|  |
| --- |
|  |
| Programming Guide for eyoung |
|  |
|  |
| **eyoung.father@gmail.com** |
| **2014/3/21** |

|  |
| --- |
|  |

目录

[1. eyoung介绍 2](#_Toc383180164)

[1.1. 为什么设计eyoung 2](#_Toc383180165)

[1.2. IPS引擎的设计 4](#_Toc383180166)

[1.2.1. 模块结构 4](#_Toc383180167)

[1.2.2. 规则加载流程 5](#_Toc383180168)

[1.3. 协议分析器的设计 7](#_Toc383180169)

[1.3.1. 协议分析器及攻击检测 7](#_Toc383180170)

[2. IPS引擎 8](#_Toc383180171)

[2.1. 主要数据结构 8](#_Toc383180172)

[2.1.1. 源代码 8](#_Toc383180173)

[2.1.2. 说明 9](#_Toc383180174)

[2.2. External API 11](#_Toc383180175)

[2.2.1. 源代码 11](#_Toc383180176)

[2.2.2. 说明 12](#_Toc383180177)

[2.3. Signature API 18](#_Toc383180178)

[2.3.1. 源代码 18](#_Toc383180179)

[2.3.2. 说明 19](#_Toc383180180)

[2.4. 二进制规则库 19](#_Toc383180181)

[2.4.1. 二进制规则 19](#_Toc383180182)

[2.4.2. Source Code 20](#_Toc383180183)

[2.4.3. 说明 21](#_Toc383180184)

[2.5. 示例 22](#_Toc383180185)

[2.5.1. demo\_http\_xss.c 22](#_Toc383180186)

[2.5.2. http\_decode.c 25](#_Toc383180187)

[2.5.3. export\_test.c 28](#_Toc383180188)

[3. 协议分析器 29](#_Toc383180189)

[3.1. Stream-mode协议词法分析器 29](#_Toc383180190)

[3.1.1. 设计 29](#_Toc383180193)

[3.1.2. 示例 30](#_Toc383180194)

[3.2. Push-Mode协议语法分析器 33](#_Toc383180195)

[3.2.1. 设计 33](#_Toc383180196)

[3.2.2. 示例 33](#_Toc383180197)

[3.3. 事件 35](#_Toc383180198)

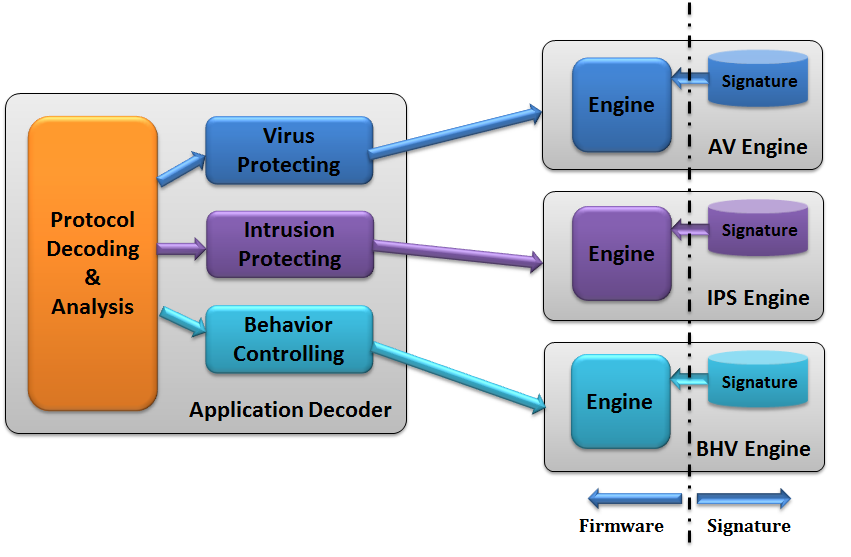
[3.3.1. 事件的类型 35](#_Toc383180199)

[3.3.2. 事件的提交 36](#_Toc383180200)

# eyoung介绍

## 为什么设计eyoung

下图是常见应用层处理架构，例如开源IPS系统Snort



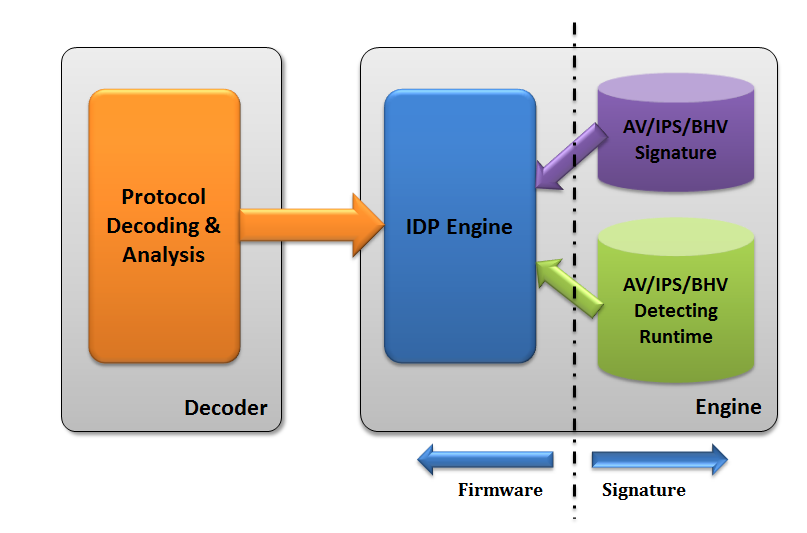
图中的虚线是Firmware（软件）与Signature（规则）的分界线。Firmware的特点是相对稳定不会频繁变化，其变换周期可能是以月计算的；Signature是通常意义的IPS规则、病毒库等会频繁更新的内容，他们的更新频率通常以天或小时计算。在这样的架构中，通常都有一个通用化的应用层协议分析器（Layer7 ApplicationDecoder）用来完成应用层协议分析的工作。此外，这样的系统中还会有一个定制化、可升级的IPS规则库。这里，IPS Engine部分是由Engine Firmware和外部IPS规则两部分组成。当前以Snort为代表的开源IPS系统，他们主要的IPS规则都是以字符串、正则表达式匹配为核心描述手段的。这种描述手段的特点是简单、适合硬件加速，但是它的弱点也是非常明显，那就是描述能力的不足：

