

13-如何利用RunLoop原理去监控卡顿？

你好，我是戴铭。今天，我来和你说说如何监控卡顿。

卡顿问题，就是在主线程上无法响应用户交互的问题。如果一个 App 时不时地就给你卡一下，有时还长时间无响应，这时你还愿意继续用它吗？所以说，卡顿问题对App的伤害是巨大的，也是我们必须重点解决的一个问题。

现在，我们先来看一下导致卡顿问题的几种原因：

- 复杂 UI、图文混排的绘制量过大；
- 在主线程上做网络同步请求；
- 在主线程做大量的IO 操作；
- 运算量过大，CPU持续高占用；
- 死锁和主子线程抢锁。

那么，我们如何监控到什么时候会出现卡顿呢？是要监视 FPS 吗？

以前，我特别喜欢一本叫作《24格》的杂志，它主要介绍的是动画片制作的相关内容。那么，它为啥叫24格呢？这是因为，动画片中1秒钟会用到24张图片，这样肉眼看起来就是流畅的。

FPS 是一秒显示的帧数，也就是一秒内画面变化数量。如果按照动画片来说，动画片的 FPS 就是24，是达不到60满帧的。也就是说，对于动画片来说，24帧时虽然没有60帧时流畅，但也已经是连贯的了，所以并不能说24帧时就算是卡住了。

由此可见，简单地通过监视 FPS 是很难确定是否会出现卡顿问题了，所以我就果断弃了通过监视FPS 来监控卡顿的方案。

那么，我们到底应该使用什么方案来监控卡顿呢？

RunLoop 原理

对于iOS开发来说，监控卡顿就是要去找到主线程上都做了哪些事儿。我们都知道，线程的消息事件是依赖于NSRunLoop 的，所以从NSRunLoop入手，就可以知道主线程上都调用了哪些方法。我们通过监听NSRunLoop 的状态，就能够发现调用方法是否执行时间过长，从而判断出是否会出现卡顿。

所以，我推荐的监控卡顿的方案是：通过监控 RunLoop 的状态来判断是否会出现卡顿。

RunLoop是iOS开发中的一个基础概念，为了帮助你理解并用好这个对象，接下来我会先和你介绍一下它可以做哪些事儿，以及它为什么可以做成这些事儿。

RunLoop这个对象，在 iOS 里由CFRunLoop实现。简单来说，RunLoop 是用来监听输入源，进行调度处理的。这里的输入源可以是输入设备、网络、周期性或者延迟时间、异步回调。RunLoop 会接收两种类型的输入源：一种是来自另一个线程或者来自不同应用的异步消息；另一种是来自预订时间或者重复间隔的同步事件。

RunLoop 的目的是，当有事件要去处理时保持线程忙，当没有事件要处理时让线程进入休眠。所以，了解 RunLoop 原理不光能够运用到监控卡顿上，还可以提高用户的交互体验。通过将那些繁重而不紧急会大量占用 CPU 的任务（比如图片加载），放到空闲的 RunLoop 模式里执行，就可以避开在 UITrackingRunLoopMode 这个 RunLoop 模式时是执行。UITrackingRunLoopMode 是用户进行滚动操作时会切换到的 RunLoop 模式，避免在这个 RunLoop 模式执行繁重的 CPU 任务，就能避免影响用户交互操作上体验。

接下来，我就通过 CFRunLoop 的源码来跟你分享下 RunLoop 的原理吧。

第一步

通知 observers：RunLoop 要开始进入 loop 了。紧接着就进入 loop。代码如下：

```
//通知 observers
if (currentMode->_observerMask & kCFRunLoopEntry )
    __CFRunLoopDoObservers(runloop, currentMode, kCFRunLoopEntry);
//进入 loop
result = __CFRunLoopRun(r1, currentMode, seconds, returnAfterSourceHandled, previousMode);
```

第二步

开启一个 do while 来保活线程。通知 Observers：RunLoop 会触发 Timer 回调、Source0 回调，接着执行加入的 block。代码如下：

```
// 通知 Observers RunLoop 会触发 Timer 回调
if (currentMode->_observerMask & kCFRunLoopBeforeTimers)
    __CFRunLoopDoObservers(runloop, currentMode, kCFRunLoopBeforeTimers);
// 通知 Observers RunLoop 会触发 Source0 回调
if (currentMode->_observerMask & kCFRunLoopBeforeSources)
    __CFRunLoopDoObservers(runloop, currentMode, kCFRunLoopBeforeSources);
// 执行 block
__CFRunLoopDoBlocks(runloop, currentMode);
```

