



# GRAPHCORE软件栈：为扩展而构建

## 简介

用于新型处理器设计的软件对于实现应用程序部署和优化性能至关重要。总部位于英国的初创公司 Graphcore 是一家用于应用程序加速的芯片供应商。该公司非常重视软件，将大约一半的工程人员配置到应对这一挑战的工作中。Graphcore 的智能处理器（IPU）利用算法的表达作为有向计算图，该公司的 Poplar 软件栈将模型和算法转换为这些计算图并加以执行。该软件使 AI 和并行计算芯片的采用变得更为简单，对于公司的成功至关重要。本文探讨了该公司软件所带来的种种益处，并讨论了这些能力是如何加快在 Graphcore IPU 上运行的应用程序的开发和部署的。

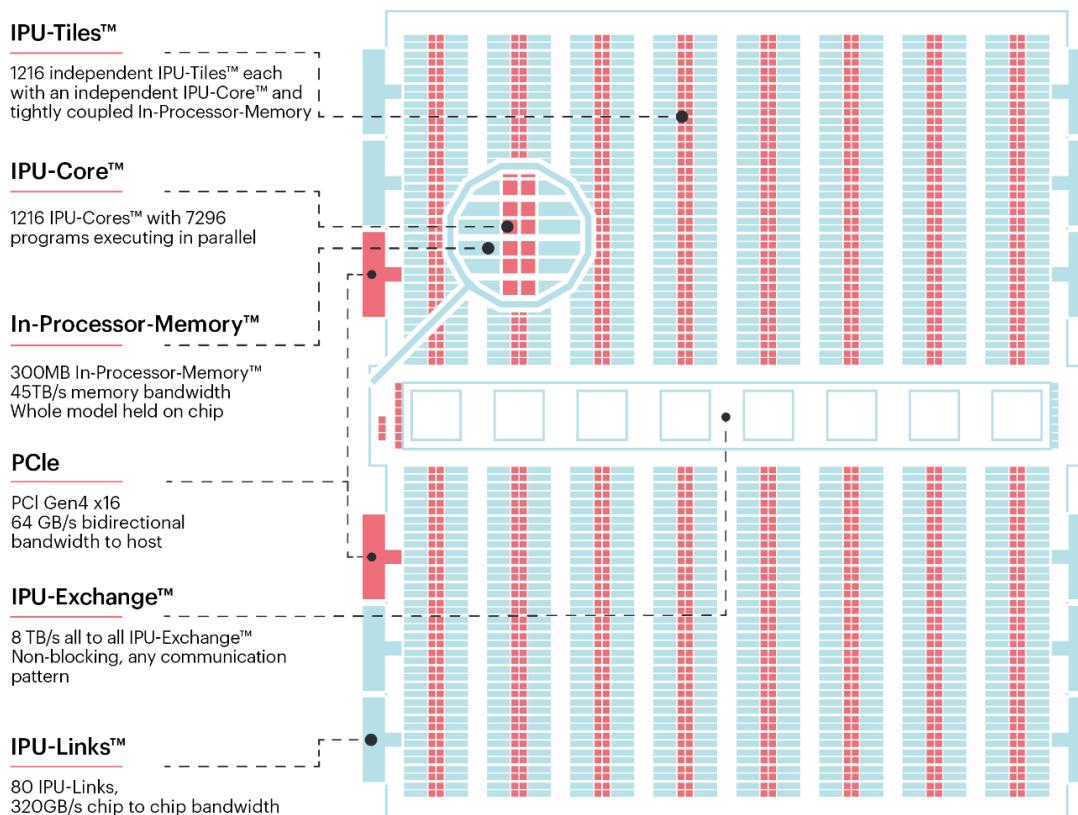
## 智能处理器的简要概述

在深入研究软件之前，简单了解一下基础硬件很有帮助。Graphcore的IPU与当今的CPU、GPU和其他AI处理器存在本质不同。IPU和Poplar软件组合并不是严格意义上的机器学习（ML）工具集。IPU是一种灵活的、可扩展的、细粒度的并行处理器，它的设计初衷是为各种计算密集型算法提供高性能。IPU和Poplar共同构成了一个计算图形编程平台，该平台支持金融、高性能计算（HPXC）、机器人技术和数据科学领域的应用程序，并支持机器智能工作负载。

