arXiv:2603.09032v1 [cs.LG] 10 Mar 2026

两个老师胜过一个：硬件-物理协同引导的分布式科学机器学习

Yuchen Yuan[*†1]、Junhuan Yang[†1]、Hao Wan[1]、、Yipei Liu[1]、、Hanhan Wu[1]、、Youzuo Lin[2], 和Lei Yang[*1]1乔治梅森大学，美国弗吉尼亚州费尔法克斯2北卡罗来纳大学教堂山分校，美国北卡罗来纳州
†这些作者贡献相同。*通讯作者。{yyuan21,lyang29}@gmu.edu

摘要

科学机器学习（SciML）正日益应用于现场处理、控制和监测；然而，广域感知、实时需求以及严格的能耗和可靠性约束，使集中式科学机器学习难以实施。大多数科学机器学习模型都假设将原始数据聚合到中心节点，导致通信延迟和能耗高得令人难以承受；然而，将为通用机器学习开发的模型分布式部署，往往会破坏关键的物理原理，导致性能下降。为应对这些挑战，我们推出EPIC——一个由硬件和物理协同引导的分布式科学机器学习框架，并以全波形反演（FWI）作为代表性任务。EPIC在终端设备上执行轻量级本地编码，并在中心节点进行物理感知解码。通过传输紧凑的潜在特征而非海量原始数据，并利用交叉注意力捕捉接收器间波场耦合，EPIC在保持物理保真度的同时显著降低了通信成本。在包含五个终端设备和一个中心节点的分布式测试平台上，以及OpenFWI的10个数据集上进行评估后发现，EPIC将延迟降低了8.9×，通信能耗降低了33.8×，并且在10个数据集中的8个上提升了重建保真度。

1 引言

随着科学人工智能的快速发展，科学机器学习（SciML）日益广泛地应用于现实场景中的监测与决策。然而，当这些科学机器学习模型从实验室研究走向现场规模部署时，就需要处理连续不断的高分辨率传感器数据流，同时满足通信带宽、能耗、延迟和可靠性方面的约束。这些实际需求对当前科学机器学习工作流程所依据的许多假设提出了挑战，也要求我们重新思考如何在现场高效地设计、部署和管理科学机器学习模型。

大多数现有的科学机器学习研究 [1–4], 无论是有意还是无意，都遵循集中式设计，假设分布式传感器采集的原始测量数据会被传输到中央服务器进行推理。在这种方案中，数据通信很容易成为性能瓶颈：高速率的传感器数据流很快就会使有限带宽饱和，而长距离数据传输则会带来显著的延迟和能耗开销。此外，这些研究理想化地假设所有数据都能到达中心节点，但单一故障点就可能危及整个模型，使其在偏远或恶劣环境中极易受到影响。因此，现有的集中式科学机器学习方法并不适用于大范围、实时的科学应用部署。

受边缘计算 [5–8], 的进展启发，将计-算移近数据源为缓解通信瓶颈提供了一种有前景的解决方案。然而，简单地应用现有的机器学习设计以分布式方式运行并非易事：许多科学机器学习任务依赖全局物理耦合，而跨设备划分计算可能会破坏这些物理依赖，严重降低推理质量。

为了更好地了解在实际场景中部署科学机器学习所面临的需求与挑战，本文以全波形反演（FWI）——一种物理规律约束的逆问题——为例，考察不同分布式计算范式的设计选择及其局限性。FWI 通过反复模拟受偏微分方程（PDE）支配的地震波传播，重建地下速度模型；它最初应用于地球科学 [9, 10]。近年来，数据驱动全波形反演应运而生，旨在解决物理驱动全波形反演计算效率低下的问题，并已应用于其他声学成像领域，例如医疗保健中的超声 [11–13]。与其他科学机器学习方法一样，数据驱动全波形反演（例如代表性方法 InversionNet [14]）假设采用集中式设计，并在理想的执行环境中运行。

本文在采用无线通信的分布式计算测试平台上评估了集中式 InversionNet [14] 。由于偏远或恶劣环境（如沙漠或公海）往往缺乏有线基础设施，因此通常需要采用无线通信。模拟 4G环境所得结果验证了我们的假设：通信会成为性能瓶颈，延迟占总延迟的比例最高可达 93%。我们还评估了两种现成的分布式机器学习方法，包括联邦学习式方法（FLA）[15] 和拆分学习式方法（SLA）[16], ，二者的性能都显著下降。我们进一步探究了根本原因，发现来自不同终端设备的分布式波形高度相关，导致 FLA 效果不佳；此外，这些相关性高度依赖位置，且不同接收器提供的信息并非同等有效，因此 SLA 也难以奏效。

为应对通信瓶颈和违反物理规律所带来的挑战，我们提出了 EPIC——一种兼容边缘设备且融入物理信息的分布式科学机器学习框架。EPIC 的核心组件是一种物理信息驱动的分布式神经网络，称为 EPIC-Net；它与额外的系统级部署与管理模块协同工作，以应对该领域的实际挑战。EPIC-Net 将反演流程拆解为两部分：在终端设备上进行轻量级本地编码以提取紧凑特征，以及在中心节点进行物理信息驱动的解码以完成重建。因此，通过传输紧凑特征（即潜变量）而非原始数据，可以解决通信瓶颈；此外，中心节点的交叉注意力机制[17]确保特征融合能够捕捉不同接收器间波场的空间耦合。

主要贡献:

- 我们设计了一种由硬件与物理共同引导的分布式科学机器学习模型 EPIC-Net，在遵循底层物理原理的同时解决通信瓶颈问题。

- 我们构建了完整的 EPIC 框架，使 EPIC-Net 能够高效部署于资源受限的分布式计算基础设施上，并在运行时保持鲁棒性。

- 我们在由终端设备组成的测试平台上验证了 EPIC 框架。结果表明，该框架显著降低了延迟和通信成本，同时在网络波动下保持较高的推理准确率和鲁棒性。

我们在包含五台终端设备和一个中心节点的分布式测试平台上评估 EPIC，并使用 OpenFWI 的 10 个数据集 [18] 测试反演保真度。结果表明，与传统的集中式基线相比，EPIC 的延迟和通信能耗分别最多可降低 8。9× 和 33。8× 值得注意的是，EPIC 还提升了 10 个数据集中 8 个数据集的重建保真度，其余 2 个数据集的表现仅略低（结构相似性指数低 0.3%）。我们认为，这一出乎意料的提升归功于硬件-物理协同指导，它将网络正则化至符合物理规律的解。

2 挑战与硬件-物理指导

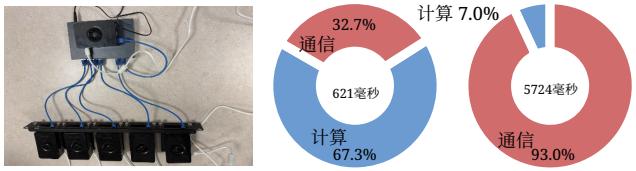
2.1 挑战：通信延迟可能成为集中式现场科学机器学习的瓶颈

我们搭建了一个由六台设备组成的硬件测试平台，如图 1(a) 所示，其中包括：(1) 五台终端设备（D_1，……，D_5），用于采集输入传感器数据；(2) 一台中心设备（D_c），用于生成最终输出速度图。该测试平台可通过千兆路由器使用 Wi-Fi 进行通信；此外，还可通过将上行带宽限制为 15 Mbps、设置 50ms 延迟和 0.5% 的丢包率，模拟高质量的 4G 蜂窝网络 [19]。

为确定性能瓶颈，我们测量了各设备上计算和通信所耗时间的比例，如图 1(b)(c) 所示，涵盖真实 Wi-Fi 环境和 4G 模拟环境。将 InversionNet [14] 部署为集中式机器学习系统后，原始波形数据由 $\mathcal{D}_1$ 至 $\mathcal{D}_5$，分布式采集，随后传输至 $\mathcal{D}_c$ 进行推理。在这两种情况下，我们都能观察到通信延迟占据了延迟分解中的很大一部分：Wi-Fi 环境下为 32.7%，4G 模拟环境下为 93%。这表明性能瓶颈在于数据通信。更糟的是，现场部署通常

需要具备实时性能，且能源预算有限。然而，由于传感产生的原始数据量很大（例如，FWI 应用中的波形数据），传输这些数据不可避免地会导致较高的通信延迟和能耗。

2.2 硬件教师：分布式机器学习可以缓解通信瓶颈，但会造成性能损失 至



(a) 硬件测试平台 (b) Wi-Fi 下的延迟 (c) 4G 下的延迟

图 1：自建测试平台及两种网络设置下集中式 InversionNet [14] 的延迟分解。

表 1：4G 场景下的延迟（毫秒）分解及结构相似性指数对比

方法	L_{edge}	L_{comm}	$L_{central}$	L_{total}	Lvs. InV	SSIM	$\Delta_{vs.\ InV}$
InV [14]	0.01	5325.52	398.42	5723.95	-	0.85	-
FLA [15]	241.72	163.07	0.95	405.74	14.1×	0.76	-10.58%
SLA [16]	93.99	163.13	179.45	436.57	13.1×	0.82	-3.52%

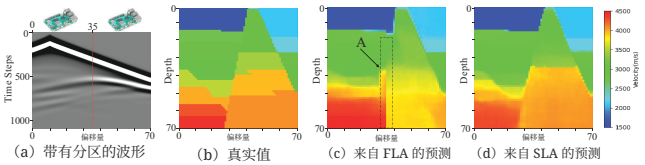
为缓解通信瓶颈，边缘计算文献中从硬件角度出发的指导原则促使我们将计算移至更靠近数据源的位置。除了集中式机器学习系统（InV） [14], 之外，我们还研究了两种成熟的分布式机器学习范式：(1) 联邦学习式方法（FLA） [15]：在每个终端设备上使用完整的机器学习模型（即本例中的 InversionNet）处理分布式原始数据，然后在中心设备上合并部分结果（即速度图）。(2) 拆分学习式方法（SLA） [16]：在每个终端设备上使用部分机器学习模型（即本例中 InversionNet 的编码器）从分布式原始数据中提取特征，然后在中心设备上合并并进一步处理这些表示（即由 InversionNet 的解码器处理），以获得最终结果（即完整的速度图）。为简单起见，本节中的所有分布式系统均采用测试平台中的 2 台终端设备接收波形数据（见图2（a））和 1 台中心设备。

通信瓶颈缓解：表1的结果表明，两种分布式机器学习范式均显著降低了4G场景下的通信成本。由于所传输的压缩表示比原始波形数据小几个数量级，FLA将通信时间从5325.52ms降至163.07ms，将能耗从2417.786mJ降至74.034mJ，降幅达96.9%。与此同时，SLA具有相近的效率，延迟为163.13ms，能耗为74.061mJ。结果验证，分布式机器学习能够有效缓解集中式实现中观察到的通信瓶颈。

性能下降：我们从结果中得到另一项重要观察：与集中式解法（InV）相比，FLA和SLA均出现不同程度的性能下降。从FLA重建速度图（VM）的可视化结果（图2（c））可见，两个终端设备结果的边界处出现了可见伪影（即图中的边界框A）。此类伪影可能对领域应用有害，并可能导致对地下地质结构的错误预测。尽管图2（d）中的SLA解法解决了这一问题，但就SSIM分数而言，与InV相比，它仍存在显著的性能下降 [20], 如表1所示。

