

# Python 在互联网广告 监测分析中的应用

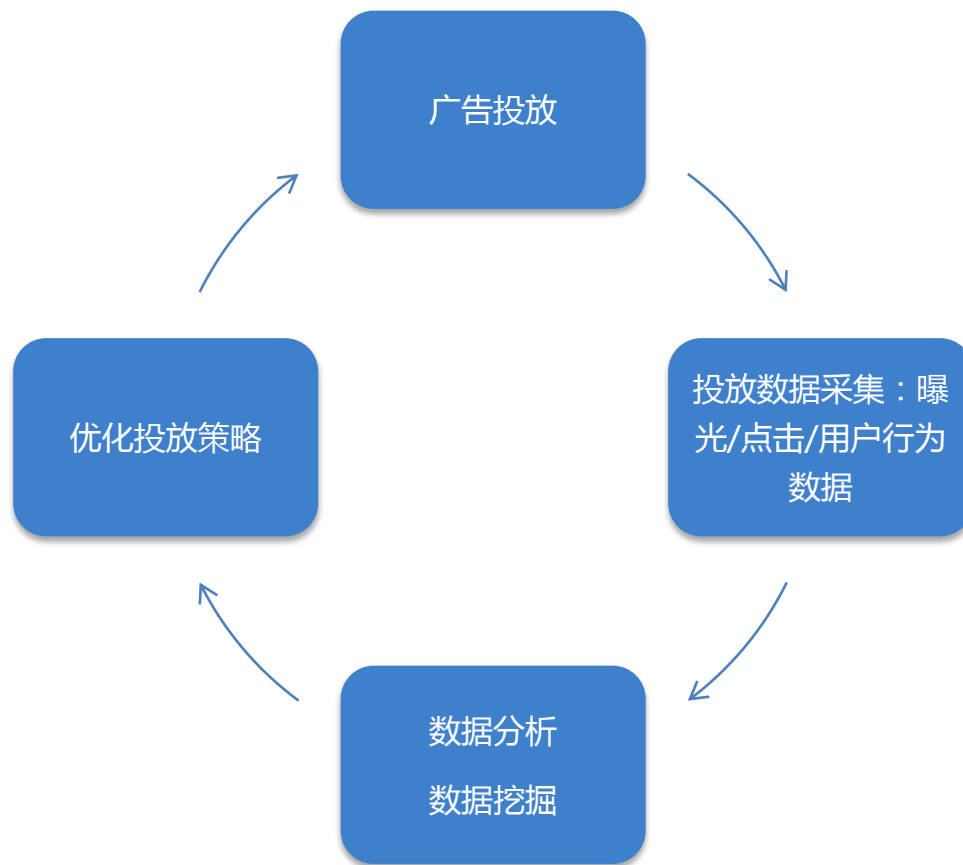
@陆丹峰 from AdMaster (精硕科技)