IPU的设计目标是解决大多数ASIC和GPU中存在的那些当前加速架构无法解决的问题。通常，这些芯片针对密集的线性代数工作负载进行了优化，这些工作负载可以在某些卷积神经网络中执行，不太适合更加变化多样的计算、通信或数据访问模式的应用。

IPU由1216个相互连接的处理tile组成。每个tile都有自己的核和本地片上SRAM内存，以使模型和数据能够驻留在IPU上，从而大大改善了内存带宽和延迟。这些tile以8TB/秒的速率通过片上结构（“IPU-Exchange”）互连，该结构也通过以320GB/秒的速度运行的“IPU-Link”进行连接，以创建芯片到芯片的结构。

图 1：GRAPHCORE 智能处理器



Graphcore智能处理器具有1216个tile，所有计算核和内存均驻留其中。这些tile都通过高速、低延迟的“Exchange”结构连接，该结构也扩展到芯片之外，从而使多芯片并行性可达到数千个芯片之多。

来源: *Graphcore*

本质上，整个系统（通常由许多 IPU 组成）在两个同步阶段执行：计算和通信。针对 IPU 的应用程序被表示为计算图。在计算图的顶点执行计算，然后根据与计算图互连的边将结果传递到相邻的顶点。通信阶段通过批量同步并行（BSP）操作实现，该操作可将数据从每个 tile 的片上 SRAM 内存有效地传输到相连的 tile 的内存。除了计算指令外，每个 IPU 核还具有专用的 tile 级指令集，用于 BSP 模型的通信阶段。

集成的交换-通信结构旨在支持通过计算图编译器实现的、用于数据和模型并行性的BSP，并可能扩展到数千个节点。根据Graphcore的说法，IPU架构的一个重要区别是其能够有效处理稀疏数据和计算图的能力，这可以提高性能，同时减少总内存需求。

Graphcore的Poplar SDK通过自动执行编译和优化工作流程来解决这种独特架构

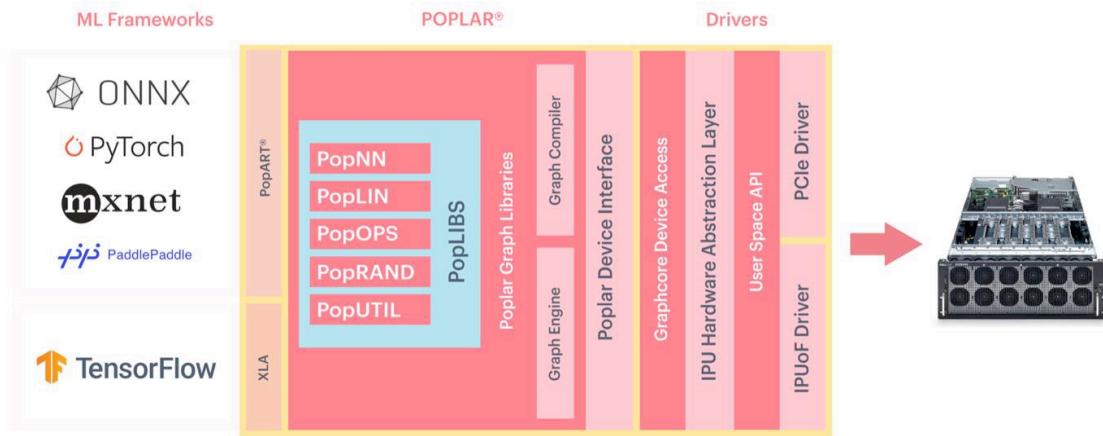
带来的编程挑战，因此无需手动调整指令级编程，即可利用IPU架构的细粒度并行性。还支持直接硬件访问。

## GRAPHCORE软件平台

### 概述

Graphcore为其芯片解决了软件开发中的两个挑战：1) 使其易于优化和运行那些以高级框架表示的现有机器软件，比如深度神经网络或DNN等；2) 实现全新细粒度并行工作负载的研发，使其可以在IPU架构上运行。后者的能力是公司战略的核心，可以使Graphcore发掘更大的细分市场。