我们将所观察到的性能下降归因于分布式策略中可能违反基本物理原理。这一观点基于更广泛的认识：科学机器学习模型必须纳入或保持物理定律，才能给出可靠且高性能的结果 [21–23], ，这与面向通用数据集（例如ImageNet、COCO）的机器学习不同。

2.3 物理教师：揭示性能下降的根本原因，为分布式机器学习设计提供指导



（a）带有分区的波形 （b）真实值 （c）来自 FLA 的预测 （d）来自 SLA 的预测

图 2：使用 2 个分布式终端设备的 FLA 和 SLA 结果可视化，并与真实值进行比较。

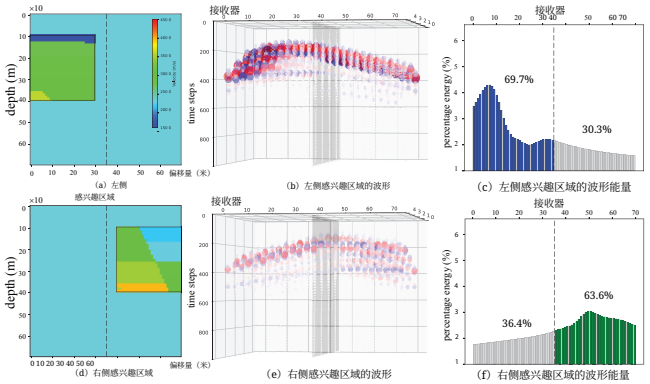


图 3：探索图 2（b）中速度图（VM）的两个感兴趣区域，并报告相应波形和能量分布的结果。

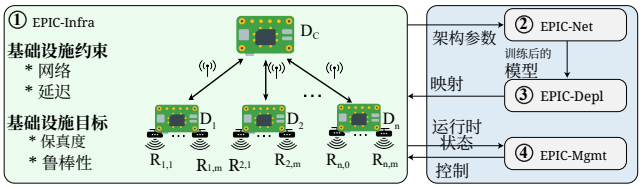
可能的根本原因1：从波动物理的角度来看，来自震源的声波会在整个介质中传播。当它到达某一区域时，携带该区域信息的波形可以传输到该区域之外的接收器，这由震源位置、该区域的位置及其结构决定。这可能是方案FLA中性能下降的根本原因：每个终端设备上的整个机器学习模型独立处理来自其对应区域的波形，而忽略了来自其他区域的波形。因此，每个分布式模型都会遭受信息损失，从而限制其性能。

根因 1 的验证： 为验证这一假设，我们利用波的叠加原理并采用正演模拟 [9] 来捕获速度图上两个感兴趣区域（ROI）所对应的波形，如图 3（a）（d）所示。为获得左侧感兴趣区域的模拟波形，我们首先计算图 3（a）中速度图所对应的波形，然后将其从与均匀速度对应的波形中减去（该均匀速度作为左侧感兴趣区域的背景）。我们对右侧感兴趣区域采用相同的步骤。最终得到的差分波形如图 3（b）（e）所示，位于跨越震源、接收器和时间步的三维空间中。

我们从这些可视化结果中得到一项概念验证观察：尽管左侧感兴趣区域仅来自地下的左侧部分，右侧部分的接收器仍能捕获大量信号。右侧感兴趣区域也表现出相同现象。这验证了我们的假设，即 FLA 因每个分布式模型的信息损失而出现性能下降。SLA 的结果进一步验证了该假设，因为在中心设备处合并潜在表示会交换来自不同区域的信息，从而相对于 FLA 带来性能提升，如表 1。

可能的根本原因 2：从物理学角度看，反射波形的强度在空间上存在变化：对于给定的地下区域，不同接收器会根据其位置捕获幅度不同的信号；同样，单个接收器记录的信号也不同

不同区域的强度。这种与位置相关的行为表明，来自不同空间位置的信号对速度模型中各个区域重建的贡献并不相同。这可能是性能下降的根本原因，



(a) EPIC 硬件　　(b) EPIC 软件

图 4：所提出的 EPIC 框架概览。

SLA 将所有潜变量视为同等重要，忽略了波传播的位置依赖性。尽管可以根据波形重建粗尺度结构，但它无法捕捉细尺度细节（见图2(d) 中的可视化结果），导致结构相似性指数较低。

对根因 2 的验证：图 3(c) 和 (f) 通过能量分布直方图量化了这种依赖位置的信号分布。对于左侧感兴趣区域的重建，左侧接收器捕获了总能量的 69.7%；与之相应，对于右侧感兴趣区域，右侧接收器捕获了 63.6% 的能量。从分布式机器学习的角度来看，这些波形会根据接收器的位置，由 SLA 中不同的编码器进行处理，从而生成潜在表示；在后续解码器的重建过程中，这些表示应得到相应的加权。

3 EPIC 框架

图4展示了所提出的边缘兼容且物理信息驱动的分布式科学机器学习框架（简称 EPIC）的概览，该框架旨在实现高保真度、低延迟和稳健的数据驱动现场全波形反演。框架包含四个组件：1 EPIC-Infra、2 EPIC-Net、3 EPIC- Depl 和4 EPIC-Mgmt。具体而言，1 EPIC-Infra 根据约束和目标构建分布式计算基础设施；2 EPIC-Net 从 EPIC-Infra 获取基础设施参数（例如终端节点的数量），以构建分布式神经网络架构；EPIC-Net 训练出的模型将传递给3 EPIC-Depl，以部署到 EPIC-Infra 上；最后，在运行时，4 EPIC-Mgmt 会在检测到可能违反实时要求时监控并控制系统。