# 互联网广告简介

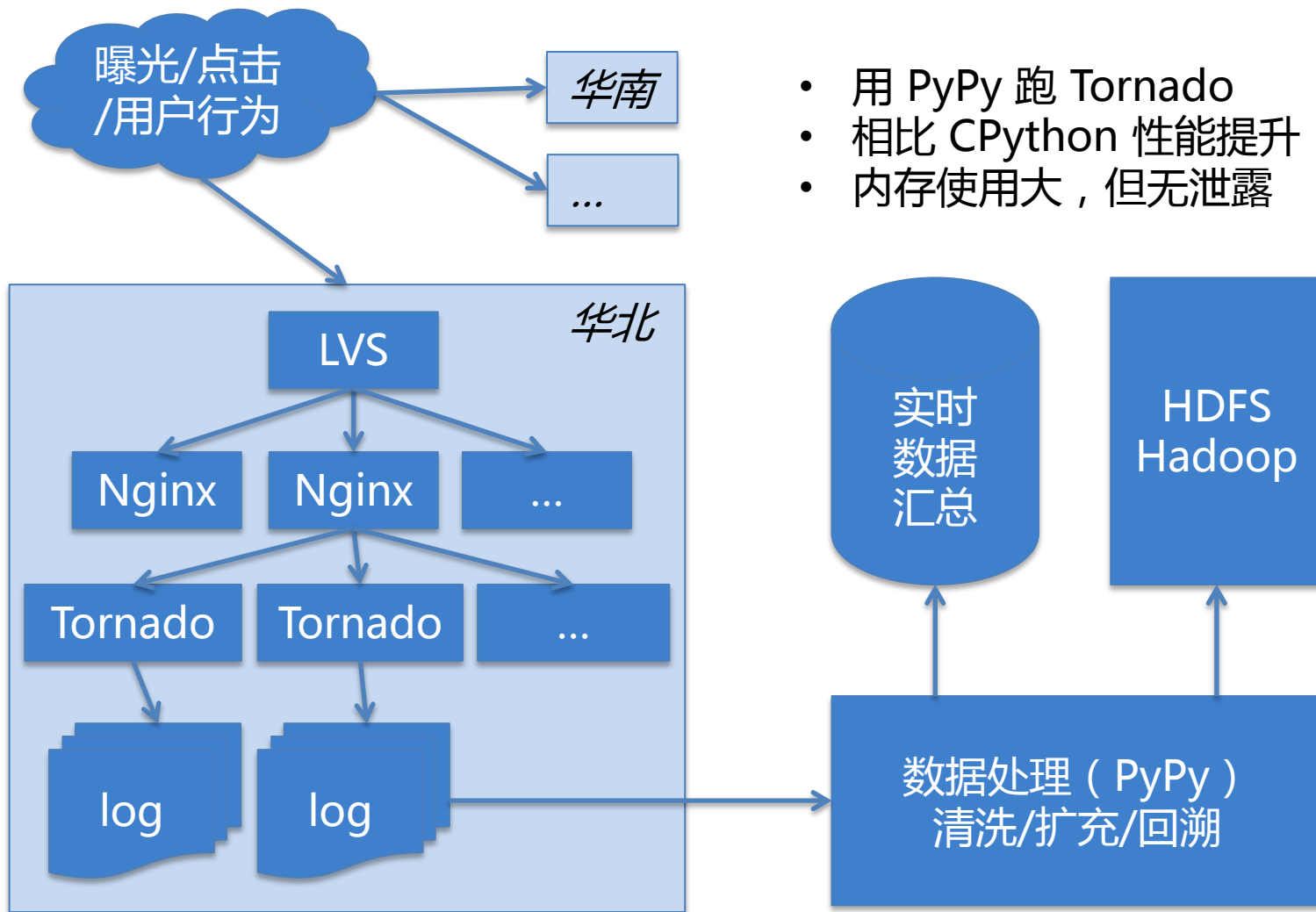
- \* 广告，和“游戏”、“电子商务”为互联网三大主要盈利途径
- \* 广告形式：文字链、静态图片、动态Flash、交互式等
- \* 投放渠道：门户、搜索、视频、微博、SNS、手机...

