中通科技

火焰图

一种非侵入式在线诊断技术

宦传建

高效协作 勇于担当

积极创新 用户至上

CONTENT

目录>>

01 / 火焰图是什么

02 / 有哪些优点？

03 / 如何使用？

04 / 生产案例

高效协作 勇于担当
积极创新 用户至上

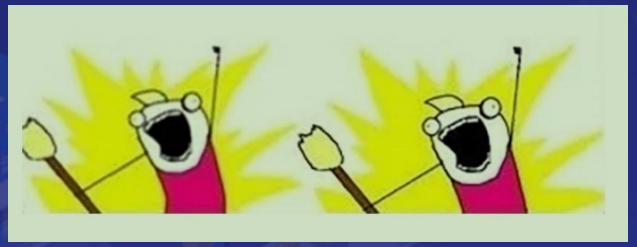
生产问题

1. 原因明显

有详细的出错堆栈信息

2. 原因不明

CPU 莫名拉高？内存居高不下？服务依旧跑、也不报错



常规方式

埋点

埋多少？侵入式修改，频繁修改上线，重启服务

摘节点

隔离生产流量，降低集群吞吐，问题复现困难

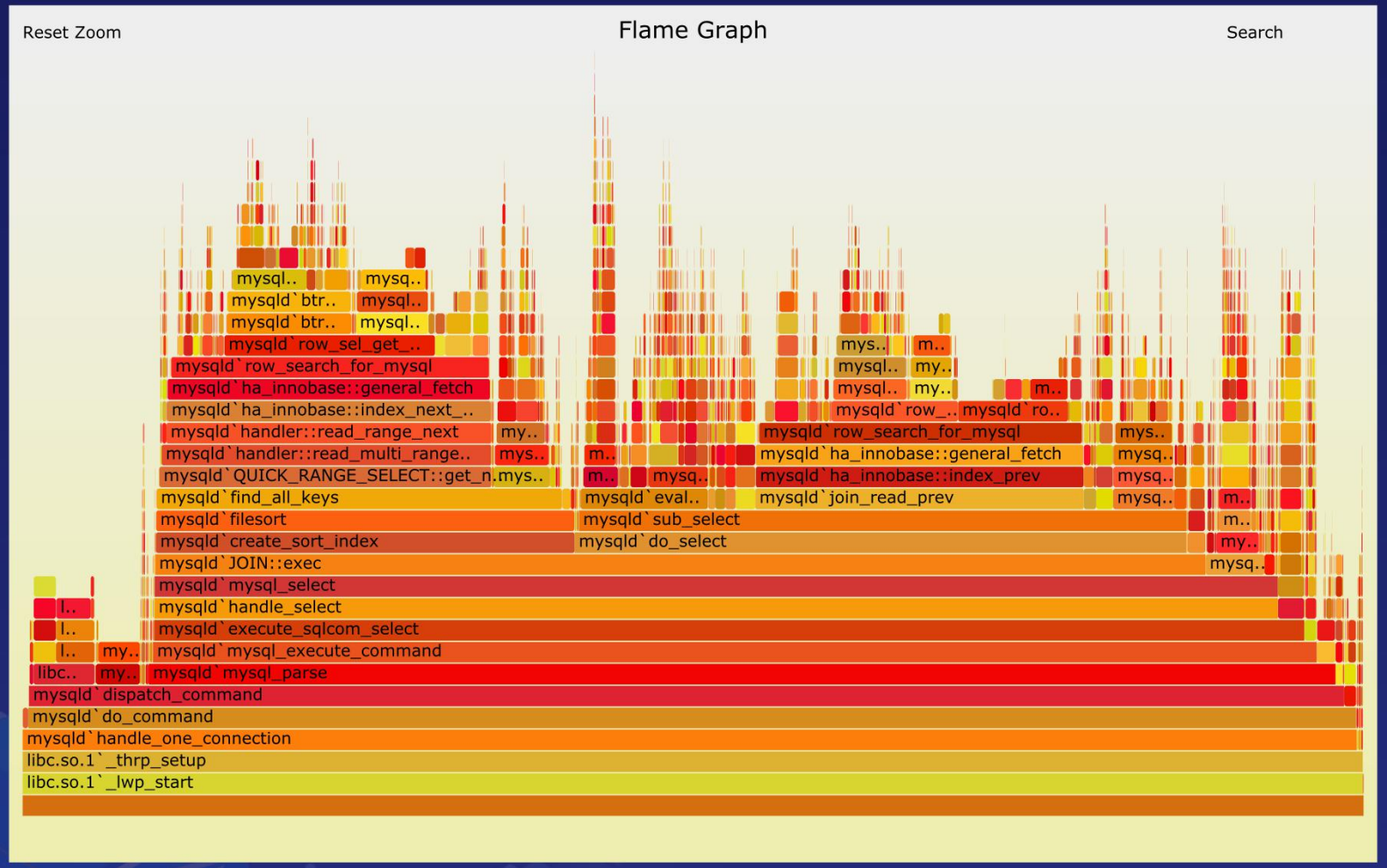
有没有一种更灵活、非侵入式的方法？

不需要修改代码、配置

不需要重启服务

不影响生产运行

.....



http://www.brendangregg.com/flamegraphs.html

基于**在线采样**生成的可视化分析工具，将时间和空间两个维度融合在一张图上，反映系统在性能方面的很多定量的统计规律。

优点

非侵入式

热插拔，随时随地，按需采集

开销很小，<3%

...

