



250014

山东省济南市历下区经十路 17703 号华特广场 B510 室 济南圣达知
识产权代理有限公司
李琳(15508679010)

发文日:

2026 年 02 月 28 日



申请号: 202610240207.9

发文序号: 2026022800624870

专利申请受理通知书

根据专利法第 28 条及其实施细则第 43 条、第 44 条的规定, 申请人提出的专利申请已由国家知识产权局受理。现将确定的申请号、申请日等信息通知如下:

申请号: 2026102402079

申请日: 2026 年 02 月 28 日

申请人: 山东省计算中心(国家超级计算济南中心), 齐鲁工业大学(山东省科学院)

发明人: 谭立状, 褚夫明, 史慧玲, 张玮, 张志远

发明创造名称: 基于数字孪生的 RoCEv2 网络水线调优方法及系统

经核实, 国家知识产权局确认收到文件如下:

权利要求书 1 份 2 页, 权利要求项数: 10 项

说明书 1 份 7 页

说明书附图 1 份 2 页

说明书摘要 1 份 1 页

发明专利请求书 1 份 5 页

实质审查请求书 文件份数: 1 份

申请方案卷号: 2026700675

提示:

1. 申请人收到专利申请受理通知书之后, 认为其记载的内容与申请人所提交的相应内容不一致时, 可以向国家知识产权局请求更正。

2. 申请人收到专利申请受理通知书之后, 再向国家知识产权局办理各种手续时, 均应当准确、清晰地写明申请号。

审查员: 自动受理
联系电话: 010-62356655

审查部门: 初审及流程管理部



权利要求书

1. 基于数字孪生的 RoCEv2 网络水线调优方法，其特征在于，包括：
基于 RoCEv2 网络水线调优需求，利用带内网络遥测采集物理域网络流状态信息；
数字域根据网络流状态信息，判断流量状态，通过蒙特卡洛方法在搜索空间内进行带倾向性的随机采样，生成 ECN 水线候选解；
基于 Fluid 流体模型对 ECN 水线候选解进行并行求解，生成不同网络收敛情况指标；
基于多目标优化函数将不同网络收敛情况指标进行归一化评估，选取出 ECN 水线全局最优解；
基于 ECN 水线全局最优解完成水线更新。
2. 如权利要求 1 所述的基于数字孪生的 RoCEv2 网络水线调优方法，其特征在于，物理域网络流状态信息至少包括流标识信息、时间戳信息、队列信息和流量统计信息中的一种；
根据单位时间上报遥测包数量和 MTU 的乘积与遥测抽样率的比值近似估计吞吐量。
3. 如权利要求 1 所述的基于数字孪生的 RoCEv2 网络水线调优方法，其特征在于，根据网络流状态信息，判断流量状态，具体表示为：
在设定长度的滑动窗口内累计得到各流累积发送字节量；
当累积发送字节量不小于预设的大流字节阈值时判为大流，当累积发送字节量小于预设的大流字节阈值但在窗口内持续活跃则判为潜在大流，否则判为小流。
4. 如权利要求 3 所述的基于数字孪生的 RoCEv2 网络水线调优方法，其特征在于，通过蒙特卡洛方法在搜索空间内进行带倾向性的随机采样，生成 ECN 水线候选解，具体包括：
基于判断的当前网络下主要承载的流量类型设置差异化偏置因子；
以基准水线为基础，基于偏置因子计算采样目标均值；
基于采样目标均值进行随机采样，生成多组 ECN 水线候选解。
5. 如权利要求 1 所述的基于数字孪生的 RoCEv2 网络水线调优方法，其特征在于，不同网络收敛情况指标包括：ECN 触发概率、瓶颈交换机的实时队列长度、各流的减速因子、各流目标速率以及各流实际速率。
6. 如权利要求 1 所述的基于数字孪生的 RoCEv2 网络水线调优方法，其特征在于，在单瓶颈多流竞争场景下，基于 Fluid 流体模型对 ECN 水线候选解进行并行求解，具体包括：
将所有指向同一接收端的活跃流划分为一个逻辑流组；
各逻辑流组的实际发送速率共同决定共享瓶颈的队列总长度变化率；
基于队列总长度计算的统一 ECN 标记概率，同时代入各逻辑流组对应的 Fluid 流体模型微分方程组中进行仿真求解；
在多瓶颈多流竞争场景下，基于 Fluid 流体模型对 ECN 水线候选解进行并行求解，具体包括：
为每个瓶颈端口分别建立队列状态和标记概率；
对于任一逻辑流组，其路径上所有瓶颈端口的标记概率共同决定该流组所感知的等效标记概率；
将等效标记概率代入该流组对应的 Fluid 流体模型微分方程组中进行仿真求解。
7. 如权利要求 1 所述的基于数字孪生的 RoCEv2 网络水线调优方法，其特征在于，多目标优化函数具体表示为：
$$\theta^* = \underset{\theta_i}{\arg \max} \text{Score}(\theta_i), \theta_i = (K_{\min,i}, K_{\max,i}, P_{\max,i}), i \in \{1, \dots, N\},$$
$$\text{Score}(\theta_i) = \alpha \cdot \text{Throughput}(\theta_i) - \beta \cdot \text{QueueLength}(\theta_i) - \gamma$$
$$\cdot \text{PacketLoss}(\theta_i);$$
式中， θ^* 表示当前网络状态下的 ECN 水线全局最优解； θ_i 表示第*i*组候选 ECN 水线配置； $\text{Throughput}(\theta_i)$ 表示在参数配置 θ_i 下，通过 Fluid 流体模型计算得到的吞吐指标； $\text{QueueLength}(\theta_i)$ 表示在参数配置 θ_i 下，对应的瓶颈端口队列长度指标； $\text{PacketLoss}(\theta_i)$ 表示在参数配置 θ_i 下，对应的累计丢包指标； α, β, γ 为预设权重系数；根据多目标优化函数计算出值最高的作为当前网络状态下的 ECN 水线全局最优解。
8. 基于数字孪生的 RoCEv2 网络水线调优系统，其特征在于，包括：
物理域状态感知模块，被配置为：基于 RoCEv2 网络水线调优需求，利用带内网络遥测采集物理域网络流状态信息；
数字域水线生成模块，被配置为：数字域根据网络流状态信息，判断流量状态，通过蒙特卡洛方法在搜索空间内进行带倾向性的随机采样，生成 ECN 水线候选解；
数字域仿真模块，被配置为：基于 Fluid 流体模型对 ECN 水线候选解进行并行求解，生成不同网络收敛情况指标；