首先，许多攻击是无法使用正则表达式描述的，最典型的就是缓冲区溢出攻击。很明显，使用正则表达式无法涵盖所有的机器指令的组合，后果是误报和漏报都非常严重。另外，DDoS也是一种典型的无法用正则表达式描述的攻击。

其次，正则表达式描述能力不够精确、全面，例如SQL注入攻击。正则表达式能够部分描述一个小的SQL片段，但是对于上下文无关特点的SQL语言，使用正则表达式理论上就是无法等价描述的，这必然会带来误报和漏报。

最后，正则表达式无法描述业务的流程。Layer7的数据变化很快，例如网页结构、表单结构、参数编码等等，这些都是跟具体的业务相关联的，而通用性的攻击防护规则，无法描述这些业务层的变化。为了能够描述业务的特点，这种架构下只能将业务数据的解析放到Layer7 Application Decoder中——这就带来了Firmware一侧设计的冗余和不通用，性能也会受拖累。在快速变化的移动互联网领域，Firmware的变化会严重滞后于业务的变化，因为Firmware的开发更复杂、更困难。

为了解决上述攻击可描述性、业务可描述性问题，eyoung将传统的Layer7处理架构调整为：



eyoung系统大致包含两个大的模块，协议分析器（Decoder）和IPS引擎（Engine）。其中，Decoder是一个单纯的Layer7协议分析器，它不理解业务相关的语义，仅仅根据RFC的定义完成Layer7协议数据的解析。Engine是一个攻击检测引擎，它负责加载eyoung规则，并根据Decoder的协议分析的结果探测攻击。eyoung IPS引擎的核心特性是规则的可编程性：

首先，eyoung规则本身可编程。eyoung支持C99兼容的语法，即规则开发这可以使用C99兼容的语法将复杂的业务描述代码或者攻击检测代码嵌入到攻击描述规则中。

其次，eyoung IPS引擎内嵌了JIT实时编译器，可以将内嵌的代码编译成机器指令，并动态连接到当前的进程镜像，从而极大地提高执行效率和检测效率。

最后，eyoung 规则支持二进制库的动态导入，例如将WebMail解析这样复杂的业务描述程序，以动态链接库.so的形式动态加载到eyoung IPS引擎中，并且可以单独升级。这样做一方面简化了Layer7 Decoder的实现，真正实现了协议个业务的分离，另一方面扩展了Engine一侧的协议描述能力，加快了对Layer7业务变化的响应速度。

eyoung的设计思路，实际上将Firmware的部分尽可能地固定下来，将可能会快速变化的部分都推送到Signature一侧，以实现动态升级和实时加载。

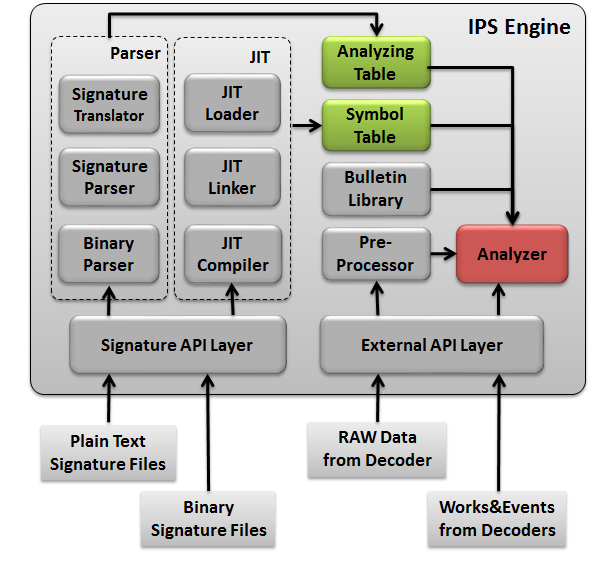
从软件模块的角度看待eyoung系统，eyoung主要分为IPS引擎和协议分析器两大功能模块，这两个模块间是严格地单向调用的关系，即协议分析器调用IPS引擎的相关功能，它们之间有清晰的API函数接口，这个API函数集被称为External API Layer。

从运行阶段的角度看待eyoung系统，eyoung主要分为静态加载阶段和动态检测阶段：

* 静态加载阶段，主要是由eyoung IPS引擎加载相关的IPS规则文件、二进制规则库等静态内容，相当于一台计算机加电启动时操作系统的加载。eyoung IPS规则支持高度的可编程性，IPS规则中调用的eyoung IPS引擎的相关功能是通过IPS引擎提供的API函数接口实现，eyoung中的起这个作用的API函数集被称为Signature API Layer。
* 动态检测阶段，主要进行攻击检测，即协议分析器接受网络数据并进行协议分析，并把协议分析的结果以结构化的方式，使用External API Layer相关的API提交给eyoung IPS引擎进行攻击检查。

