

## PACT-2015 PC 会议记录与硅谷工业界访问之行

国重编译组 刘磊

国际并行体系结构与编译技术会议(International Conference on Parallel Architectures and Compilation Techniques, PACT)是计算机系统结构领域重要的学术会议。自1992年起, PACT每年一届, 已经召开了23届。PACT关注领域内基础性的、实质性的创新型研究成果, 同时也对新事物、新方向保持着紧密的关注。PACT对业界展现了旺盛的生命力, 基本上每年的会议都会获得ACM/IEEE及有影响力的大公司的鼎力赞助。

国重自2010年起至2014年, 基本每年都有一到两篇文章被PACT录用(2011年空缺, 2012年有两篇), 其中三篇的第一作者均来自于编译组。经北美学术界的朋友推荐, 我加入了PACT-2015的程序委员会(Program Committee), 能为学术界作一些力所能及的服务, 有机会参与评审和讨论高水平学术会议论文的全过程。PACT-2015的程序委员会(Program Committee Meeting)在芝加哥召开, 会议中心是美国著名的Big Ten Conference, 距离奥黑尔国际机场只有十分钟的车程。程序委员会的主席是UIUC的Marc Snir教授。“老骥伏枥, 志在千里”, Marc教授虽然已经快70岁了, 但还在为学术界服务, 还在积极参与各项学术活动。

PACT-2015的程序委员会会议在当地时间7月25日上午8:30准时开始, 会议有其非常严格的流程和操作规范。除了低于某个基准线的文章未被提上日程之外, 剩下的文章都是逐篇讨论的, 讨论的形式也绝非走个过场, 而是非常犀利的硬碰硬。讨论的时候经常会出现双方僵持不下的情况, 那么主席就会宣布这个文章存在争议, 并被暂时搁置, 而不是草率了事。此外, 如果被讨论文章的某个评委没有在现场, 还有可能会通过电话语音介入评审过程。我留意了一下, 讨论最长的

一篇文章用时超过12分钟，其余的都基本都在5~10分钟的范围。被搁置的暂未做出决定的文章在评审过程中还会被拿出来讨论，并与接下来有录用倾向的文章相比较。整个过程不能不说是非常“纠结”的，但评委们的纠结是为了保证会议录取论文的质量，并且，严谨的评审流程是对每篇论文背后的作者的尊重。

需要讨论的文章有90余篇，基本上占了投稿数量的一半。每篇文章开始讨论的时候首先要将与本文作者有冲突的PC请到隔壁喝咖啡，之后参与本文评审的PC将会逐一对本文的内容进行评价，给出自己的看法，讨论的过程将直接决定文章的命运。PC在会议上有可能会更改自己的意见。本次上会讨论的一篇文章原先分数靠前，但是在讨论的过程中一个PC又发现了一点技术上的问题，当时就将分数从2（接受）调整为-1（弱拒绝），并且其他评委也提不出有建设性的意见来反对。这篇排名靠前的文章成为本次会议第一篇被拒绝的文章。普遍认为综合分数排名靠前的文章基本上被拒绝的可能不大，现在看来上会讨论才是最为关键的步骤。遇到评审意见好坏参半的文章的时候，PC们就要开始“纠结”了，发言的内容五花八门。有的PC会说虽然本文的实验作的不是很完善，但是想法观点很新；有的PC则会列举出种种例外的情况，说明这篇文章考虑的还不够完善；也有的PC会质疑本文能不能吸引到更多人的关注，建议是否推荐到Poster的环节，等等。可见，在投稿之前作者需要作大量的工作、思考研究的方方面面，而行文又必须谨慎考究。首先做到让自己满意，挑不出毛病，才有可能在严格的评审过程中脱颖而出。

我本次评阅了9篇文章（每个PC需要评阅8~10篇文章不等），仅有一篇文章的综合评价排在30名以内，也仅有这一篇文章被录用。坦率的讲，这篇将Cache划分技术细化到指令一级的文章仅在模拟环境下取得了平均小于10%性能的提

升，但想法还算比较有新意，实验部署的也比较合理，撰写也相对谨慎考究。我评审的其它文章都或多或少的能看出明显的缺陷。其他评委的情况也基本如此，有的评委评阅的文章一篇都没有被录用，我仅见过一个评委手上的文章有3篇被录用，属于最好的情况了。会议一直持续到下午的5:30，所有的PC基本都人困马乏，精力疲力竭。Marc主持了全程的讨论，中途非但没有一次离场休息，而且需要不间断的思考以便为文章的录用与否做出最后的决断。会议结束的时候只录取了27篇文章，另外还有9篇文章被标记为“准备接受”，但相关的评委或者没有到现场开会，或者是相持不下不能做出决断，但无论如何，Marc希望能够在会后通过各种方式联系到这些未到场的PC，再征求一下他们的意见，之后再作最后的决定。这里我想再次给Marc Snir教授点赞!“靡不有初，鲜克有终”，Marc严谨踏实、善始善终的行事风格给我留下很深的映像，值得我辈学习。良好的学术规范是构建良好学术氛围的必要保证。通过不断的与国内外优秀学者的交流，青年学者也将不断的获得提升与进步。参加学术活动及学术服务是一种好方式，希望我在学术界的职业生涯能为业界做好服务、多做贡献。