高效协作 勇于担当
积极创新 用户至上

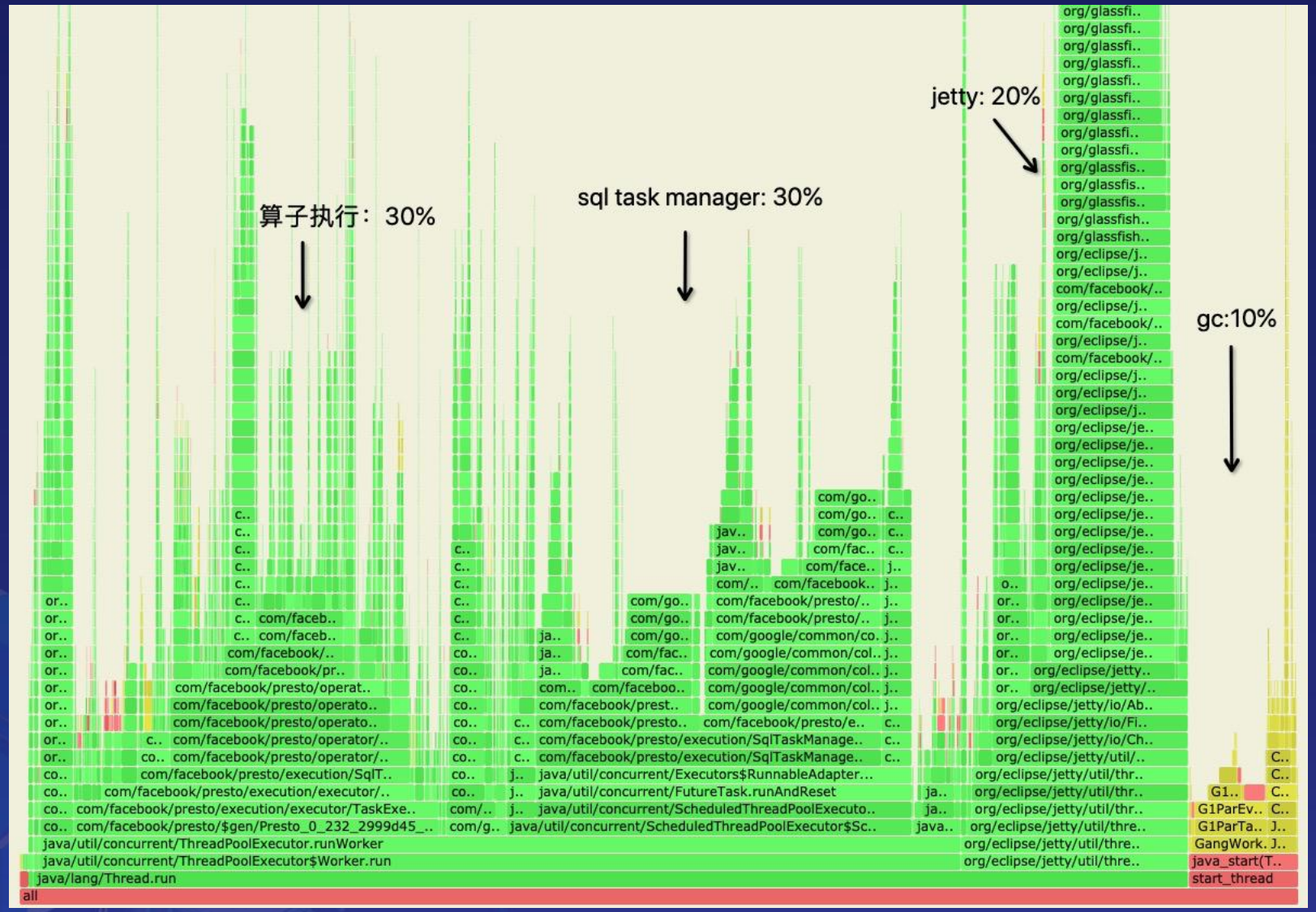
如何生成？ Arthas

```
[arthas@283911]$ help
 NAME         DESCRIPTION
 help         Display Arthas Help
 keymap       Display all the available keymap for the specified connection.
 sc           Search all the classes loaded by JVM
 sm           Search the method of classes loaded by JVM
 classloader  Show classloader info
 jad          Decompile class
 getstatic    Show the static field of a class
 monitor      Monitor method execution statistics, e.g. total/success/failure count, average rt, fail rate, etc.
 stack        Display the stack trace for the specified class and method
 thread       Display thread info, thread stack
 trace        Trace the execution time of specified method invocation.
 watch        Display the input/output parameter, return object, and thrown exception of specified method invocation
 tt           Time Tunnel
 jvm          Display the target JVM information
 perfcounter  Display the perf counter information.
 ognl         Execute ognl expression.
 mc           Memory compiler, compiles java files into bytecode and class files in memory.
 redefine     Redefine classes. @see Instrumentation#redefineClasses(ClassDefinition...)
 dashboard    Overview of target jvm's thread, memory, gc, vm, tomcat info.
 dump         Dump class byte array from JVM
 heapdump     Heap dump
 options      View and change various Arthas options
 cls          Clear the screen
 reset        Reset all the enhanced classes
 version      Display Arthas version
 session      Display current session information
 sysprop      Display, and change the system properties.
 sysenv       Display the system env.
 vmoption     Display, and update the vm diagnostic options.
 logger       Print logger info, and update the logger level
 history      Display command history
 cat          Concatenate and print files
 echo         write arguments to the standard output
 pwd          Return working directory name
 mbean        Display the mbean information
 grep         grep command for pipes.
 tee          tee command for pipes.
 profiler     Async Profiler. https://github.com/jvm-profiling-tools/async-profiler
 stop         Stop/Shutdown Arthas server and exit the console.
```

```
[arthas@283911]$ profiler start
Started [cpu] profiling
[arthas@283911]$ profiler status
[perf] profiling is running for 2 seconds
[arthas@283911]$ profiler getSamples
1480
[arthas@283911]$ profiler stop
OK
profiler output file: /home/presto/data232/arthas-output/20200917-110107.svg
[arthas@283911]$
```

