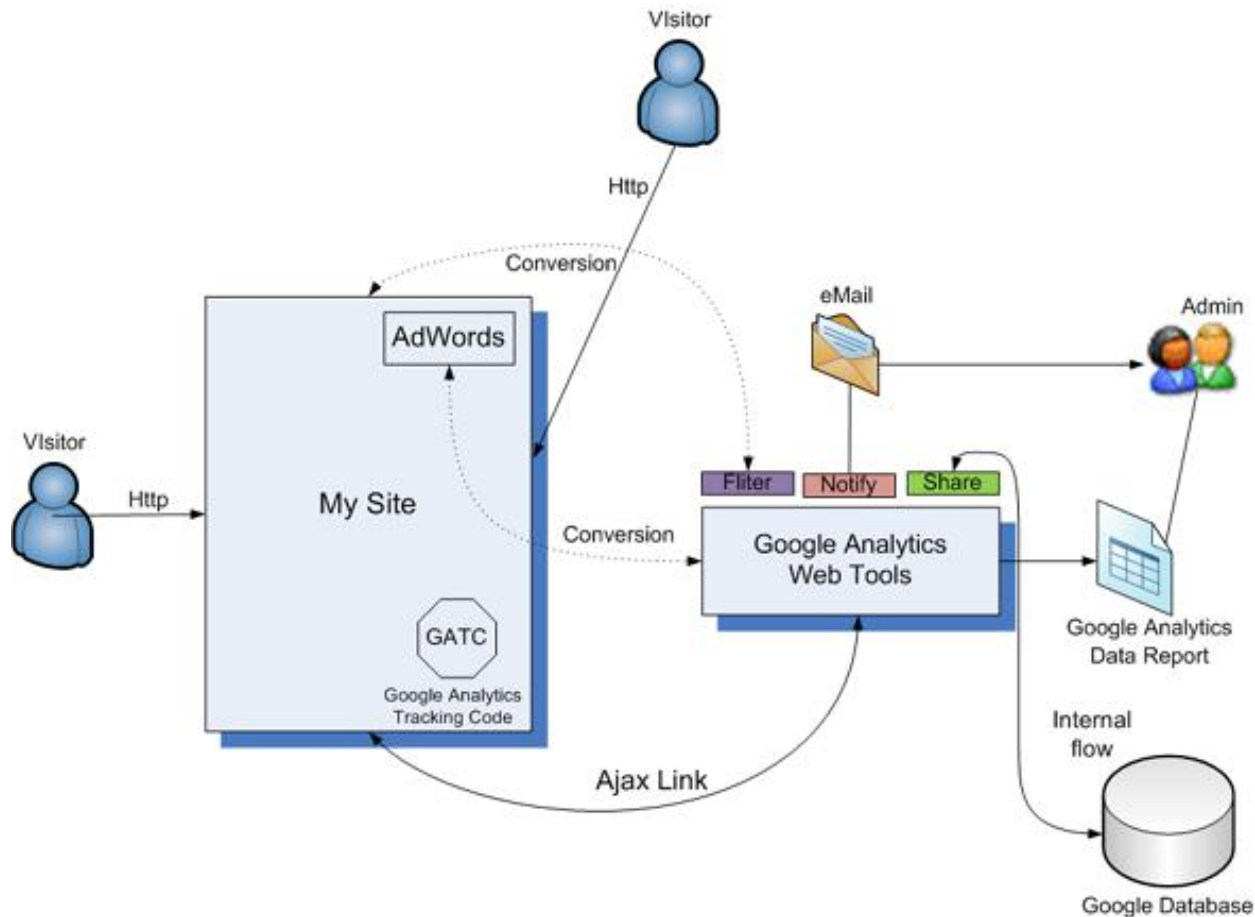


Google Analytics数据统计的原理

GA整体架构

- Google Analytics功能非常强大，只要在网站的页面上加入一段代码，就可以提供的丰富详尽的图表式报告。
- Google Analytics的数据采集功能是采用AJAX技术来实现，整体架构如图所示：