权 利 要 求 书

数字域最优解选取模块，被配置为：基于多目标优化函数将不同网络收敛情况指标进行归一化评估，选取
出 ECN 水线全局最优解；

物流水线配置模块，被配置为：基于 ECN 水线全局最优解完成水线更新。

9. 一种计算机可读存储介质，其上存储有计算机程序，其特征在于，该程序被处理器执行时实现如权利
要求 1-7 中任一项所述的基于数字孪生的 RoCEv2 网络水线调优方法中的步骤。

10. 一种计算机设备，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，其特征
在于，所述处理器执行所述程序时实现如权利要求 1-7 中任一项所述的基于数字孪生的 RoCEv2 网络水线
调优方法中的步骤。

基于数字孪生的 RoCEv2 网络流水线调优方法及系统

技术领域

[0001] 本发明属于流水线调优技术领域，具体涉及一种基于数字孪生的 RoCEv2 网络流水线调优方法及系统。

背景技术

[0002] 这里的陈述仅提供与本发明相关的背景技术，而不必然地构成现有技术。

[0003] RoCEv2 (RDMA over Converged Ethernet version 2) 是一种在以太网上实现远程直接内存访问的技术，它依赖于 ECN (显式拥塞通知) 机制来实现网络的无损传输。ECN 通过设置流水线阈值 (包括 K_{min} 、 K_{max} 和 P_{max} 等参数) 来标记拥塞，从而触发源端的拥塞控制算法 (如 DCQCN) 进行降速。流水线参数的设置直接影响到网络的吞吐量、延迟和丢包率。若流水线设置过高，可能导致拥塞控制滞后而产生丢包；若设置过低，则可能过早触发 ECN 标记，导致源端降速过快，无法充分利用带宽。

[0004] 目前，ECN 流水线的配置主要依赖于经验或设备厂商的静态推荐配置，或者通过分析业务模型并搭建测试环境进行人工调优。前者往往保守，无法达到最优性能；后者调优成本高，且效率低下。现有技术存在以下不足：

(1) 依赖真实物理网络顺序试参，搜索效率低：现有部分方案需要在物理网络上逐组修改 ECN 参数，并等待网络动态收敛后再评估效果，属于顺序试错过程。每次试参都需要较长的观测与稳定时间，效率低且对线上业务造成扰动。

[0005] (2) 交换机本地调优易造成全网不收敛：现有部分方案基于单点交换机的局部观测数据进行参数调整。不同交换机的局部最优配置可能存在冲突，容易引发交换机间的恶性竞争，导致全网 ECN 参数持续震荡，难以收敛到全局最优。

[0006] (3) 可解释性与可控性不足：现有部分方案依赖于深度强化学习等黑盒模型进行参数搜索，其搜索过程难以解释和验证，容易出现盲目搜索或结果不可预期的问题。

发明内容

[0007] 本发明的目的是克服上述现有技术中存在的不足，提供一种基于数字孪生的 RoCEv2 网络流水线调优方法及系统，通过引入带内网络遥测、数字孪生架构、模型驱动仿真并行计算以及多目标协同的全局寻优决策机制，有效克服了传统单点调优导致的参数竞争与网络震荡问题，显著提升了 RoCEv2 网络的吞吐量与稳定性。