接下来，触发 Source0 回调，如果有 Source1 是 ready 状态的话，就会跳转到 handle_msg去处理消息。代码如下：

```
if (MACH_PORT_NULL != dispatchPort ) {
    Boolean hasMsg = __CFRunLoopServiceMachPort(dispatchPort, &msg)
    if (hasMsg) goto handle_msg;
}
```

第三步

回调触发后，通知 Observers：RunLoop的线程将进入休眠（sleep）状态。代码如下：

```
Boolean poll = sourceHandledThisLoop || (0ULL == timeout_context->termTSR);
if (!poll && (currentMode->_observerMask & kCFRunLoopBeforeWaiting)) {
    __CFRunLoopDoObservers(runloop, currentMode, kCFRunLoopBeforeWaiting);
}
```

第四步

进入休眠后，会等待 mach_port 的消息，以再次唤醒。只有在下面四个事件出现时才会被再次唤醒：

- 基于 port 的 Source 事件；
- Timer 时间到；
- RunLoop 超时；
- 被调用者唤醒。

等待唤醒的代码如下：

```
do {
    __CFRunLoopServiceMachPort(waitSet, &msg, sizeof(msg_buffer), &livePort) {
        // 基于 port 的 Source 事件、调用者唤醒
        if (modeQueuePort != MACH_PORT_NULL && livePort == modeQueuePort) {
            break;
        }
        // Timer 时间到、RunLoop 超时
        if (currentMode->_timerFired) {
            break;
        }
    } while (1);
```

第五步

唤醒时通知 Observer：RunLoop 的线程刚刚被唤醒了。代码如下：

```
if (!poll && (currentMode->_observerMask & kCFRunLoopAfterWaiting))
    __CFRunLoopDoObservers(runloop, currentMode, kCFRunLoopAfterWaiting);
```

第六步

RunLoop 被唤醒后就要开始处理消息了：

- 如果是 Timer 时间到的话，就触发 Timer 的回调；
- 如果是 dispatch 的话，就执行 block；
- 如果是 source1 事件的话，就处理这个事件。

消息执行完后，就执行加到 loop 里的 block。代码如下：

```
handle_msg:
// 如果 Timer 时间到，就触发 Timer 回调
if (msg-is-timer) {
    __CFRunLoopDoTimers(runloop, currentMode, mach_absolute_time())
}
// 如果 dispatch 就执行 block
else if (msg_is_dispatch) {
    __CFRUNLOOP_IS_SERVICING_THE_MAIN_DISPATCH_QUEUE__(msg);
}

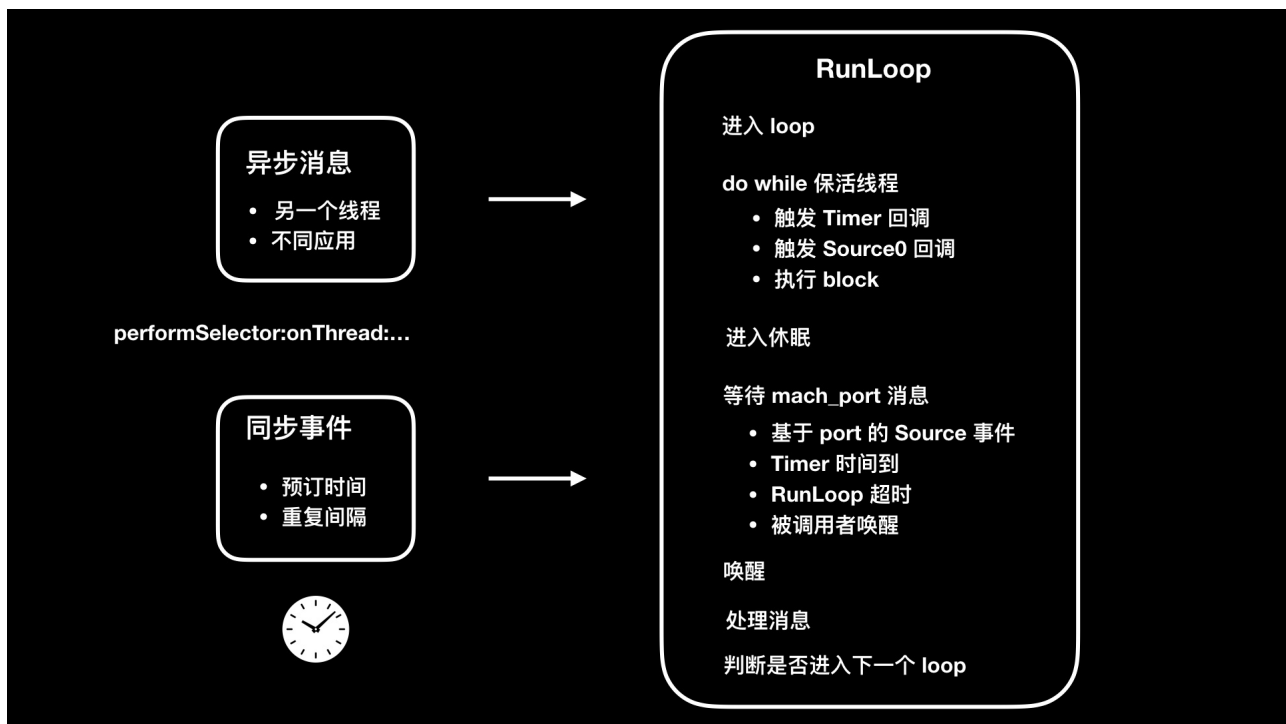
// Source1 事件的话，就处理这个事件
else {
    CFRunLoopSourceRef source1 = __CFRunLoopModeFindSourceForMachPort(runloop, currentMode, livePort);
    sourceHandledThisLoop = __CFRunLoopDoSource1(runloop, currentMode, source1, msg);
    if (sourceHandledThisLoop) {
        mach_msg(reply, MACH_SEND_MSG, reply);
    }
}
```

