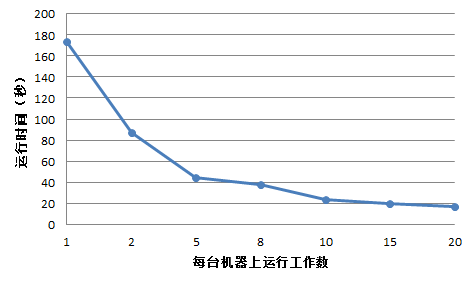
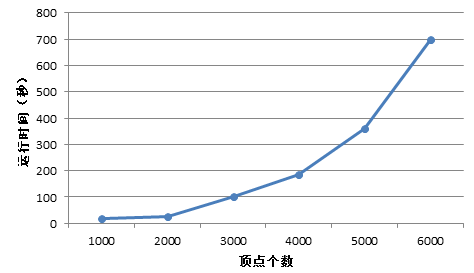
我们在两台虚拟机上做了单源最短路径实验。使用了对数正态分布的随机图(为了研究扩展性)进行实验。我们使用如下公式生成随机图，其中，，因此图的平均出度为127.1。

下图表示计算有1000个节点的二叉树图的单源最短路径问题，运行时间随着每台机器上的任务总数增多的改变情况。可以看出任务总数从1增加到20，时间从174s减少到17.3s，加速比为10。



为了显示Giraph的扩展性，我们生成了不同大小的图进行试验，其中每台机器上进行20个任务。运行时间从17.3s增加到了702s，证明了在图平均度数较小的时候，运行时间是随着顶点增多而线性增长的。



由以上实验可以看出Giraph的扩展性非常好，只要增加机器数量，就能够减少运行时间。随着顶点数的增加，运行时间呈线性增长。

实际生活中的图一般为稀疏图，很少有稠密的，这也是Giraph的假设之一。Giraph也存在一定问题。第一，Giraph目前考虑的算法都是在内存中完成，没有考虑到图如果大到一定程度，如何在外存中调度问题；第二，需要人工的设置数据划分，保证负载平衡；第三，就是之前提到的问题，如果有非常稠密的图，应该如何处理消息传递的问题。