**图 2: GRAPHCORE 软件平台**



Graphcore 软件平台采用了许多流行的 AI 框架中的模型，并优化了代码，以在 IPU 上执行。它还支持构建优化代码的自定义顶点或小程序，以开发新的神经网络和其他可表示为计算图的算法。

来源: *Graphcore*

如图2所示，Poplar由计算图和元素编译器、优化的库以及用于运行时管理和调度的计算图引擎组成。机器学习框架通过用于开放神经网络交换（ONNX）的Poplar高级运行时（PopART）接口以及用于基于TensorFlow的模型的加速线性代数（XLA）编译器向Poplar堆栈进行馈送。Graphcore表示，到2020年底将为PyTorch提供直接支持。

让我们看一下Graphcore的软件平台是如何促进用于IPU的应用程序开发的。具体来说，我们将研究这些主要组件，并考虑它们在加速性能和采用方面的潜力：框架，编译器，运行时支持和库。

### 使用开放框架移植和开发模型



为了评估新的AI加速器，深度学习科学家将首先使用TensorFlow或PyTorch等标准框架移植、训练和测试现有DNN，然后将结果（训练时间和准确性）与已知平台进行比较。Graphcore的软件团队简化了在IPU上训练和运行这些神经网以及为新解决方案建立原型的过程。开发人员可以使用自定义的层类型和新的库功能来扩展标准网络层，然后将其包括在开源框架中。此功能还可以增强或改进现有的层。

对于超出常见DNN框架范围的新应用程序，Graphcore已为IPU构建了自定义框架。Graph框架无意与现有框架竞争。相反，它可以帮助全新的并行工作负载在其计算图架构上运行。

Graphcore打算在来年将其软件开源，并且IPU SDK下载已经包含了Poplar库源代码。这对Graphcore而言是一个关键方向，因为开源策略有助于扩大行业积极参与到IPU平台上的研发工作，使社区能够开发和向上整合新的库和层。

### 优化执行：*Poplar* 编译器，库和计算图引擎

#### 计算图和计算图元素编译器

开发过程从计算图编译器开始，计算图编译器构建用于部署的模型，然后提取在IPU上运行的计算图元素（核或“小代码”）。这里有两个阶段，编译计算（计算图顶点）和生成代码以实现BSP通信（计算图的边缘）。

对于运行时部署和优化，关键在于Graph编译器。Graph编译器的开发工作已经进行了五年以上。它优化了代码重用，最大程度地减少了数据移动，并利用了数据本地性。这种方法有助于IPU在稀疏模型中有效开展工作，而稀疏模型会对其他更适合大型、连续内存访问模式的架构造成负担。Graphcore设计了编译器，以简化IPU的编程，尤其是在跨多个IPU部署工作的情况下。重要的是，编译器减轻了开发人员管理数据或模型并行性的负担。Poplar计算图编译器不是在一个服务器中对16个不同的ASIC或GPU进行编程，而是针对单个“Multi-IPU（多IPU）”，使开发人员可以专注于其数据和算法。与大多数其他加速器相比，这是一个明显的优势。

Poplar元素编译器充当后端，将代码计算编译为元素，以在顶点处运行。因此，工作流程如下所示：

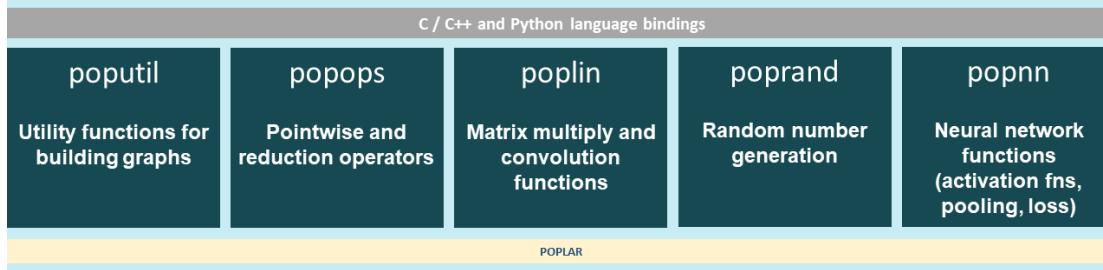
TensorFlow前端 -> XLA -> Poplar计算图编译器 -> Poplar元素编译器