第七步

根据当前 RunLoop 的状态来判断是否需要走下一个 loop。当被外部强制停止或loop 超时，就不继续下一个 loop 了，否则继续走下一个 loop。代码如下：

```
if (sourceHandledThisLoop && stopAfterHandle) {
    // 事件已处理完
    retVal = kCFRunLoopRunHandledSource;
} else if (timeout) {
    // 超时
    retVal = kCFRunLoopRunTimedOut;
} else if (__CFRunLoopIsStopped(runloop)) {
    // 外部调用者强制停止
    retVal = kCFRunLoopRunStopped;
} else if (__CFRunLoopModeIsEmpty(runloop, currentMode)) {
    // mode 为空，RunLoop 结束
    retVal = kCFRunLoopRunFinished;
}
```

整个 RunLoop 过程，我们可以总结为如下所示的一张图片。



这里只列出了 CFRunLoop 的关键代码，你可以点击[这个链接](#)查看完整代码。

loop 的六个状态

通过对RunLoop原理的分析，我们可以看出在整个过程中，loop的状态包括6个，其代码定义如下：

```
typedef CF_OPTIONS(CFOptionFlags, CFRunLoopActivity) {
    kCFRunLoopEntry , // 进入 loop
    kCFRunLoopBeforeTimers , // 触发 Timer 回调
    kCFRunLoopBeforeSources , // 触发 Source0 回调
    kCFRunLoopBeforeWaiting , // 等待 mach_port 消息
    kCFRunLoopAfterWaiting , // 接收 mach_port 消息
    kCFRunLoopExit , // 退出 loop
    kCFRunLoopAllActivities // loop 所有状态改变
}
```

如果RunLoop的线程，进入睡眠前方法的执行时间过长而导致无法进入睡眠，或者线程唤醒后接收消息时间过长而无法进入下一步的话，就可以认为是线程受阻了。如果这个线程是主线程的话，表现出来的就是出现了卡顿。

所以，如果我们要利用RunLoop原理来监控卡顿的话，就是要关注这两个阶段。RunLoop在进入睡眠之前和唤醒后的两个 loop 状态定义的值，分别是 kCFRunLoopBeforeSources 和 kCFRunLoopAfterWaiting，也就是要触发 Source0 回调和接收 mach_port 消息两个状态。

接下来，我们就一起分析一下，如何对loop的这两个状态进行监听，以及监控的时间值如何设置才合理。

如何检查卡顿？

要想监听 RunLoop，你就首先需要创建一个 CFRunLoopObserverContext 观察者，代码如下：

```
CFRunLoopObserverContext context = {0,(__bridge void*)self,NULL,NULL};
runLoopObserver = CFRunLoopObserverCreate(kCFAllocatorDefault,kCFRunLoopAllActivities,YES,0,&runLoopObserve
```

将创建好的观察者 runLoopObserver 添加到主线程 RunLoop 的 common 模式下观察。然后，创建一个持续子线程专门用来监控主线程的 RunLoop 状态。

