

## 与Jeff Dean聊ML for EDA，最佳论文花落伯克利：EDA顶级会议 DAC 2021 精彩回顾

本文作者：我在思考中

2021-12-20 11:34

导语：谈线下DAC会议：大佬云集。



作者 | 谢知遥

校对 | 王晔

编辑 | 青暮

12月9日，第58届EDA首要会议DAC ( Design Automation Conference ) 的线下部分在旧金山落下帷幕，此次会议为期五天。

受新冠疫情影响，近年来多数学术会议都转到了线上进行。延期半年后，此次以线下形式进行的DAC会议给从业者提供了难得的见面与交流机会。而部分作者未能到场的论文将在之后的virtual session进行分享。

除学术交流外，DAC一直以来也是全球EDA工具、Foundry、IP提供商的盛会。在两层的展区中可以看到众多EDA公司提供的精彩展示，其产品内容涵盖芯片设计流程中几乎所有的步骤。在EDA三大家之外的很多名不见经传的小公司的产品也能让人眼前一亮。在展厅中，一些公司甚至使用了飞刀杂耍以及脱口秀式的宣传方式，营造了难得的热闹场面。

### — 1 —

#### 主题演讲：最前沿的EDA技术

作为顶尖的EDA会议，DAC每年所邀请的演讲嘉宾自然而然地成为了全场关注的重点。这次DAC邀请到了不少传奇人物来分享关于EDA行业的研究观点和趋势观察。



我在思考中  
编辑

发私信

#### 当月热门文章

吴恩达：告别，大数据

本科学历马斯克当选美国工程院院士！张宏江、萨蒂亚：“我们都美好的未来”

清华博士后用10分钟讲解AlphaCode背后的技术原理，程序员不是那么容易被取代的！

中国首次！清华刘奕群团队获得WSDM 2022唯一最佳论文奖，中文获得「时间检验奖」

突发！TensorFlow技术主管皮特登离职，重返斯坦福读博：我在歌“太难了”

0

#### 最新文章

谷歌 AI 加入蛋白质解析大军 ProtENN 模型助增 680 万个蛋白质注释词条，登顶 Nature 刊

参数量翻了10倍！Meta AI 凭 100亿参数的“新SEER”，为宇宙铺路

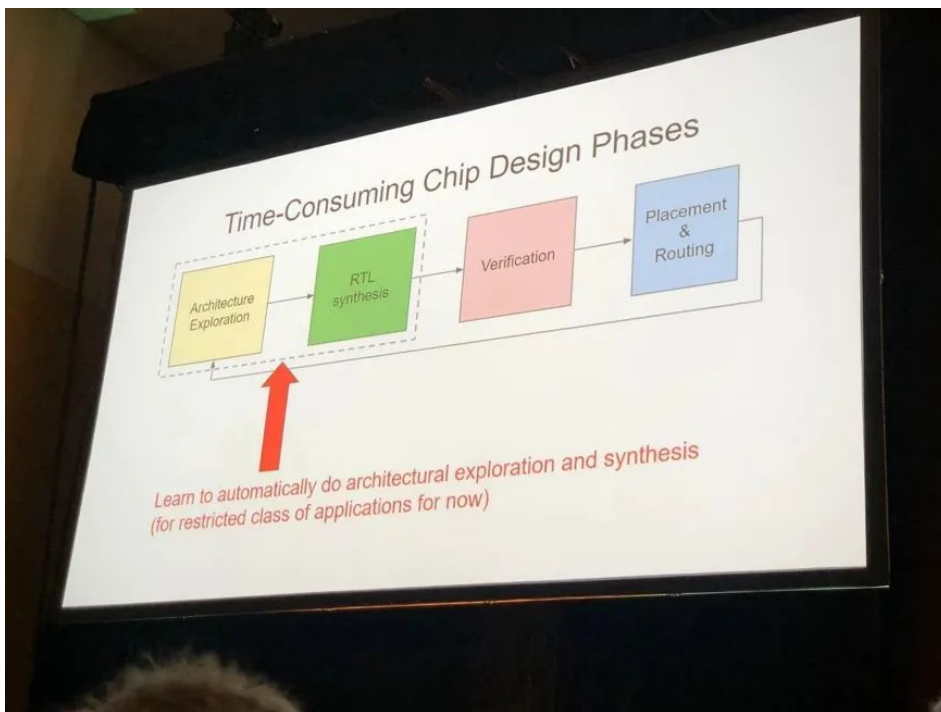
普林斯顿邓嘉学生亲述：一定博士学位？不，我本科生也能大厂当应用科学家



Jeff Dean

第一天，Google大神、Google AI的领导人Jeff Dean，进行了题为“机器学习在硬件设计中的潜力”的主题演讲。近年来，谷歌研究了不少深度学习在EDA方面的应用，其中最著名的是他们去年发表在Nature上的工作，通过强化学习自动进行macro placement，并真正应用于Google的硬件加速器TPU的设计过程。

