

Exaflop 是衡量超级计算机性能的单位，表示该计算机每秒可至少进行百亿亿次浮点运算。

为了解决这个时代最复杂的问题，比如如何治疗像新冠肺炎和癌症这样的疾病、以及如何缓解气候变化等。计算机的计算量正在不断增加。

所有这些重大挑战将计算带入了现今的百亿亿次级时代，顶级性能通常以 exaflops 来衡量。

### 什么是 Exaflop ?

Exaflop 是衡量超级计算机性能的单位，表示该计算机每秒可以至少进行  $10^{18}$  或百亿亿次浮点运算。

Exaflop 中的 exa-前缀表示“百亿亿”，即 10 亿乘以 10 亿或 1 的后面有 18 个零。同样，单个 exabyte 的内存子系统可以储存百亿亿字节的数据。

exaflop 中的“flop”是浮点运算的缩写。exaflop/s 是表示系统每秒浮点运算次数的单位。

浮点是指所有数字都用小数点表示的计算方法。

### 1000 Petaflop = 1 Exaflop

前缀 peta- 表示  $10^{15}$ ，即 1 的后面有 15 个零。因此 1 exaflop 等于 1000 petaflop。

### 1 exaflop

的计算量到底有多大？相当于十亿人中的每个人都拿着十亿个计算器。

如果他们同时按下等号，就是进行了 1 个 exaflop。

拥有 Big Red 200 和其他几台超级计算机的印第安纳大学表示，exaflop

计算机的速度相当于一个人每秒钟进行一次计算，并一直计算 31,688,765,000 年。

## Exaflop 简史

在超级计算发展史的大部分时间里，一次浮点运算就是一次，但随着工作负载引入 AI，这种情况也发生了变化。

人们开始使用最高的精度格式来表示数字，这种格式被称为双精度，由 IEEE 浮点运算标准定义。它之所以被称为双精度或 FP64，是因为计算中的每个数字都需要以 64 位用 0 或 1 表示的数据块表示，而单精度为 32 位。

双精度使用 64 位确保每个数字都精确到很细微的部分，比如  $1.0001 + 1.0001 = 2.0002$ ，而不是  $1 + 1 = 2$ 。

这种格式非常适合当时的大部分工作负载，比如从原子到飞机等全部需要确保模拟结果接近于真实的模拟。

因此，当 1993 年全球最强大的超级计算机榜单 TOP500 首次发布时，衡量 FP64 数学性能的 LINPACK 基准（又称 HPL）自然成为了默认的衡量标准。

## AI 大爆炸

十年前，计算行业发生了 NVIDIA 首席执行官黄仁勋所说的 AI 大爆炸。

这种强大的新计算形式开始在科学和商业应用上展现出重大成果，而且它运用了一些非常不同的数学方法。

深度学习并不是模拟真实世界中的物体，而是在堆积如山的数据中筛选，以找到能够带来新洞察的模式。

这种数学方法需要很高的吞吐量，所以用经过简化的数字（比如使用 1.01 而不是 1.0001）进行大量计算要比用更复杂的数字进行少量计算好得多。

因此 AI 使用 FP32、FP16 和 FP8 等低精度格式，通过 32 位、16 位和 8 位数让用户更快地进行更多计算。

## 混合精度不断发展

AI 使用 64 位数就如同在周末外出时带着整个衣柜。

研究人员一直在积极地为 AI 寻找理想的低精度技术。

例如首个 NVIDIA Tensor Core GPU——Volta，它使用了混合精度，并以 FP16 格式执行矩阵乘法，然后用 FP32 累积结果以获得更高的精度。

### Hopper 通过 FP8 加速

最近，NVIDIA Hopper 架构首次发布了速度更快的低精度 AI 训练方法。Hopper Transformer Engine 能够自动分析工作负载，尽可能采用 FP8 并以 FP32 累积结果。

在进行计算密集度较低的推理工作，比如在生产中运行 AI 模型时，TensorFlow 和 PyTorch 等主要框架通过支持 8 位整数实现快速性能，因为这样就不需要使用小数点来完成工作。

好消息是，NVIDIA GPU 支持上述所有精度格式，因此用户可以实现每个工作负载的最优加速。

去年，IEEE P3109 委员会开始为机器学习中使用的精度格式制定行业标准。这项工作可能还需要一到两年的时间才能完成。

### 一些模拟软件在低精度工作中大放异彩

虽然 FP64 在模拟工作中仍然很受欢迎，但当低精度数学能够更快提供可用结果时，许多人会使用后者。

### 影响 HPC 应用程序性能的因素各不相同

例如，研究人员用 FP32 运行广受欢迎的汽车碰撞模拟器——Ansys LS-Dyna。基因组学也倾向于使用低精度数学。

此外，许多传统的模拟开始在部分工作流程中采用 AI。随着越来越多的工作负载使用 AI，超级计算机需要支持较低的精度才能有效运行这些新兴应用。

## 基准与工作负载同步发展

在认识到这些变化后，包括 Jack Dongarra (2021 年图灵奖得主和 HPL 的贡献者) 在内的研究人员在 2019 年首次发布了 HPL-AI，这项新基准更适合测量新的工作负载。

Dongarra 在 2019 年的博客中表示：“无论是技术不断优化的传统模拟，还是 AI 应用，混合精度技术对于提高超级计算机的计算效率越来越重要。正如 HPL 实现了对双精度能力的基准测试一样，这种基于 HPL 的新方法可以对超级计算机的混合精度能力进行大规模基准测试。”

尤利希超级计算中心主任 Thomas Lippert 同意了这一观点。

他在去年发表的一篇博客中表示：“我们使用 HPL-AI 基准是因为它既能够准确地衡量日益增加的 AI 和科学工作负载中的混合精度工作，也能反映准确的 64 位浮点计算结果。”

## 现今的 Exaflop 系统

在 6 月的一份报告中，全球 20 个超级计算机中心提交了 HPL-AI 结果，其中有三个中心提供了超过 1 exaflop 的性能。

在这些系统中，橡树岭国家实验室的超级计算机在 HPL 上的 FP64 性能也超过了 1 exaflop。

## 2022 年 6 月 HPL-AI 结果的采样器

两年前，一非传统系统首次达到 1 exaflop。这台由 Folding@home 联盟组装的众源超级计算机在呼吁帮助抵御新冠疫情后，达到了这一里程碑，到现在已有超过 100 万台计算机加入其中。

## 理论和实践中的 Exaflop

许多组织从那时起就已开始安装理论峰值性能超过 1 exaflop 的超级计算机。需要注意的是，TOP500 榜单同时发布 Rmax (实际) 和 Rpeak (理论) 分数。

Rmax 指计算机实际表现出的最佳性能。

Rpeak 是一切系统都处于高水平运行时的最高理论性能，而这几乎从未发生过。该数值的计算方法通常是将系统中的处理器数量乘以其时钟速度，然后再将结果乘以处理器在一秒钟内可执行的浮点运算数。

因此，如果有人说他们的系统达到 1 exaflop，请问他说的是 Rmax ( 实际 ) 还是 Rpeak ( 理论 )。

### Exaflop 时代的众多指标

这也是新百亿亿次时代的众多细微变化之一。

值得注意的是，HPL 和 HPL-AI 属于合成基准，即它们衡量的是数学程序的性能，而不是真实世界的应用。MLPerf 等其他基准则基于真实世界中的工作负载。

最后，衡量系统性能的最佳标准当然是它运行用户应用程序的情况。该衡量标准不是基于 exaflop，而是基于投资回报率。