高效协作 勇于担当
积极创新 用户至上

如何看图？



x轴表示 stack 采样数量，越宽的表示出现的频率越高

y轴表示 stack 的调用深度

颜色的深浅不重要，是随机表示的

宽平顶通常意味着潜在的热点

高效协作 勇于担当
积极创新 用户至上

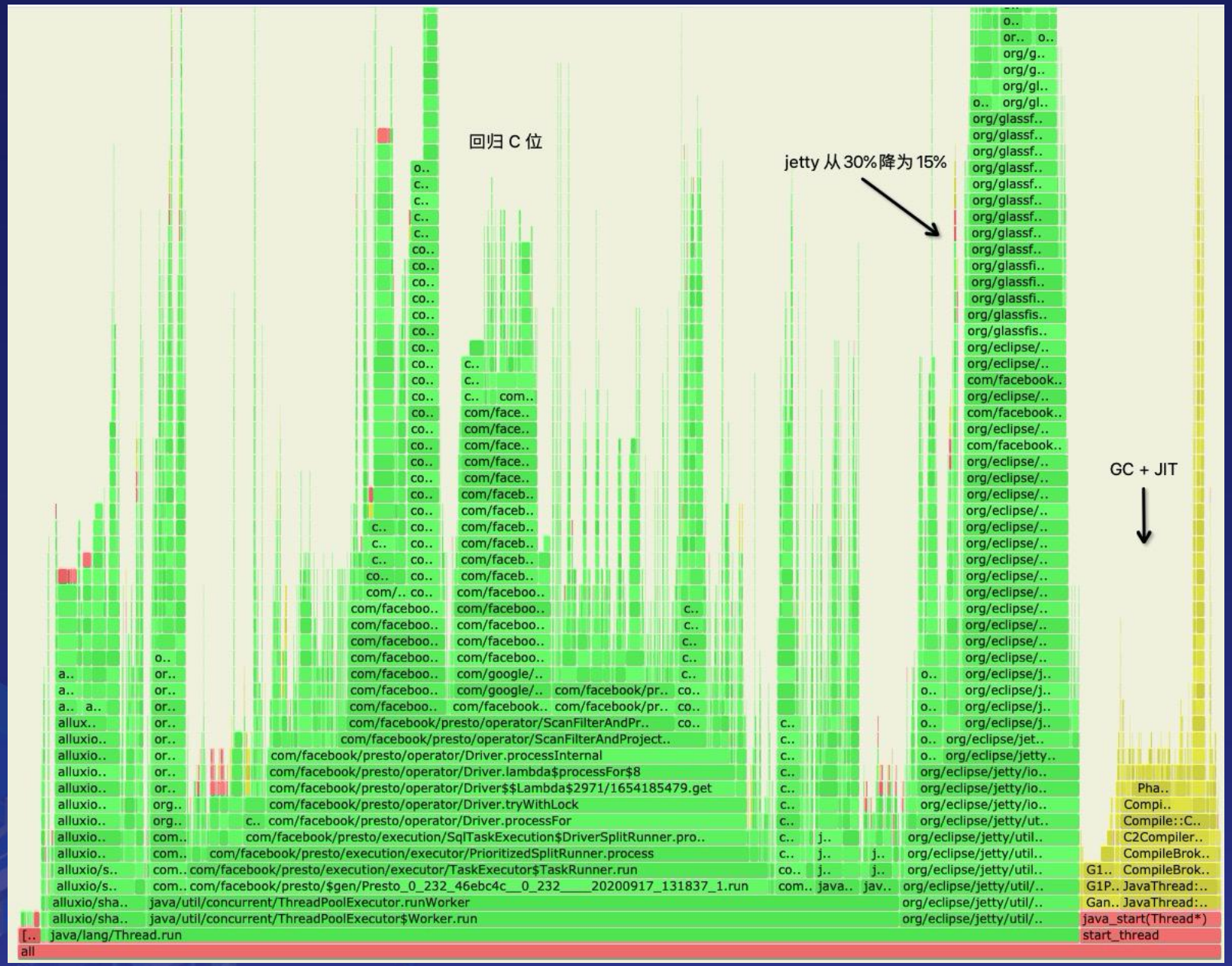
task manager 优化

task manager 定时任务间隔从 200ms 降低为 5s 一次

task info 过期时间从 15分钟改为 5分钟

高效协作 勇于担当
积极创新 用户至上

http 请求优化



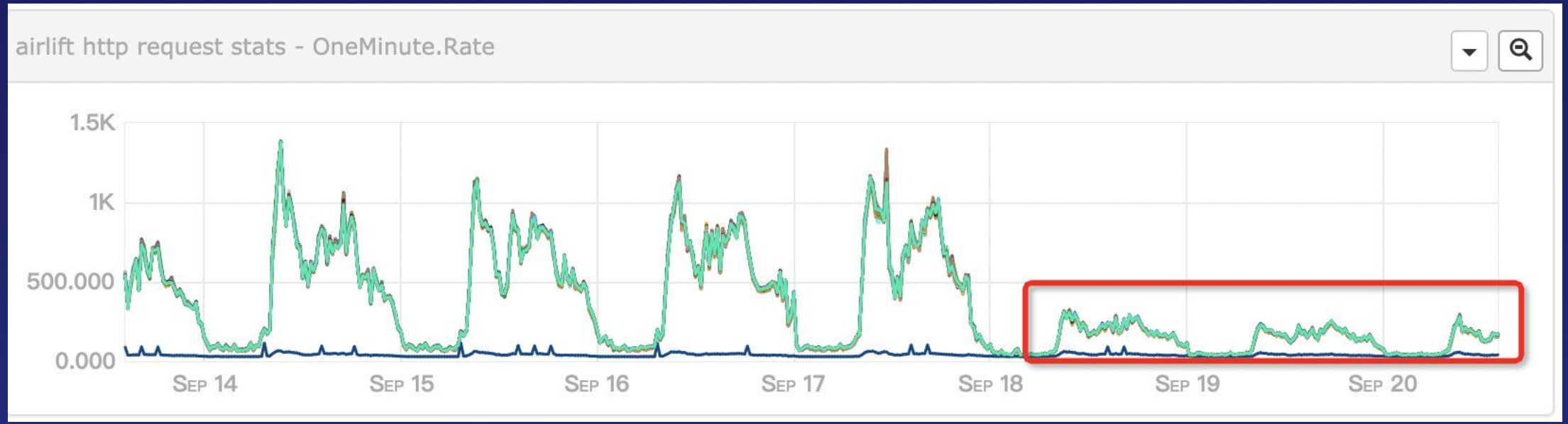