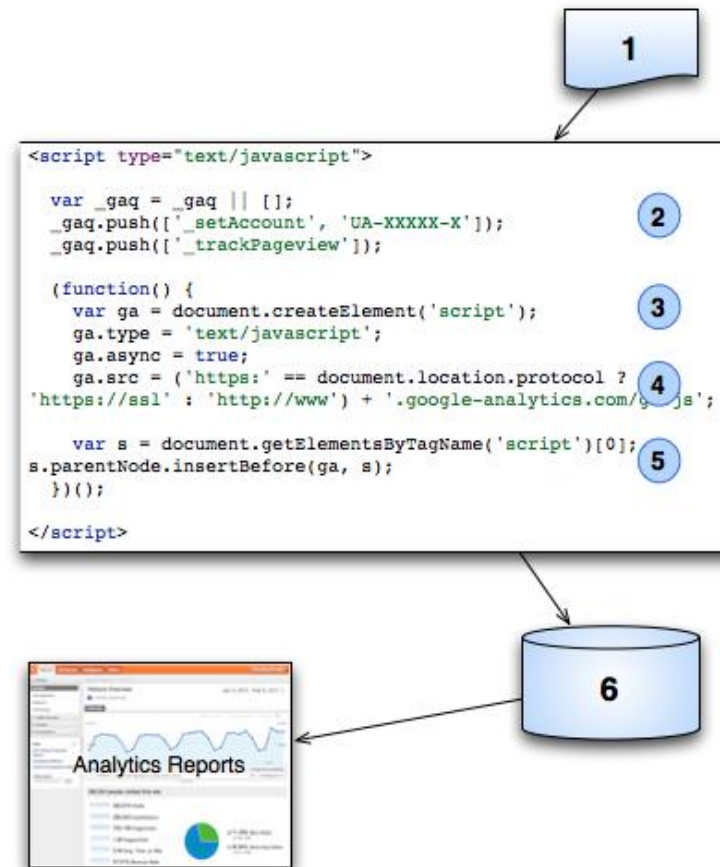


跟踪代码（GATC） 工作方式

- 我们都知道在使用Google Analytics需要在网站上加上跟踪代码（Google Analytics Tracking Code，简称GATC），下面我们就来研究下跟踪代码是如何进行工作的。

谷歌分析跟踪代码(GATC) 检索网页数据如下：

- 1、浏览器请求的网页包含跟踪代码。
- 2、GATC创建并初始化一个对象的属性与网络相关的跟踪ID在代码中。
- 3、GATC执行你自定义任何跟踪方法。
- 4、跟踪代码初始化管理以下信息：跟踪检索查看是否包括广告系列；收集HTTP请求中的各种用户信息的到Google GATC。
- 5、将访问者HTTP请求包含GATC跟踪信息收集到参数名单。



跟踪代码（GATC） 工作过程

- 向Google服务器发送请求
 - 当我们访问带有Google Analytics追踪代码的页面时，页面中的GA追踪代码被执行，然后会向Google服务器发送一个1像素的图片请求。
 - http://www.google-analytics.com/__utm.gif
- GATC将数据作为变量附在请求上
 - 并将所收集到的数据作为请求__utm.gif图片链接的变量一起发送回google服务器。
- Google服务器处理数据
 - 请求然后经过google服务器的处理发布到我们的数据报告里
- GATC的作用
 - Google Analytics统计代码中的Javascript的主要作用是给请求URL赋值

URL格式与含义（一）

- 例如：当访问www.douban.com首页的时候，Google Analytics就会发回这样一条数据给Google服务器。
 - `http://www.google-analytics.com/__utm.gif?utmwv=4.8.9&utmhn=www.douban.com&utmcs=utf-8&utmsr=1280×768&utmcs=32-bit&utmfl=zh-cn&utmje=1&utmfl=10.2%20r152&utmdt=%E8%B1%86%E7%93%A3&utmhid=974895699&utmr=-&utmp=%2F&utmac=UA-7019765-1&utmcc=__utma%3D30149280.845782037.1298725704.1298782111.1299505037.3%3B%2B__utmz%3D30149280.1298725704.1.1.utmcsr%3D(direct)%7Cutmccn%3D(direct)%7Cutmcmd%3D(none)%3B&utmu=qBM`
- 参数含义
 - `http://www.google-analytics.com/__utm.gif` 请求文件地址
 - `?utmwv=4.8.9` Google Analytics追踪代码的版本
 - `&utmhn=www.douban.com` utm.gif的唯一ID编号，防止GIF图像缓存
 - `&utmhn=www.douban.com` 用户访问的主机名（就是我域名）
 - `&utmcs=UTF-8` 用户浏览器语言编码
 - `&utmsr=1280×768` 用户屏幕分辨率
 - `&utmcs=32-bit` 用户的屏幕颜色
 - `&utmfl=zh-cn` 用户浏览器语言设置
 - `&utmje=1` 用户浏览器是否支持JAVA
 - `&utmfl=10.2%20r152` Flash的版本
 - `&utmdt=%E8%B1%86%E7%93%A3` 当前网页的标题（title里的内容），在这里是经过编码的字符串。
 - `&utmhid=974895699` 随机数字
 - `&utmr=-` 流量来源，这里为直接来源。
 - `&utmp=%2F` 当前页面产生的请求，记录你在这个页面的点击行为及其他自定义的返回值。
 - `&utmac=UA-7019765-1` 帐户字符串。在所有请求中显示。
 - `&utmcc=__utma%3D30149280.845782037.1298725704.1298782111.1299505037.3%3B%2B__utmz%3D30149280.1298725704.1.1.utmcsr%3D(direct)%7Cutmccn%3D(direct)%7Cutmcmd%3D(none)%3B` 用户Cookie数据。
 - `&utmu=qBM`

