

---

## 互联网广告推荐迎变革，雪湖科技联合赛灵思发布推荐算法解决方案

12月3日，2019年XDF-亚洲站在北京国家会议中心召开，在这场年度技术盛宴上，数以千计的技术工程师与专家携广大合作伙伴、行业领导者集聚一堂，带来最新的技术解决方案与灵感洞察。4日上午，在DCG数据中心分论坛上，国内顶尖的FPGA应用开发企业雪湖科技技术副总裁赵小吾受邀出席，与赛灵思共同推出了由Alveo U200加速卡支持实现的“Wide and Deep广告推荐算法加速解决方案”。这一产品是双方面向互联网企业提升广告和内容CTR（点击率）的需求而开发，也是雪湖科技数据中心解决方案的代表作。相较于CPU服务器，基于FPGA加速器打造的这一解决方案将吞吐量提高了3~5倍，性价比是CPU的5.5倍。

“雪湖科技为赛灵思Alveo加速卡产品在互联网公司广告、内容推荐方面提供商业化解决方案”，赛灵思数据中心资深产品行销经理梁晓明表示，“在数据爆炸时代，以出色算力提升企业运作效率的FPGA芯片已博得许多互联网公司的青睐，推荐算法广泛用在个性化广告，信息流等高商业价值的应用，很期待雪湖科技为互联网公司的CTR（点击率）和营收带来新的增长。”

赵小吾介绍道，Wide and Deep算法由谷歌发布于2016年，被广泛应用于提升互联网公司广告推荐和内容推荐的CTR（点击率）。由于部署于GPU的经济成本高，这一算法往往使用CPU来加速，但性能并不理想。在此背景下，雪湖科技与赛灵思针对数据中心广告推荐算法加速推出“Wide and Deep算法推荐系统”高性能解决方案。

得益于赛灵思加速卡Alveo U200高速度、低时延的优异性能，相较于CPU服务器，雪湖的FPGA加速器把吞吐量提高了3~5倍，加上功耗更低，其TOC（总体拥有成本）是CPU

---

云的 5 倍以上。一组广告预估 CTR 数据显示，综合成本、性能来看，这一 FPGA 加速器产品相对 CPU 呈现出了绝对优势。

考虑到大型推荐系统的上线都是通过云端部署，同时用在线和离线方式更新模型。雪湖科技 FPGA 开发团队还把以 Wide and Deep 为基础网络的模型移植到阿里云 FPGA 服务器 F3 (FPGA: 赛灵思 VU9P) 上，用户可以通过镜像文件部署。根据最近的更新数据显示，模型精度损失可控制在十万分之二。“且雪湖 FPGA 加速器拥有灵活性，当模型更新时，通过雪湖科技提供的专有工具可直接载入模型参数，可做到一键式更新模型参数。”赵小吾表示

赵小吾还透露道，鉴于这一方案的前沿性，雪湖科技已吸引到数家国内外知名互联网公司达成合作意向，并有合作项目进行到了落地阶段。

新浪微博广告业务部技术总监李东升对这一算法的应用表示期待，“Wide and Deep 是推荐算法中普及率最高的一种，不论是内容推荐还是广告推荐，它在视频媒体、生活服务、电子商务类媒体上的应用都极为广泛。当然，广告营收对互联网公司的重要性使得雪湖科技的这一算法加速系统意义惊人，其带来的营收增速助力，甚至会对应用此方案的互联网公司行业身份产生影响，也将推动国内互联网企业广告售价接轨 Facebook 这样的国际一流企业”。

据了解，本次此次 XDF-亚洲站上，雪湖科技还携其数据中心解决方案：目标检测神经网络加速器家族 Yolo 系列亮相展区，这家成长于摩尔定律终结时代的 FPGA 应用开发企业成长速度惊人，目前已与赛灵思、阿里云、浪潮、上汽，美图等行业巨头建立合作，其创始人兼 CEO 张强表示，我们很荣幸能与这些优秀的大公司一起站在行业前沿，以算力赋能未来的 5G、物联网、AI 时代。