a network switch, and the reliability of distributed collective communication can be improved by each destination end replying with the unicast packet to the source end.

(57) 摘要: 公开了一种基于RDMA的组播方法及相应的节点设备，适用于涉及源端和N个目的端的分布式集合通信。根据所述方法的一个示例，源端基于第一UD类型QP发送组播报文，组播报文携带第一虚拟端点ID以及第一组播ID；而每个目的端可基于第二UD类型QP向源端发送单播报文，单播报文携带第二虚拟端点ID以及第二组播ID。其中，源端的第一组播ID与N个目的端的第二组播ID均相同，且源端的第一虚拟端点ID以及N个目的端的第二虚拟端点ID相互不同。这样，源端发送的组播报文可经由网络交换机的组播功能转发至所有目的端，并可通过每个目的端向源端回复单播报文来提高分布式集合通信的可靠性。