[0008] 为了实现上述目的，本发明是通过如下的技术方案来实现：

第一方面，本发明的技术方案提供了 1. 基于数字孪生的 RoCEv2 网络流水线调优方法，包括：

基于 RoCEv2 网络流水线调优需求，利用带内网络遥测采集物理域网络流状态信息；

数字域根据网络流状态信息，判断流量状态，通过蒙特卡洛方法在搜索空间内进行带倾向性的随机采样，生成 ECN 流水线候选解；

基于 Fluid 流体模型对 ECN 流水线候选解进行并行求解，生成不同网络收敛情况指标；

基于多目标优化函数将不同网络收敛情况指标进行归一化评估，选取出 ECN 流水线全局最优解；

基于 ECN 流水线全局最优解完成流水线更新。

[0009] 在至少一个实施例中，物理域网络流状态信息至少包括流标识信息、时间戳信息、队列信息和流量统计信息中的一种；

根据单位时间上报遥测包数量和 MTU 的乘积与遥测抽样率的比值近似估计吞吐量。

[0010] 在至少一个实施例中，根据网络流状态信息，判断流量状态，具体表示为：

在设定长度的滑动窗口内累计得到各流累积发送字节量；

当累积发送字节量不小于预设的大流字节阈值时判为大流，当累积发送字节量小于预设的大流字节阈值但在窗口内持续活跃则判为潜在大流，否则判为小流。

[0011] 在至少一个实施例中，通过蒙特卡洛方法在搜索空间内进行带倾向性的随机采样，生成 ECN 流水线候选解，具体包括：

基于判断的当前网络下主要承载的流量类型设置差异化偏置因子；

以基准流水线为基础，基于偏置因子计算采样目标均值；

基于采样目标均值进行随机采样，生成多组 ECN 流水线候选解。

[0012] 在至少一个实施例中，不同网络收敛情况指标包括：ECN 触发概率、瓶颈交换机的实时队列长度、各流的减速因子、各流目标速率以及各流实际速率。

[0013] 在至少一个实施例中，在单瓶颈多流竞争场景下，基于 Fluid 流体模型对 ECN 流水线候选解进行并行求解，具体包括：

说明书

将所有指向同一接收端的活跃流划分为一个逻辑流组；

各逻辑流组的实际发送速率共同决定共享瓶颈的队列总长度变化率；

基于队列总长度计算的统一 ECN 标记概率，同时代入各逻辑流组对应的 Fluid 流体模型微分方程组中进行仿真求解；

在多瓶颈多流竞争场景下，基于 Fluid 流体模型对 ECN 水线候选解进行并行求解，具体包括：

为每个瓶颈端口分别建立队列状态和标记概率；

对于任一逻辑流组，其路径上所有瓶颈端口的标记概率共同决定该流组所感知的等效标记概率；

将等效标记概率代入该流组对应的 Fluid 流体模型微分方程组中进行仿真求解。

[0014] 在至少一个实施例中，多目标优化函数具体表示为：

$$\theta^* = \underset{\theta_i}{\operatorname{arg\,max}} \operatorname{Score}(\theta_i), \theta_i = (K_{\min,i}, K_{\max,i}, P_{\max,i}), i \in \{1, \dots, N\},$$
$$\operatorname{Score}(\theta_i) = \alpha \cdot \operatorname{Throughput}(\theta_i) - \beta \cdot \operatorname{QueueLength}(\theta_i) - \gamma$$

[0015]

$$\cdot \operatorname{PacketLoss}(\theta_i);$$

[0016] 式中， θ^* 表示当前网络状态下的 ECN 水线全局最优解； θ_i 表示第*i*组候选 ECN 水线配置； $\operatorname{Throughput}(\theta_i)$ 表示在参数配置 θ_i 下，通过 Fluid 流体模型计算得到的吞吐指标； $\operatorname{QueueLength}(\theta_i)$ 表示在参数配置 θ_i 下，对应的瓶颈端口队列长度指标； $\operatorname{PacketLoss}(\theta_i)$ 表示在参数配置 θ_i 下，对应的累计丢包指标； α, β, γ 为预设权重系数；根据多目标优化函数计算出值最高的作为当前网络状态下的 ECN 水线全局最优解。

[0017] 第二方面，本发明的技术方案还提供了基于数字孪生的 RoCEv2 网络水线调优系统，包括：

物理域状态感知模块，被配置为：基于 RoCEv2 网络水线调优需求，利用带内网络遥测采集物理域网络流状态信息；

数字域水线生成模块，被配置为：数字域根据网络流状态信息，判断流量状态，通过蒙特卡洛方法在搜索空间内进行带倾向性的随机采样，生成 ECN 水线候选解；

数字域仿真模块，被配置为：基于 Fluid 流体模型对 ECN 水线候选解进行并行求解，生成不同网络收敛情况指标；

数字域最优解选取模块，被配置为：基于多目标优化函数将不同网络收敛情况指标进行归一化评估，选取最优 ECN 水线全局最优解；

物流水线配置模块，被配置为：基于 ECN 水线全局最优解完成水线更新。

[0018] 第三方面，本发明的技术方案还提供了一种计算机可读存储介质，其上存储有计算机程序，该程序被处理器执行时实现如第一方面所介绍的基于数字孪生的 RoCEv2 网络水线调优方法中的步骤。

[0019] 第四方面，本发明的技术方案还提供了一种计算机设备，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，该处理器执行上述程序时实现如第一方面所介绍的基于数字孪生的 RoCEv2 网络水线调优方法中的步骤。

[0020] 上述本发明的技术方案的有益效果如下：

1) 本发明的基于数字孪生的 RoCEv2 网络水线调优方法能够利用数字域的并行计算能力替代传统的物理域顺序试错，将参数搜索效率显著提升，克服了传统的单点调优难以收敛；同时基于可解释的 Fluid 流体模型进行驱动，也避免了 AI 调优模型的不可控风险，实现了 RoCEv2 网络性能的确定性保障。

[0021] 2) 本发明通过数字孪生架构将 ECN 水线参数的试错过程从物理域转移到数字域，利用数字域的并行计算能力，能够同时对大规模候选参数进行仿真评估，显著提高了参数搜索效率，避免了物理域顺序试错对线上业务的扰动。

[0022] 3) 本发明采用改进的蒙特卡洛方法，根据实时流量特征生成带倾向性的候选参数集，提高了搜索的针对性和收敛速度；同时，基于可解释的 Fluid 流体模型进行仿真，避免了黑盒模型带来的不可控风险，保证了调优过程的可解释性和结果的可靠性。

[0023] 4) 本发明通过多目标效用函数进行全局寻优，能够综合考虑吞吐量、队列长度和丢包率等关键性能指标，得到全局最优的 ECN 水线配置，避免了单点调优导致的参数竞争和网络震荡问题，提升了 RoCEv2 网络的整体性能和稳定性。

附图说明

[0024] 构成本发明的一部分的说明书附图用来提供对本发明的进一步理解，本发明的示意性实施例及其说明用于解释本发明，并不构成对本发明的不当限定。

[0025] 图 1 是本发明实施例 1 公开的基于数字孪生的 RoCEv2 网络水线调优方法示意图；

说明书

图 2 是本发明实施例 1 公开的基于数字孪生的 RoCEv2 网络水线调优方法架构图；

图 3 是本发明实施例 1 公开的数字域流体仿真及水线决策示意图。

具体实施方式

[0026] 应该指出，以下详细说明都是例示性的，旨在对本发明提供进一步的说明。除非另有指明，本发明使用的所有技术和科学术语具有与本发明所属技术领域的普通技术人员通常理解相同含义。