worker 的 http 请求主要用于 task 之间交换数据

集群不宜过大，主节点会为每个 task 创建两个轮询任务：轮询 task 执行状态、定时更新 task info

小查询居多，减少不必要的 task

buffer 不宜过大，启发式公式会根据 buffer 的剩余容量发起更多的 http 请求

高效协作 勇于担当
积极创新 用户至上

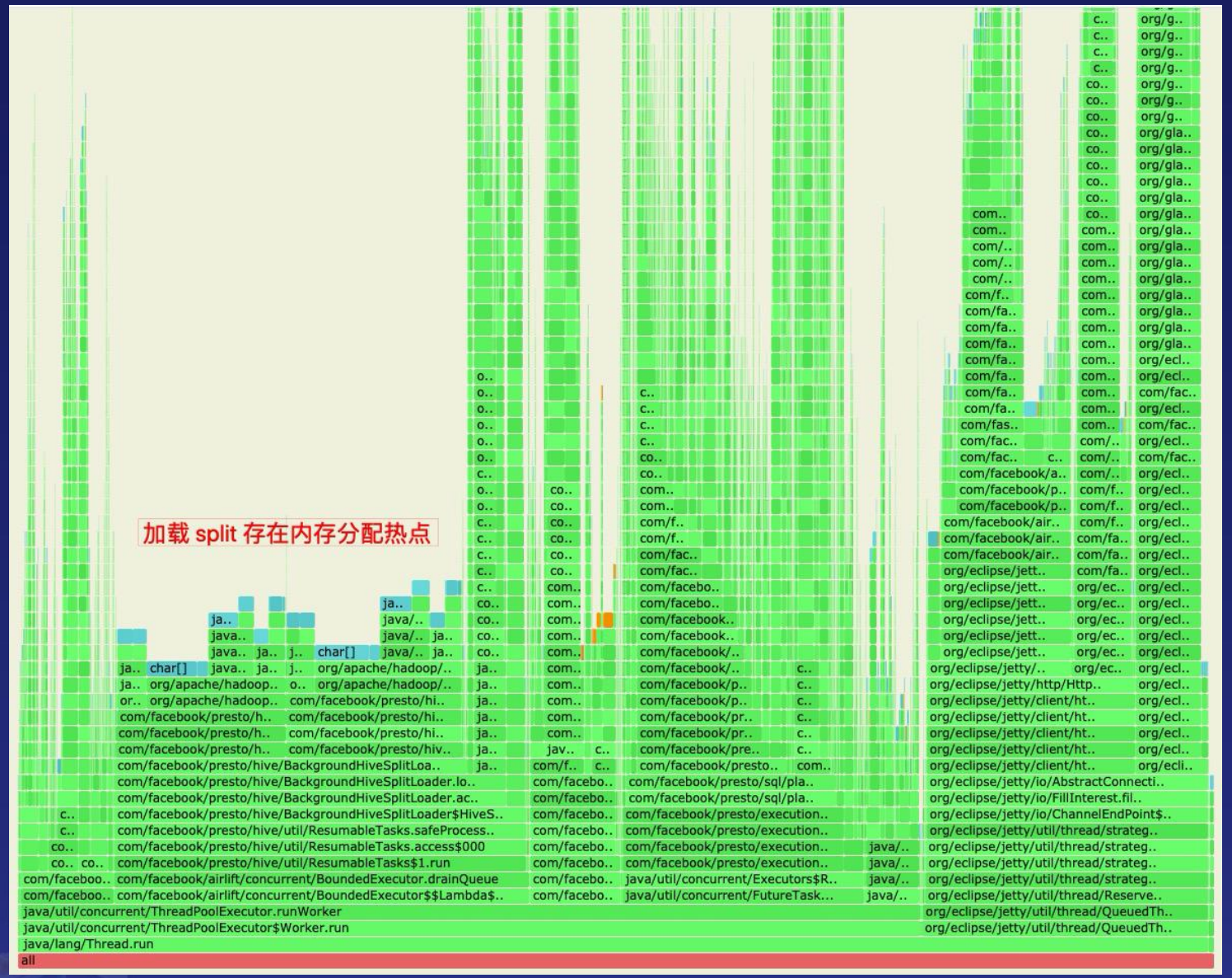


高效协作 勇于担当
积极创新 用户至上

https://github.com/jvm-profiling-tools/async-profiler

高效协作 勇于担当
积极创新 用户至上

内存火焰图



./profiler.sh -e alloc -i 1024k -d 60 -f fileName.svg PID

高效协作 勇于担当
积极创新 用户至上

```
public static void copy(Configuration from, Configuration to)  from: "Configuration: "  to: "Configuration: "
{
    for (Map.Entry<String, String> entry : from) {  from: "Configuration: "
        to.set(entry.getKey(), entry.getValue());
    }
}
```

ConfigurationUtils › copy()