一旦发现进入睡眠前的 kCFRunLoopBeforeSources 状态，或者唤醒后的状态 kCFRunLoopAfterWaiting，在设置的时间阈值内一直没有变化，即可判定为卡顿。接下来，我们就可以 dump 出堆栈的信息，从而进一步分析出具体是哪个方法的执行时间过长。

开启一个子线程监控的代码如下：

```
//创建子线程监控
dispatch_async(dispatch_get_global_queue(0, 0), ^{
    //子线程开启一个持续的 loop 用来进行监控
    while (YES) {
        long semaphoreWait = dispatch_semaphore_wait(dispatchSemaphore, dispatch_time(DISPATCH_TIME_NOW, 3
        if (semaphoreWait != 0) {
            if (!runLoopObserver) {
                timeoutCount = 0;
                dispatchSemaphore = 0;
                runLoopActivity = 0;
                return;
            }
            //BeforeSources 和 AfterWaiting 这两个状态能够检测到是否卡顿
            if (runLoopActivity == kCFRunLoopBeforeSources || runLoopActivity == kCFRunLoopAfterWaiting) {
                //将堆栈信息上报服务器的代码放到这里
            } //end activity
        } // end semaphore wait
        timeoutCount = 0;
    } // end while
});
```

代码中的 NSEC_PER_SEC，代表的是触发卡顿的时间阈值，单位是秒。可以看到，我们把这个阈值设置成了3秒。那么，这个3秒的阈值是从何而来呢？这样设置合理吗？

其实，触发卡顿的时间阈值，我们可以根据 WatchDog 机制来设置。WatchDog 在不同状态下设置的不同时间，如下所示：

- 启动（Launch）：20s；
- 恢复（Resume）：10s；
- 挂起（Suspend）：10s；
- 退出（Quit）：6s；
- 后台（Background）：3min（在iOS 7之前，每次申请10min；之后改为每次申请3min，可连续申请，最多申请到10min）。

通过WatchDog 设置的时间，我认为可以把启动的阈值设置为10秒，其他状态则都默认设置为3秒。总的原

则就是，要小于 WatchDog的限制时间。当然了，这个阈值也不用小得太多，原则就是要优先解决用户感知最明显的体验问题。

如何获取卡顿的方法堆栈信息？

子线程监控发现卡顿后，还需要记录当前出现卡顿的方法堆栈信息，并适时推送到服务端供开发者分析，从而解决卡顿问题。那么，在这个过程中，如何获取卡顿的方法堆栈信息呢？

获取堆栈信息的一种方法是直接调用系统函数。这种方法的优点在于，性能消耗小。但是，它只能够获取简单的信息，也没有办法配合 dSYM 来获取具体是哪行代码出了问题，而且能够获取的信息类型也有限。这种方法，因为性能比较好，所以适用于观察大盘统计卡顿情况，而不是想要找到卡顿原因的场景。

直接调用系统函数方法的主要思路是：用 signal 进行错误信息的获取。具体代码如下：

```
static int s_fatal_signals[] = {
    SIGABRT,
    SIGBUS,
    SIGFPE,
    SIGILL,
    SIGSEGV,
    SIGTRAP,
    SIGTERM,
    SIGKILL,
};

static int s_fatal_signal_num = sizeof(s_fatal_signals) / sizeof(s_fatal_signals[0]);

void UncaughtExceptionHandler(NSException *exception) {
    NSArray *exceptionArray = [exception callStackSymbols]; //得到当前调用栈信息
    NSString *exceptionReason = [exception reason];          //非常重要，就是崩溃的原因
    NSString *exceptionName = [exception name];              //异常类型
}

void SignalHandler(int code)
{
    NSLog(@"signal handler = %d",code);
}

void InitCrashReport()
{
    //系统错误信号捕获
    for (int i = 0; i < s_fatal_signal_num; ++i) {
        signal(s_fatal_signals[i], SignalHandler);
    }

    //oc未捕获异常的捕获
    NSSetUncaughtExceptionHandler(&UncaughtExceptionHandler);
}

int main(int argc, char * argv[]) {
    @autoreleasepool {
        InitCrashReport();
        return UIApplicationMain(argc, argv, nil, NSStringFromClass([AppDelegate class]));
    }
}
```