[0027] 正如背景技术所介绍的，本发明的目的是克服上述现有技术中存在的不足，提供一种基于数字孪生的 RoCEv2 网络水线调优方法及系统，通过引入带内网络遥测、数字孪生架构、模型驱动仿真并行计算以及多目标协同的全局寻优决策机制，有效克服了传统单点调优导致的参数竞争与网络震荡问题，显著提升了 RoCEv2 网络的吞吐量与稳定性。

[0028] 实施例 1

本发明的一种典型实施方式中，图 1 至图 3 所示，本实施例公开了基于数字孪生的 RoCEv2 网络水线调优方法，具体包括以下步骤：

S1. 基于 RoCEv2 网络水线调优需求，利用带内网络遥测采集物理域网络流状态信息；

S2. 数字域根据网络流状态信息，判断流量状态，通过蒙特卡洛方法在搜索空间内进行带倾向性的随机采样，生成 ECN 水线候选解；

S3. 基于 Fluid 流体模型对 ECN 水线候选解进行并行求解，生成不同网络收敛情况指标；

S4. 基于多目标优化函数将不同网络收敛情况指标进行归一化评估，选取出 ECN 水线全局最优解；

S5. 基于 ECN 水线全局最优解完成水线更新。

[0029] 本实施例提出的基于数字孪生的 RoCEv2 网络水线调优方法，通过构建物理域与数字域的孪生映射，将复杂的网络调优任务解耦为物理数据采集与数字并行仿真两部分，利用带内网络遥测实现全网状态的实时感知，通过改进的蒙特卡洛方法在数字域生成大规模 ECN 水线候选解，同时结合 Fluid 流体模型的高效演算能力，动态并行演算不同 ECN 配置下的网络收敛状态，满足数据中心应对 ECN 参数的调优需求。下面结合具体实施方式对上述基于数字孪生的 RoCEv2 网络水线调优方法进行详细说明。

[0030] S1. 基于 RoCEv2 网络水线调优需求，利用带内网络遥测采集物理域网络流状态信息。

[0031] 在本步骤中，由物理域状态感知及水线配置子系统持续对数据中心网络进行跟踪，通过带内网络遥测实时跟踪活跃流的吞吐和延迟等端到端网络性能状态信息并上报状态信息至数字域。当物理网络承载活跃流时，交换机对经过的数据包封装 INT 元数据，采集的元数据包括但不限于流标识信息、时间戳信息、队列信息以及流量统计信息。

[0032] 为了简化带内网络遥测开销，根据遥测得到的队列长度，吞吐采用上报遥测包数量进行推测，从而获取吞吐和队列长度两个指标。其中，吞吐可近似估计为单位时间上报遥测包数量和 MTU 的乘积与遥测抽样率的比值，具体表示为：

$$\text{Throughput} = \frac{\text{Nums}_{\text{TelemetryPackets}} \times \text{AvePacketSize}}{\text{SamplingRate}};$$

[0033] 式中， Throughput 表示吞吐量； $\text{Nums}_{\text{TelemetryPackets}}$ 表示单位时间上报遥测包数量；

AvePacketSize 表示 MTU，即最大传输单元； SamplingRate 表示遥测抽样率。

[0034] 物理域状态感知及水线配置子系统提取出数据中心网络中每条活跃流的端到端传输性能，并将这些网络性能状态信息并上报至数字域，同时物理域会根据数字域下发的水线配置参数设置交换机 ECN 水线。

[0035] S2. 数字域根据网络流状态信息，判断流量状态，通过蒙特卡洛方法在搜索空间内进行带倾向性的随机采样，生成 ECN 水线候选解。

[0036] 在本步骤中，数字域首先根据物理域上报的活跃流的端到端传输性能，判断流量状态。具体的，根据带内网络遥测上报活跃流的端到端传输性能，在监控周期内计算各流发送字节量，并在长度为 δ 的滑动窗口内累计得到 $\Phi(f)$ ，具体可表示为：

$$\Phi(f) = \sum_{i=1}^N b_i(f);$$

[0037] 式中， $\Phi(f)$ 表示累积发送字节量； $b_i(f)$ 表示流 f 在第 i 个监控周期内通过带内网络遥测上报的字节数； N 表示监控周期数。

[0038] 预设一个大流字节阈值 t ，当累积发送字节量 $\Phi(f) \geq t$ 判为大流，当 $\Phi(f) < t$ 但在窗口内持续活跃则判为潜在大流，否则判为小流，有效降低了降低拥塞场景下的大小流误判。

[0039] 然后，根据当前网络状态的流量特征，通过改进蒙特卡洛方法生成带倾向性的 ECN 配置参数空间，

说明书

避免了无效解空间的冗余计算，提高了求解的收敛速度和精度，使最优解的搜索过程更加高效合理。

[0040] 蒙特卡洛方法是一种以概率统计理论为指导的数值计算方法，其核心思想是通过多维解空间进行大规模的随机抽样来逼近问题的全局最优解。鉴于 ECN 参数与网络性能之间存在复杂的非线性及非凸关系，传统的梯度下降法极易陷入局部最优陷阱，且顺序搜索效率低下，更关键的是，对蒙特卡洛方法的随机性进行改进，传统随机性是盲目随机生成的，而本步骤采用改进的蒙特卡洛方法，根据流量状态生成的是带倾向性的解构造，根据流量特征，将盲目的全空间随机搜索转化为带倾向性的高效搜索，其生成的每一个随机样本点，即每一组 ECN 配置。

[0041] 具体的，改进的蒙特卡洛方法引入流量类型偏置因子，以基准水线值为基础，构建差异化的采样目标均值，表示为：

$$\mu_{target} = \lambda \mu_{base};$$

[0042] 式中， μ_{target} 表示采样目标均值； λ 表示流量类型偏置因子； μ_{base} 表示基准水线值。

[0043] 针对不同的流特征，施加差异化的偏向性策略，当判断出当前网络下主要承载大流时，系统设定偏置因子 $\lambda = 1.5$ ，此时构造的 ECN 水线解向高位偏移；当判断出当前网络下主要承载潜在大流时，系统设定偏置因子 $\lambda = 1.25$ ，此时构造的 ECN 水线解向中高位偏移；当根据判断出当前网络下主要承载小流时，系统设定偏置因子 $\lambda = 0.8$ ，此时构造的 ECN 水线解向低位偏移。

