IPLib库析构大map优化

1. 优化前方法

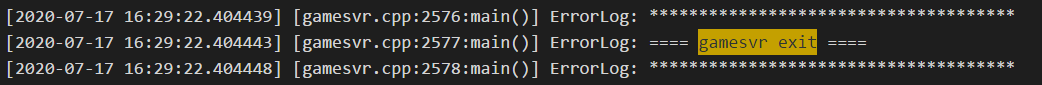
m\_ProvinceNameMap用map<int, string>维护，映射IP地址和对应的省份名字。其中key为IP地址的前24位，因此map需要存储256\*256\*256=16777216个键值对。

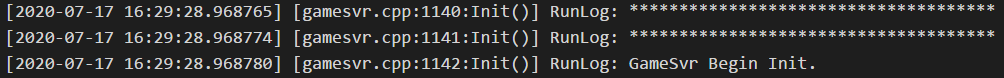


Gamesvr进程占用的内存: 约2.8G



此方法造成的问题：map太大，析构时间过长，私服上测试gamesvr进程重启时间约6秒





1. 方法优化
2. 方法一：m\_ProvinceNameMap改为数组

为了减少析构时间，可以把m\_ProvinceNameMap改为char\*数组， 数组的下标代表IP地址的前24位，数组中每个元素为一个64B 的字符串。



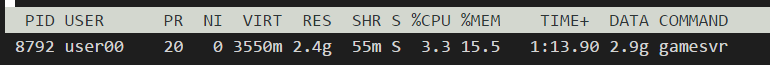
1. 内存占用评估：

m\_ProvinceNameMap大小：256\*256\*256\*8 B = 134217728 B = 128 MB

m\_ProvinceNameMap中存储的szProvince大小：256\*256\*256\*64 B = 1024 MB

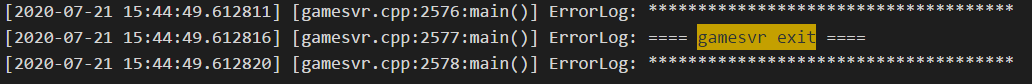
占用总内存约为128MB + 1024MB = 1152 MB

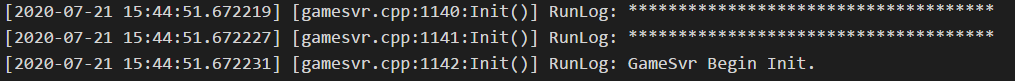
Gamesvr进程实际占用内存：



1. 进程重启时间评估：

私服测试进程重启时间约2秒。





1. 存在的问题：

16777216个IP地址映射了约3000个省份的字符串，其中有大量重复存储的字符串，造成数据的冗余。

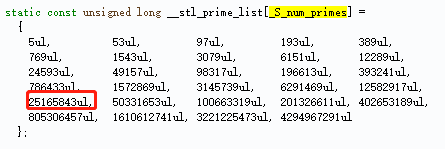
1. 方法二：m\_ProvinceNameMap改为哈希表

C++提供的哈希表结构有C++11的std::unordered\_map和\_\_gnu\_cxx::hash\_map，这里使用的是\_\_gnu\_cxx::hash\_map。



1. 内存占用评估

\_\_gnu\_cxx::hash\_map的桶的数目是根据插入数据量动态增长的，当插入数据个数>桶数时，hash\_map就把桶数扩大。hash\_map源码中定义了一个素数表，下一次增长的桶数从素数表中选择



当插入数据为16629255个时，hash\_map的桶数就增长到25165843个。

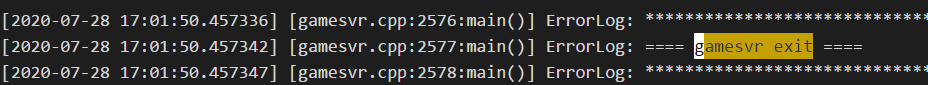


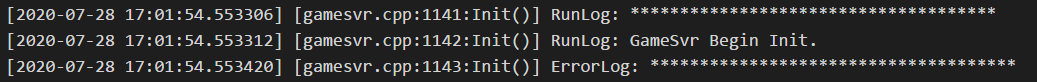
所以m\_ProvinceNameMap大致需要的内存为（sizeof(string)=8；string长度按64B计算）： 25165843\*8 + 16629255\*(4 + 8 + 64) B = 1397MB

