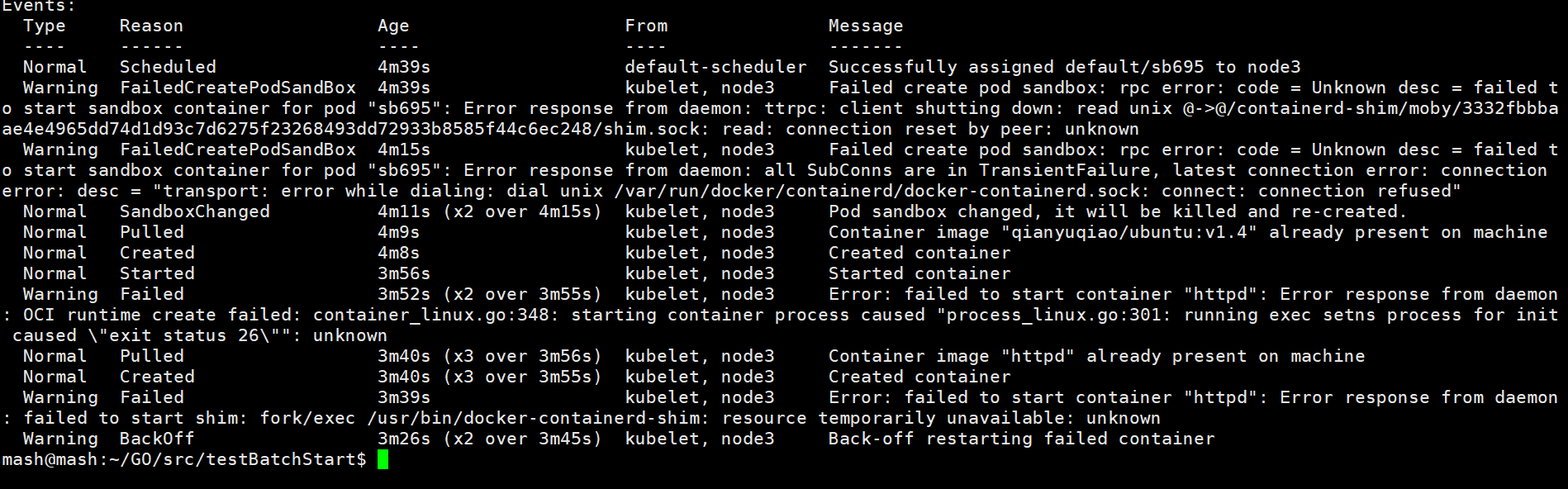
1. 修改cidr

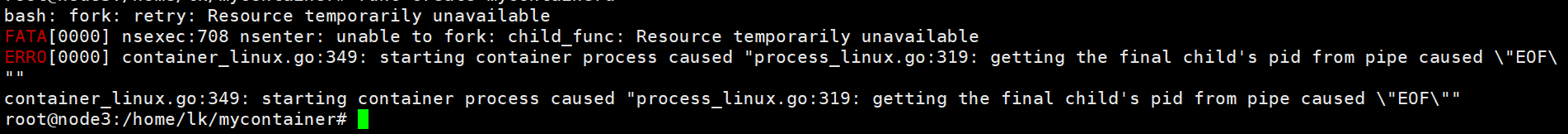
https://github.com/qianyuqiao/K8S/blob/master/%E6%94%B9%E5%8F%98node%E4%B8%8A%E7%9A%84podCIDR.md

K8s集群每个节点默认最大开启pod数量为110个。其配置文件位于路径Ubuntu /etc/systemd/system/kubelet.service.d/10-kubeadm.conf，在kubelet启动参数中加入--max-pods=1200，然后重启kubelet。

以pod形式启动，每个pod中有两个容器，提供不同的服务，发现在启动600多个之后pod无法继续正常启动，其错误提示为

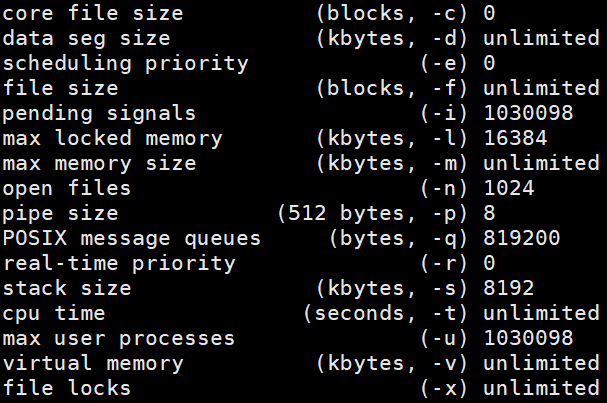


其源码对应位置为process\_linux.go的301行，下载<https://github.com/opencontainers/runc>源代码编译安装测试发现，



