缓存命中率分析

缓存命中率分析报告：

1. 简单模式缓存命中是根据：

命中率=缓存命中/缓存命中+非缓存命中。

预抓取服务器从spv表里取出url，根据点击率排行高并且最近修改时间大于3个小时原则优先抓取(agent的缓存失效时间为3个小时)，通过抓取处理，存入缓存memcache中。

现有缓存服务器10台10.50.3.165-----10.50.3.174 每台memcache 大小8G

简单模式预抓取服务器 10.50.3.165-----10.50.3.173 (9台) 启动线程数20个，每台抓取30个url， myattion新闻列表内容 预抓取服务器 10.50.3.174启动线程数20个 ，url根据天天关注新闻预抓取提供。

2．分析数据：现网缓存命中为56.26%

以新浪为例：抓取<http://www.sina.com.cn> 两层共花40分钟左右，大小110M,要想提高命中率重要的两个因素，1.增加预抓取服务器台数，2.增加memcache 大小。