

第一届燕山大学数据可视化挑战赛

赛题介绍

Windows 内置的性能监视器 (perfmon) 始终是系统管理人员进行性能校调时的必备工具。通过它, 我们不仅可以找出操作系统本身的运作问题, 对于许多安装在该系统之上的应用服务, 像是熟知的 SQL Server、Hyper-V 及 Vmware 等解决方案的性能症结, 也都可以通过它来辅助您发现背后真正的问题所在。在系统默认的状态下, 性能监视器可以对所有 CPU 的核心进行监视, 您可以加入相关计数器对象如 PhysicalDisk, 其目的是观察现行运行进程的磁盘变化信息。在面对不同连接流量冲击时, 性能监视器可以呈现 CPU 负载与物理磁盘队列长度的状态变化。当然您也可以加入更多不同的计数器, 来动态观察它们之间的相互关系。例如, 您可以比对目前 chrome 浏览器打开的网站连接要求数量 (Request Current), 以及它对于每一个连接要求的运行时间 (Request Execution Time), 如果经常性保持响应的时间过长, 这时候就可以观察 Processor Time 计数器, 或是 Memory 的 Available MBytes 计数器, 或是 PhysicalDisk 的 Disk Time 计数器变化。如果这三者之一的计数器呈现持续维持在高点, 这就表示相对的硬件资源性能不足, 有替换或升级的必要性。但如果这些硬件资源并没有相对呈现高点, 则可能的问题通常会应用程序本身或是后端数据库结构的设计有关。

第一届燕山大学数据可视化挑战赛 (ChinaVis Data Challenge 2023) 竞赛主体是所有参加《数据可视化技术》选修课的同学, 分析确定导致系统瓶颈的原因, 并相应地调整系统及应用程序的性能, 能使计算机在性能优化的状态中高效地工作。

参赛作品要求以 windows 性能监视器的数据收集器集为基础，利用可视分析相关技术和方法，探索并发现每个进程隐藏的漏洞和缺陷。包括以下主题：

(1) 计算机运行状态分析：利用可视分析技术，分析当前进程所占资源的分布情况，包括 cpu，内存，硬盘，网络，评估进程资源变化状态。

(2) 计算机故障分析：利用可视分析技术，分析当前是否存在进程存在**内存泄漏**等问题，并分析是否需要更高级别的硬件配置。

(3) 自选主题：参赛者可自行选择感兴趣的主题，利用可视分析技术解决计算机进程底层相关的问题。如果认为提供的进程监控数据难以满足希望的分析需求，也可以自己用性能监视器（perfmon）获取数据并做分析，前提是完成(1)(2)的任务后。

参赛须知

(1) 报名方式

所有参加《数据可视化技术》选修课的同学以组队的方式报名参赛，每队参赛者 3-5 人，每个参赛队排名第一的参赛者为队长，负责沟通事宜。参赛队命名自取（不违法），报名请填写在线文档【腾讯文档】
<https://docs.qq.com/sheet/DY2d2eXpoVEJUcXhO?tab=BB08J2>

(2) 作品提交要求

作品提交内容包括：(1) 项目代码；(2)作品文档；(3)视频；(4)作品代表性图片。竞赛作品提交采取发送邮箱的方式（121162433@qq.com），请将最后提交内容打包成一个文件，文件以“队号+队名”命名。

建议参考 ChinaVis 往年优秀作品，具体见“历届挑战赛优秀作品回顾”

<https://chinavis.org/2022/challenge.html>; 也可以参考 IEEE VIS Conference 组织的 IEEE VAST Challenge 的入选作品，作品库见：
<http://cs.umd.edu/hcil/varepository/benchmarks.php>。

竞赛作品提交材料要求

(1) 作品说明文档：参赛者根据提供的文档推荐模板，用图文并茂的方式介绍作品，以 Word 或 PDF 格式提交。

(2) 视频：参赛者制作带解说视频，围绕作品解释其可视分析流程，视频总长度不超过 5 分钟，视频数量 1 个，限 MP4 格式，视频大小严格要求控制在 50M 以内。

(3) 作品代表性图片：请提供高清版本 1 张，限 JPG 格式，多图请拼接，图片大小不超过 20M。

(3) 评分规则

评审将重点评价参赛作品的**主题导向和应用价值，以及参赛作品在交互设计、数据利用程度、社会效益、分析思路与方法等方面的有效性、新颖性和艺术性。**

请在作品说明文档中注明每位队员对可视化作品的主要工作内容和贡献度。
最终每位同学的成绩由队伍整体成绩结合队员贡献度给出。

(4) 作品等级设置

课程老师将根据评审结果，按比例评选出多项作品，在课程最后环节进行获奖作品现场汇报和颁奖环节。

数据简介

监视系统资源为 5 秒 1 次，监视 10 分钟为 4 个 csv 文件（分别是 cpu，内存，磁盘，网络），总计 1 小时。提示：利用 python 的 pandas 库处理数据，务必仔细阅读《数据格式》内容和文档最后提供的链接。

监视进程列表：

Chrome：浏览器

clash for windows：代理软件

code：vscode IDE

Ding Talk：办公软件

Everything：本地搜索软件

Mysqld：mysql 数据库进程

Notion：笔记软件

QQ：社交软件

Snipaste：截图软件

Ssh：远程服务

Steam：游戏平台

ToDesk：远程控制软件

Vmware：虚拟机软件

WeChat：社交软件

等等

其他说明

(1) 参赛作品不得违反国家相关法律法规，不得侵犯任何第三方知识产权或者其他权利。作品如引起知识产权异议和纠纷，其责任由参赛者承担；

(2) 参赛者可以使用开源或商业数据分析与可视化软件，例如：DataV、Tableau、R 和 Excel 等等，鼓励参赛者使用软件开发工具设计并实现自己的可视分析方案或艺术可视化作品，常见的可视化开发工具有 ECharts、D3、AntV 和 Processing 等等。请参赛者在作品说明文档中明确写出使用的开发工具和使用的开源或商业软件；

3) 入选参赛队至少要有 1 人注册 ChinaVis 2022；

4) 入选参赛队要求将自己的参赛作品制作为海报，并参加 ChinaVis 2022 的海报张贴环节，具体要求参照海报环节；

5) 参赛作品的知识产权归参赛者所有。大会组织方有权将参赛作品、作品相关资料、参赛团队信息用于宣传品、授权媒体发布、官方网站浏览及下载、展览（含巡展）等活动项目；

6) 参赛作品已经提交后，不得更改参赛者姓名及顺序。