

安启华：数据云URL过滤技术思科认证 PDF转换可能丢失图片或格式，建议阅读原文

https://www.100test.com/kao_ti2020/644/2021_2022__E5_AE_89_E5_90_AF_E5_8D_8E_EF_c101_644087.htm 当今，互联网给我们带来了生活方式、生产方式上很多的便利与快捷。庞大的网络信息资源，使人们通过浏览器就能便捷的获取到生活或生产过程中所需要的信息。但是，过分或不适当的网络资源访问不仅给企业带来了生产力和网络带宽的损失，还严重威胁着企业的网络安全架构和信息系统，甚至网络上的不适当或非法内容还极大危害着企业员工个人的身心健康甚至给企业带来法律问题。面临以上种种问题，该如何解决？全部关闭网络的大门不符合现代人类文明的发展与进步，也不利于合理商业业务的开展。单纯依靠使用者自我约束的网络使用制度也已经无法执行。于是，诞生了专门针对URL进行过滤的技术，以此达到对员工上网浏览内容的控制与管理。下面我们将对URL过滤技术的进化做一个简单的介绍。20世纪90年代中期，URL过滤解决方案依靠企业内部IT人员人工建立、更新与编辑站点黑白名单。这一做法的缺点是，所有分类由一个或少数几个人自由决定，对于这种资源密集型而且缺乏客观性的站点分类方法，不仅会使许多被认可的网站被封锁或被禁止的网站允许通过，而且随着web站点的快速增长与相关技术的日益复杂，这种方案很难实现客观、细粒度的URL分类，显然不能成为企业或单位有效的URL过滤方案。20世纪90年代末，出现了专门对URL进行收集、分类的厂商。URL过滤技术开始采用本地数据库分类引擎。URL及其内容在根据预先定义的类别下通过分类引擎进行相应的关键字

查找分析与分类(如赌博、色情及网上购物等)，分类好的URL存储在一个集中的主数据库中，然后通过更新复制一份副本移交到客户本地数据库中。这种URL过滤方案的缺点是，随着网页数量的激增，由于一刀切的关键字分类技术和本地分类数据库的限制，无法实现更高、更准确的覆盖率和更广泛的URL分类。2000年初，URL过滤解决方案试图采用启发式内容分析的方法，这种动态的分类技术，通过智能分析网站标题和网页html主体中相关内容的概率来确定URL类别。从理论上讲，相比前两类URL分类方案，这是一种很好的分类方法，然而在实际中它本身却存在问题，很多基于启发式的Web内容分析结果没有相关的配套技术实时地发送给终端用户，而且采用的仍然是传统的本地数据库进行存储。但是当今web2.0时代，web数据是一个不定数据且日益多样化的集合体，而每个用户的需求却独特且具体，基于本地URL数据库的过滤技术，只能过滤存储本地用户需要的数据，不能存储所有相关及最新数据，以执行快速和准确的监测，因此这种传统的URL过滤技术也无法应对高度复杂且快速发展壮大的web2.0网络。据Google调查，互联网上的网页数量以每天一亿的数量急速增长。以上三种URL分类方法已经不能够准确有效的收集、分类所有的URL类别。数据存储和处理要求也已经远远超出了本地数据库能力。于是在2009年，业界出现了数据云的URL过滤技术，这类厂商有CommTouch、Anchiva等。这种数据云URL过滤机制，基于云技术的URL收集、分类处理及发放策略，并不依赖于本地数据库有限的资源进行分析与检测，也不依赖于数据库更新最新的URL分类，利用的是专门的分类服务器群，根据实际网络的使用与

普及方式对网页内容及语义进行全面分析后的分类。与传统的云不同的是，真正做到了云的客户端自动主动地去云的服务端获取所需的数据，而不是单纯的基于云服务端的定时推送更新方法。下面我们将借助Anchiva的数据云URL过滤机制对数据云URL过滤技术进行简单的介绍。 Anchiva(安启华)数据云的URL分析过滤技术由两部分组成：部署在企业网络边界处的Anchiva web安全网关(SWG)和Anchiva基于云的URL分类中心。网关设备和URL分类中心实时通信获取最新的URL分类。不仅打破了传统本地数据库的限制，而且与其他基于云的技术不同的是Anchiva web安全网关中具有URL缓存技术，这个本地的缓存为每一个独立的用户存储最为相关的URL，这些设备本地的URL类别，有效地确保了URL匹配的最佳性能，并且采用的是一种本地缓存自动学习的机制，随着企业用户使用时间的增长，这个本地缓存的URL库会更加贴近每一个用户的实际需求，将能给客户更加精准的URL过滤。 100Test 下载频道开通，各类考试题目直接下载。详细请访问 www.100test.com