以太网的双边效率：五十年后重新校准

Metcalfe 和 Boggsv0.3— 2026年3月19日

Metcalfe 和 Boggs 50周年纪念，《*ACM* 通讯》 19(7)，1976年7月

Paul Borrill DÆDÆLUS — 开放原
子以太网项目 旧金山，加利福尼亚州，美
国 paul@daedaelus.com

arXiv:2603.19406v1 [cs.DC] 19 Mar 2026

摘要

1976年7月，Metcalfe 和 Boggs 在《*ACM* 通讯》上发表了关于以太网的奠基性论文。他们的效率模型——$E = (P/C)/(P/C + W \cdot T)$——衡量的是在竞争状态下，承载有效前向数据包的以太信道时间比例。五十年来，这一模型一直定义着网络社区对以太网性能的理解。我们认为，该模型虽然适用于其预定目的，但只衡量前向 信道，并未回答现代分布式系统中至关重要的问题：双边事务效率——即链路时间中用于在发送方与接收方之间产生已提交协议的比例。

我们指出，Metcalfe 和 Boggs 本人直觉上就理解了这种区分。他们的 EFTP“末端延迟”协议（原论文第 7.2.2 节）是一种三阶段双边握手，旨在实现对传输完成的共同认知——而这恰恰是他们的效率模型无法捕捉的特性。我们将这一观察与开放原子以太网的双边事务原语、采用完美信息反馈的背靠背香农信道表述，以及物理学中的双态矢量形式主义（TSVF）联系起来。后者为理解为何必须同时指定事务的两个边界条件——发送方和接收方——才能使事务具有确定值，提供了理论框架。

对 Metcalfe 和 Boggs 的表1所做的修正，并不是换了一组数字，而是换了一个问题。

1 引言

在五十周年纪念之际，Metcalfe 和 Boggs 的 1976年论文 [1] 仍是计算机发展史上最具影响力的出版物之一。它引入的不只是一项技术，更是一种设计理念：分布式控制、无源共享介质、统计仲裁，以及将数据包传输与可靠交付分离。论文中的表1“以太网效率”，成为几代工程师思考网络性能的范本。

然而，产生表1的效率模型包含一个因习以为常而变得不可见的无声假设：它只测量正向吞吐量。分数 E 告诉我们，以太信道时间中有多大比例承载着从源端发往目的地的有效数据包。它没有说明这些数据包是否到达、接收方是否处理它们，或者发送方和接收方是否同意一次传输已经完成。

这并非模型的缺陷——而是适用于1976年的适用范围限制，当时亟待解决的问题是共享同轴电缆能否在竞争环境下支持有效流量。但这一适用范围限制却被误认为是完整答案。现代以太网运行在截然不同的环境中——点对点交换链路、全双工、800 Gbps SerDes——然而，性能问题仍然沿用原始模型所继承的前向吞吐量表述。

我们认为，2026年的正确问题并不是“链路时间中有多大比例承载着有效的正向数据包？”，而是“链路时间中有多大比例促成了通信站点之间已提交的双边协议？”本文分四步展开这一论证：(1) 我们说明，Metcalfe 和 Boggs 本人在其 EFTP 末端延迟协议中就埋下了双边事务的种子；(2) 我们将正向效率与双边效率之间的区别形式化；(3) 我们把这一形式化与背靠背香农信道模型以及物理学中的双态矢量形式主义联系起来；(4) 我们说明，最初的以太网碰撞检测机制本身就是一种非 FITO 双边原语，只是一直隐藏在显而易见之处。

2 末端延迟：1976年的一项双边事务

Metcalfe 和 Boggs 的第 7.2.2 节介绍了 EFTP 中用于完成文件传输的“End”协议。我们认为，这段文字值得大段引用，因为它是整篇论文中最被低估的一段。

> 该协议分三个阶段进行。首先，发送方发送一个序列号连续递增的 END 数据包。其次，接收方以匹配的 ENDREPLY 作出回应，然后等待 “一段相当长的时间（10 秒）”。第三，发送方发送一个回显 ENDREPLY，并且“可以放心离开，因为文件已成功传输。等待中的接收方随后收到回显的 endreply，也放心离开。”——Metcalfe 和 Boggs [1], §7.2.2

“放心 ”一词出现了两次——一次指发送方，一次指接收方。这并非偶然。Metcalfe 和 Boggs 力图实现**共同认知**：双方都知道传输已成功，而且也知道对方知道这一点。10 秒的末端延迟，是对开放原子以太网通过双边事务以确定性方式实现这一点的基于超时的近似。