另一种方法是，直接用 [PLCrashReporter](#)这个开源的第三方库来获取堆栈信息。这种方法的特点是，能够定

位到问题代码的具体位置，而且性能消耗也不大。所以，也是我推荐的获取堆栈信息的方法。

具体如何使用 PLCrashReporter 来获取堆栈信息，代码如下所示：

```
// 获取数据
NSData *lagData = [[[PLCrashReporter alloc]
                    initWithConfiguration:[PLCrashReporterConfig alloc] initWithSign

// 转换成 PLCrashReport 对象
PLCrashReport *lagReport = [[[PLCrashReport alloc] initWithData:lagData error:NULL];
// 进行字符串格式化处理
NSString *lagReportString = [PLCrashReportTextFormatter stringValueForCrashReport:lagReport withTextFormat:
//将字符串上传服务器
NSLog(@"lag happen, detail below: \n %@",lagReportString);
```

搜集到卡顿的方法堆栈信息以后，就是由开发者来分析并解决卡顿问题了。

在今天这篇文章中，我们用到的从监控卡顿到收集卡顿问题信息的完整代码，你都可以点击[这个链接](#)查看。

小结

今天我给你介绍了使用 RunLoop 监控卡顿的方案，我还跟你说了下 RunLoop 的原理，希望能够帮助你更好地理解 RunLoop 监控卡顿的方案。

读到这里你可能会想，为什么要将卡顿监控放到线上做呢？其实这样做主要是为了能够更大范围的收集问题，如果仅仅通过线下收集卡顿的话，场景无法被全面覆盖。因为，总有一些卡顿问题，是由于少数用户的数据异常导致的。

而用户反馈的卡顿问题往往都是说在哪个页面卡住了，而具体是执行哪个方法时卡主了，我们是无从得知的。在碰到这样问题时，你一定会感觉手足无措，心中反问一百遍：“我怎么在这个页面不卡，测试也不卡，就你卡”。而且，通过日志我们也很难查出个端倪。这时候，线上监控卡顿的重要性就凸显出来了。

有时，某个问题看似对 App 的影响不大，但如果这个问题在某个版本中爆发出来了就会变得难以收场。所以，你需要对这样的问题进行有预见性的监控，一方面可以早发现、早解决，另一方面在遇到问题时能够快速定位原因，不至于过于被动。要知道，面对问题的响应速度往往是评判基础建设优劣的一个重要的标准。

以上就是我们今天的内容了。接下来，我想请你回顾一下你都碰到过哪些卡顿问题，又是如何解决的呢？

感谢你的收听，欢迎你在评论区给我留言分享你的观点，也欢迎把它分享给更多的朋友一起阅读。

iOS 开发高手课

从原理到实战，带你解决 80% 的开发难题

戴铭

前滴滴出行技术专家



新版升级：点击「👤 请朋友读」，10位好友免费读，邀请订阅更有**现金**奖励。

精选留言：

- Keep-Moving 2019-04-09 06:59:25

PLCrashReporter怎么和卡顿检测结合起来呢？我理解它是收集崩溃信息的，但卡顿又不是一定会崩溃 [8赞]

作者回复2019-04-09 14:31:32

卡顿后收集造成卡顿的原因

- Geek_a03ab1 2019-04-09 10:22:53

为什么监控 kCFRunLoopBeforeSources、kCFRunLoopAfterWaiting这两个事件就能判断出卡顿呢？为什么不是kCFRunLoopBeforeWaiting、kCFRunLoopAfterWaiting这两个事件呢？没想明白，老师能展开说一下吗？ [7赞]

- WeZZard 2019-04-10 11:14:52

單純監控 FPS 確實不行，因為很多卡頓其實是 FPS 變化率的大幅變化（既 FPS 一階導數的波動）所造成的。

比如說點 A 花一秒勻速位移至點 B，秒間平均幀率 60FPS，但是前 59/60 秒才跑了一幀出來，剩下 59 幀都在後 1/60 秒跑出來，那麼用戶必然看到的是點 A 花了 59/60 秒停在 $A + ((B - A)/60)$ 這個位置上，然後突然「飛」到了 B 點上。

所以監控 FPS > 24 的同時還要監控 FPS 變化率的波動。而這個波動的成因其實就是主線程上計算任務的性能衝擊。 [3赞]

- Wechat Team 2019-04-10 20:37:48

文章提到可以『直接调用系统函数』来获取堆栈，但是通过注册那几个信号的方式获取堆栈，一般不是在闪退的情况下才可以触发回调吗？

请老师指点 [1赞]