# 互联网广告目的：在线营销

广告监测/数据分析：优化广告主的在线营销策略，提升广告转化率

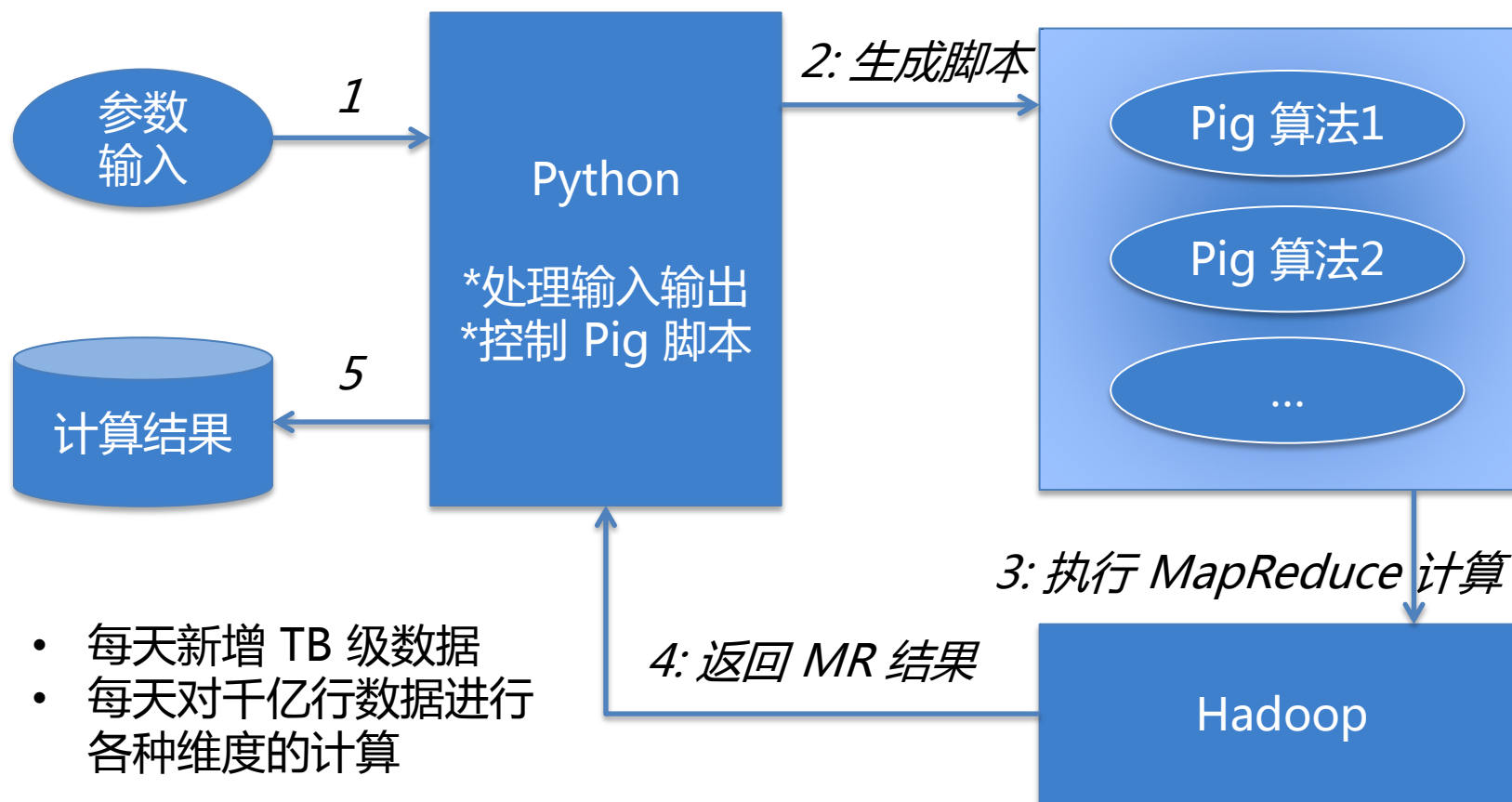


# AdMaster 广告数据采集分析架构



- 用 PyPy 跑 Tornado
- 相比 CPython 性能提升 > 30%
- 内存使用大, 但无泄露

# Python 在 Hadoop 处理中的角色



- 每天新增 TB 级数据
- 每天对千亿行数据进行各种维度的计算

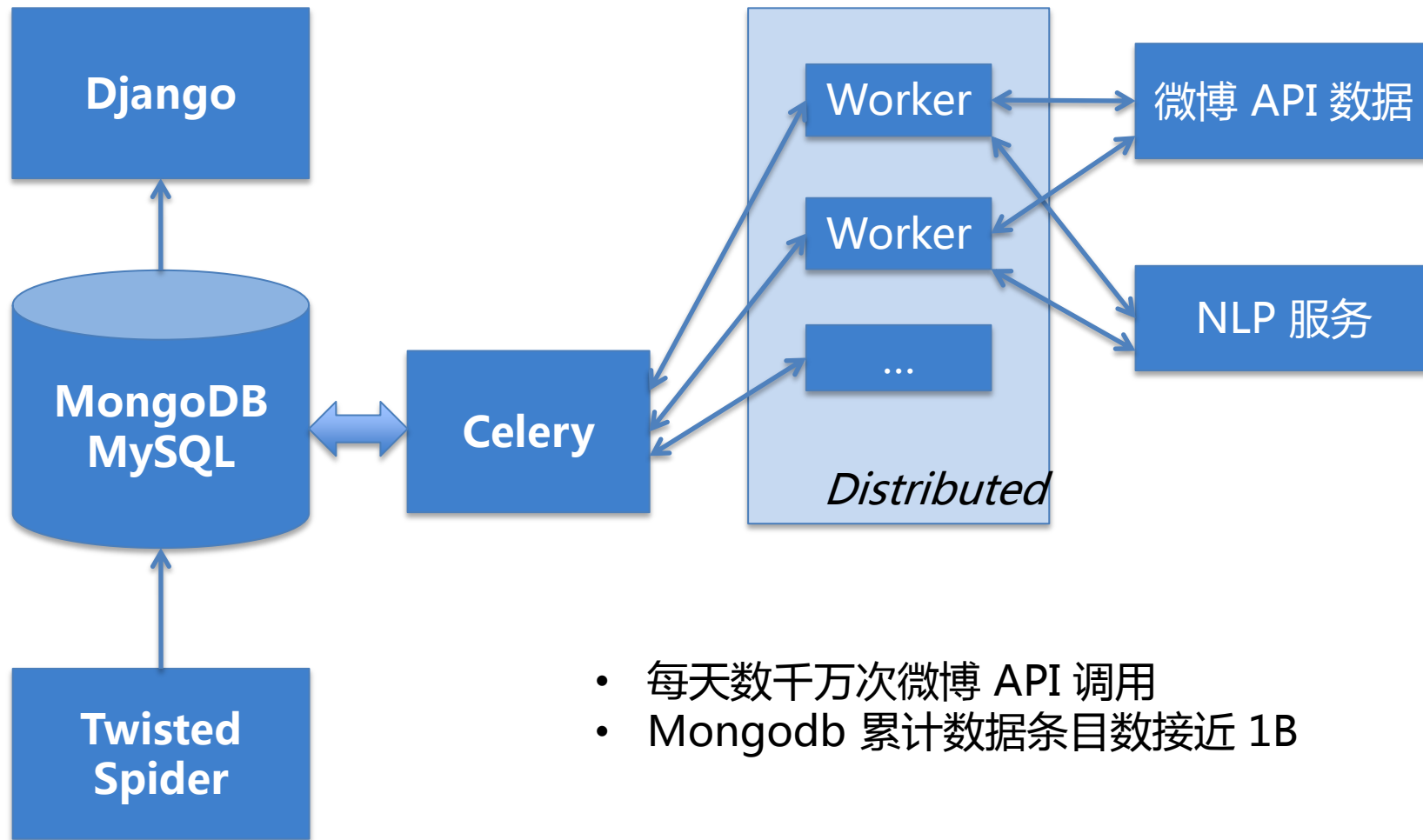
# 微博分析工具 WeiboMaster

The screenshot displays the WeiboMaster interface. On the left is a sidebar with the AdMaster logo and navigation menu items: 首页, 营销效果分析, 微博舆情分析, 市场营销, BenchMark, and Social Listening. The main area shows a '微博列表' (Weibo List) for the keyword '伦敦奥运' (London Olympics). It features filters for '今天', '昨天', '最近7天', '最近30天', and a date range '2012-10-17 - 2012-10-17'. Below the filters are tabs for '实时' and '热门微博', and a '所有' filter. A sorting dropdown is set to '发布时间'. The list contains three tweets:

- 负面** (Negative): 陈代朴: 【奥运撑杆跳比赛中杆突然断成三节】8月8日讯, 奥运会男子撑杆跳预赛正在进行中, 古巴选手在一次试跳时出现惊险一幕, 他在半空中时所撑的杆突然断成三节, 所幸运动员本人还没有被断了的杆伤到。可见这个杆也真有年头了, 是该更换了。伦敦奥运, 真可谓让人揪心啊! (2012年10月17日 11:01 新浪微博, 曝光: 73 评论: 0 转发: 0)
- 中性** (Neutral): 繁星行官方微博: 【美跳水名将梦幻婚礼】美国跳水名将、伦敦奥运会男子单人10米台金牌得主大卫DaveBoudia 上周五在家乡与女友索尼-布莱德步入婚姻殿堂, 上百位布迪亚和布莱德的亲戚与好友出席了他们的婚礼仪式。(2012年10月17日 11:00 皮皮时光机, 曝光: 10484 评论: 0 转发: 0)
- 正面** (Positive): 英国奢侈品: 英国伦敦Movement Pop Up Cafe坐落在格林威治DLR (城际轻轨) 附近, 看台设计的咖啡厅非常引人注目, 在伦敦奥运会前夕, 仅用了16天建成。该咖啡厅由英国艺术设计师 Morag Myerscough和诗人朋友 Len Sissay一起设计完成。不知道奥运会后此咖啡厅是否还在, 在英国的朋友如果感兴趣可以去看看。(2012年10月17日 11:00 皮皮时光机, 曝光: 202670 评论: 0 转发: 0)