① 资源受限且多目标的 EPIC-Infra

该基础设施包含一个分布式感知系统，地震数据的采集和处理在多个设备之间按空间划分。我们将 $\mathrm{D} = \{\mathrm{D}_1, \mathrm{D}_2, \ldots, \mathrm{D}_n\}$ 表示为由 n 个终端设备组成的集合，D_c 为中心设备。每个终端设备 D_i 管理一组接收器 $\mathrm{R}_i = \{\mathrm{R}_{i,1}, \ldots, \mathrm{R}_{i,m}\}$，用于采集其指定空间区域内的地震波形。除基本的传感和计算组件外，EPIC-Infra 还有两项约束。首先，网络 $\mathrm{N} = b, l, p$ 定义为一个元组，其中 b 表示上行带宽，l 表示网络时延，p 表示丢包率。其次，科学应用通常具有时间约束 T。

基于 EPIC-Infra，我们可以构建一个分布式 FWI 系统。以集中式 InV（见第 2 节）为例，每个终端设备 $\mathcal{D}_i \in \mathcal{D}$ 都会将所有接收器的全部原始波形数据 $\forall \mathcal{R}_{i,j} \in \mathcal{R}_i$ 传输到中心设备 $\mathcal{D}_c$。训练好的 InV 模型部署在 $\mathcal{D}_c$ 上，用于 FWI 任务。如表 1 所示，较高的通信延迟可能导致违反时间约束 T。

在本研究中，我们的目标不仅是在 EPIC-Infra 上执行全波形反演任务，以满足约束 T，还要实现高保真度

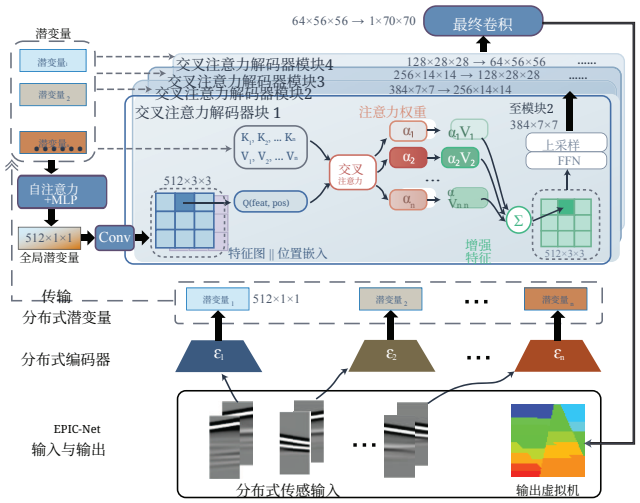


图 5：EPIC-Net 概览。

结果，并在网络可能发生故障时提供稳健的计算。我们正式定义该问题如下：给定一个包含中心设备 D_c、终端设备 D、与终端设备 $D_i \in D$ 关联的接收器 R_i、网络 N 和时间约束 T 的 EPIC-Infra，我们的目标是：

- 设计分布式神经网络 NN，
- 将 NN 部署到 D_c 和 D，
- 通过网络 N 管理 D_c 和 D 上 NN 的执行，

使系统能够容忍部分终端设备遭遇网络故障，同时满足时间约束 T，并最大限度地提高以结构相似性指数衡量的保真度。

② 硬件与物理协同引导的 EPIC-Net

实现我们的最终目标，关键在于设计一个分布式神经网络。如动机部分（第2节）所示，不考虑物理因素的神经网络设计（即 FLA 和 SLA ）能够实现出色的运行速度，但会牺牲保真度，而保真度是全波形反演任务等科学应用中最重要的指标之一。在本模块中，我们提出一种由硬件与物理共同引导的神经网络，即 EPIC-Net，如图5所示。

受第2节性能下降分析的启发，EPIC-Net 沿用 SLA 的分布式编码器设计，并在设计集中式解码器时融入物理引导。如图5所示，EPIC-Net 主要包含三个组件：（1）分布式编码；（2）自注意力融合；（3）位置感知交叉注意力解码器。

分布式编码。根据包含 n 个终端设备 D 的 EPIC-Infra，将部署分布式编码器 $\{\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n\}$。每个编码器 ε_i 在设备 D_i 上处理来自相关接收器 R_i 的波形，并通过逐步下采样的卷积块生成紧凑的潜在表示。这种策略使资源受限的终端设备能够并行地将波形编码为紧凑的潜在表示 $L = \{latent_1, latent_2, \ldots\ldots, latent_n\}$，其中 $latent_i \in \mathbb{R}^{512\times1\times1}$。随后，这些潜变量将被传输到中心设备，以通过解码过程进行重建。与传输原始波形相比，这种方式可以显著降低通信成本。

自注意力融合。在从分布式编码器接收到 n 潜变量后，EPIC-Net应用自注意力模块 [24] 聚合这些潜变量的信息以生成全局潜变量 gl。通过这种方式，我们可以避免第2 节中根本原因1所导致的性能下降（即FLA的信息损失）。

位置感知交叉注意力解码器。EPIC-Net 的最后一个组成部分旨在防止第 2 节中根本原因 2 导致的性能下降。EPIC-Net 不仅使用全局潜变量 gl，还利用 $\mathcal{L}$ 实现位置感知交叉注意力[17]，，根据空间位置自适应地为分布式潜变量分配权重。具体而言，对于给定分辨率下的解码器特征，我们为 70×70输出空间维护可学习的位置嵌入，并对其进行双线性插值，使其与各解码器分辨率相匹配。位置感知查询通过拼接解码器特征与插值后的位置嵌入来计算：$Q(F, E_{\text{pos}}) = \text{线性}([F; E_{\text{pos}}])$，其中线性$(\cdot)$表示可学习的线性投影。键和值由分布式编码器潜变量生成：$[K_i, V_i] = \text{线性}(\text{潜变量}_i)$。随后，交叉注意力会为每个局部特征计算注意力权重 α_i：

$$\alpha_i = \text{softmax}\left(\frac{QK_i^T}{\sqrt{d_k}}\right), \quad \text{Output} = \sum_{i=1}^{n} \alpha_i V_i \tag{1}$$