[0044] 通过本步骤可以一次性生成 N 组差异化的 ECN 水线候选解构成候选 ECN 配置参数集，相较于传统方法的逐一尝试和盲目随机，该方法生成的解空间更加合理有效，避免陷入局部最优。

[0045] S3.基于 Fluid 流体模型对 ECN 水线候选解进行并行求解，生成不同网络收敛情况指标。

[0046] 在本步骤中，基于 Fluid 流体模型的并行仿真演算，通过并行执行 Fluid 流体模型对生成的 N 组 ECN 水线候选解进行演算，计算出每组 ECN 水线的稳态指标与目标值。

[0047] 在本步骤中，基于 Fluid 流体模型通过求解核心微分方程组来模拟网络状态的演进。具体的，在瓶颈交换机处，将 ECN 触发概率与实时队列长度的关系建模为分段线性函数，具体表示为：

$$p_{ecn}(t) = \begin{cases} 0, & q(t) < K_{min} \\ \frac{q(t) - K_{min}}{K_{max} - K_{min}} P_{max}, & K_{min} \leq q(t) \leq K_{max} \\ 1, & q(t) > K_{max} \end{cases};$$

[0048] 式中， $p_{ecn}(t)$ 表示 ECN 触发概率； $q(t)$ 表示实时队列长度； K_{min} 表示交换机拥塞队列的长度的最小阈值，当队列长度低于此值时，拥塞标记的概率为 0； K_{max} 表示交换机拥塞队列的长度的最大阈值，当队列长度超过此值时，拥塞标记的概率为 1； P_{max} 表示 ECN 最大标记概率。

[0049] 该 ECN 触发概率模型描述了瓶颈交换机队列长度随时间的非线性变化，在未开启 PFC 的情况下，队列的变化率 $\frac{dq}{dt}$ 仅取决于聚合后的总输入流速率与链路物理带宽 C 之间的差值，具体表示为：

$$\frac{dq(t)}{dt} = \sum_{i=1}^n R_C^i(t) - C;$$

[0050] 式中， n 表示汇聚的流数量； $R_C^i(t)$ 表示单条流的当前发送速率； C 表示出端口的链路带宽。当聚合后的总输入流速率大于链路物理带宽时，队列积压，反之队列排空，具体表示为：

$$\begin{cases} \sum_{i=1}^n R_C^i(t) > C & \text{队列积压} \\ \sum_{i=1}^n R_C^i(t) \leq C & \text{队列排空} \end{cases};$$

[0051] 对于 DCQCN 算法中的参数 α_i ，其更新算法包含两种，分别是 $\alpha_i = (1 - g)\alpha_i + g$ 或 $\alpha_i = (1 - g)\alpha_i$ ，其中前者更新条件为发送端收到有效 CNP，后者更新条件为发送端持续未收到 CNP。因此，各流的减速因子 α_i 演变趋势可表示为：

$$\frac{d\alpha_i}{dt} = \frac{g}{\tau'} \left(\left(1 - (1 - p_{ecn}(t - \tau^*))^{\tau' R_C^i(t - \tau^*)} \right) - \alpha_i(t) \right);$$

说明书

[0052] 式中， \mathcal{G} 表示更新增益； τ' 表示 α_i 参数的更新时间间隔； t 表示时间变量； τ^* 表示反馈时延。

[0053] 各流目标速率 R_T^i 演变包含两种情况，分别对应 DCQCN 减速阶段和加速阶段，其中加速阶段按照不同加速策略可以分为快速恢复阶段和加性增加阶段。在减速阶段， R_T^i 变化量就是 $R_T^i - R_C^i$ ；在快速恢复阶段， R_T^i 变化量为 0；在加性增加阶段， R_T^i 变化量为 R_{AI}^i 。其中，加性增加阶段触发条件是 Timer 和 Byte Counter 不全大于 F 或者不全小于 F ， F 表示快速恢复阶段的连续增速次数。

[0054] 目标速率 R_T^i 演变，综合了快速恢复和加性增加阶段的动态变化逻辑，因此目标速率 R_T^i 演变趋势可表示为：

$$\begin{aligned} \frac{dR_T^i}{dt} = & -\frac{R_T^i(t) - R_C^i(t)}{\tau} \left(1 - (1 - p_{ecn}(t - \tau^*))^{\tau R_C^i(t - \tau^*)} \right) \\ & + R_{AI}^i R_C^i(t - \tau^*) \frac{1 - p_{ecn}(t - \tau^*)}{F_B p_{ecn}(t - \tau^*) ((1 - p_{ecn}(t - \tau^*))^{-B} - 1)} ; \\ & + R_{AI}^i R_C^i(t - \tau^*) \frac{1 - p_{ecn}(t - \tau^*)}{F_T R_C^i(t - \tau^*) p_{ecn}(t - \tau^*) ((1 - p_{ecn}(t - \tau^*))^{-T R_C^i(t - \tau^*)} - 1)} \end{aligned}$$

[0055] 式中， F_T 表示字节定时器阈值； T 表示速率增加的时间定时器阈值。

[0056] 各流实际速率 R_C^i 演变趋势可表示为：

$$\begin{aligned} \frac{dR_C^i}{dt} = & -\frac{R_C^i(t) \alpha_i(t)}{2\tau} \left(1 - (1 - p_{ecn}(t - \tau^*))^{\tau R_C^i(t - \tau^*)} \right) \\ & + \frac{R_T^i(t) - R_C^i(t)}{2} \frac{R_C^i(t - \tau^*) p_{ecn}(t - \tau^*)}{(1 - p_{ecn}(t - \tau^*))^{-B} - 1} ; \\ & + \frac{R_T^i(t) - R_C^i(t)}{2} \frac{R_C^i(t - \tau^*) p_{ecn}(t - \tau^*)}{(1 - p_{ecn}(t - \tau^*))^{-T R_C^i(t - \tau^*)} - 1} \end{aligned}$$