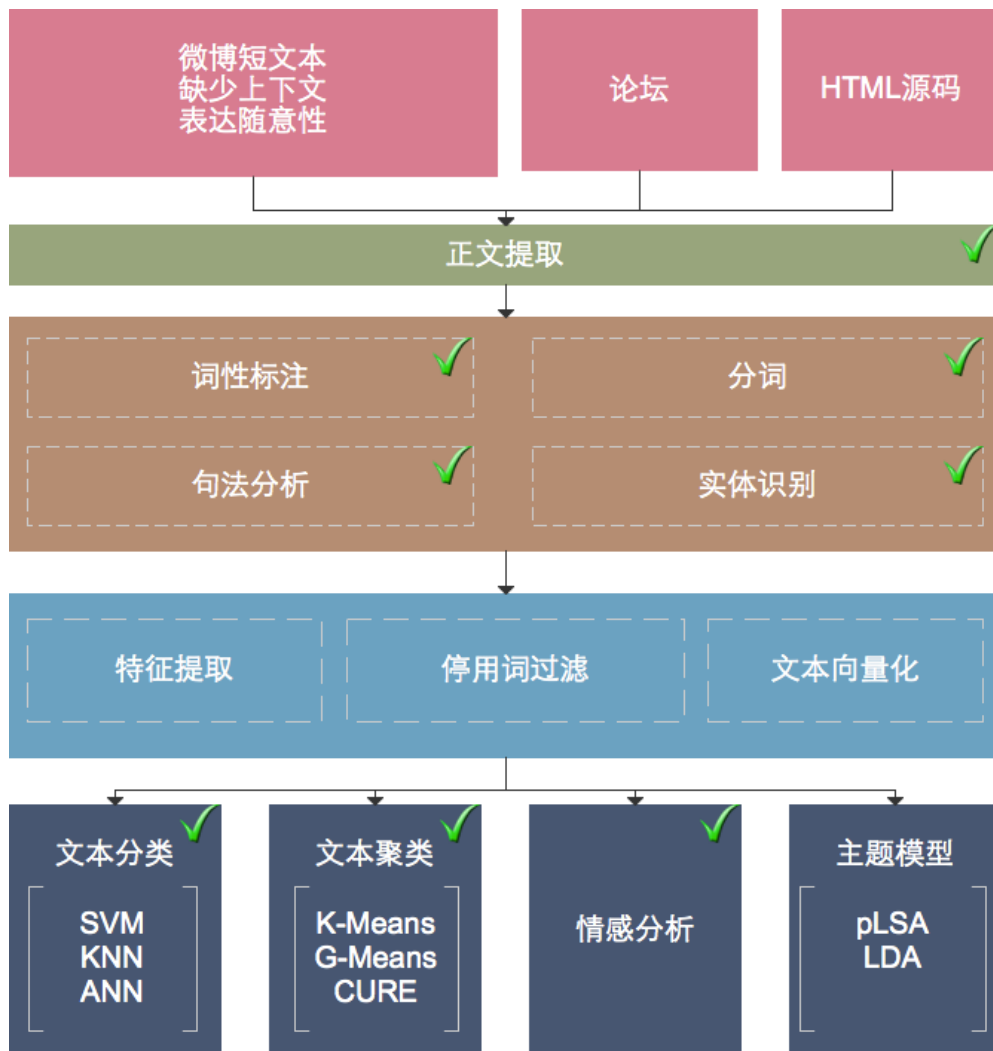
- 企业级
- 微博营销分析管理平台
- 新浪 + 腾讯
  
- 管理发布
- 指标分析
- 舆情监控

# 微博分析工具 WeiboMaster 架构



- 每天数千万次微博 API 调用
- MongoDB 累计数据条目数接近 1B

# NLP 服务介绍



- RESTful API: Tornado
- 进程控制: Supervisor
- libsvm, SciPy , NumPy
- 服务被多个产品线使用



# AdMaster 介绍

- **中国领先的独立第三方**数字营销全流程解决方案供应商；提供包括媒体审计、受众分析、社交媒体评估、舆情监测、在线调研等全方位的效果评估解决方案；
- AdMaster 成立于 2006 年，设有**上海、北京、广州** 3 家公司；
- 宝洁、海尔、LVMH、大众等全球 500 强公司，累计近 **3000** 个数字营销项目；
- 已完成 **B 轮融资**。



**We are hiring!**

Python          PHP  
                        Ruby          GIT          Hadoop  
HTML5  
                        MongoDB      Erlang

Thank you!

新浪微博：@陆丹峰