## IPS引擎的设计

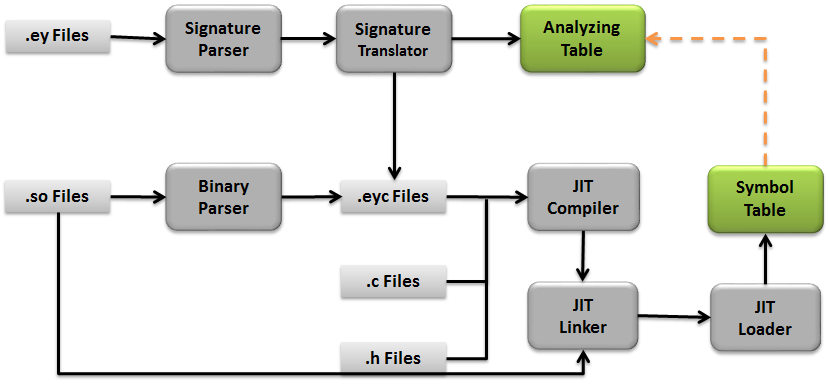
### 模块结构



IPS引擎从结构上主要分为：

1. IPS规则解析器（Parser），它的作用主要负责解析规则文件，既包括Plain Text的规则，也包括二进制规则库。它主要完成：
   1. 规则书写的合法性检查，探测并报告语法错误
   2. 对规则文件.ey进行预处理，转换成中间代码文件.eyc。eyoung IPS规则中间代码是采用C99兼容格式的C代码源文件。
   3. 对规则文件中表达的信息，例如事件的定义、注册的函数接口、规则的表示，转换成数据结构，保存在符号表（Symbol Table）和分析表（Analyzing Table）中。
2. IPS规则运行时编译器（JIT），它的作用是把经过解析器预处理之后的中间代码文件进行运行时刻编译、链接和加载，将全局函数入口地址、全局变量地址保存到符号表中。
3. IPS引擎函数库（Bulletin Library），主要提供基础的处理功能，包括内存管理、基础数据结构封装等等。
4. 预处理器（Pre-Processor）。预处理器是一个固定的接口层，通过对其不同的实现，可以将异构、异步的IPS引擎和规则集成到eyoung系统中。预处理器通常不对协议分析器的输出进行处理，它的处理对象通常是未进行协议分析的原始数据（RAW Data）。
5. 分析器（Analyzer）。分析器负责对协议分析器的分析结果进行攻击检测，是运行时刻整个eyoung IPS引擎的核心。

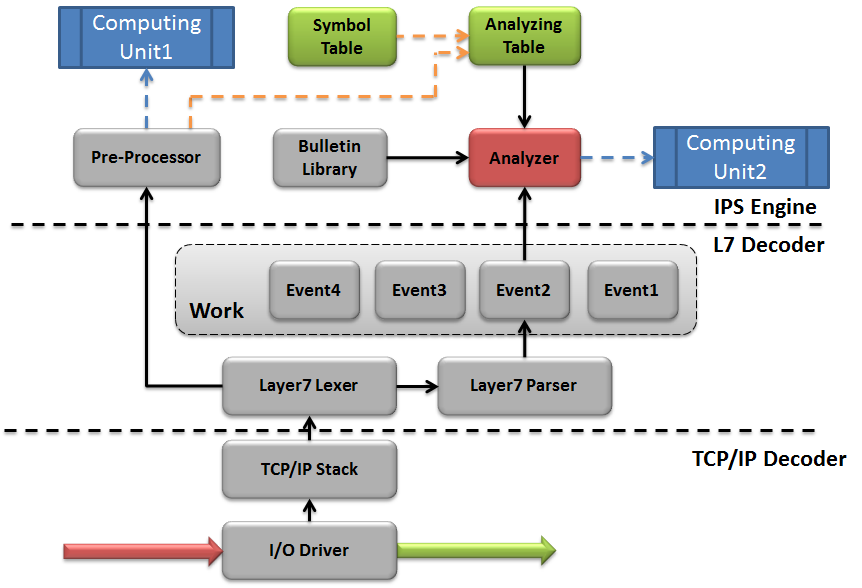
### 规则加载流程



1. 对于Plain Text的规则文件：
   1. 规则解析器（Signature Parser）进行规则合法性检查
   2. 规则解释器（Signature Translator）进行规则翻译，将规则翻译成中间代码.eyc文件（前文说过中间代码是采用了C99的兼容格式）；规则解释器还把规则文件中的规则语义，转换成分析表中的表项。
   3. 规则解释器还会将规则文件中记录的file、work、event等元素的init、finit函数入口，记录到分析表中。相关的详细说明在本文档的2.3中介绍。
2. 对于二进制规则.so文件：
   1. 二进制文件解析器（Binary Parser）负责读取动态链接库.so文件中的.eyoung\_ident和.eyoung\_type段（如果存在的话）的内容，将其中的内容也保存在中间代码文件.eyc中。
   2. 二进制文件解析器解析.eyoung\_init段内容，并执行.eyoung\_init段中包含的初始化入口函数，完成二级制规则的初始化工作。
   3. 二进制文件解析器还负责解析.eyoung\_finit段的内容，并记录下段中保存的finit析构函数，用于在二级制规则库卸载时清理资源。
3. 经过1和2，所有的规则文件都转换成中间代码文件.eyc，而后实时编译器JIT接管规则加载的操作：
   1. 实时编译器（JIT Compiler）负责编译中间代码.eyc文件以及其它附加的.h头文件、.c源文件，在内存中完成文件单元编译。
   2. 实时连接器（JIT Linker）负责将JIT Compiler编译好的文件单元进行链接。通过JIT Linker的处理，不同规则文件中的全局符号可以完成重定位和相互引用。
   3. 实时加载器（JIT Loader）将链接后的代码加载到当前程序的运行镜像，同时将本进程中的全局符号与规则代码动态链接。这个过程可以使得规则文件中的代码直接调用当前进程中开放的函数接口和全局变量，增加Firmware和Signature的交互，有效提升规则的描述能力。例如，某些复杂耗时的协议分析功能可能仅仅是在特定条件下需要完成，没有必要无差别地调用。此时，可以把这种协议分析的功能以全局函数的形式实现在协议分析器中，但协议分析器并不调用，而是在IPS规则中在一定条件下调用该功能的函数。
   4. 加载后的全局符号及其地址都保存在符号表（Symbol Table）中。