Jeff在演讲中提到了Google使用深度学习优化整个芯片设计流程的工作，主要分为三个部分，对于芯片设计的三个主要阶段。如下图所示，演讲包括使用深度学习加速1.架构搜索和RTL综合，2. 验证，3. 芯片布局布线。



在架构搜索阶段，Google提出了叫做FAST的架构自动对硬件加速器的设计进行优化，他们使用了Google自己的黑盒优化器Vizier进行搜索。对于验证阶段的工作，Google提出了使用图神经网络(GNN)对RTL阶段的芯片设计进行分析处理。对于布局布线部分，重点自然就是发表在Nature的macro placement工作。

正式Keynote结束后，我们也和Jeff就ML for EDA进行了讨论。Jeff肯定了现有的商业EDA工具的表现。当我们问到在EDA方面，是否直接生成结果的强化学习方法将会取代仅进行预测的ML模型时，他认为两者在未来都将发挥重要作用。

数据集拥有自己的世界观？不  
其实还是人的世界观

斯隆奖新晋得主宋舒然：从视  
出发，打造机器人之「眼」

AAAI 2022大奖出炉！中科院  
州扑克程序AlphaHoldem获  
论文奖

## 热门搜索

高通

Instagram

Windows 10

直播

蚂蚁金服

叫兽发言

天猫

电信

Evernote

李明

莫须有

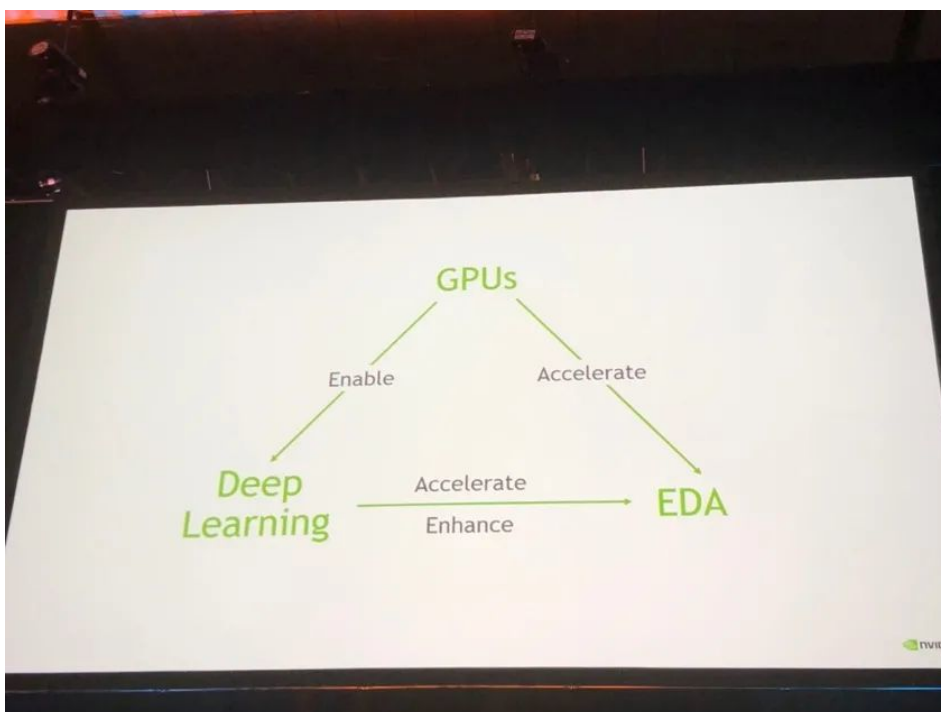


Bill Dally

第二天的keynote演讲者是Nvidia首席科学家Bill Dally，他的演讲题目为“GPUs, Machine Learning, and EDA”。Bill Dally的演讲结构清晰，概括了Nvidia如何使用GPU帮助ML，同时如何用GPU和ML帮助EDA发展。

