

关于志愿计算的报告

作者：唐金潼 学号：1600012610 指导老师：陈斌

【摘要】自 1995 年世界上第一个志愿计算项目启动起，志愿计算实现了让普通人参与科研，帮助科学家快速进行数据分析，缩短了科研消耗的时间。本文将简要介绍志愿计算的概况，列举一些志愿计算的经典案例，从而对志愿计算有更全面的认识。

【关键词】志愿计算 分布式计算 BOINC 科研

一、志愿计算简介

志愿计算是通过互联网让全球的普通大众志愿提供计算资源（如处理能力和存储能力），参与科学计算或数据分析的一种分布式计算方式。在具体实施中，项目方把大的计算任务分割成小块（任务单元），通过互联网分发给志愿者进行计算，志愿者计算完成后再通过网络把各自的计算结果返回到项目方的服务器。对计算结果统一合并，得出结论。

志愿计算采用的是分布式计算。分布式计算是利用网络把成千上万台计算机连接起来，组成一台虚拟的超级计算机，完成单台计算机无法完成的超大规模问题求解的计算方式。^[1]这样的计算方式能够充分发挥每一台计算机的计算潜力。

自上个世纪末出现以来，志愿计算逐渐显示了其在解决大规模、复杂问题上的优势。志愿计算拥有庞大的个人电脑群体，其总体计算能力甚至超过当今世界上一些最优秀的超级计算机。^[2]通过志愿计算，不少耗时长久的科研项目加快了步伐，降低了成本，还调动了大众参与科研的积极性，向大众普及了科学知识，激发了人们探索的好奇心。

例如，Einstein@Home 项目从 118 个未识别类脉冲星源样本中发现了 17 颗新伽马射线脉冲星。Einstein@Home 是一个分布式计算项目，利用全世界志愿者计算机的空闲计算时间帮助研究人员快速分析从费米伽玛射线太空望远镜上收集的伽玛射线数据。^[3]

二、志愿计算平台——BOINC

BOINC（全称 Berkeley Open Infrastructure for Network Computing，伯克利开放式网络计算平台）是目前世界上最为主流的分布式计算平台，最初开发是为了支持 SETI@home 项目而开发的，之后逐渐为众多的数学、物理、化学、生命科学、地球科学等学科类别的项目所使用，结束了各个志愿计算平台“各自为政”的局面。目前有 56 个科学项目在该平台上进行志愿计算。

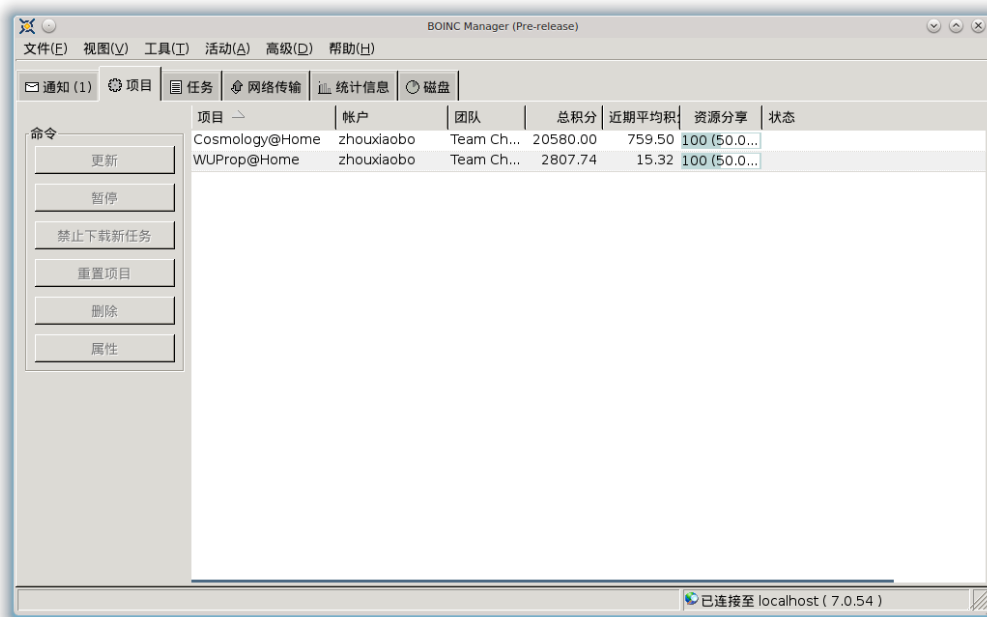


图 1 BOINC 界面

三、志愿计算案例

1、寻找梅森素数——GIMPS。

世界上最早的志愿计算项目 GIMPS（全称 Great Internet Mersenne Prime Search，因特网梅森素数大搜索），始于 1995 年。1995 年底~1996 年初美国数学家及程序设计师乔治·沃特曼编制了一个梅森素数计算程序，并把它放在网页上供数学家和数学爱好者免费使用。梅森素数是形如 $2^p - 1$ （ p 为素数）的素数。