## 协议分析器的设计

### 协议分析器及攻击检测

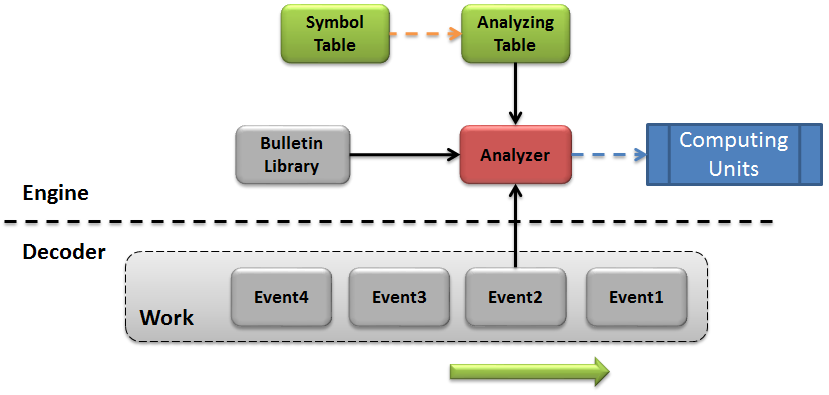


1. TCP/IP Decoder是低层协议栈，负责链路层、网络层、传输层数据的解析、重组和代理等工作。eyoung系统中不包括此内容，他们的通常是操作系统核心的网络部分功能，或是类似Intel DPDK等用户态数据转发架构。
2. 应用层协议分析器（Layer7 Decoder），用来分析应用层数据。分两部分：
   1. 应用层数据词法分析器（Layer7 Lexer），同低层协议栈交互并维护应用层数据缓冲区。它将应用层数据按照词法规则分解成一系列应用层“词元”，并将它们提交给Layer7 Parser用于分析；同时，应用层数据词法分析器还将低层获得的数据，通过eyoung IPS引擎的External API Layer提交给eyoung IPS引擎的Pre-Processor进行预处理。应用层数据词法分析器采用Stream-Mode进行解析，具体方法本文3.1节详细介绍。
   2. 应用层数据语法分析器（Layer7 Parser），将接收到的应用层数据“词元”序列，按照应用层语法进行“自底向上”的语法分析，从逻辑上对应用层数据进行解释，并将语法分析的结果以work、event等抽象数据形式，提交给eyoung IPS引擎。Work、event等概念在本文2.1节、3.3节中详细介绍。应用层数据语法分析器采用Push-Mode方式，具体方法在本文3.2一节介绍。
3. 预处理器（Pre-Processor），它可以兼容第三方检测引擎和规则格式。eyoung期望预处理器可以异步地在其它的计算单元上执行，例如其它CPU、ASIC芯片或GPU等等。预处理器的处理结果，将被记录在分析表（Analyzing Table）中，并被分析器（Analyzer）利用。
4. 分析器（Analyzer），它是eyoung IPS引擎在检测阶段的核心部件。它结合eyoung IPS规则、分析表的数据、预处理器等信息，对应用层协议分析器提交的work、event等对象进行攻击检测。

# IPS引擎

## 主要数据结构

eyoung IPS引擎的核心数据结构定义在include/libengine\_type.h头文件中。在概念上，该文件中定义的数据结构主要有三个：engine\_t、engine\_work\_t和engine\_work\_event\_t。其中，engine\_t是一个静态的概念，表示eyoung IPS引擎本身；engine\_work\_t和engine\_work\_event\_t都属于动态概念，用于表示被检测的对象。



### 源代码

#include "ey\_queue.h"

**typedef** **void\*** engine\_t**;**

**typedef** **int** **(\***file\_init\_handle**)(**engine\_t eng**);**

**typedef** **int** **(\***file\_finit\_handle**)(**engine\_t eng**);**

**typedef** **struct** engine\_work

**{**

TAILQ\_ENTRY**(**engine\_work**)** link**;**

engine\_t engine**;**

**void** **\***priv\_data**;** */\*for libengine itself\*/*

**void** **\***predefined**;** */\*for protocol parser\*/*

**void** **\***user\_defined**;** */\*for signature writer\*/*

**}**engine\_work\_t**;**

**typedef** TAILQ\_HEAD**(**engine\_work\_list**,** engine\_work**)** engine\_work\_list\_t**;**

**typedef** **int** **(\***work\_init\_handle**)(**engine\_work\_t **\***work**);**

**typedef** **int** **(\***work\_finit\_handle**)(**engine\_work\_t **\***work**);**

**typedef** **struct** engine\_work\_event

**{**

engine\_work\_t **\***work**;**

**void** **\***event**;**

**void** **\***predefined**;** */\*for protocol parser\*/*

**void** **\***user\_defined**;** */\*for signature writer\*/*

engine\_action\_t **\***action**;** */\*OUTPUT\*/*

**}**engine\_work\_event\_t**;**

**typedef** **int** **(\***event\_init\_handle**)(**engine\_work\_event\_t **\***work\_event**);**

**typedef** **int** **(\***event\_finit\_handle**)(**engine\_work\_event\_t **\***work\_event**);**

**typedef** **int** **(\***event\_preprocess\_handle**)(**engine\_work\_event\_t

**\***work\_event**);**

**typedef** **int** **(\***event\_condition\_handle**)(**engine\_work\_t **\***engine\_work**,**

engine\_work\_event\_t **\***work\_event**);**

**typedef** **int** **(\***event\_action\_handle**)(**engine\_work\_t **\***engine\_work**,**

engine\_work\_event\_t **\***work\_event**);**

### 说明

1. engine\_t

eyoung IPS引擎实例对外的类型抽象，eyoung IPS引擎之外的模块需要操作IPS 引擎时，例如加载规则文件、卸载规则文件、注册事件等操作时，相关API都需要提供engine\_t实例。

1. file\_init\_handle、file\_finit\_handle

函数指针类型，其对应的函数需要使用2.3.1中定义的ey\_add\_file\_init、ey\_add\_file\_finit两个宏向engine中注册，也可以通过规则文件中的%file-init和%file-finit两个关键字在规则中向engine中注册。

eyoung IPS引擎可以加载多个.ey规则文件，每一个规则文件中都可以定义至多一个file\_init\_handle和至多一个file\_finit\_handle函数。对于file\_init\_handle函数，其调用的时机是在所有规则文件都加载完毕，由eyoung IPS引擎按照规则文件加载顺序依次调用各个文件中的init函数，完成文件级的初始化工作。而对于file\_finit\_handle函数，是在eyoung IPS引擎要卸载规则时，按照当初的规则加载顺序自动调用相关的finit函数，完成文件级的资源释放工作。

