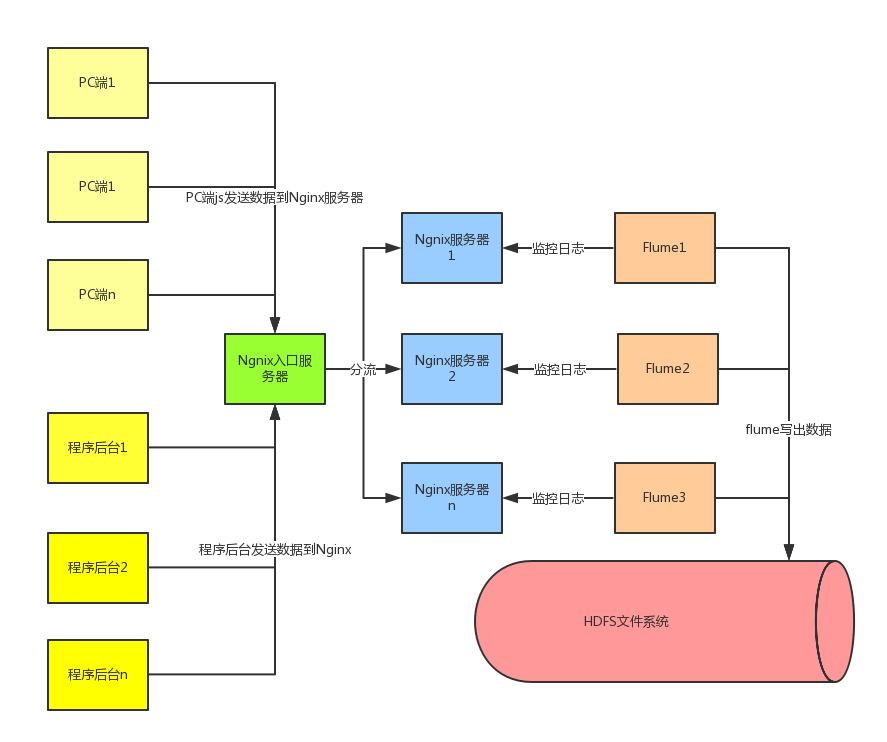
# 日志收集系统

## 收集系统总述

我们需要将用户浏览的数据采集到我们的存储系统(hdfs)中，我们只收集pc端数据和程序后台的部分数据。在pc端我们通过集成js来收集用户浏览行为数据；在程序后台，通过集成java的jar文件来收集我们需要的数据。在这里只考虑java开发环境。js/jar将收集的数据发送到nginx，然后flume监控nginx日志，将数据写入到hdfs中。

## 收集系统结构



## java sdk

对应项目中需要在后台调用的程序，通过Java SDK来实现。

#### JavaSDK设计规则

JavaSDK提供事件触发方法，我们在java sdk中通过一个单独的线程来发送线程数据，这样可以减少对业务系统的延时性。

备注：LinkedBlockingQueue队列，HttpURLConnection

## Js sdk

用户浏览网页，主要通过js SDK，来获取pc端不同的数据

#### JsSDK设计规则

在js sdk中我们需要收集launch、pageview、chargeRequest和eventDuration四种数据，所以我们需要在js中写入四个方法来分别收集这些数据，另外我们还需要提供一些操作cookie和发送数据的的公用方法。