GPU对于ML的促进作用是大家最为熟悉的，Bill首先介绍了GPU对深度学习的架构优化与Nvidia开发的深度学习加速器。而近年来出现了不少使用GPU加速解决EDA问题的工作，最著名的就是19年由UT Austin与Nvidia合作，同时获得DAC与TCAD最佳论文的DREAMPlace。Bill也提到了用GPU加速timing simulation。

在ML for EDA方面，近年来Nvidia也做出了大量工作，包括使用不同ML模型对IR drop、功耗、寄生参数进行预测。除了这些预测工作，Nvidia也提出了NVcell，使用强化学习方法直接生成优化的standard cell设计。纵观整个keynote，可以说在Bill的领导下，Nvidia Research对EDA方面的科研工作是比较充分的。





Joe Costello

第三天的keynote演讲者是EDA传奇人物、Cadence第一任CEO Joe Costello。他的演讲技术内容较少，主要从商业角度鼓励EDA业界拥抱变化。谈到的变化包括使用云计算平台，使用新的商业模式，使用开源生态系统，支持后摩尔定律时代的架构设计，以及熟悉政策变化。

另外值得一提的是，他大力批评了美国政府对中国的贸易战以及半导体产业的制裁，认为这反而激发了中国对支持半导体产业的共识与巨额投资。他表示由于中国近年出现的上千家硬件初创公司，五年之后中国将成为EDA的最大市场。

第四天的keynote由UC Berkeley教授，SqueezeNet的作者Kurt Keutzer提供。他的演讲主要回顾了深度学习的发展史，同时区分了人工智能，机器学习，和深度学习的概念。他鼓励EDA从业者应用和模型层面探索高效率的ML方法。

除了正式的keynote，DAC还提供了三场skytalk，类似于较小规模的keynote。第一天由微软Azure介绍他们为芯片设计与ML提供的云服务。他们认为云计算在安全和扩展性上展现了巨大优势。第二天由IBM介绍他们在深度学习加速器方面的探索，尤其是超低精度下的模型训练和预测方法。第三天由AMD介绍先进封装技术，例如chiplet对于未来计算硬件的重要性。

另外，大会也邀请了各大公司通过大量的presentation和poster来分享他们最新的研究进展和趋势观察。这种学界与工业界的紧密结合与交流体现了EDA行业的特点，同时也是DAC会议的优秀传统。

## — 2 —

### 研究论文：最佳论文花落谁家？

本次DAC一共收录了215篇研究论文，涵盖的内容非常广泛。受篇幅所限，我们只能够对获得最佳论文与提名的文章进行简单介绍。在本次线下活动中，共有三篇论文获得最佳论文提名。

该奖项今天刚刚揭晓！

UC Berkeley的"Gemmini: Enabling Systematic Deep-Learning Architecture Evaluation via Full-Stack Integration"榜上有名。这篇文章作者众多，也可以看出充分的工程投入。值得一提的是，该工作也成为了UC Berkeley基于RISC-V的硬件开源生态chipyard的一部分。而这个生态也包括著名的Rocket Chip以及Chisel。

## Gemmini: Enabling Systematic Deep-Learning Architecture Evaluation via Full-Stack Integration

Hasan Genc\*, Seah Kim\*, Alon Amid\*, Ameer Haj-Ali\*, Vighnesh Iyer\*, Pranav Prakash\*, Jerry Zhao\*, Daniel Grubb\*,  
Harrison Liew\*, Howard Mao\*, Albert Ou\*, Colin Schmidt\*, Samuel Steffl\*, John Wright\*, Ion Stoica\*,  
Jonathan Ragan-Kelley†, Krste Asanovic\*, Borivoje Nikolic\*, Yakun Sophia Shao\*  
\*UC Berkeley, †MIT  
hngenc@berkeley.edu

论文链接 : <https://people.eecs.berkeley.edu/~ysshao/assets/papers/genc2021-dac.pdf>

根据文章介绍,大部分已有的深度学习硬件的生成器(generator)只考虑加速器本身的性能,而没有考虑整个系统层级的性能。

本文提出Gemmini,这是一种开源的全栈式DNN加速器设计框架。使用Gemmini生成的硬件加速器已经被成功流片,并且取得了与商业加速器NVDLA接近的性能。在Gemmini中,设计师不仅能选择不同的加速器结构,同时也能配置整个搭载了加速器的基于RISC-V的SoC,并且这个SoC提供软件支持。设计者可以在OS上直接运行需要优化的DNN应用。