1. engine\_work\_t

简称为work，是一个动态概念的类型抽象。work是由一系列event组成，用来抽象表示一个IPS检测过程，例如一条TCP连接可以被视为一个work。

* work通过link成员组成一个链表，该链表记录在engine\_t对象中
* engine成员用来指向本work隶属的engine\_t对象
* priv\_data成员，用来指向work在engine中的数据对象，该成员不允许外部修改其内容
* predefined成员，指向一块由协议分析器开发者定义的内存对象。因为work对象是由协议分析开发者创建，是一个“连接”级概念，所以如果协议开发人员期望记录一部分“连接”级的数据（例如连接上的计数信息），就可以通过predefined成员来实现。predefined具体的格式定义由协议开发人员自行解释，engine对象不能理解其结构。
* user\_defined成员，指向一块由规则开发者定义的内存对象。由于应用层协议分析开发和检测规则开发是分离的，所以在协议解析开发阶段不可能预期所有的检测需求。这就需要有一部分“连接”级的信息需要在规则开发阶段，由规则开发人员自行定义。user\_defined就是为了满足这样的需要，其格式的定义、数据对象的操作都是由规则开发人员自己完成，eyoung IPS引擎不能理解其结构（只关心它分配、释放的入口地址）。

1. work\_init\_handle、work\_finit\_handle

是两个函数指针类型，其对应的函数实例在work创建、销毁的时候被eyoung IPS引擎对象调用。其主要用来分配、释放predefined和user\_defined的数据。这些函数的实例分别由协议分析开发人员和规则开发人员定义，并通过2.3.1中的相关API注册到eyoung IPS引擎对象中。具体地，work中predefined成员的创建和销毁功能的函数实例，由协议开发人员通过使用两个宏ey\_set\_predefine\_work\_init和ey\_set\_predefine\_work\_finit注册到eyoung IPS引擎对象中；work中user\_defined成员的创建和销毁功能的函数实例，由规则开发人员通过ey\_set\_userdefine\_work\_init、ey\_set\_userdefine\_work\_finit两个宏向eyoung IPS引擎对象注册，或者在规则文件中通过使用%work-init和%work-finit两个关键字向eyoung IPS引擎对象注册。

1. engine\_work\_event\_t

简称为event，该类型是对检测元素的数据类型抽象。同work一样，event也是一个动态的概念，work是由一系列event组成。一个event既可以是一个实实在在的应用层协议成员，例如HTTP URI，也可以是一个抽象的事件，例如一次非法的DNS请求。event的创建和销毁都由协议分析开发人员决定，目前不支持规则开发人员在规则中创建和销毁event。

* work成员，用来指向event所隶属的work对象。
* event成员，用来指向event定义时的类型对象。该类型对象，属于一个静态的概念，是在规则加载时由规则中的%event关键字创建的。该成员的解释、使用都由engine对象完成，event之外不能修改该成员。
* predefined成员，指向一块由协议分析器开发者定义的内存对象。如同work对象中的predefined成员一样，这里predefined成员的具体定义也是由协议开发人员自行解释并操作，eyoung IPS引擎对象不能理解其结构。
* user\_defined成员，指向一块由规则开发者定义的内存对象。同work对象中user\_defined成员一样，该成员的设立是为了对协议分析的结果进行补充。例如，对于HTTP URI，在经过协议分析之后被简单解释成字符串之后，规则中可能需要对URI中某一个特定的参数进行过滤。URI中的参数可以被解释成一个“键-值”对数组，如果一开始在协议分析阶段就完成“键-值”对的分析，很可能会无谓地浪费性能。此时，可以在规则匹配阶段，针对特殊的URI单独做参数的解析。定制化的解析功能由规则开发人员提供，解析的结果保存在user\_defined成员中。

1. event\_init\_handle、event\_finit\_handle、event\_preprocess\_handle

三个函数指针类型，其对应的函数实例在event创建、销毁、检测的时候被eyoung IPS引擎对象调用。其主要用来分配、释放、处理event中的predefined和user\_defined数据。这些函数的实例分别由协议分析开发人员和规则开发人员定义，并通过**错误!未找到引用源。**中的相关API注册到eyoung IPS引擎中。具体地，event对象中predefined成员的创建、销毁、处理对应的函数实例，由应用层协议分析器的开发人员通过使用2.3.1中定义的ey\_set\_predefine\_event\_init，ey\_set\_predefine\_event\_finit，ey\_set\_predefine\_event\_preprocess三个宏注册到engine对象中；event中user\_defined成员的创建、销毁、处理功能的函数实例，由规则开发人员通过使用三个宏ey\_set\_userdefine\_work\_init、ey\_set\_userdefine\_work\_finit、ey\_set\_userdefine\_event\_preprocess向eyoung IPS引擎对象注册，或在规则文件中通过%event-init、%event-finit、%event-preproces三个关键字向eyoung IPS引擎对象注册。

1. event\_condition\_handle、event\_action\_handle

这是两个由engine内部使用的函数指针类型，分别对应于规则中condition和action两部分转换之后的函数类型。可以参考《Writing Signatures for eyoung》中3.2一节的叙述。

## External API

定义在include/libengine\_function.h头文件中。

### 源代码

#include "libengine\_type.h"

#include "libengine\_export.h"