梅森素数的探究在当代已有了十分丰富的意义，数论研究、计算技术、密码学等领域都得益于对梅森素数的寻找。

计算机时代之前，人类只能通过手工计算寻找梅森素数，前进极为艰难；进入计算机时代后，寻找梅森素数的步伐加快，但运用超级计算机寻找梅森素数，参与者有限，且成本极其高昂。GIMPS 启动后，梅森素数的寻找工作如虎添翼，在 21 年的时间内找到了 15 个梅森素数。目前，全球联网计算机数超过 144 万台 [4]。

2、搜寻地外文明——SETI@home

SETI（搜寻地外文明）是一个寻找地球以外智慧生命的科学实验，由美国加州大学伯克利分校创立，中心平台设立在伯克利空间科学实验室。使用射电望远镜来监听太空中的窄带无线电信号。如果能探测到非自然产生的信号，就可以证明地外文明的存在 [5]。

1995 年，该项目由专用超级计算机分析改为使用由全球联网的大量计算机所组成的虚拟超级计算机来进行计算，即今天的 SETI@home。志愿者可以通过运行一个免费程序下载并分析从射电望远镜传来的数据来加入这个项目。SETI@home 的原始目标有两个，第一个是要证明分布式计算概念的可行性和实用性。第二个是本身 SETI 的有益科学研究，以科学方法分析、探索地球以外的智慧生命。

遗憾的是，该项目至今没有收到地外文明的信号，且一直面临资金短缺的问

题。

3、中科院志愿计算平台——CAS@home

CAS@home 是由中国科学院高能物理研究所计算中心所搭建的，基于BOINC 志愿计算软件的志愿计算平台，用来给中国的科学家们使用。CAS@home 收集利用志愿者们所贡献的计算资源用于中国科学院以及中国其它的研究机构使用，它能够提供大量免费的计算资源来帮助科学家们来完成科学计算任务。

2014 年 2 月 12 日，研究人员宣布，在志愿者的热心帮助下，该项目已经成功完成了第一批的计算任务，成功破解了 5 万条新蛋白质序列的结构。研究人员称，基于这些计算结果，生物学家在利用海藻提炼石油的研究方面已经得出重要的科学结论^[6]。

中国目前已经有志愿计算爱好者组建的团队参与各个科研项目，并在国际竞赛中取得较好成绩^[7]。但志愿计算的发展在中国还处于起步阶段，远远落后于欧美发达国家。

四、志愿计算的缺点

首先，志愿计算的应用范围有限。志愿计算只应用于可能通过计算来解决全部或部分问题，而且其计算过程必须比较容易被分割成个人计算机的处理能力能够接受的大小的科学研究。

其次，志愿计算也存在一定的安全性问题。只有可靠的项目方才能保证志愿者计算机上的隐私数据不会恶意的取走、修改等。

另外，志愿计算的参与者多为个人计算机。虽然分布式计算的计算程序一般运行在最低优先级，不会对志愿者的日常使用造成影响，但计算程序全负荷运行时仍会对计算机的各个部件造成一定压力。

对项目方来说，志愿者毕竟不是项目方自己的人员，并不是全体可信任，因此必须引入一定的冗余计算机制，才防止计算错误、恶意作弊等^[8]。

另外，自 2014 年起，BOINC 的活跃用户比例不断下降。截至今年 2 月，BOINC 注册用户数接近 400 万，但活跃用户不足 30 万，仅占总注册用户数的约 6%。^[9]

综上所述，志愿计算在需要进行大规模计算的问题上具有成本低、速度快等优势，也起到了良好的科普效果，但也有些问题需要解决。国内志愿计算近年来取得了迅速发展，但普通大众对于志愿计算的了解还较少。中国计算机用户数量庞大，志愿计算在国内仍然有巨大的上升空间。

【参考文献】

- [1] 周晓峰，王志坚. 分布式计算技术综述[J]. 计算机时代，2004，12： 3.
- [2][9] US supercomputer needs more people power [EB/OL]. [2017-02-22]. <https://www.natureindex.com/news-blog/us-supercomputer-needs-more-people-power>
- [3] Einstein@Home 项目发现 17 个新伽马射线脉冲星 [EB/OL]. [2017-01-16]. <http://digi.163.com/17/0116/21/CAUBO6OO001687H3.html>
- [4] PrimeNet Activity Summary [EB/OL]. [2017-03-04]. <https://www.mersenne.org/primenet/>
- [5] 孙卫华. 基于互联网的合作探究项目——在家搜寻外星人[J]. 中国教育技术装备，2009，33： 94.
- [6] CAS@home [EB/OL]. [2014-02-12]. <http://casathome.ihep.ac.cn/>

[7] BOINC Pentathlon Gesamtwertung | seti-germany.de [EB/OL] [2016-06-19]

https://www.seti-germany.de/boinc_pentathlon/statistiken/pentathlon.php

[8] 新手指南:分布式计算常见问答 - 中国分布式计算总站[EB/OL]. [2013-06-24].

<http://www.equn.com/wiki/%E6%96%B0%E6%89%8B%E6%8C%87%E5%8D%97:%E5%88%86%E5%B8%83%E5%BC%8F%E8%AE%A1%E7%AE%97%E5%B8%B8%E8%A7%81%E9%97%AE%E7%AD%94>