其中，d_k 是键向量的维度。该设计使解码器能够根据分布式潜变量与重建区域之间的空间对应关系，对其进行加权。

③ 自动化模型到设备部署模块：EPIC-Depl

EPIC-Net 定义了神经网络架构，EPIC-Infra 指定了硬件基础设施，而 EPIC-Depl 则提供了一条自动化流水线，可将训练好的模型映射到分布式设备上。每个编码器 ε_i 都分配给终端设备 D_i，用于处理来自接收器 $R_i = \{R_{i,1}、\ldots\ldots、R_{i,m}\}$ 的波形；解码器则运行在中心设备 D_c 上，负责全局重建。EPIC-Depl 可自动完成：(1) 所有设备的环境配置；(2) 将模型权重分发到相应设备；以及 (3) 通过网络 N 建立通信信道。这使得全波形反演系统能够快速部署到现场环境中。

在通信组件方面，EPIC-Depl 使用带有哈希映射缓冲区的乱序数据包处理机制，在终端设备与中心节点之间建立连接，使中心节点一旦收到来自终端设备的潜变量 n，便可立即处理样本。

④ 延迟感知运行时管理器：EPIC-Mgmt

在运行时，终端设备可能会遇到网络延迟或暂时性故障，导致数据阻塞在哈希映射缓冲区中，并违反 EPIC-Infra 要求的时间约束 T。我们提出了一种运行时管理器，即 EPIC-Mgmt，用于监控每个终端设备 D_i 的传输状态，并在其延迟超过时间约束 $T - T_d$ 时，从哈希映射缓冲区中释放流水线阻塞，其中 T_d 是中心设备上执行 EPIC-Net 解码器的延迟，EPIC-Mgmt 可在运行时之前通过一次性性能分析获得该值。

算法1 展示了 EPIC-Mgmt 的工作流程，该流程包含两个阶段：**阶段 1**，每个终端设备执行编码，并在计时阈值 $T - T_d$ 下**阶段 2**，中央解码器仅使用在 $T - T_d$ 时间范围内成功接收的潜变量执行自适应重构。得益于位置感知交叉注意力机制，解码器能够自动重新分配

算法1EPIC-Mgmt 运行时控制

输入： 波形分片 $\{D_1, \ldots, D_n\}$，超时阈值 $T - T_d$
确保： 速度图 $\hat{V}$

// **阶段 1：带超时监控的边缘编码**
对每个终端设备 D_i **并行执行**
  潜变量$_i \leftarrow \varepsilon_i(D_i)$
  如果超时**则**
    触发 EPIC-Mgmt 释放中心节点缓冲区中的同步块
结束 for
// **阶段 2：中心设备上的自适应重构**
收集在 $T - T_d$内成功接收的$\{\text{latent}_i\}$
$\hat{V} \leftarrow$ 解码器($\{\text{latent}_i\}$)

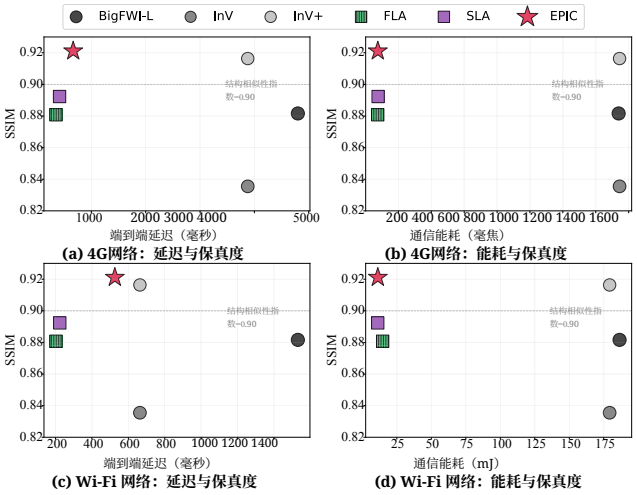


图6：通信效率与保真度之间的权衡。

注意力权重，即使部分设备发生故障，也能实现稳定重构。

4 实验

4.1 实验设置

我们使用图1（a）中的分布式测试平台来评估所提出的 EPIC 框架。该测试平台由六台设备组成：五台树莓派 5（每台配备 16GB 随机存取存储器和 128GB 存储空间），作为分布式终端设备；以及一台作为中心节点。测试平台在两种网络条件下运行：（1）通过千兆路由器连接的真实 Wi-Fi 环境；（2）模拟的 4G 蜂窝网络，其上行带宽为 15Mbps、延迟为 50ms、丢包率为 0.5%。在两种网络条件下，我们将预训练模型部署到这些设备上进行推理，以测量延迟和能耗。

在评估 EPIC-Net 时，我们采用 OpenFWI 数据集 [18]，该数据集由三个地球物理地下结构系列组成：(1)速度系列，包含四个关于平面/曲线速度变化的数据集（记为 FlatVel-A、FlatVel-B、CurveVel-A、CurveVel-B）；(2) 断层系列，包含四个关于地质断层和不连续面的数据集（记为 FlatFault-A、FlatFault-B、CurveFault-A、CurveFault-B）；(3) 风格系列，包含两个关于风格迁移生成的复杂模式的数据集（记为 Style-A 和 Style-B）。在这里，每个系列都包括难度逐渐增加的 A/B 变体。所有样本的维度均保持一致：地震波形 [5, 1000, 70]（5 个震源、