**extern** engine\_t ey\_engine\_create**(const** **char** **\***name**);**

**extern** **void** ey\_engine\_destroy**(**engine\_t engine**);**

**extern** **int** ey\_engine\_load**(**engine\_t engine**,**

**char** **\***files**[],** **int** files\_num**);**

**extern** **int** ey\_engine\_find\_event**(**engine\_t engine**,**

**const** **char** **\***event\_name**);**

**extern** engine\_work\_t **\***ey\_engine\_work\_create**(**engine\_t engine**);**

**extern** **void** ey\_engine\_work\_destroy**(**engine\_work\_t **\***work**);**

**extern** engine\_work\_event\_t **\***ey\_engine\_work\_create\_event**(**

engine\_work\_t **\***work**,**

**unsigned** **long** event\_id**,**

engine\_action\_t **\***action**);**

**extern** **int** ey\_engine\_work\_detect\_data**(**engine\_work\_t **\***work**,**

**const** **char** **\***data**,**

size\_t data\_len**,**

**int** from\_client**);**

**extern** **int** ey\_engine\_work\_detect\_event**(**

engine\_work\_event\_t **\***event**,**

**void** **\***predefined**);**

**extern** **void** ey\_engine\_work\_destroy\_event**(**engine\_work\_event\_t **\***event**);**

**extern** **int** debug\_engine\_parser**;**

**extern** **int** debug\_engine\_lexier**;**

**extern** **int** debug\_engine\_init**;**

**extern** **int** debug\_engine\_compiler**;**

**extern** **int** debug\_engine\_runtime**;**

### 说明

1. ey\_engine\_create

**用途**：

该函数用于创建一个eyoung IPS引擎。

**参数**：

name：是该ips engine的名称字符串，以’\0’结尾。系统不检查名称是否重名。

**返回值**：

创建成功则返回engine\_t类型的指针，否则返回NULL。

**示例**：

参见2.5.1节

1. ey\_engine\_destroy

**用途**：

该函数用于销毁一个eyoung IPS引擎。

**参数**：

**engine**：ey\_engine\_create的返回值。

**返回值**：

该函数没有返回值。

**示例**：

参见2.5.1节

1. ey\_engine\_load

**用途**：

eyoung ips engine创建之后，必须通过该函数加载相关IPS规则文件。

**参数**：

**engine**：ey\_engine\_create的返回值。

**files：**字符指针数组，用来记录需要加载的规则文件路径。

**file\_num**：指示files数组的长度。

**返回值**：

成功返回0；否则返回-1。

**示例**：

参见2.5.1节

1. ey\_engine\_find\_event

**用途**：

该函数通常由协议分析器调用，用来根据“事件名称”，查找“事件索引值”。在eyoung中，“事件”的定义是通过规则文件，在规则加载的时候完成，具体的事件定义格式参照《Writing Signatures for eyoung》的2.11一节。事件有名字，在一个eyoung IPS引擎实例中，事件的名称不允许重复。在eyoung系统中，规则、IPS引擎、协议分析器都是通过事件名称的约定来完成数据的传递。为了方便传递，eyoung IPS引擎使用事件的索引值来代替事件名称进行信息传递。ey\_engine\_find\_event是通过“事件名称”查找“事件索引值”的API。

**参数**：

**engine**：ey\_engine\_create的返回值。

**event\_name**：事件名称字符串，以’\0’结尾。

**返回值**：

成功则返回非负值，表示对应的事件索引值；失败则返回-1。

**示例**：

libdecoder/http/http\_server.y中：

**void** http\_server\_register**(**http\_decoder\_t **\***decoder**)**

**{**

assert**(**decoder**!=**NULL**);**

assert**(**decoder**->**engine**!=**NULL**);**

engine\_t engine **=** decoder**->**engine**;**

**int** index **=** 0**;**

**for(**index**=**0**;** index**<**YYNNTS**;** index**++)**

**{**

**const char \*name = yytname[YYNTOKENS + index];**

**if(!**name **||** name**[**0**]==**'$' **||** name**[**0**]==**'@'**)**

**continue;**

**yytid[YYNTOKENS + index]** **=**

**ey\_engine\_find\_event(engine, name);**

**if(**yytid**[**YYNTOKENS **+** index**]** **>=** 0**)**

http\_debug**(**debug\_http\_server\_parser**,**

**"server event id of %s is %d\n",**

name**,** yytid**[**YYNTOKENS **+** index**]);**

**else**

http\_debug**(**debug\_http\_server\_parser**,**

**"failed to register server event %s\n",** name**);**

**}**

**}**

*/\*bison所生成的语法分析代码中，yytname数组记录着所有符号的名称，YYNTOKENS宏记录着非终结符的起始ID。在eyoung中，协议分析器在“移动-规约”的语法分析过程中，将归约出来的非终结符依次提交给IPS引擎检测。所以，协议分析器中的非终结符，实际就是IPS引擎中的“事件”，所有的非终结符都需要在IPS规则中书写，以使IPS引擎感知到“事件”的存在。该函数在规则加载之后，HTTP协议分析器初始化的时候被调用，通过依次对所有非终结符调用ey\_engine\_find\_event查找IPS引擎中的事件索引值，并保存在yytid数组中。\*/*

1. ey\_engine\_work\_create

**用途**：

创建一个engine\_work\_t实例。在engine\_work\_t实例创建的过程中，通过ey\_set\_predefine\_work\_init、ey\_set\_userdefine\_work\_init、%work-init注册的初始化函数，会被eyoung IPS引擎依次调用以完成相关的初始化工作。

**参数**：

**engine**：ey\_engine\_create的返回值

**返回值**：

成功则返回一个engine\_work\_t的实例指针；否则返回NULL。

**示例**：

参见2.5.2节。

1. ey\_engine\_work\_destroy

**用途**：

销毁一个engine\_work\_t的实例。在engine\_work\_t实例销毁的过程中，通过ey\_set\_predefine\_work\_finit、%work-finit 、ey\_set\_userdefine\_work\_finit注册的销毁函数，会被eyoung IPS引擎依次调用以完成相关的资源释放工作。

**参数**：

**work**：ey\_engine\_work\_create的返回值

**返回值**：

无。

**示例**：

参见2.5.2节。

1. ey\_engine\_work\_create\_event

**用途**：

该函数通常由协议分析器调用，用来创建一个事件实例向IPS引擎提交检测。在engine\_work\_event\_t实例创建的过程中，通过ey\_set\_predefine\_event\_init、ey\_set\_userdefine\_event\_init、%event-init注册的初始化函数，会被eyoung IPS引擎依次调用以完成相关的初始化工作。

