

面向网络、高性能计算和云的专门硬件

客座编辑导言 • Art Sedighi, 自由顾问 • 2013 年 8 月



上世纪九十年代和本世纪前十年的大部分年头风行的是商品硬件，即广泛可用且价格不算贵的通用产，当时 X86 架构一统市场。后来风向变了，专门硬件渐成趋势，专门硬件是为专门目的而设计和建造的硬件平台，一典型例子是超级计算机 BlueGene。就像当年通用硬件代替专门硬件一样，后者重整旗鼓、卷土重来，背后的原因是：

- 节约成本：为什么要为用不上的东西破费呢？例如，高性能计算环境真的需要 PCI 桥吗？
- 节省能源：持续膨胀的基础设施的能耗已经难以为继，数据中心能耗已经见顶，还别说空间占用
- 信息饥饿：贪婪是一切事物背后的魔鬼。有关数据分析表明，随着处理器和服务器变得越来越快，我们需要的和消费的信息越来越多

一句话，我们希望比以前更快、更好、更便宜的硬件，这就导致市场转向专用处理器（ARM、FPGA 和 GPU），它们更小、能耗相对低，可广泛应用于解决网络、高性能计算和云环境中的更专门的一小群问题。

万用硬件不再必要

采用到处可以得到的商品级组件，现在几乎每个人都能根据给定应用的性能、成本和能耗约束来设计建造专用硬件解决方案。“今日计算”本月主题选择的三篇文章展示了这个还在上升阶段的新趋势。

在《ARM 集群的能耗和成本效率分析》中，Zhonghong Ou 和他的同事对 X86 和 ARM 处理器上运行的负载进行了比较。特别地，他们评估了 4 个双核 ARM A9 和一个 4 核 Core2-Q9400 的性价比。从性能角度看，可以期望 Core2 能够战胜 ARM 处理器，但在成本和能耗方面，ARM 更高一筹。典型的 ARM 处理器的功耗是 1~5 瓦，而 X86 处理器却要 40 到 60 瓦（如果加上硬盘驱动器等外设，会超过 100 瓦）。

作者用性能和功耗之比来表示能效（energy efficiency, EE），这样两种技术路线就有可比性了。随后作者考察了三个不同的用例：

- 采用 SQLite 的内存数据库
- 使用 Nginx 和 httpd 的 Web 应用
- 使用 HD-VideoBench 的视频转码

为了比较，作者用完成同样工作量的 ARM 处理器个数和 X86 处理器个数来计算两者的能效比（EE ratio）。内存数据库的能效比（#ARM/#Intel）是 2.6-9.5，

而成本只有 Intel 方案的 2/3。其它应用的能效比也大于 1，但内存数据库对网络和云架构来说更典型，因为它们都经常用到数据管理功能。

在评估《面向高性能计算的 ARM 集群的性能和能耗评估》一文中，Edson L. Padoin 及同事又进了一步，他们用 ARM 主板构造了一个超大规模高性能计算环境。他们采用了的货架产品来自 PandaBoard.org (ARM A9 2-core)和 BeagleBoard.org (ARM A8)，板上运行的都是 4.5 版的 Ubuntu Linux，并使用支持 ARM ISA 指令第 7 版的 GCC (GNU Compiler Collection) 编译器。

两种板都加载两种负载：1000 x1000 的整数矩阵和浮点矩阵。采用 A9 芯片和双核的 PandaBoard 板在两种负载下都表现突出：755MFlops 对 24MFlops。虽然 A9 的时钟频率和性能都更高，但其能效还是高达 92 MFlops/Max Watt，而 A8 相形见绌，只有 20 MFlops/Max Watt。也就是说，计算性能的提升并不意味着能耗一定要同步增长。

最后一篇主题文章是《论提升消息速度》，Holger Froning 及其同事讨论了他们在 EXTOLL 计划中的工作，该计划的目标是建造一种比常规 NIC 卡消息速度更快的专门硬件。网络和集群环境都运行在下层消息和通讯基础设施之上，诸如 MPP (大规模并行处理器) 等登上 [Top500](#) 榜单的多数大型高性能计算系统都采用特别紧耦合的和高度集成的网络层来支持处理器之间的通讯，从而使得性能超过集群和网络。后者往往采用通用货架产品，因此比昂贵的 MPP 要便宜很多。

Extoll 是一款 6 口的专门 NIC 卡，采用 FPGA 来支持环网的动态配置和重配置。Extoll 的主要特色是通过硬件级的虚拟化通信引擎来实现极低的开销和最小化的内容占用，从而更好地支持多核环境。下层网络协议和布局的虚拟化简化了到应用的集成，该团队已经把 Extoll 库集成到了 [OpenMPI](#)。

Froning 和同事的报告显示，Extoll 的性能超越了 Infiniband Quad Data Rate 40G 和 10G 以太网，作者还在考虑用专用集成电路 (ASIC) 重建 Extoll 引擎，希望比 FPGA 版本的性能提高 5 到 10 倍。

结论

本月主题的焦点是应用于网络和云环境的 ARM 处理器和 FPGA。推介论文表明这样一个事实：专门计算硬件已不再是大投入项目的“专利品”，ARM 处理器、FPGA 和通用目的 GPU 正日益普及，性价比越来越好，其原因在于硬件生产成本越来越低，支持专门硬件的软件工具也越来越多。随着 ARM、FPGA 和 GPU 生态系统的成熟，它们的身影也会出现在关键任务应用和产品环境中。



Art Sedighi 是一位自由职业的顾问，专长与大型基础设施的设计与实现，是 Computing Now 的编委。他从伦斯勒理工学院获得计算机科学硕士学位，专业兴趣是调度和博弈论。他的联系方式是 sediga@alum.rpi.edu 和 <http://phd.artsedighi.com>。

(黄铁军 译)