2026年3月，Metcalfe 在私人通信中证实了这一解读：“据我回忆，末端延迟的设计初衷是加快进程。就像 NAK 通常发挥的作用一样” [15]。这种比较

很能说明问题。基于 NAK 的协议比 ACK 超时协议更快，原因恰恰在于反向流动的信息——接收方报告失败——能让系统更快达成一致。发送方不必默默等待可能永远不会到来的 ACK，接收方会主动将结果告知发送方。反向信息畅通流动时，系统会更加高效，而非效率更低。

这是以太网发明者本人的深刻见解。Metcalfe 在 1976年就已明白，双向反馈——信息从接收方回传给发送方——并非需要尽量减少的开销，而是值得加以利用的结构。结束延迟之所以能加快进程，是因为它比保持沉默更快地弥合了发送方与接收方之间的认知鸿沟。然而，他在同一篇论文中发表的效率模型只衡量前向信道，将反向路径视为不存在。指标与协议设计优化的目标并不相同。

关键在于，EFTP 协议会在数据传输阶段（第 7.2.1 节）与完成阶段（第 7.2.2 节）之间切换模式：

> • **数据传输（7.2.1）**是单向且带有概率性的。数据包向前发送，确认则零星地返回。发送方发出数据后便不再过问；接收方则会悄然丢弃受损的数据包。在传输过程中，不存在某个时刻能让双方同时就传输状态达成一致。• **末端延迟（7.2.2）**力求实现双向且确定的交互。三阶段握手（END → ENDREPLY → 回显的 ENDREPLY）旨在实现论文所说的：“文件的发送方和接收方将就文件是否已正确传输达成一致。”

数据阶段针对前向吞吐量进行优化。完成阶段则针对双向一致进行优化。这些是根本不同的目标，而第 6 节的效率模型只衡量前者。

3 效率模型衡量什么——以及它不衡量什么

Metcalfe 和 Boggs 通过考察交替出现的 Ether 时间段来建立其性能模型：传输间隔，即 Ether 承载成功数据包的期间；以及竞争间隔，由重传时隙组成，在此期间站点竞争访问权。效率为：

$$E = \frac{P/C}{P/C + W \cdot T} \tag{1}$$

其中，P 是以位为单位的数据包大小，C 是以每秒位数为单位的峰值信道容量，T 是时隙持续时间（一个往返传播延迟），$W = (1 - A)/A$ 是平均竞争时隙数，而 $A = (1 - 1/Q)^{Q-1}$ 则是恰好由 Q 个持续排队站点中的一个成功获得访问权的概率。

该模型对于其既定目的而言既简洁又正确。但它所衡量的只有一件事：**Ether 时间中承载前向传输有效数据包的比例**。它是一个纯粹的 FITO（仅前向时间）吞吐量指标。它统计成功传输期间线路上的比特数，并除以包括竞争时隙在内的线路总时间。

模型没有说明：

(1) **接收方处理**。 目的站点是否实际接收、缓冲并处理了数据——按照 Semantic Arrow 系列 [2] 的记法，这正是物理交付（T_4）与语义一致（T_6）之间的差距。

(2) **双边完整性**。 发送方和接收方是否确认 传输已完成——这正是末端延迟协议试图提供的特性。

(3) **end-dally 开销**。END/ENDREPLY/echosequence 所消耗的多次往返；表1完全忽略了这些开销，因为它们发生在不同的协议层。

(4) **重传成本**。 因冲突、噪声或接收方过载而丢失的数据包必须重传；数据阶段 ACK 超时和重传开销在竞争模型中不可见。

(5) **语义正确性**。 已送达的数据 对应用程序是否有意义——这一区别在于“比特已到达”和“事务已提交”，而现代 RDMA 系统总是系统性地将其略去。

Metcalfe 本人对 NAK 的比较（第2节）揭示了这一差距。他意识到，反向流动的信息能让系统 更高效——结束延迟让事务比保持沉默更快完成。但效率模型 没有纳入反向信息这一项。公式（1）计算正向数据包传输时间和竞争时隙；其中没有任何变量代表接收方对系统性能的贡献。尽管 Metcalfe 本人指出 NAK 能加快处理速度，但在 E 中，NAK 是不可见的。

简言之，Metcalfe 的 E 回答的是：“在有 Q 个站点争用的情况下，以太信道时间中有多少比例用于传输有效的前向数据包？”对于 2026 年，我们需要回答的问题是：“在有 Q 个站点进行交易的情况下，有多少比例的链路时间能产生已提交的双边协议？”