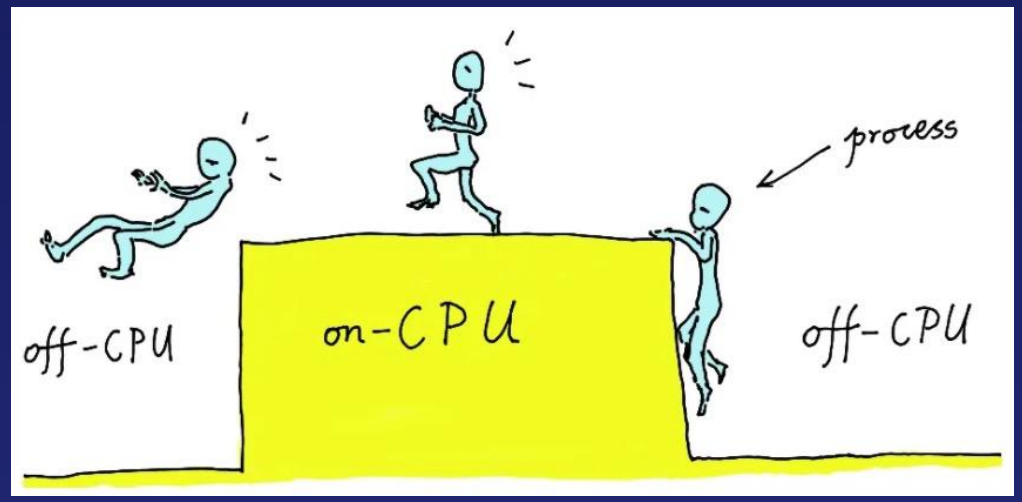
```
loadDefaults = false
updatingResource = {ConcurrentHashMap@20016} size = 438
properties = {Properties@20017} size = 438
  "io.bytes.per.checksum" -> "512"
  "fs.s3a.connection.timeout" -> "50000"
  "hadoop.security.kms.client.encrypted.key.cache.expiry" -> "43200000"
  "dfs.image.transfer.bandwidthPerSec" -> "0"
  "hadoop.security.sensitive-config-keys" -> "password$,fs.s3.*[Ss]ecret.?[Kk]ey,fs.azure.account.key.*,dfs.webhdfs.oauth2.[a-z]+.token,hadoop.securit
  "dfs.block.access.token.lifetime" -> "600"
  "dfs.namenode.resource.du.reserved" -> "104857600"
  "hadoop.proxyuser.hive.hosts" -> "*"
  "fs.AbstractFileSystem.file.impl" -> "org.apache.hadoop.fs.local.LocalFs"
  "fs.du.interval" -> "600000"
  "fs.AbstractFileSystem.har.impl" -> "org.apache.hadoop.fs.HarFs"
  "dfs.domain.socket.path" -> "/var/run/hdfs-sockets/dn"
  "dfs.balancer.max-no-move-interval" -> "60000"
  "hadoop.ssl.keystores.factory.class" -> "org.apache.hadoop.security.ssl.FileBasedKeyStoresFactory"
  "fs.s3a.multipart.size" -> "104857600"
  "hadoop.http.authentication.kerberos.keytab" -> "${user.home}/hadoop.keytab"
```

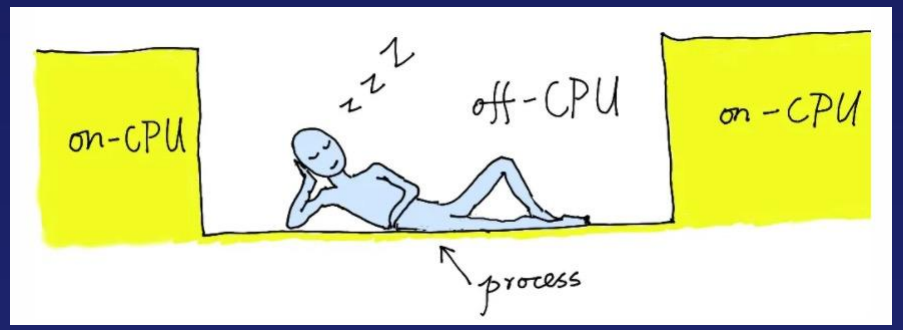
```
public void set(String name, String value, String source) {
  Preconditions.checkArgument(
      expression: name != null,
      errorMessage: "Property name must not be null");
  Preconditions.checkArgument(
      expression: value != null,
      errorMessage: "The value of property " + name + " must not be null");
  name = name.trim();
  DeprecationContext deprecations = deprecationContext.get();
  if (deprecations.getDeprecatedKeyMap().isEmpty()) {
    getProps();
  }
  getOverlay().setProperty(name, value);
  getProps().setProperty(name, value);
  String newSource = (source == null ? "programatically" : source);

  if (!isDeprecated(name)) {
    updatingResource.put(name, new String[] {newSource});
    String[] altNames = getAlternativeNames(name);
    if(altNames != null) {
      for(String n: altNames) {
        if(!n.equals(name)) {
          getOverlay().setProperty(n, value);
          getProps().setProperty(n, value);
          updatingResource.put(n, new String[] {newSource});
        }
      }
    }
```