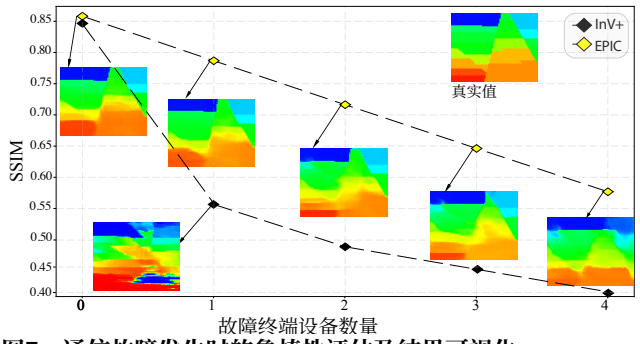


图7：通信故障发生时的鲁棒性评估及结果可视化。

70 个接收器、1000 个时间步），以及与之配对的速度图 [70, 70]，其范围为 1500–4500 m/s。

我们将 EPIC-Net 与多种基线方法进行了比较。BigFWI-L [25] 代表当前最先进的集中式 FWI 模型。InversionNet（InV）[14] 的结果直接引用自 OpenFWI 基准论文，其中模型训练了 120 个 epoch。为在我们的训练协议下进行公平比较，我们还报告了在本地环境中重新训练 InversionNet 400 个 epoch 后的 **InV+** 结果，且使用与 EPIC-Net 相同的超参数。此外，我们还与第 2.2 节分析的两种分布式机器学习范式进行比较：FLA [15]和 SLA [16]。所有模型均使用配备 40GB 显存的 NVIDIA A100 GPU，并采用 AdamW 优化器 [26] 进行训练，初始学习率为 1×10^{-3}。我们采用相同的损失函数，并对平均绝对误差和均方误差项赋予相同权重。训练批大小设为 32，共训练 400 个 epoch。

4.2 EPIC 框架评估

EPIC 在硬件/软件性能方面实现了更好的权衡。图6 展示了所有十个数据集上软件平均性能（结构相似性指数（SSIM）得分）与硬件性能（延迟和能耗）之间的权衡。测试时，我们采用了两种通信模式：模拟的 4G 网络和真实环境中的 Wi-Fi 网络。理想结果位于左上角，即重建保真度高、通信成本低。

从图 6可以看出，与集中式方法相比，EPIC 始终取得最高的 SSIM。此外，与分布式方法相比，EPIC-Net 的能耗更低；与集中式方法相比，其能耗则低了一个数量级。换言之，如图 6(b)(d) 所示，在结构相似性指数与能耗权衡方面，EPIC 优于所有其他方案。在延迟方面，与所有集中式基线相比，EPIC 将端到端延迟降低了数倍；但与另外两种分布式方法相比，其延迟更高。这是因为交叉注意力给中心节点带来了计算开销。在我们的测试平台中，中心节点与边缘设备相同；如果使用服务器级计算机，这一延迟差距就会缩小。这些结果表明，EPIC-Net 在所有部署条件下均表现出色，在实现最高重建精度的同时，也达到了最低通信成本，验证了其作为系统与物理协同设计分布式方案的有效性。

EPIC具有很高的鲁棒性。 现场部署中的一项实际要求是对缺失或延迟的边缘节点具有鲁棒性。

表2：EPIC-Net与竞争方法在Open-FWI数据集上的比较。最佳结果以粗体标示，次佳结果加下划线。

数据集	BigFWI-L	InV	InV+	FLA	SLA	EPIC-Net
FlatVel-A	0.9965	0.9895	0.9976	0.9938	0.9944	**0.9987**
FlatVel-B	0.9756	0.9461	0.9817	0.9750	0.9784	**0.9872**
CurveVel-A	**0.9199**	0.8223	0.9153	0.8827	0.8855	0.9198
CurveVel-B	0.8134	0.6727	0.8225	0.7621	0.7759	**0.8351**
FlatFault-A	**0.9918**	0.9798	0.9872	0.9719	0.9824	0.9882
FlatFault-B	0.8033	0.7208	**0.8621**	0.7964	0.8178	0.8601
CurveFault-A	**0.9801**	0.9566	0.9719	0.9457	0.9608	0.9687
CurveFault-B	0.6790	0.6163	0.7586	0.6883	0.7258	**0.7604**
Style-A	0.9136	0.8910	0.9731	0.9507	0.9732	**0.9782**
Style-B	0.7429	0.7599	0.8935	0.8410	0.8290	**0.9147**

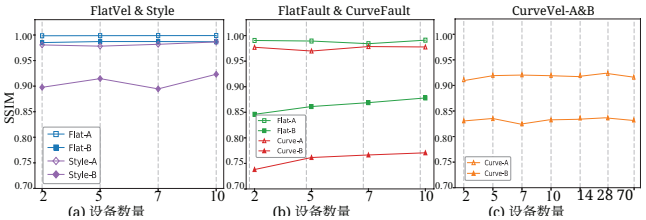


图8：EPIC-Net关于终端设备数量的可扩展性测试结果。

图 7 在FlatFault-B数据集上，通过模拟5个边缘节点中丢失0至4个节点的情况，评估了EPIC相对于集中式方法的鲁棒性。从结果可以看出，InV+ 的性能即使仅缺失一个节点也会急剧下降至0.56。相比之下，当1、2和3个节点不可用时，EPIC的平均结构相似性指数分别从0.86（无节点失效）降至0.79、0.72和0.65。我们还展示了不同情况下重建的地下剖面的可视化结果。与真实值相比，EPIC的结果即使缺失四个潜变量，仍能提取层状特征与断层特征；另一方面，集中式方法的可视化结果即使仅缺失一个节点也会严重失真。EPIC的鲁棒性源于EPIC-Net与EPIC-Mgmt的协同效应。总体而言，EPIC能够在中断情况下持续运行，其方式是避免全局停滞并满足截止时间 T，这是分布式科学系统中实时监测的一项关键要求。