4 双边效率：一种新指标

我们将**双边效率** E_B 定义为：

$$E_B = \frac{\text{committed transactions}}{\text{total link-seconds}} \tag{2}$$

这里的“已提交”是指双方都已获得成功的相互认知——这正是 Metcalfe 试图通过 end-dally 实现的条件。

对于最初的 1976年以太网，$E_B \ll E$。即使有 256 个排队站点，大数据包（$P = 4096$ 位）的正向效率E 仍保持在 98% 以上。但双边效率还必须计入：

> • 数据阶段 ACK 往返（每个数据包至少一次）。• 三阶段 end-dally 握手。• 作为最坏情况开销的 10 秒 dally 超时。
>
> • 由丢失的 ACK 或丢失的 END/ENDREPLY 数据包触发的重传。

对于现代交换式全双工以太网，E 与 E_B 之间的差距甚至更为严重。这里不存在共享介质竞争，因此$E \approx 1$ 轻而易举——线速转发是基准。但每个 RDMA“完成”都会在物理交付时报告成功（T_4），而应用程序可能并未达成语义一致（T_6），这就造成了 SemanticArrow 系列所称的系统性语义损坏 [2]。正向效率完美无缺；双边效率却可能差到任意程度。

5 背靠背香农信道与完美信息反馈

开放原子以太网（OAE）规范 [3] 将链路重新表述为两个同时运行的香农信道——前向和反向，并采用逐片反馈。OAE 不再采用只有一个前向信道、事后才附加确认的做法，而是将返回路径视为与前向路径并发运行的一等信息信道 。

OAE 定义了香农式知识深度的四个递进级别，每个级别都由切片确认（SACK）确认：

SACK 00 — 信息（惊异度）
无错误地检测到第一个 8 字节切片。二元问题：“硬件是否检测到了有效位？”这确认了线缆正常工作且串行器/解串器功能正常。这是纯粹的香农信息，用于区分接收与未接收。

SACK 01 — 知识（捕获的信息）
在网络物理层中捕获循环冗余校验以及预期的头字段。数据已识别并安全存储。

SACK 10 — 语义（含义）
已接收 32 字节；可进行部分解码。判定协议或消息类型。确认缓冲区槽位和环形描述符可用。已加载正确的状态机。

SACK 11 — 理解（句法）
已接收完整的 64 字节，并确认其符合有效的帧布局。消息已准备好发送到 PCIe 总线或内部环形缓冲区。无需进一步进行二层修复。

这是 Norm Abramson 的完美信息反馈（PIF） [4], 在物理层上的实现。发送方可以逐片看到自己发送的数据包从接收方返回，从而无需校验和、循环冗余校验码、前向纠错或奇偶校验，就能检测信道错误。双向信道不只是确认——它还共同构建这笔事务。

这种双向信道的效率并非

$$E = P\ C \quad P\ C + (\ /\ )/(\ /\ \cdot)。$$

$W\ \ T \qquad\qquad W\ =\ 0$
这里不存在竞争（在交换式点对点链路上），而且反向信道的“成本”为零，因为它利用的是全双工以太网中本已存在的返回路径。双向效率为：

$$E_B^{\mathrm{OAE}} = \frac{N_{\mathrm{committed}}}{N_{\mathrm{attempted}}} \cdot \frac{P_{\mathrm{eff}}}{P_{\mathrm{eff}} + \Delta T_{\mathrm{commit}}} \tag{3}$$

其中， $N_{\mathrm{committed}}/N_{\mathrm{attempted}}$ 是事务成功率， P_{eff} 是有效载荷持续时间， $\Delta T_{\mathrm{commit}}$ 是实现双边承诺所需的边际时间成本，超出仅完成前向传输所需的时间。对于 OAE， $\Delta T_{\mathrm{commit}} \approx 0$ ，因为 SACK 反馈与前向数据流水传输——双边承诺是在传输期间实现的，而非传输之后。

6 TSVF 联系：双方边界均为必需

由 Aharonov、Bergmann 和 Lebowitz [5] 提出，并由 Aharonov 和 Vaid- man [6], 拓展的双态矢量形式主义（TSVF），通过一个向前演化的态矢量 $|\psi$（由初始条件制备）和一个向后演化的态矢量 $\phi|$（由最终测量确定）来描述量子系统。

可观测量 A 的“弱值”为：

$$\langle A\rangle_w = \frac{\langle\phi|A|\psi\rangle}{\langle\phi|\psi\rangle} \tag{4}$$