文章最后提供了两个使用Gemmini的进行设计的例子,分别是探索虚拟地址转换的设计方式,与探索内存资源的分配方式。

除此之外,获得最佳论文提名的还有Maryland University的"A Resource Binding Approach to Logic Obfuscation"。

## A Resource Binding Approach to Logic Obfuscation

Michael Zuzak, Yuntao Liu, and Ankur Srivastava  
Department of Electrical and Computer Engineering, University of Maryland, College Park, MD USA  
{mzuzak, ytlui, ankurs}@umd.edu

论文链接 : <https://eprint.iacr.org/2021/252.pdf>

根据文章介绍,设计者为了保护IP设计,避免恶意的foundry对IP进行窃取或者逆向工程,需要引入额外的设计给IP上锁,使得IP的功能取决于设定的密码。这个过程叫做logic locking或者obfuscation。然而,现有的方法无法兼顾多种安全需求。

为了解决这一缺陷,相比于多数在gate-level才进行上锁的工作,本文提出在更高层的high-level synthesis的resource binding步骤中,利用架构层面的知识来对整个IP进行上锁。结果表明,通过对binding与上锁进行协同设计,这种方法获得了上锁效果的巨大提升。

另一篇获得最佳论文提名的是UT Austin与Intel合作的"DNN-Opt: An RL Inspired Optimization for Analog Circuit Sizing using Deep Neural Networks"。

## DNN-Opt: An RL Inspired Optimization for Analog Circuit Sizing using Deep Neural Networks

Ahmet F. Budak<sup>1\*</sup>, Prateek Bhansali<sup>2</sup>, Bo Liu<sup>3</sup>, Nan Sun<sup>1</sup>, David Z. Pan<sup>1</sup> and Chandramouli V. Kashyap<sup>2†</sup>  
<sup>1</sup>ECE Department, The University of Texas at Austin <sup>2</sup>Intel Corp. <sup>3</sup>James Watt School of Eng., University of Glasgow  
\* ahmetfarukbudak@utexas.edu, †chandramouli.v.kashyap@intel.com

论文链接 : <https://arxiv.org/pdf/2110.00211.pdf>

文章提出了一种高效的对于模拟电路进行gate-sizing优化的方法。借鉴于强化学习方法,作者同时训练了两个深度学习模型,其中critic-network负责评估每一次gate-sizing的效果,而actor-network负责选择效果最好的sizing方式。但这种方法依然是监督式学习而并不是强化学习。

另外为了减小搜索空间,文章提出了分析每种优化操作对于最终目标的影响(sensitivity)。对于影响小于阈值的优化操作不进行搜索。实验证明,无论是在较小的电路设计还是大规模工业界的电路设计中,本

文的方法都能大幅减少需要的搜索次数，对应更少的设计时间。

本文作者是杜克大学博士生谢知遥。他以第一作者获得了今年的MICRO最佳论文。他将在2022年加入香港科技大学并正在积极寻找ML for EDA方向的博士学生。欢迎有兴趣的同学发送邮件至 zhiyao.xie@duke.edu

雷峰网 (公众号：雷峰网)

雷峰网原创文章，未经授权禁止转载。详情见[转载须知](#)。

## AI研习社年度会员核心特权



会员课程免费学



每月课程优惠券



海量资源免费下载



专享新课折扣价



专属答疑



VIP身份标识

点击了解详情

1人收藏

分享：



## 相关文章

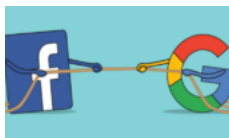
EDA首要会议DAC

Jeff Dean

Bill Dally



2021谷歌年度AI技术总结 | Jeff Dean执笔万字展望人



谷歌接二连三申请AI专利，Pytorch该不该想想侵权的



搜狗并入腾讯幕后：核心高管只余一人，PCG全力备战



华为的「军团」组织模式：破茧重生，还是作死？

[联系我们](#) [关于我们](#) [意见反馈](#) [投稿](#)

[申请专栏作者](#)

