本周在ailab服务器上进行了多次单站点和多站点并行的压力测试，发现采集器中的数个问题，并及时解决和调优。现在作如下总结：

1. **压力测试情况：**压力测试暂时以功能性测试为侧重点，单站点从预检查阶段开始，一直到解析所有fetchable字段为1的节点为止全部通过，其中一些解析困难难以统一的细节均设置异常捕获机制并提供应急处理；多站点并行测试，以天涯，猫扑两个站点为测试（其他站点尚未配置解析参数或工具类不完善），并行效果明显，速度有提高。但测试未涉及暂停，终止，取消，恢复等功能的测试。
2. **较严重的问题：**mongoDB在32位操作系统中仅支持最大的2G的存储空间，在64位系统中无限。
3. **待解决的问题：**fetchSuccNum与真实值不符，原因正在核查；更多解析以及爬取过程中难以克服的细节异常捕获和应急处理；曾经遇到访问过量ip被封2小时的情况（天涯论坛），考虑加入线程调度任务轮休机制；暂时未遇到内存溢出的问题。
4. **采集器效率：**经过测试，在单站点采集的过程中，平均每分钟解析下载8个节点（约24个页面），在多站点（2站点）并行的情况下，平均每分钟解析下载12个节点（约36个页面）峰值能够达到20各节点（约60个页面），但如果每个站点要处理5000各节点（共12000各节点，余额36000个页面），则需要16~18个小时，采集速度严重依赖网速，令人担忧。
5. **采集器最后遗留的工作：**（1）采集器页面与内核的对接，按钮触发的逻辑功能编写以及测试（2）更多站点参数的配置（3）更多类型网站爬取解析类的补充（4）采集器的最后性能调优和健壮性增强