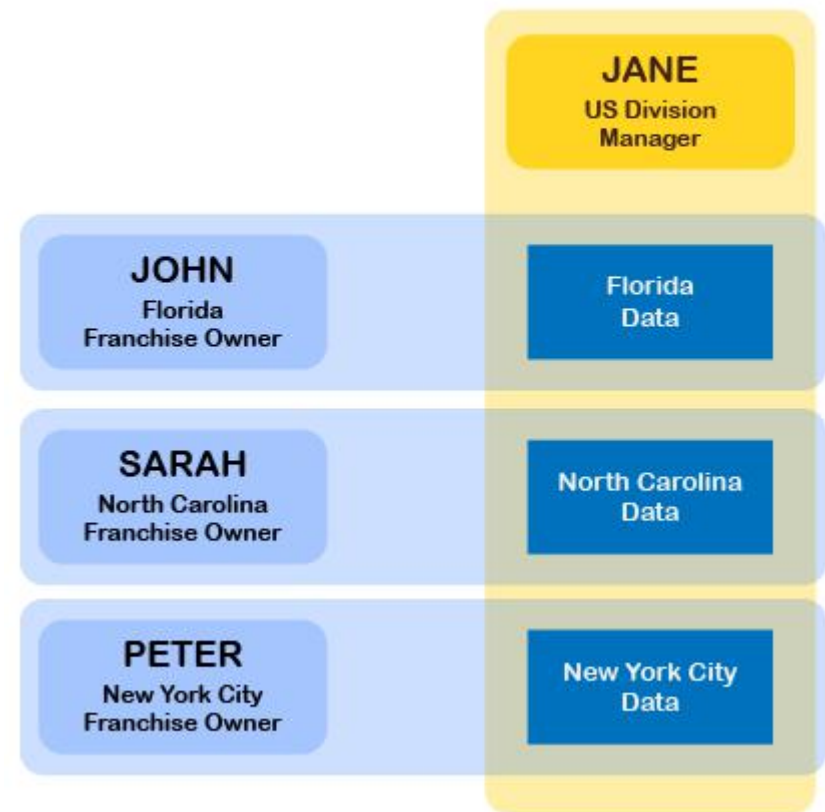
URL格式与含义（二）

- `&utmcc=__utma%3D30149280.845782037.1298725704.1298782111.1299505037.3%3B%2B__utmz%3D30149280.1298725704.1.1.utmcsr%3D(direct)%7Cutmccn%3D(direct)%7Cutmcmd%3D(none)%3B`
 - 以上是Cookie里存储的数据，__utma用来分析唯一用户，访问次数，停留时间，新访或回访等等。__utmz用来记录流量来源。__utmv存储细分用户所定义的值。
 - 以上这些并不是全部的返回数据，只是Google Analytics追踪的基本数据
- 如果你的网站开通了电子商务追踪功能或是你自定义了事件追踪，那么在返回Google服务器的链接中就会看到更多的变量值
 - &utme 事件追踪数据
 - &utmipc 用户购买的产品编号
 - &utmipn 用户购买的产品名称
 - &utmipr 用户购买的产品单价
 - &utmtsp 运费
 - &utmttx 税款

网站配置文件

- 使用Google Analytics分析时需要你创建 Google Analytics的用户“网站配置文件”
- 创建“网站配置文件”的目的是为了方便网站多个管理员从各种对网站不同的关注角度对网站的访问进行分析

Google Analytics Profile Architecture

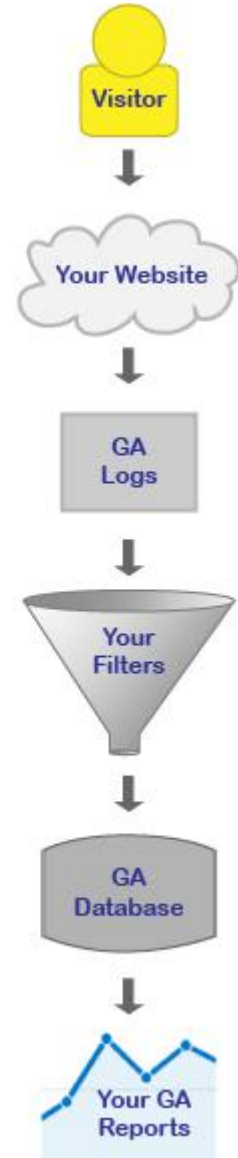


GA工作过程

- 访问者访问网站
- 触发到 GATC 代码，根据GATC采集用户信息
- 将信息放入Google Analytics 日志
- 通过你创建的过滤器对需要进行过滤的数据进行筛选
- 将采用和过滤后的信息放入Google Analytics 的数据库中
- 管理者通过Google Analytics 的界面查看分析数据

How Google Analytics Works

(*Basically)



GA基础指标

指标	含义
访问数	网站的访问次数
浏览量	查看过的页面总数。单面的重复查看次数也计算在内。
每次访问页数	在每次访问网站期间平均查看的页数。单页的重复查看页数也计算在内。
平均网站停留时间	网站的平均访问时长
跳出率	单页访问次数（即访问者从进入页面离开网站的访问次数）所占的百分比
唯一身份访问者	网站的唯一身份访问者数量
新访问次数百分比	以前从未访问过网站的用户所占的访问次数百分比