该表达式只有在明确指定双方边界条件时才具有明确定义的值。仅有 $|\psi$（正向状态）或仅有 $\phi|$（反向状态）都无法确定弱值。双方边界都是必需的。

与双向网络的类比是精确的：

- $|\psi$ 是发送方的知识沿信道向前传播。这正是 Metcalfe 的 E 所衡量的内容：正向信道占用率。• $\phi|$ 是接收方验证向后传播。这正是结束延迟协议以及 OAE 的 SACK 机制所提供的内容：确认正向状态已被接收、处理并接受。

• 弱值 A_w 是事务值——只有在双方边界都得到解析时才存在的已提交结果。FITO 效率模型只看到 $|\psi$，便宣称信道效率为 98%。双向模型要求 $\phi|\psi \neq 0$——内积必须非零，这意味着接收方验证必须与发送方意图相容。

只测量正向吞吐量的网络效率模型，就像只指定制备过程、却忽略后选择的量子测量。它会产生一个系综平均，而这个平均值可能与任何单个事务的结果都毫无关联。双态矢量形式主义告诉我们，单个结果需要两个边界条件——单个事务也是如此。

7 碰撞检测：原始的非 FITO 基元

以太网的历史中有一个深刻的讽刺：让 Metcalfe 竞争模型得以成立的机制——CSMA/CD 碰撞检测——本身就是一种双边物理交互，而不是 FITO 交互。

当两个站点同时传输并检测到冲突时，它们是通过物理交互而非消息传递，获得共同认知。双方都知道冲突发生在同一物理时刻（传播延迟 τ范围内）。因果闭合条件 $T \geq 2\tau$——最小时隙时间——是一种**双边约束**：它确保两个站点都有足够的时间在自身传输完成前检测到对方的信号。

效率模型利用这一双向物理机制生成竞争时隙，但随后测量的却只有由此产生的前向吞吐量。碰撞检测机制是 OAE 完美信息反馈的基础：双方都利用了这样一个事实——发送方可以观察到自身信号如何受到接收方（或干扰方）活动的影响。在原始以太网中，这种双向观察被用于破坏性目的——检测失败。在 OAE 中，它则被用于建设性目的——逐片确认成功。

因果闭合条件 $T \geq 2\tau$ 有一个 OAE 类比：最小事务时间必须至少为 2τ（一个往返传播时延），以便双向 SACK 能够在事务窗口关闭之前传播回来。Metcalfe 在 1976年 为冲突检测推导出了这一约束；我们在 2026年 为双向承诺重新推导了它。物理学是相同的，诠释则发生了转变。

8 表1的修正

Metcalfe 和 Boggs 的表1列出了其实验性以太网（$C = 3$ Mbps，$T = 16\ \mu$s）在数据包大小为 $P \in$ {48、512、1024、4096} 位、站点数量为 $Q \in$ {1、2、……、256} 时的效率值。这些数值对于正向效率 E 是正确的。

我们认为，五十周年纪念正是为表1补充一个双边对应项的恰当时机：表2，在相同参数范围下展示 E_B，但纳入完整的事务生命周期——不仅是数据包的正向传输，还包括双边承诺。

表1的修正并不是换了一组数字，而是换了一个问题。五十年来，我们一直在问“线路传输数据包的效率有多高？”并为超过98%的数字而自鸣得意。双边问题——“链路生成已提交协议的效率有多高？”——从未被提出过，因为 FITO假设使它隐而不见。

Metcalfe 在第 7.2.2 节中已经提出了这个问题的雏形。端到端协议是一种双边握手，试图通过基于超时的近似来实现相互知晓。OAE 以确定性双边事务取代这种近似，以因果闭合条件取代超时，并以逐片的完美信息反馈取代发后不管数据阶段。

9 结论：五十年后

Metcalfe 和 Boggs 的 1976年论文之所以卓越，在于它既包含 FITO 模型，也埋下了自我修正的种子。第 6 节的效率公式纯粹是 FITO——在共享介质上的正向吞吐量。但第 7.2.2 节的 end-dally 协议追求的是 FITO 无法表达的东西：发送方与接收方之间的双向一致。

如今，Metcalfe 本人已经证实了这一解读。他将末端延迟与 NAK [15]相比较，揭示了一个被五十年来的正向吞吐量优化所掩盖的原则：**反向信息自由流动时，系统会变得更高效，而非更低效**。NAK 之所以能加快速度，是因为它弥合了发送方与接收方之间的认知鸿沟。末端延迟加快速度，原因也相同。OAE 逐片段进行的 SACK 反馈再次以同样的原因加快了速度。它们各自都是同一双边原则的体现：同时承载双向信息的信道，比依赖正向传递和期盼沉默的信道收敛得更快。