实际占用内存约2.5G，比数组还大。



1. 进程重启时间





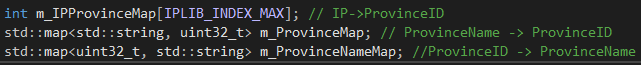
进程重启约4秒，比数组慢。

1. 存在的问题

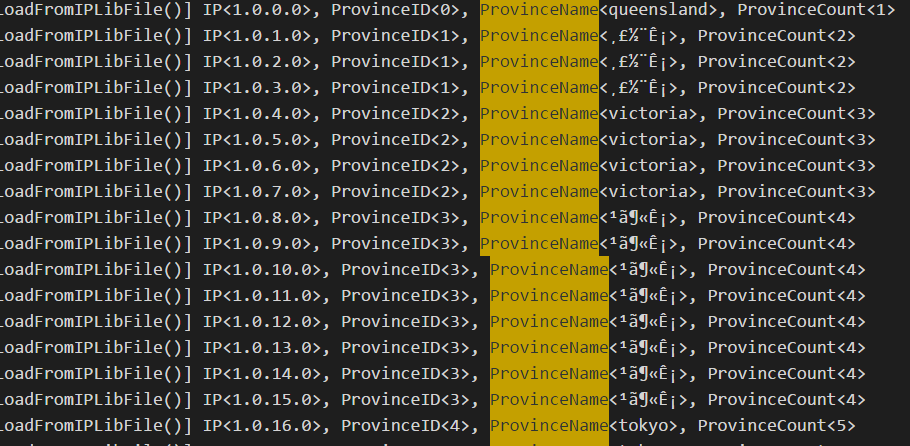
桶的数目扩充太多，造成内存的浪费和析构时间的增长；而且也没有解决省份字符串存储冗余的问题。

1. 方法三(最优)：<IP，省份名>映射改为<IP，省份编号>映射

为了尽量减少数据冗余，可参考IP到国家的映射方式，为每个省份标记一个ProvinceID，把m\_ProvinceNameMap改为int数组，其中下标为IP地址前24位，元素为对应的ProvinceID。



另外再使用m\_ProvinceMap和m\_ProvinceNameMap，分别存储<ProvinceID, ProvinceName>映射和<ProvinceName, ProvinceID>映射。这两个map的读取和ProvinceID的赋值可以在读取IP库文件的时候完成。



（1） 内存占用评估：

总共省份数：2852个； sizeof(map)=48；sizeof(string)=8；string长度按64B计算（实际小于64B）

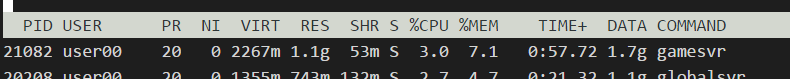
m\_IPProvinceMap: 256\*256\*256\*4 B = 64MB

m\_ProvinceMap: 48 + 2852\*(8 + 4 + 64) B = 0.2MB

m\_ProvinceNameMap: 48 + 2852\*(8 + 64 + 4) B = 0.2MB

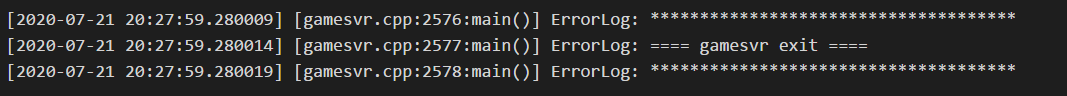
占用总内存：约65MB

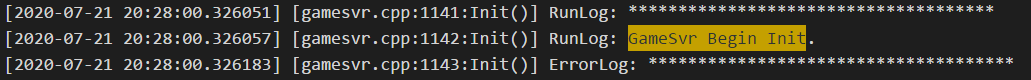
Gamesvr进程实际占用内存：



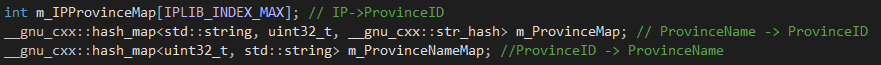
（2）进程重启时间

私服测试时间约1秒





补充：把m\_ProvinceMap和m\_ProvinceNameMap改为hash\_map



1. 内存占用评估：（bucket\_size: 3079）

m\_IPProvinceMap: 256\*256\*256\*4 B = 64MB

m\_ProvinceMap: 3079\*8 + 2852\*(4 + 8 + 64) B = 0.2MB

m\_ProvinceNameMap: 3079\*8 + 2852\*(8 + 64 + 4) B = 0.2MB

实际占用内存：



1. 进程重启时间：约1秒

内存占用和重启时间与map差不多。

核心源码：