在会议结束之后，我应邀访问了位于加州硅谷的华为北美研究中心和VMware，其中VMware是第二次造访。这次访问又认识了不少值得信赖的新朋友。让我惊讶的是，他们中的很多人都直接或间接的与计算所有关系。计算所能够影响国内的工业界这并不奇怪，但她距离加州那么远，居然还有那么多同根同源师兄师弟，这不能不让人感到激动、充满底气。这是一种身为计算所人的自豪感。

在VMware公司访问期间我并不进行学术报告，但每天都会和相关方向的工程师、研究员进行深度沟通，分享各自的一些思考（并不涉及公司机密）。文

化的传播需要经历时间的砥砺，其过程也往往甚为迂曲。学术思想的传播也大致如此。分布式处理的思想自上个世纪70年代就已经出现了；基于“页着色”（Page-Coloring）的资源管理与优化的思想大约诞生于1985年前后，与分布式共享内存（DSM）思想提出的年代大致相仿；关于操作系统内核形态的问题，在经历了宏内核、微内核的讨论和演进之后，业界又将目光移向诸如面向“物联网”等特殊设备环境下的内核形态。我向大家介绍了在大数据环境下，在将要到来的“全民编程”的时代，编程、编译、资源、优化、异构和整合的一体化系统设计思想，以及对能耗、吞吐量、服务质量、可编程性、易用性等关键指标的影响。我们的很多讨论甚至都延续到了用餐的时间，大家普遍认为今后还会有更多的合作机会。



图1：VMware内部的Poster交流会，地点设在公司的体育馆里，你可以为出色的Poster投票。

值得一提的是，一天午饭后，他们介绍我去参加一个企业内部的学术Poster交流会。展出的Poster全部由在VMware暑期实习的学生完成，这些学生基本都是来自美国各个大学的博士研究生。大概有五十几个展板，实习生站在展板前，随

时回答大家提出的各种问题。这个形式和学术会议的Poster基本一样。美国大学生的暑期实习是受到老板大力支持的，企业里也非常欢迎这些新面孔。根据公司的政策，他们实习的内容基本是一些预研类的项目，研究成果属于公司，实习生能获得的是与项目一起成长的经历，以及在大公司真实环境下研究真实问题的经验。每个实习生均配有研究经理负责指导和管理，其中一个经理是本次与我交流的华人研究员，他告诉我，每年有很多实习生应征报名，实习期大概三四个月，一般能完成一到两个小的研究项目。由于公司制度上的保证，他们对实习生的管理比较严格，每个实习生对自己的要求也都很高，因此，这种暑期实习机制往往能带来双方共赢的成果。我们组也招收不少实习生，如何管理和培养们？如何有效的利用实验室资源？如何能让他们真实的提升自己各方面的素质和能力？VMware有很多值得我和其他同事借鉴的地方，相信在今后的工作中，我们会建立更为有效的机制。



图 2: 在访问华为美研院期间的照片。蒋奕（左）、叶寒栋（中）和刘磊。我们三人都是计算所的毕业生，而且都是编译组吴承勇老师的学生。

北美华为研究院也在硅谷，距离VMware大约半小时的车程。美研院里的华人员工超过大半，他们大都毕业于美国的高校。给人映像深刻的是，还有很多老先生也奋战在第一线，与青年员工一起做设计、写代码。据说，华为70%左右的利润来自于海外，这让人听起来很提气！我在华为做了一场学术报告，题目是”Rethinking Memory Management in Modern Operating System -- Memory Optimization by Leveraging Hardware Features”，内容涵盖了我近5年的主要科研成果，报告从上午10:30到12:00点，与系统相关的研究人员基本都来了，这里面就有几个是国重编译组的毕业生。师兄弟见面，大家格外高兴。在报告之后，我与大家进行了深入的交流。有位研究经理对我们采用监控操作系统页表的方式来获得程序访存行为的系统工具SYSMON比较感兴趣，大家认为可以合作将此工具扩展到集群任务调度等领域，以优化数据中心的资源利用率；另一位从事数据库开发的研究员则表示，通过深度挖掘体系结构特征来优化系统性能的方法有可能对提高DBMS的性能有帮助，等等。此外，我询问了这些长期在企业里工作的研究人员他们的经历，我发现他们大部分都不止一次的转换过研究内容和工作方向。他们博士阶段从事的研究未必能在工作中派上用场，而在工作中由于项目需求导向，则又有可能同时开展不止一项研究，并且内容也可能不太相关。再之，企业也非常盼望能和高水平的研究机构合作，双方形成优势互补，实现双赢。通过与他们的沟通，我能体会到他们“上下而求索”的艰辛过程。