4.3 EPIC-Net评估

硬件-物理协同指导提升软件性能。表2 比较了EPIC-Net与最先进的基线方法在全部十个OpenFWI基准数据集上的表现。EPIC-Net在六个数据集上取得最佳性能，并在其余四个数据集中的三个上取得第二佳性能，展现出具有竞争力的重建质量。值得注意的是，与集中式基线InV相比+，EPIC-Net在10个数据集中的8个上提升了结构相似性指数（SSIM）分数。与分布式基线方法FLA和SLA相比，EPIC-Net在B变体数据集上的优势尤为明显，这些数据集具有更复杂的地质

其地质结构优于对应的 A变体。这些结果表明，EPIC-Net 不仅能够避免分布式设置导致的性能下降，还能在具有多样地质条件的数据集上超越集中式方法，展现了硬件-物理协同引导设计在分布式科学机器学习中的有效性。

可扩展性。 EPIC 的一个关键特性是其相对于终端设备数量的可扩展性。图 8(a)(b) 评估了 EPIC-Net

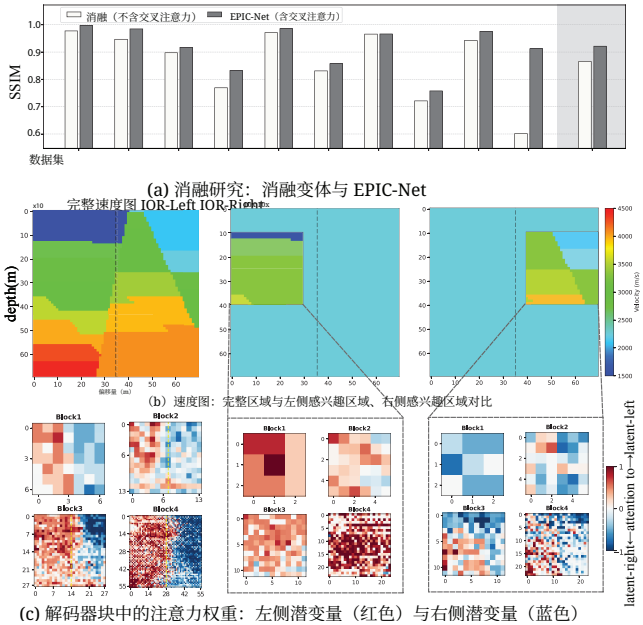


图 9：交叉注意力有效性的消融研究：（a）有无交叉注意力的消融研究；（b）重建的速度图和感兴趣区域；（c）两个感兴趣区域在各解码器块中的注意力权重分布。

在不同配置下，对所有数据集从 2 到 10 个终端设备进行评估。在所有情况下，结构相似性指数保持高度稳定，仅有轻微波动。图 8(c) 进一步将规模扩展到 70，表明每个接收器配备一个终端设备；与配备 28 个终端设备的配置相比，其略有下降。这些结果表明，EPIC 在分布式 EPIC-Infra 上仍保持良好的可扩展性。

消融研究。为评估交叉注意力模块的作用，图9（a）对比了 EPIC-Net 与移除中央解码器中交叉注意力模块后的消融变体。在全部十个数据集上，EPIC-Net 的结构相似性指数始终更高，提升幅度从中等（例如 FlatVel-A/B）到在 CurveVel-B 和 Style-B 等具有挑战性的数据集上显著提升不等。移除交叉注意力会导致性能明显下降，因为解码器无法再融合不同分区的信息，从而损失全局波场相干性。这些结果证实了物理引导对于科学机器学习设计的重要性。

物理引导的注意力验证。此外，我们进一步可视化了 EPIC-Net 如何利用不同空间区域的信息。图 9(b)(c) 展示了两个感兴趣区域重建得到的速度图及其对应的交叉注意力权重。对于每个感兴趣区域，我们提取了四个解码器块中的注意力图。每张热力图都表示在重建相应感兴趣区域时，解码器对左侧区域的潜变量（红色）与右侧区域的潜变量（蓝色）分别赋予了多大的权重。可视化结果揭示了清晰的空间对应关系：左侧感兴趣区域的重建主要关注来自左侧传感器组的潜变量，而右侧感兴趣区域的重建则更多依赖来自右侧传感器组的潜变量。这种行为符合全波形反演的物理规律：附近接收器记录的地震波包含了重建局部地下结构最相关的信息。这些结果

证实，EPIC-Net 的交叉注意力机制学会了以符合物理规律的方式融合信息，并自然地使注意力模式与地震数据的空间分布相匹配。

5 结论

本文介绍了 EPIC，这是一种利用硬件和物理知识这两个互补的老师，实现科学机器学习实际现场部署的分布式框架。以全波形反演作为代表性科学机器学习任务，EPIC 将计算分解为轻量级边缘编码和物理感知的中心解码。在由五台树莓派终端设备组成的测试平台上，分别在 Wi-Fi 和仿真 4G 条件下进行了评估。结果表明，与集中式方法相比，EPIC 可将端到端延迟降低 8.9×，将通信能耗降低 33.8×，同时在 10 个 OpenFWI 数据集中的 8 个上实现相当或更高的重建保真度。我们证明，硬件-物理协同指导能够构建更优的分布式科学机器学习系统，其性能甚至超过集中式科学机器学习。