INSIGHTS & STRATEGY

IPU具有完整而灵活的指令集，实质上允许tile运行任何代码或算法，而且计算图元素可以通过使用基于LLVM的编译器被写入C/C++，或直接写入IPU汇编中。所有的Poplar库都是使用这一编译工具集构建的。

图 3: GRAPHCORE 的 POPLAR 库



Graphcore为常用的功能和运算符构建了50多个基元。用户可以添加新的库以支持正在开发中的新工作负载，并将新库贡献给开源社区。

来源: *Graphcore*

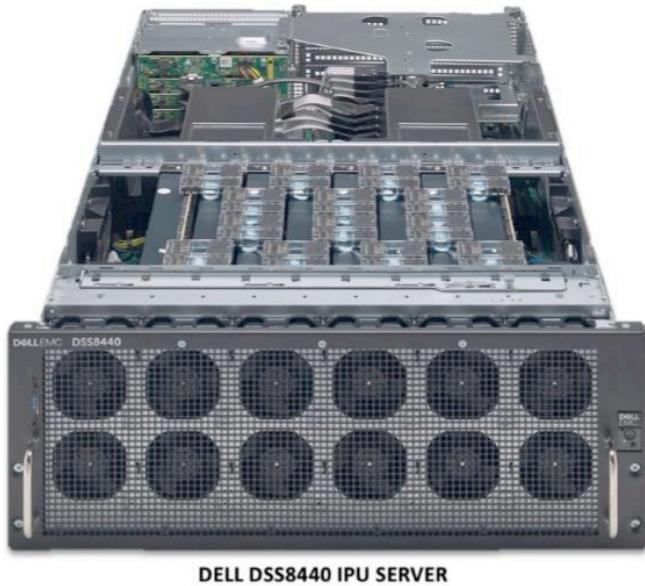
### Poplar库

Poplar库包含50多个高度优化的基元和构件块，这些基元和构件块针对的是诸如线性代数、公共神经网络功能以及其他机器智能模型中使用的常见操作。通过组合来自包含750多种高性能计算元素或小代码在内的扩展库的构件块，每项功能得以正常开展工作。Poplar库使用用户应用程序中的高级计算图描述，并构建针对IPU所需的大规模并行计算图。该过程包括工作和数据分区，以IPU的分布式处理器和内存架构为目标。Poplar库为每个操作建立了自定义布局，这些操作经过了高度调整，可以在IPU上执行。库还提供了一种可能性，即为那些可以在IPU架构上运行的新应用程序，发布和运行全新的内核和DNN层。

### 计算图引擎

Poplar计算图引擎为IPU提供运行时支持。这种支持包括将主机CPU上的软件与IPU上运行的软件连接起来，管理数据移动，并管理IPU设备本身以进行I/O，应用程序加载，调试和分析。优化来自CPU的数据流对于实现吞吐量最大化至关重要，而且计算图引擎编排数据I/O管道以实现连续执行。计算图引擎由IPU上的自定义硬件支持，有助于调试和分析。

图 4: GRAPHCORE 生态系统支持



STANDARD ECOSYSTEM SUPPORT

DEPLOYMENT



ORCHESTRATION



CONTAINERS



VIRTUALIZATION & SECURITY



Graphcore对部署能力的支持在创业公司中实为少见。

来源: *Graphcore*

### 部署: 行业标准工具和平台

Graphcore软件扩展到在公共云和本地基础设施中有效部署生产工作负载。具体而言，戴尔技术公司销售和支持针对服务器的双IPU Graphcore C2 PCIe卡，而微软Azure和Cirrascale支持IPU云实例(在Azure预览版)。对于管理和编排，Graphcore支持虚拟化、安全性和编排工具，这些工具已成为整个行业的标准，包括对Kubernetes、Docker和微软Hyper-V的支持。Graphcore是目前已知的唯一一家将其产品扩展到囊括了如此庞大的部署软件和基础设施的初创公司。

