淘宝网
Taobao.com

一个JVM crash的分析解决过程

阿里巴巴核心系统研发部
梅路峣/云达

关于我

- 2011.7 开始在淘宝实习，从事JVM相关工作
- 2012.3 硕士毕业（北邮计算机专业）
- 脑残G粉
- 对编程语言的实现很感兴趣，目前关注重点在JVM的实现

- http://weibo.com/u/1063244843

主要内容

- 什么是JVM crash
- 阿里和百度都碰到的一个JVM crash

什么是JVM crash

- Java进程意外消失
- 和Java层面的Exception、Error不是一个概念
- 通常伴随着
  - crash log：`hs_err_pid<pid>.log`（重要）
  - core dump: `core.<pid>`
  - 或系统日志
    - 主要是 `/var/log/messages` 里的异常信息

制造JVM crash

- 最简单的方式：

  jinfo -flag FLSLargestBlockCoalesceProximity <pid>

```
#
# A fatal error has been detected by the Java Runtime Environment:
#
#  Internal Error (globals.cpp:153), pid=32269, tid=1118550336
#  Error: ShouldNotReachHere()
#
# JRE version: 6.0_32-b05
# Java VM: Java HotSpot(TM) 64-Bit Server VM (20.7-b02 mixed mode
linux-amd64 compressed oops)
# An error report file with more information is saved as:
# /home/yunda.mly/test/hs_err_pid32269.log
#
# If you would like to submit a bug report, please visit:
#   http://java.sun.com/webapps/bugreport/crash.jsp
#
Aborted(core dumped)
```

今天要讲的JVM crash

- 阿里和百度先后遇到
- 共同特点：使用了反向JNI，即用native（C/C++）代码调用Java代码，并申请了大量内存
- 人为制造：

  在Java launcher（java.c）初始化JVM代码前加上：

```
size_t K = 1024;
size_t len = 3*K*K*K;
char * test_polling_page = (char *)malloc(len);
```

百度碰到的crash


```
int main(int argc, char **argv) {
    if (argc != 2) {
        fprintf(stderr, "Usage: hdfs_write <mem>\n");
        exit(-1);
    }
    tSize m = strtoul(argv[1], NULL, 10);

    fprintf(stderr, "allocate %ld bytes\n", m * 1024 * 1024 * 1024L);

    char* memmem = (char*)malloc(m * 1024*1024*1024L);
    fprintf(stderr, "memmem %p \n", memmem);
    fprintf(stderr, "before\n");

    hdfsFS fs = hdfsConnect("default", 0);
    if (!fs) {
        fprintf(stderr, "Oops! Failed to connect to hdfs!\n");
        exit(-1);
    }

    fprintf(stderr, "after\n");

    return 0;
}
```

Crash log

- 地址：https://gist.github.com/4602677
- 出错的地方：

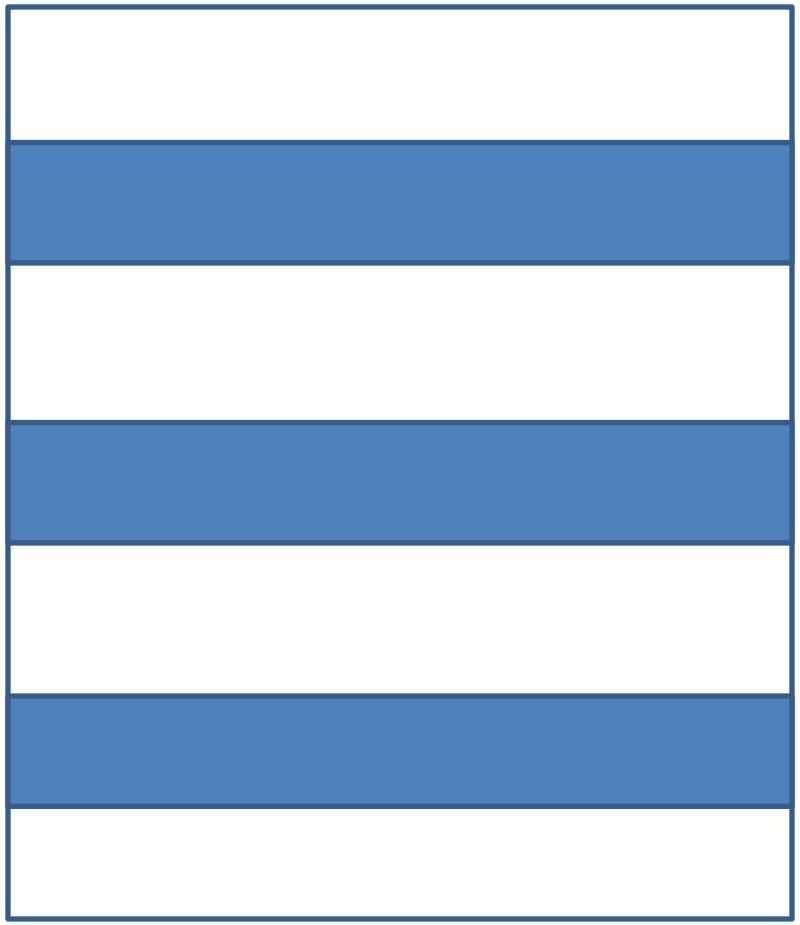
0x00007f943ff492ad: test %eax,-0x3bd062b3(%rip)
0x00007f94 04243000
; {poll_return}
0x00007f943ff492b3: retq

- rip：0x00007f943ff492b3

TaobaoJVM
特有功能

什么是poll_return

- JIT过的方法在返回前进行的polling page访问判断
- 主要用于safepoint（如GC，线程dump等）



Gdb 调试core文件

- $ readelf -sW libjvm.so|grep polling_page

42370: 00000000011cfb90 8 OBJECT LOCAL DEFAULT 29 _ZN2os13_polling_pageE

- (gdb) p &_ZN2os13_polling_pageE
  $1 = (<data variable, no debug info> *) 0x7f95040f7788
- (gdb) x /2 0x7f95040f7788
  0x7f95040f7788 <_ZN2os13_polling_pageE>:
  0x04243000 0x00007f95

0x00007f95 04243000

JIT生成polling_page地址的代码

- src/cpu/x86/vm/x86_64.ad

```
    emit_d32_reloc(cbuf, os::get_polling_page());

void emit_d32_reloc(CodeBuffer& cbuf, address addr) {
  address next_ip = cbuf.insts_end() + 4;
  emit_d32_reloc(cbuf, (int) (addr - next_ip),
                 external_word_Relocation::spec(addr),
                 RELOC_DISP32);
}
```

有趣的现象

- 当申请1G、4G、8G空间的时候没有问题，申请2G、3G、7G、10G内存的时候会crash
  - 有偶然性，以4G为单位结果相同
- 这个crash 在2.6.18内核上不会发生，在2.6.32上会稳定重现
  - 和mmap系统调用实现有关

参考资料

- 撒迦关于JVM crash的PPT：
  http://www.slideshare.net/RednaxelaFX/java-crash
- 百度遇到的JVM crash的讨论：
  http://hllvm.group.iteye.com/group/topic/35498
- OpenJDK6源码：
  http://download.java.net/openjdk/jdk6/
- TaobaoJVM主页：
  http://jvm.taobao.org/index.php
- safepoint介绍：
  http://blog.csdn.net/raintungli/article/details/7162468

淘宝网
Taobao.com

Thank you

Q&A section