作者回复2019-04-13 15:19:02

获取堆栈的方法任何时候都可以调的。闪退只是一个时机

- 80后空巢老肥狗 2019-04-09 09:33:17

这个3秒是不是太长了，1秒60帧，每帧16.67ms。runloop会在每次sleep之前去刷新ui，这样的话如果掉了30帧，就是500ms左右，用户的体验就已经下去了，能感觉到卡顿了。

[1赞]

- 鹏哥 2019-04-09 08:04:53

老师，请问下，如果我在用户滑动界面的时候不去加载图片，等停止滑动的时候再去加载图片，这个场景用runloop或者scrollview的代理来实现，和使用sdwebimage异步下载图片有什么区别，这几种方式貌似都没有影响用户滑动体验！[1赞]

作者回复2019-04-13 16:53:37

图片少感觉不出，当图片大、列表长、快速滚动时会有区别

- 戒惜舍得 2019-04-09 07:48:46

大神 可否来一篇TDD oc 和swift 希望都有 [1赞]

- 子嘉 2019-04-11 18:45:06

系统事件是通过source1传递激活app的 然后通过source0分发吗 还是直接通过source0分发？比如点击屏幕 到响应 如果是第二个 那么应该激活不了啊..

- drunkenMouse 2019-04-11 15:15:58

触发Source0回调，如果有Source1是ready状态的话，就会跳转到handle_msg去处理消息。为啥？？拿sources0当备胎？其实是为了消息分流，让sources1也跟着处理吗？

基于port的Source事件？这是什么事件与什么原因。。

- drunkenMouse 2019-04-11 14:44:30

想起以前遇到过的卡顿效果，真的是坑啊。

印象最深的一次卡顿：在TableView的Cell里访问了数据库。

其次是在一个答题系统里，多选单选与判断题，一个Cell从最少两个选项、一个选项三到N行，到最多九个选项。然后使用了TableViewCell的高度自动计算，后来换成了自己手动计算并缓存。

还有Cell的样式太多，后来把尽可能相同的都放到一个Cell里，也就是：一个Cell有七行，如果有四行不显示，那就移除四行，如果需要九行就再加两行。还有就是网络请求同步操作，结果数据过大，然后就做了异步和无数据的默认显示。

其实我就是从这几个方面下手的：页面重叠部分、简单的动画与显示用CALayer自己画、对象释放与创建的次数、高度自动缓存、autorelease使用、static使用部分（如单列对象）、UI刷新前的数据处理部分、页面刷新次数、通知和监听的移除、数据的不合理读取、复杂操作尽量异步、

最后想说下：线程的消息传递依赖于NSRunLoop，函数的调用依赖runtime。感觉二者挺像的。

- kidzss 2019-04-11 11:08:45

老师你好，我之前做了一个卡顿监控，也是使用了PL库，但是发现PL库有Crash的问题。

Out of memory。。。。。

GoodTransport +[PLCrashReportTextFormatter stringValueForCrashReport:withTextFormat:] (PLCrashR

eportTextFormatter.m:)

- Geek 2019-04-11 10:06:34

卡顿的原因讲的可以，但是讲的RunLoop让我听的头大，有几个疑问请教老师，

1，线程与run loop是什么关系。

2，mach_port是什么，第一次进入runloop，也是mach_port触发的吗，

3.进入休眠状态的runloop为什么要等mach_port.

4,系统会有多个runloop吗，不同runloop之间是串行还是并行执行的？

- 庞佳星 2019-04-10 08:34:30

老大，啥时候讲讲，resct native 和iOS的关系。谢谢啦，

作者回复2019-04-13 16:26:54

后面会专门说

- Ant 2019-04-09 19:54:08

今天没作业真是爽歪歪

作者回复2019-04-13 16:28:21

有啊，😁

我想请你回顾一下你都碰到过哪些卡顿问题，又是如何解决的呢？

- 大官人 2019-04-09 08:01:50

一直跟着听，受益匪浅，学到很多也了解很多，如果只听不做，感觉都会了，都知道，一动手，卡住了…
…，动手才会思考，

- Geek_cc73f2 2019-04-09 07:27:22

老师你好，从文章真的受益匪浅，以前只会用第三方的，现在能从本质看出端倪，另外能不能从源头给一些建议呢，比如怎么避免卡顿，哪些操作容易卡顿，然后怎么处理呢，感谢

作者回复2019-04-09 14:30:55

文章最开头部分又说到容易造成卡顿的原因

- 🐼 Ceezy🐼 2019-04-09 02:35:59

对于那个 GCD 的 dispatchPort，我看源码貌似是 GCD 提交给主线程的任务，这一事件源能称为 source1 吗？