**参数**：

**work**：ey\_engine\_work\_create的返回值

**event\_id**：由ey\_engine\_find\_event返回的合法的事件索引值。

**action**：输出参数，用来描述事件的检测结果。

**返回值**：

成功则返回engine\_work\_event\_t的实例指针；否则返回NULL。

**示例**：

libdecoder/http/decode/http\_decode.c

**int** http\_element\_detect**(**http\_data\_t **\***http\_data**,**

**const** **char** **\***event\_name**,** **int** event\_id**,** **void** **\***event**,**

**char** **\***cluster\_buffer**,** size\_t cluster\_buffer\_len**)**

**{**

**……**

engine\_action\_t action **=** **{**ENGINE\_ACTION\_PASS**};**

engine\_work\_t **\***work **=** http\_data**->**engine\_work**;**

engine\_work\_event\_t **\***work\_event **=**

**ey\_engine\_work\_create\_event(work, event\_id, &action);**

**if(!**work\_event**)**

**{**

http\_debug**(**debug\_http\_detect**,**

**"create event for %s failed\n",** event\_name**);**

**return** 0**;**

**}**

**ey\_engine\_work\_detect\_event(work\_event, event);**

**ey\_engine\_work\_destroy\_event(work\_event);**

http\_debug**(**debug\_http\_detect**,**

**"detect %s[%d], get actoin %s\n",**

event\_name**,** event\_id**,**

ey\_engine\_action\_name**(**action**.**action**));**

**if(**action**.**action**==**ENGINE\_ACTION\_PASS**)**

**return** 0**;**

**return** **-**1**;**

**}**

*/\*该函数在HTTP协议分析语法分析器中被调用，调用时机是语法分析器在规则加载时被自动调用，相关调用方式参照3.2\*/*

1. ey\_engine\_work\_destroy\_event

**用途**：

该函数通常由协议分析器调用，用来销毁一个事件实例，并释放相关资源。在engine\_work\_event\_t实例销毁的过程中，通过ey\_set\_predefine\_event\_finit、ey\_set\_userdefine\_event\_finit、%event-finit注册的销毁函数，会被eyoung IPS引擎依次调用以完成相关的销毁和资源释放工作。

**参数**：

**event**：ey\_engine\_work\_create\_event的返回值

**返回值**：

无

**示例**：

参考ey\_engine\_work\_create\_event

1. ey\_engine\_work\_detect\_event

**用途**：

该函数通常由协议分析器调用，用来将一个创建出来的event实例，提交给IPS引擎进行检测。通过ey\_set\_predefine\_event\_preprocess、%event-preprocessor、ey\_set\_userdefine\_event\_preprocess注册的事件预处理函数，会被eyoung IPS引擎依次调用以完成事件数据的预处理工作。

**参数**：

**event**：ey\_engine\_work\_create\_event的返回值

**predefined**：event的初始值。

**返回值**：

0表示事件是干净的，-1表示事件是包含攻击数据的。

**示例**：

参考ey\_engine\_work\_create\_event

1. ey\_engine\_work\_detect\_data

**用途**：

该函数通常由协议分析器调用，与ey\_engine\_work\_detect\_event不同的是，ey\_engine\_work\_detect\_data函数提交的数据参与的是IPS引擎预处理器的检查。预处理器可以被视为与eyoung IPS引擎平行的检测单元，其检测结果可以视为eyoung IPS引擎事件分析的前提。这个函数可以是异步完成的（目前的实现是同步的）。

**参数**：

**work**：ey\_engine\_work\_create的返回值

**data**：要检测的缓冲区起始地址

**data\_len**：要检测的缓冲区长度

**from\_client**：缓冲区是否来自客户端一侧的连接

**返回值**：

目前固定返回0

**示例**：

libdecoder/http/decode/http\_decode.c

**int** http\_decode\_data**(**http\_work\_t work**,**

**const** **char** **\***data**,** size\_t data\_len**,**

**int** from\_client**,** **int** last\_frag**)**

**{**

assert**(**work **!=** NULL**);**

http\_data\_t **\***http\_data **=** **(**http\_data\_t**\*)**work**;**

**if(**http\_data**->**engine\_work**)**

**ey\_engine\_work\_detect\_data(http\_data->engine\_work,**

**data, data\_len, from\_client);**

**if(**from\_client**)**

**return** parse\_http\_client\_stream**(**http\_data**,**

data**,** data\_len**,** last\_frag**);**

**else**

**return** parse\_http\_server\_stream**(**http\_data**,**

data**,** data\_len**,** last\_frag**);**

**}**

*/\*http\_decode\_data函数是在Socket上收到HTTP数据之后首先被调用的。该函数先调用ey\_engine\_work\_detect\_data函数，用来完成相关预处理器的检查，而后才调用parse\_http\_client\_stream和parse\_http\_server\_stream完成HTTP的协议分析工作。\*/*

1. debug\_engine\_init

这是一个debug开关，用来控制eyoung IPS引擎初始化阶段的信息，包括ey\_engine\_create、ey\_engine\_destroy和ey\_engine\_find\_event的debug信息。

1. debug\_engine\_lexier

这个debug开关用来控制eyoung IPS引擎加载规则文件时，eyoung IPS引擎词法分析器的debug信息。这些debug信息位于API ey\_engine\_load函数中。

1. debug\_engine\_parser

这个debug开关用来控制eyoung IPS引擎加载规则文件时，eyoung IPS引擎语法分析器的debug信息。这些debug信息位于API ey\_engine\_load函数中。

1. debug\_engine\_compiler

这个debug开关用来控制eyoung IPS引擎加载规则文件时，JIT编译器、预处理器的debug信息。这些debug信息位于API ey\_engine\_load函数中。

1. debug\_engine\_runtime

这个debug开关用来控制eyoung IPS引擎在进行攻击检测阶段的debug信息。包括ey\_engine\_work\_create、ey\_engine\_work\_create\_event、ey\_engine\_work\_detect\_data、ey\_engine\_work\_detect\_event、ey\_engine\_work\_destroy\_event、ey\_engine\_work\_destroy中的debug信息。

## Signature API

eyoung signature中可以直接调用的IPS Engine的功能函数API，都定义在头文件include/libengine\_export.h中。Signature API主要是为了方便规则开发人员，注册file、engine\_work、engine\_work\_event等对象的开始、结束、预处理等动作。