如何与企业合作？经过这次访问与交流，我希望简要的总结一下我的体会。总的来说，分为以下三点。

**第一，看需求，勿照搬。**工业界往往有自己明确的需求，工程师也往往有让自己头痛的很多问题。因此，将学术界的东​​西生搬硬套硬塞给工业界往往是行不通

的。重要的是基于对本质上的需求和应用架构的思考，进而去优雅的设计系统，使其更具普遍性和推广价值，能够解决实际的问题，而不是实验室里的玩具。例如，VMware公司的一位资深工程师告诉我，他们内部经常会读来自学术界的优秀论文，但这些很漂亮的idea对他们来说基本都没用。我无意于跟他纠结孰是孰非，赶忙请他解释一下。他举了VMware系统软件中内存资源管理部分为例。

“资源回收与垃圾清理算是系统管理的经典问题了吧？”

-- “是的，”

“VMware曾经花了很大的人力物力来研究并实现这部分机制，”

-- “哦，那是应该的。”

“但是，很遗憾，这部分代码很少被trigger，目前看来，将来用到的机会会越来越来少，”

-- “....”

“现在的服务器动辄就是几个T的内存容量，构成集群的话，容量就会更大，内存资源根本就用不完，”

-- “.....”

-- “就算是内存资源不够，我们通过添置硬件的方式，成本会更低，也好过十几个人年的人力投入。另外，看似漂亮的算法可能会给系统的稳定性带来隐患。你知道，很多当初设计这部分的人都已经离职了。”

我们的谈话并不涉及VMware的技术机密。但总的来讲，可扩展性则是他们更为关注的问题。如果在这方面有所研究，那么会更容易找到共同话题。因此，了解到什么是企业真正需要的，或许应该是问题的开始？

**第二，多合作，促合力。**很多有影响力的创新和产品都源自于工业界。有些观点认为，只有工业界推动的研究才是有价值的，因为工业界掌握着大量的真实数据，了解用户需求，具备产品生产和推广的要素，并且具备市场引导能力。例如，大型互联网公司海量的数据和用户请求，而且实际运行着大规模的平台，每分每秒都处理着来自四面八方的各种数据。面对真实的数据信息，大公司有机会能够大展拳脚，而这些条件都是学术界的高校和科研院所不具备的。但由于工业界产品导向的根源，很多存在一定风险性的、必须长期投入才有可能有所收获的研究性项目往往不能开展。工业界可以和学术界形成优势互补。学术界的优势在于理论的深度、科研机制的良性运转、以及在前沿问题上的实质探索。因此，如果能将企业里发现的有价值的问题拿到科研机构中进行研究，对于双方来说都会有好处。

**第三，走近点，多沟通。**学术界也有其固有的问题，学院派看待问题的角度往往不会涉及到企业的真实应用架构，也往往欠缺系统的设计思路（这包括从逻辑上的设计思路和物理上的系统架构和程序）。另外，企业与研究所工作的目标也不尽相同。在企业里，生产系统面临的压力通常是非常大的，需要响应的请求类型也多种多样。因此，对于企业来说，一个系统最复杂的地方往往不在于取得高性能，而是实现大压力下的稳定性。例如，以操作系统方面的研究工作为例，操作系统需要解决的问题是采用什么方式来合理的调度资源，进而最大化硬件资源的利用率，同时还必须保持系统的稳定性。而实验室中对操作系统的研究通常追求的都是高性能，而欠缺对稳定性的系统化的考量。操作系统核心机制的演进相对比较保守，实验室中新的研究成果距离真正被使用通常还有一定的距离。本次在北美华为研究院和VMware的访问让我有机会能够真实的感受到，由于产品导



向的原则，企业对能够“落地”的科研成果的迫切渴望；同时，如果不多交流，企业也无法理解我们在研究问题上的思考的深度。

## 后记

本次北美之行有不少的收获，我走访了老朋友并初步发现了一些新的合作契机，这一周的时间过的很充实。平时难有写文章的时间，我写这篇文章的时候已经是9月3号了，这天恰好是举国同庆的阅兵日。我是军史迷，最让人感动的莫过于看到高举抗战中英雄部队番号的功勋部队方阵正步走过天安门广场，这恰如几个月前，高举在苏联卫国战争中各方面军军旗的部队跨过莫斯科红场。这种精神力量让人感到一种“纯粹”，“纯粹”是诗章，伟大的中华民族是诗章！