[0057] 式中， $p_{ecn}(t - \tau^*)$ 为经过反馈时延 τ^* 后的 ECN 标记概率； τ 为往返时延； \mathcal{G} 为更新增益； R_{AI}^i 为加性增速率； F_B 为字节计数器； F_T 为字节定时器阈值。

[0058] 对于 $M:N$ 多流竞争场景，在复杂流量解耦与叠加机制下，可将活跃流按照接收端划分为 N 个逻辑的 $X:1$ 流组 G_j ，每个流组内部的各流 α_i 、 R_T^i 和 R_C^i 等状态变量满足上述 Fluid 微分演进方程；第 j 个流组 G_j 内包含若干条指向同一接收端的活跃流，记该组内第 i 条流的各流组之间通过网络队列与 ECN 反馈概率产生耦合。具体分为单瓶颈与多瓶颈两类情况。

[0059] 在单瓶颈多流竞争场景下，当 N 个接收端位于不同端口且各自对应不同上行链路时，各流组之间不共享带宽资源，可对各流组独立开展仿真推演并行计算；当 N 个接收端共享同一上行链路带宽 C ，例如瓶颈交换机同一出口端口，各流组共享同一队列，总队列长度 $Q_{total}(t)$ 的变化率由所有流组输入速率之和与链路容量之差决定，具体可表示为：

$$\frac{dQ_{total}(t)}{dt} = \sum_{j=1}^N \sum_{i \in G_j} R_{C,j,i}(t) - C;$$

[0060] 式中， $R_{C,j,i}(t)$ 表示第 j 个流组中的第 i 条流在时刻 t 的发送速率。该全局统一的 $p_{ecn}(t - \tau^*)$ 将同时代入到各流对应的 $\frac{d\alpha_{j,i}}{dt}$ 、 $\frac{dR_{T,j,i}}{dt}$ 和 $\frac{dR_{C,j,i}}{dt}$ 演进方程中，从而形成方程形式不变、输入变量耦合的建模方式；即各流分别按同形方程演进，但其输入变量 p_{ecn} 由共享队列动态统一决定，进而驱动各流在共享瓶颈条件下进行速率自适应调整并最终收敛至稳定。

[0061] 为便于理解，作一种示例性说明，当 $M:N = 5:3$ 时，可按接收端分成 $k = 3$ 个逻辑流组，组内流条数分别为 $(x_1, x_2, x_3) = (2, 2, 1)$ ，对应 $(2:1, 2:1, 1:1)$ 三个子模型。若三组共享同一瓶颈交换机出口链路容量 $C = 100Gbps$ ，则三组的输入速率将共同决定同一总队列长度 $Q_{total}(t)$ 的演化，并由统一的 ECN 触发概率 $p_{ecn}(t)$ 对三组的演进方程产生耦合，从而在保持方程结构不变的前提下实现对 $M:N$ 多流竞争的精确推演。

说明书

[0062] 在多瓶颈多流竞争场景下，当网络中存在多条潜在拥塞链路时，必须为每个瓶颈端口分别建立队列状态与标记概率。设 \mathcal{E} 为瓶颈端口集合，端口 $e \in \mathcal{E}$ 的链路容量为 C_e ，其队列长度与标记概率分别记为 $Q_e(t)$ 与 $p_e(t)$ 。同时，设每个逻辑流组 G_j 的路径为 P_j ，并令该流组包含的流条数为 x_j 。采用流组建模方式：令 $R_{C,j}(t)$ 表示第 j 个流组的发送速率，则该流组对任一端口队列的聚合输入为 $x_j R_{C,j}(t)$ 。在此条件下，多瓶颈下的队列耦合满足：

$$\frac{dQ_e(t)}{dt} = \sum_{j:e \in P_j} x_j R_{C,j}(t) - C_e, \quad \forall e \in \mathcal{E};$$

[0063] 针对大规模数据中心网络中的多瓶颈场景，进一步将耦合变量由单一端口状态扩展为全局状态向量集合，表示为：

$$\{Q_e(t), p_e(t) \mid \forall e \in \mathcal{E}\};$$

[0064] 对于任意流组 G_j ，其演进方程中的标记概率项由其遍历路径 P_j 上所有瓶颈端口共同决定。定义流组 G_j 在时刻 t 看到的等效 ECN 标记概率为：

$$p_j^{(eff)}(t) = 1 - \prod_{e \in P_j} (1 - p_e(t - \delta_{e,j}));$$

[0065] 式中， $\delta_{e,j}$ 表示 ECN 反馈信号从端口 e 传播至流组 G_j 源端的传播延迟。通过将 $p_j^{(eff)}(t)$ 实时代入该流组的 $\frac{d\alpha_j}{dt}$ 、 $\frac{dR_{T,j}}{dt}$ 、 $\frac{dR_{C,j}}{dt}$ 微分演进方程或进一步代入组内各流的 $\frac{d\alpha_{j,i}}{dt}$ 、 $\frac{dR_{T,j,i}}{dt}$ 、 $\frac{dR_{C,j,i}}{dt}$ 方程，即可实现多瓶颈条件下的全局耦合推演建模，从而精确刻画多流、多路径竞争场景中的速率收敛行为、队列稳定工作点以及 ECN 参数配置对系统吞吐与稳定性的影响。

[0066] 本步骤通过并行执行 Fluid 流体模型，对各组候选 ECN 水线参数进行并发迭代求解，计算出所有候选水线参数的网络状态指标。借助大规模并行计算，能够一次性获得 N 组不同 ECN 配置下的网络性能预期，为后续的多目标求解最优 ECN 水线提供量化依据。