Presto 调度阶段访问 HDFS

高效协作 勇于担当
积极创新 用户至上



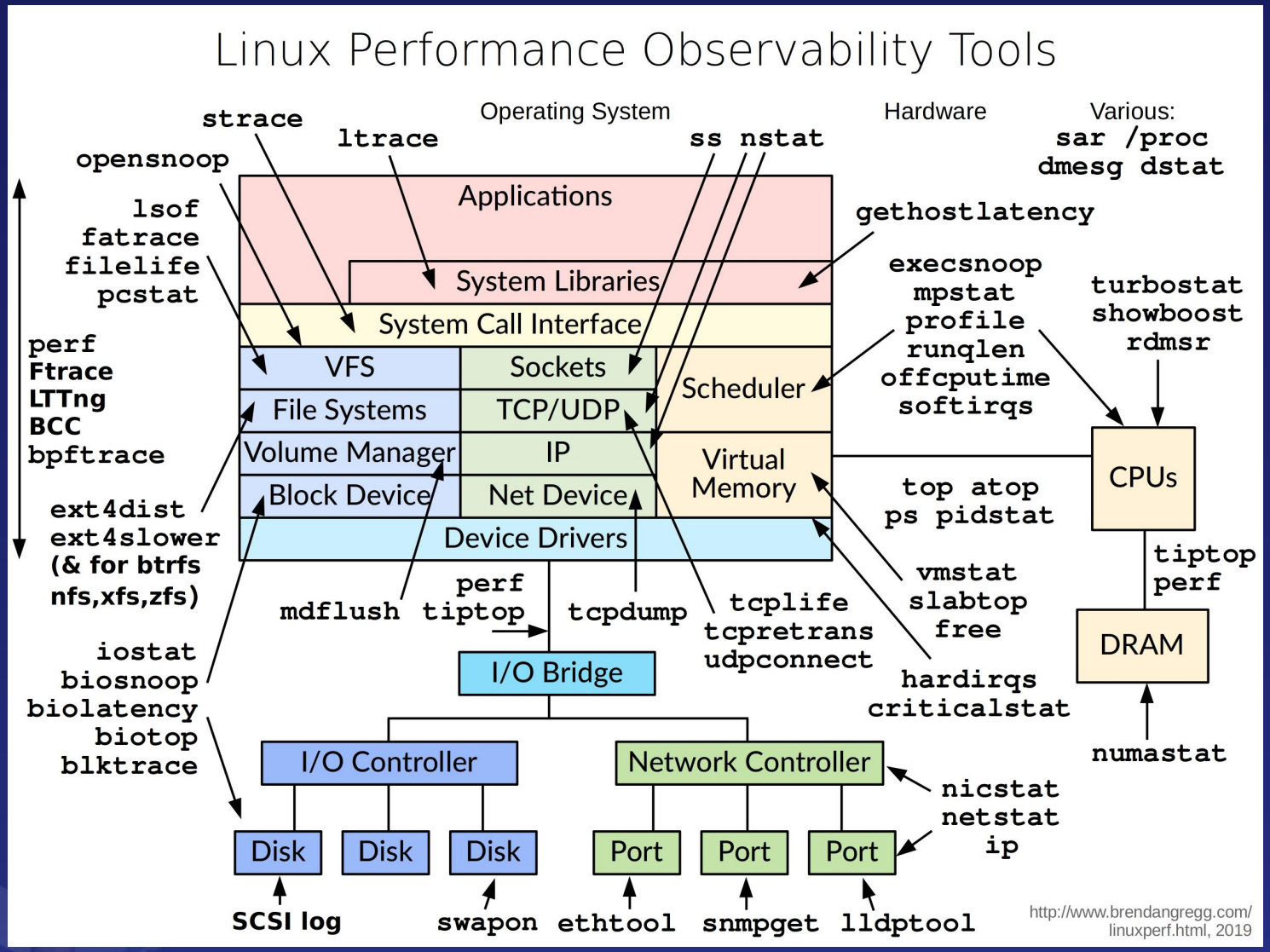


* off-CPU 维度的火焰图：进程/线程不运行在任何 CPU 上的时间，锁、竞争、I/O 等待。。。

高效协作 勇于担当
积极创新 用户至上

动态追踪 or 静态分析？

高效协作 勇于担当
积极创新 用户至上



高效协作 勇于担当
积极创新 用户至上

Q 有什么想知道的？

What would You like to know?

我想知道... A

I wank to know...

高效协作 勇于担当
积极创新 用户至上

谢谢

THANK YOU

中通科技