参考文献

[1] Maziar Raissi、Paris Perdikaris 和 George E Karniadakis。物理信息神经网络：一种用于求解涉及非线性偏微分方程的正问题与逆问题的深度学习框架。《计算物理学杂志》，378:686–707，2019年。[2] Jian Sun、Kristopher A. Innanen 和 Chao Huang。采用混合训练和不确定性分析的物理引导深度学习地震反演。《地球物理学》，86(3):R303–R317，2021年03月。[3] Jian Zhang、Hui Sun、Gan Zhang、Xingguo Huang 和 Li Han。基于深度学习的物理与模型引导同步地震反演。*IEEE*《地球科学与遥感汇刊》，62:1–11，2024年。[4] Junhuan Yang、Yuzhou Zhang、Yi Sheng、Youzuo Lin 和 Lei Yang。一种用于成对地球科学数据生成的新型扩散模型，适用于不平衡训练数据集。载于*AAAI*人工智能会议论文集，第39卷，第21965–21973页，2025年。[5] Weisong Shi、Jie Cao、Quan Zhang、Youhuizi Li 和 Lanyu Xu。边缘计算：愿景与挑战。*IEEE*《物联网期刊》，3(5):637–646，2016年。[6] Mahadev Satyanarayanan。边缘计算的兴起。《计算机》，50(1):30–39，2017年。[7] Junhuan Yang、Yi Sheng、Yuzhou Zhang、Weiwen Jiang 和 Lei Yang。设备端无监督图像分割。载于*2023*年第*60*届*ACM/IEEE*设计自动化会议（*DAC*），第1–6页。IEEE，2023年。[8] Junhuan Yang、Yuzhou Zhang、Yi Sheng、Youzuo Lin、Weiwen Jiang 和 Lei Yang。Digit：用于设备端多模态数据生成的基于扩散的模块化地球物理工具包。*ACM*《嵌入式计算系统汇刊》，2025年。[9] J. Virieux 和 S. Operto。勘探地球物理学中全波形反演概述。《地球物理学》，74(6):WCC1–WCC26，2009年。[10] Albert Tarantola。声学近似下地震反射数据的反演。《地球物理学》，49(8):1259–1266，1984年。[11] Thomas P. Matthews、Kun Wang、Cuiping Li、Neb Duric 和 Mark A. Anastasio。用于全波超声计算机断层成像的正则化对偶平均图像重建。*IEEE*《超声学、铁电学与频率控制汇刊》，64(5):811–825，2017年。[12] Yi Sheng、Junhuan Yang、Youzuo Lin、Weiwen Jiang 和 Lei Yang。迈向公平的超声计算机断层成像：挑战、解决方案与展望。载于*2024*年大湖区超大规模集成电路研讨会论文集，第748–753页，2024年。[13] Junhuan Yang、Yi Sheng、Yuzhou Zhang、Hanchen Wang、Youzuo Lin 和 Lei Yang。利用基于扩散的生成式人工智能增强科学人工智能——超声计算机断层成像案例研究。载于*2024*年大湖区超大规模集成电路研讨会论文集，第754–759页，2024年。[14] Yue Wu 和 Youzuo Lin。InversionNet：一种高效且准确的数据驱动全波形反演方法。*IEEE*《计算成像汇刊》，6 433，2019 419: – 。[15] Brendan McMahan、Eider Moore、Daniel Ramage、Seth Hampson 和 Blaise Aguera y Arcas。基于去中心化数据的深度网络通信高效学习。载于人工智能与统计国际会议，第1273–1282页。PMLR，2017年。[16] Otkrist Gupta 和 Ramesh Raskar。多个智能体上的深度神经网络分布式学习。《网络与计算机应用杂志》，116 8，2018 1: – 。[17] Hezheng Lin、Xing Cheng、Xiangyu Wu 和 Dong Shen。CAT：视觉变换器中的交叉注意力。载于2022年IEEE多媒体国际会议和展（*ICME*），第1–6页，2022年。[18] Chengyuan Deng、Shihang Feng、Hanchen Wang、Xitong Zhang、Peng Jin、Yinan Feng、Qili Zeng、Yinpeng Chen 和 Youzuo Lin。OpenFWI：用于全波形反演的大规模多结构基准数据集。载于神经信息处理系统进展，第35卷，第6007–6020页，2022年。[19] Junxian Huang、Feng Qian、Alexandre Gerber、Z. Morley Mao、Subhabrata Sen 和 Oliver Spatscheck。深入考察4G LTE网络的性能和功耗特征。载于第*10*届移动系统、应用与服务国际会议论文集，MobiSys '12，第225–238页，美国纽约州纽约市，2012年。计算机协会。[20] Zhou Wang、Alan C Bovik、Hamid R Sheikh 和 Eero P Simoncelli。图像质量评估：从误差可见性到结构相似性。*IEEE*图像处理汇刊，13(4):600–612，2004年。[21] Tom Beucler、Michael Pritchard、Stephan Rasp、Jordan Ott、Pierre Baldi 和 Pierre Gentine。在模拟物理系统的神经网络中施加解析约束。《物理评论快报》，126:098302，2021年3月。[22] Jared Willard、Xiaowei Jia、Shaoming Xu、Michael Steinbach 和 Vipin Kumar。将科学知识与机器学习相结合，应用于工程与环境系统。*ACM*计算综述，55(4)，2022年11月。[23] Junhuan Yang、Hanchen Wang、Yi Sheng、Youzuo Lin 和 Lei Yang。Edgeo：用于边缘设备地球物理监测的物理引导生成式人工智能工具包。载于第*61*届*ACM/IEEE*设计自动化会议（*DAC*）论文集，第1–6页，2024年。[24] Ashish Vaswani、Noam Shazeer、Niki Parmar、Jakob Uszkoreit、Llion Jones、Aidan N Gomez、Ł ukasz Kaiser 和 Illia Polosukhin。《注意力就是你所需要的一切》。载于I. Guyon、U. Von Luxburg、S. Bengio、H. Wallach、R. Fergus、S. Vish- wanathan 和 R. Garnett 编，《 style id='22'>神经信息处理系统进展，第30卷。Curran Associates, Inc.，2017年。[25] Peng Jin、Yinan Feng、Shihang Feng、Hanchen Wang、Yinpeng Chen、Benjamin Consolvo、Zicheng Liu 和 Youzuo Lin。大规模数据驱动全波形反演的实证研究。《科学报告》，14(1):20034，2024年。[26] Ilya Loshchilov 和 Frank Hutter。解耦权重衰减正则化。载于国际学习表征会议，2019年。