[0067] S4.基于多目标优化函数将不同网络收敛情况指标进行归一化评估，选取出 ECN 水线全局最优解。

[0068] 在本步骤中，采用多目标导向寻优策略，集成一个多目标优化函数对仿真的输出结果进行归一化评估。

[0069] 根据设定的调优目标，构造的多目标优化函数表示为：

$$\theta^* = \arg \max_{\theta_i} \text{Score}(\theta_i), \theta_i = (K_{min,i}, K_{max,i}, P_{max,i}), i \in \{1, \dots, N\},$$

$$\text{Score}(\theta_i) = \alpha \cdot \text{Throughput}(\theta_i) - \beta \cdot \text{QueueLength}(\theta_i) - \gamma$$

[0070]

$$\cdot \text{PacketLoss}(\theta_i);$$

[0071] 式中， θ^* 表示当前网络状态下的 ECN 水线全局最优解； θ_i 表示第 i 组候选 ECN 水线配置； $\text{Throughput}(\theta_i)$ 表示在参数配置 θ_i 下，通过 Fluid 流体模型计算得到的吞吐指标； $\text{QueueLength}(\theta_i)$ 表示在参数配置 θ_i 下，对应的瓶颈端口队列长度指标； $\text{PacketLoss}(\theta_i)$ 表示在参数配置 θ_i 下，对应的累计丢包指标； α, β, γ 为预设权重系数。系统根据目标函数计算出值最高的作为当前网络状态下的 ECN 水线全局最优解，即系统将 θ^* 作为当前网络状态下的全局最优 ECN 水线配置。

[0072] S5.基于 ECN 水线全局最优解完成水线更新。

[0073] 在本步骤中，水线配置下发至物理域，将多目标优化函数计算出的 ECN 水线全局最优解，下发至物理 RoCEv2 交换机集群，完成水线更新。

[0074] 实施例 2

本发明的一种典型实施方式中，本实施例公开了基于数字孪生的 RoCEv2 网络水线调优系统，包括：物理域状态感知模块，被配置为：基于 RoCEv2 网络水线调优需求，利用带内网络遥测采集物理域网络流状态信息；

数字域水线生成模块，被配置为：数字域根据网络流状态信息，判断流量状态，通过蒙特卡洛方法在搜索空间内进行带倾向性的随机采样，生成 ECN 水线候选解；

数字域仿真模块，被配置为：基于 Fluid 流体模型对 ECN 水线候选解进行并行求解，生成不同网络收敛情况指标；

数字域最优解选取模块，被配置为：基于多目标优化函数将不同网络收敛情况指标进行归一化评估，选取

说明书

出 ECN 水线全局最优解；

物流水线配置模块，被配置为：基于 ECN 水线全局最优解完成水线更新。

[0075] 实施例 3

本发明的一种典型实施方式中，本实施例提供了一种计算机可读存储介质，其上存储有计算机程序，该程序被处理器执行时实现如实施例 1 所介绍的基于数字孪生的 RoCEv2 网络水线调优方法中的步骤，该步骤包括：

S1.基于 RoCEv2 网络水线调优需求，利用带内网络遥测采集物理域网络流状态信息；

S2.数字域根据网络流状态信息，判断流量状态，通过蒙特卡洛方法在搜索空间内进行带倾向性的随机采样，生成 ECN 水线候选解；

S3.基于 Fluid 流体模型对 ECN 水线候选解进行并行求解，生成不同网络收敛情况指标；

S4.基于多目标优化函数将不同网络收敛情况指标进行归一化评估，选取出 ECN 水线全局最优解；

S5.基于 ECN 水线全局最优解完成水线更新。

[0076] 实施例 4

本发明的一种典型实施方式中，本实施例提供了一种计算机设备，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，该处理器执行上述程序时实现如实施例 1 所介绍的基于数字孪生的 RoCEv2 网络水线调优方法中的步骤，该步骤包括：

S1.基于 RoCEv2 网络水线调优需求，利用带内网络遥测采集物理域网络流状态信息；

S2.数字域根据网络流状态信息，判断流量状态，通过蒙特卡洛方法在搜索空间内进行带倾向性的随机采样，生成 ECN 水线候选解；

S3.基于 Fluid 流体模型对 ECN 水线候选解进行并行求解，生成不同网络收敛情况指标；

S4.基于多目标优化函数将不同网络收敛情况指标进行归一化评估，选取出 ECN 水线全局最优解；

S5.基于 ECN 水线全局最优解完成水线更新。

[0077] 以上所述仅为本发明的优选实施例而已，并不用于限制本发明，对于本领域的技术人员来说，本发明可以有各种更改和变化。凡在本发明的精神和原则之内，所作的任何修改、等同替换、改进等，均应包含在本发明的保护范围之内。