五十年来，网络社区一直在优化 E——这一正向指标。其成果便是现代以太网：800 Gbps 线路速率、近乎完美的正向效率，以及在网络线上层层发生的系统性语义损坏。RDMA 以线路速率传输字节，并在应用程序达成一致之前就宣告成功。TCP 会不断重传，直到收到 ACK，并称之为“可靠”，却从未验证语义正确性。末端延迟对相互保证的追求，已被越来越快的单向传输所取代。

开放原子以太网重新拾起了 Metcalfe 在 1976年所理解、而 FITO时代所遗忘的观点：网络的目的不是传递数据包，而是达成协议。双边事务才是基本单元。配备完美信息反馈的背靠背香农信道是其实现机制。双态矢量形式主义提供了理论框架：事务如同弱测量，只有在双方的边界条件——发送方和接收方——都已确定时，才具有确定的值。

值此以太网五十周年之际，我们纪念 Metcalfe 和 Boggs 的方式，不是替换表1，而是对其加以扩展——从正向效率扩展到双边效率，从 $|\psi$ 扩展到 $\phi|A|\psi$，从数据包送达扩展到事务提交。

AI 披露：本文在 Claude（Anthropic）的协助下完成，Claude 被用作研究和起草工具。所有技术内容、论点和结论均为作者本人所作。

参考文献

[1] R. M. Metcalfe 和 D. R. Boggs，“以太网：用于局域计算机网络的分布式分组交换”，《ACM 通讯》，第 19 卷，第 7 期，第 395–404 页，1976年7月。doi | pdf [2] P. Borrill，“语义时间箭头，第三部分：RDMA 与完成谬误”，arXiv:2603.04774 [cs.DC], 2026。arXiv | pdf [3] P. Borrill，“《开放原子以太网：规范与架构》”，DÆDÆLUS技术报告，2026年。[4] N. Abramson，“卫星分组交换”，AFIPS 会议论文集，第 42 卷，NCC，第 695–702 页，1973年。doi [5] Y. Aharonov、P. G. Bergmann 和 J. L. Lebowitz，“量子测量过程中的时间对称性”，《物理评论》，第 134 卷，1416，1964 6 1410，第 B 期，第 B –B 页。doi [6] Y. Aharonov 和 L. Vaidman，“双态矢量形式主义：更新综述”，《物理学讲义》，第 734 卷，第 399–447 页，2008年。doi | pdf [7] P. Borrill，“分布式计算的错误之处：将设计选择变成自然法则的范畴错误”，arXiv:2602.18723 [cs.DC], 2026。arXiv | pdf [8] P. Borrill，“双同步先进先出队列与 FITO 范畴错误：经硅验证的分布式协调交互原语”，arXiv:2603 03470 [ ], 。2026 cs.DC 。arXiv | pdf [9] P. Borrill，“语义时间箭头，第一部分：从 Eddington 到以太网”，arXiv:2603.01440 [cs.DC], 2026。arXiv | pdf [10] P. Borrill，“不具时间方向的消息传递：约束语义与 FITO 范畴错误”，arXiv:2603.01405 [cs.DC], 2026。arXiv | pdf [11] P. Borrill，“链路之争：语义危机。争论结束了吗，还是才刚刚开始？”arXiv:2603.07391 [cs.DC], 2026。arXiv | pdf [12] P. Borrill，“语义时间箭头，第五部分：莱布尼茨桥——迈向语义时间统一理论”，arXiv:2603.04826 [cs.DC], 2026。arXiv | pdf [13] R. M. Metcalfe，“分组通信”，博士论文，哈佛大学，MAC项目 TR-114，1973年12月。论文 [14] C. E. Shannon，“通信的数学理论”，《贝尔系统技术期刊》，第 27 卷，第 379–423、623–656 页，1948年。doi | pdf [15] R. M. Metcalfe 于2026年3月10日致作者的个人通信。

表1：效率范式比较。

	Metcalfe E (1976年)	**双边效率 E_B (2026年)**
措施	正向数据包吞吐量	已承诺的双边交易
信道	单向（仅正向）	双向（正向 + 反向）
成功	“以太网上无冲突的数据包”	“双方相互确认”
模型	$E = (P/C)/(P/C + W \cdot T)$	$E_B = N_c/N_a \cdot P_e/(P_e + \Delta T_c)$
边界	仅发送方 ($\|\psi$)	双方（$\phi\|A\|\psi$）
冲突	失败（中止并重试）	反馈（双方观测）
末端延迟	隐性开销	交易成本的组成部分
物理学	FITO（仅向前时间）	双态矢量形式主义（时间对称）