现实最终代码出错位置在nsenter文件夹下的nsexe.c的708行出错，出错函数为clone\_parent，作用为开启一个容器的进程，于是判断因为资源限制导致无法创建更多的进程。

查看node3资源限制ulimit –a



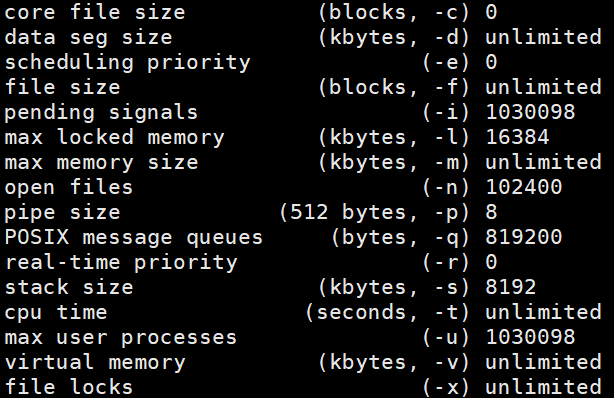
初步判断应当时最大用户进程的限制和最大打开文件数量。

虽然此处最大用户进程数为1030098，但是由于系统总限制，使得用户进程数不能达到该数量，于是在/etc/sysctl.conf中写入kernel.pid\_max=1030098并重启机器使其生效，但是问题没有得到解决，考虑到原来该值就已经很大，所以判断应当修改最大打开文件数量。修改方法为在/etc/security/limits.conf文件中加入两行内容：

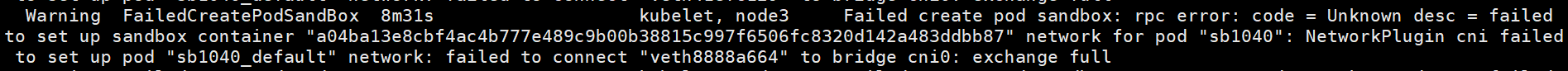
root soft nofile 102400

root soft nofile 102400

重启机器再次运行ulimit –a发现已生效

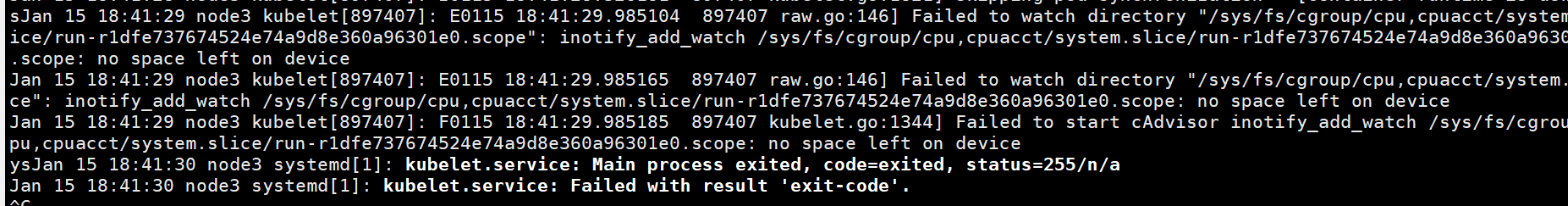


此时，pod可以继续创建至1000个，但是超过1000多个时依旧出现问题，查看得



判断是网桥最大端口数量限制，该参数是由内核中参数BR\_MAX\_PORTS 和BR\_PORT\_BITS控制，需要修改参数并编译内核。待解决。

将pod中的容器数量改为5个，发现在800个左右时出现错误



运行sudo sysctl fs.inotify.max\_user\_watches=1048576，可以继续运行1000个pod

参考：<https://www.ibm.com/cloud/blog/docker-insane-scale-on-ibm-power-systems>

<https://zgljl2012.com/failed-to-start-cadvisor-inotify_add_watch-sys-fs-cgroup-devices-no-space/>