说明书附图

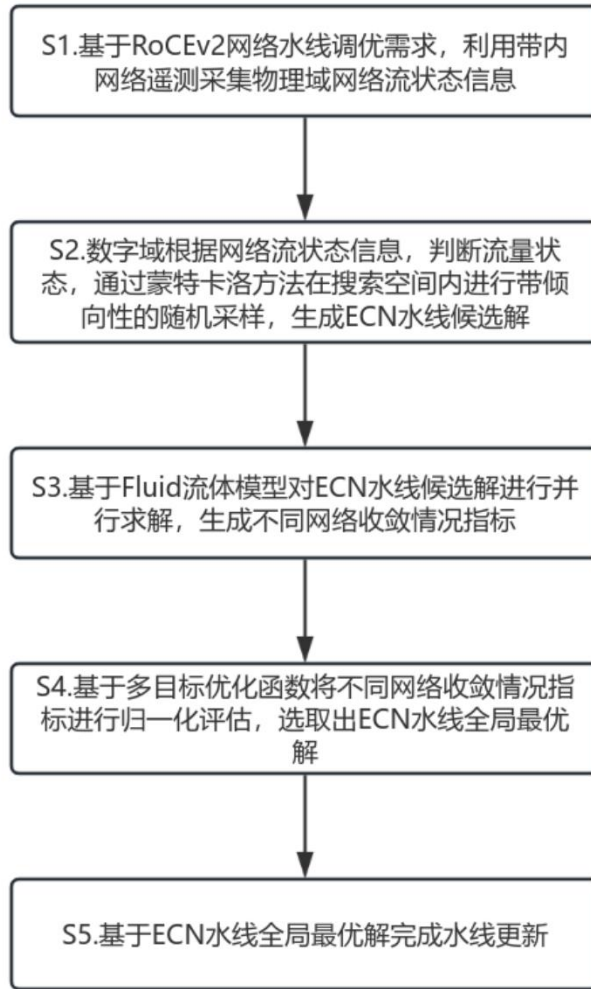


图 1

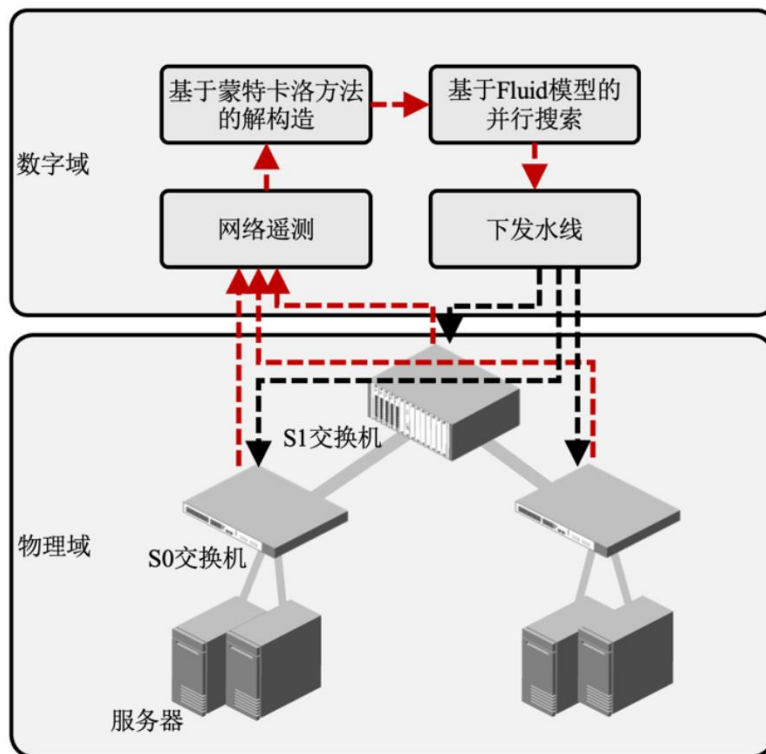


图 2

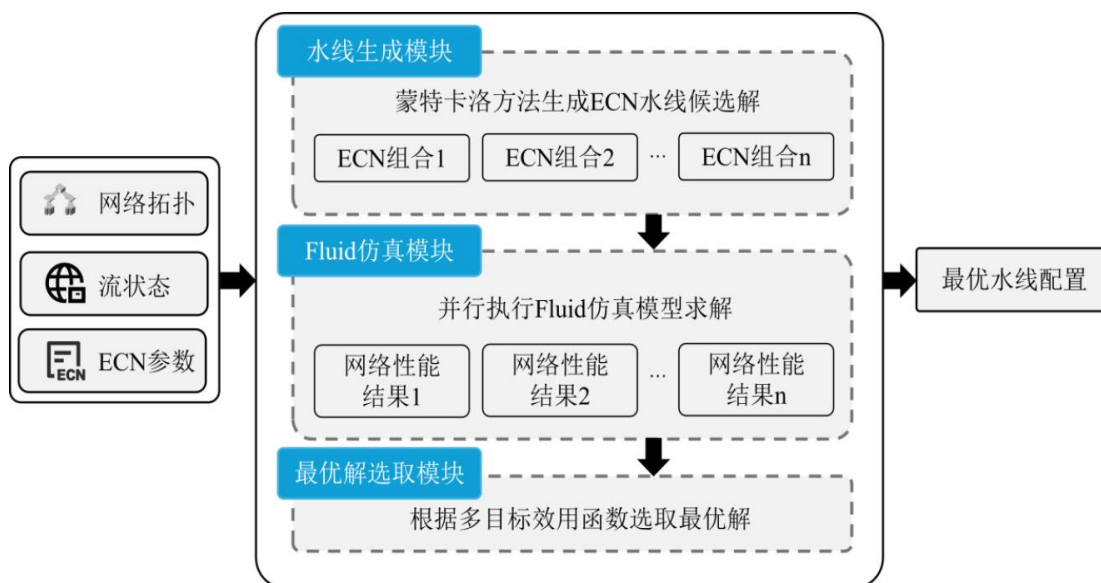


图 3

说明书摘要

本发明公开了基于数字孪生的 RoCEv2 网络水线调优方法及系统，属于水线调优技术领域，该方法包括：基于 RoCEv2 网络水线调优需求，利用带内网络遥测采集物理域网络流状态信息；数字域根据网络流状态信息，判断流量状态，通过蒙特卡洛方法在搜索空间内进行带倾向性的随机采样，生成 ECN 水线候选解；基于 Fluid 流体模型对 ECN 水线候选解进行并行求解，生成不同网络收敛情况指标；基于多目标优化函数将不同网络收敛情况指标进行归一化评估，选取出 ECN 水线全局最优解；基于 ECN 水线全局最优解完成水线更新。本发明克服了传统单点调优导致的参数竞争与网络震荡问题，显著提升了 RoCEv2 网络的吞吐量与稳定性。