### 结论: GRAPHCORE是为大规模生产而构建的

在训练神经网络时，开发人员需要快速、可扩展的加速器，这些加速器可以处理较大模型(如自然语言处理和对话界面)所需的大量计算负荷。更广泛地讲，计算密集型应用程序正在出现，这些应用程序需要一个高速并行处理器，该处理器必须具有超越神经网络常见的矩阵乘法运算的能力。Graphcore似乎已经开发出了满足这些需求的一款通用、高性能硬件和软件平台。

Graphcore的软件堆栈在以下几个方面脱颖而出：

1. Poplar计算图编译器实现了“Multi-IPU(多IPU)”并行性，同时实现了有效的内存使用和数据移动。
2. Graph框架支持新的工作负载，尤其是从传统机器学习框架领域之外的领域出



现的新算法（计算图）。

3. 通过使用可执行新层和核的自定义“小程序”，用户可以扩展对开源ML框架的全面支持和优化。

4. Graphcore的开源策略应有助于该公司扩大其跨多个学科领域的研究人员的范围。

Graphcore 在管理软件方面也迈出了新的一步，提供了容器化、编排、安全性和虚拟化功能，数据中心更加依赖这些功能，也更安心无忧。结合该公司的硬件设计和软件堆栈，随着更多应用程序在 Graphcore 平台上部署，这些步骤将简化采用过程。

Graphcore面临的挑战之一是如何将其平台与替代产品进行比较。这是将目标专注于一些新型工作负载上的必然结果，这些负载是指未被广泛采用的新工作负载，或其他平台未能很好调整加以支持的新工作负载。尽管Graphcore声称IPU在诸如Resnet之类的“简单”模型上运行的速度可能快2-4倍，但Graphcore在其网站上发布的结果显示，新兴工作负载的性能提高了一个数量级<sup>1</sup>。魔术就此开始。

---

<sup>1</sup> [https://cdn2.hubspot.net/hubfs/729091/assets/pdf/Benchmarks\\_slides\\_May2020-comp.pdf](https://cdn2.hubspot.net/hubfs/729091/assets/pdf/Benchmarks_slides_May2020-comp.pdf)



## 关于此文的重要信息

### 投稿人

Karl Freund, Moor Insights & Strategy高级分析师

### 出版方

Patrick Moorhead, Moor Insights & Strategy公司创办人，总裁和首席分析师

### 垂询

如果您有意讨论此报告，请与我们联系，Moor Insights & Strategy会迅速做出回应。

### 引用

认证媒体和分析人员可引用该论文，但必须在引用中体现上下文语境，显示作者的姓名、职务以及“Moor Insights & Strategy”。非新闻和非分析人士必须获得Moor Insights & Strategy的事先书面许可，方可进行引用。

### 许可

本文档（包括所有支持材料）归Moor Insights & Strategy所有。未经Moor Insights & Strategy事先书面许可，不得以任何形式复制、分发或共享本出版物。

### 披露

本文由Graphcore委托编写。Moor Insights & Strategy为本文提到的许多高科技公司提供研究、分析、建议和咨询服务。该公司的员工不持有任何在本文档中引用的公司的股权。

### 免责声明

本文档中提供的信息仅供参考，可能包含技术上的不准确性、遗漏和排版错误。Moor Insights & Strategy对此类信息的准确性、完整性或充分性不承担任何保证，并且对于此类信息的错误、遗漏或不充分性不承担任何责任。本文档包含Moor Insights & Strategy的观点，不应解释为事实陈述。本文所表达的观点如有更改，恕不另行通知。

Moor Insights & Strategy提供的预测和前瞻性陈述只是方向性指标，而不是对未来事件的精确预测。尽管我们的预测和前瞻性陈述代表了我们当前对未来前景的判断，但仍存在可能导致实际结果出现重大差异的风险和不确定性。提醒您不要过分依赖这些预测和前瞻性陈述，这些预测和前瞻性陈述仅反映了本文发布之日我们的观点。请注意，我们没有义务根据新信息或未来事件修改或公开发布对这些预测和前瞻性陈述的任何修改结果。

©2020 Moor Insights & Strategy.公司和产品名称仅用于提供信息，可能是其各自所有者的商标。