### 源代码

#include "libengine\_type.h"

#define ey\_add\_file\_init(eng, func) \

\_ey\_add\_file\_init(eng,#func,func,\_\_FILE\_\_,\_\_LINE\_\_)

#define ey\_add\_file\_finit(eng, func) \

\_ey\_add\_file\_finit(eng,#func,func,\_\_FILE\_\_,\_\_LINE\_\_)

**extern** **int** \_ey\_add\_file\_init**(**engine\_t engine**,**

**const** **char** **\***function**,** file\_init\_handle address**,**

**const** **char** **\***filename**,** **int** line**);**

**extern** **int** \_ey\_add\_file\_finit**(**engine\_t engine**,**

**const** **char** **\***function**,** file\_finit\_handle address**,**

**const** **char** **\***filename**,** **int** line**);**

#define ey\_set\_userdefine\_work\_init(eng,func) \

\_ey\_set\_work\_init(eng,1,#func,func,\_\_FILE\_\_,\_\_LINE\_\_)

#define ey\_set\_predefine\_work\_init(eng,func) \

\_ey\_set\_work\_init(eng,0,#func,func,\_\_FILE\_\_,\_\_LINE\_\_)

**extern** **int** \_ey\_set\_work\_init**(**engine\_t engine**,** **int** type**,**

**const** **char** **\***function**,** work\_init\_handle address**,**

**const** **char** **\***filename**,** **int** line**);**

#define ey\_set\_userdefine\_work\_finit(eng,func) \

\_ey\_set\_work\_finit(eng,1,#func,func,\_\_FILE\_\_,\_\_LINE\_\_)

#define ey\_set\_predefine\_work\_finit(eng,func) \

\_ey\_set\_work\_finit(eng,0,#func,func,\_\_FILE\_\_,\_\_LINE\_\_)

**extern** **int** \_ey\_set\_work\_finit**(**engine\_t engine**,** **int** type**,**

**const** **char** **\***function**,** work\_finit\_handle address**,**

**const** **char** **\***filename**,** **int** line**);**

#define ey\_set\_userdefine\_event\_init(eng,ev,func) \

\_ey\_set\_event\_init(eng,#ev,1,#func,func,\_\_FILE\_\_,\_\_LINE\_\_)

#define ey\_set\_predefine\_event\_init(eng,ev,func) \

\_ey\_set\_event\_init(eng,#ev,0,#func,func,\_\_FILE\_\_,\_\_LINE\_\_)

**extern** **int** \_ey\_set\_event\_init**(**engine\_t engine**,**

**const** **char** **\***event**,** **int** type**,**

**const** **char** **\***function**,**

event\_init\_handle address**,**

**const** **char** **\***filename**,** **int** line**);**

#define ey\_set\_userdefine\_event\_finit(eng,ev,func) \

\_ey\_set\_event\_finit(eng,#ev,1,#func,func,\_\_FILE\_\_,\_\_LINE\_\_)

#define ey\_set\_predefine\_event\_finit(eng,ev,func) \

\_ey\_set\_event\_finit(eng,#ev,0,#func,func,\_\_FILE\_\_,\_\_LINE\_\_)

**extern** **int** \_ey\_set\_event\_finit**(**engine\_t engine**,**

**const** **char** **\***event**,** **int** type**,**

**const** **char** **\***function**,**

event\_finit\_handle address**,**

**const** **char** **\***filename**,** **int** line**);**

#define ey\_set\_userdefine\_event\_preprocess(eng,ev,func) \

\_ey\_set\_event\_preprocessor(eng, #ev, 1, \

#func, func, \_\_FILE\_\_, \_\_LINE\_\_)

#define ey\_set\_predefine\_event\_preprocess(eng,ev,func) \

\_ey\_set\_event\_preprocessor(eng, #ev, 0, \

#func, func, \_\_FILE\_\_, \_\_LINE\_\_)

**extern** **int** \_ey\_set\_event\_preprocessor**(**engine\_t engine**,**

**const** **char** **\***event**,** **int** type**,**

**const** **char** **\***function**,**

event\_preprocess\_handle address**,**

**const** **char** **\***filename**,** **int** line**);**

### 说明

本节所涉及的API函数，在2.2.2中都有讲述，这里不再多做叙述。可以从libdecoder/http/testsuite/http\_xss.ey、libdeocder/pop3/testsuite/pop3.ey、libdecoder/html/testsuite/html.ey等例子中看到相关API的应用。

## 二进制规则库

### 二进制规则

eyoung支持规则的高度可编程性，其中重要的一点就是支持由eyoung IPS规则文件导入外部动态链接库.so文件。这个特性与Linux内核的可加载内核模块.ko的运行方式比较相近。eyoung在设计这个特性的时候，要解决的核心问题是功能的可扩展性、可升级性与执行效率之间的平衡。主要有三个相关的问题：

1. 符号的共享，即当前eyoung引擎运行环境要能感知到足够的外部动态链接库中的符号，包括全局变量（名称、加载后地址和变量类型）、全局函数（名称、入口地址、返回类型和参数列表）、类型定义等等信息。对于这个问题，在eyoung的解决方案中，既可以通过#include头文件的方式向外界发布，也可以通过eyoung内置的EY\_EXPORT\_IDENT和EY\_EXPORT\_TYPE两个宏，将相关定义以字符串的形式写入到外部动态链接库ELF文件中。
2. 动态库可以动态加载和卸载，eyoung规则语法中，支持%import关键字，用来动态导入外部动态链接库。动态库的卸载（通常是在规则升级的过程中），则通过在重新加载规则的过程中完成。
3. 外部的动态链接库的自动初始化操作，例如动态库中一部分全局变量的资源分配、初始化操作。eyoung借鉴了Linux内核可加载内核模块.ko文件的设计思想，将相关的init、finit函数的入口信息使用EY\_EXPORT\_INIT和EY\_EXPORT\_FINIT两个宏，写入到动态链接库自身的ELF文件中。在执行库加载的时候，eyoung分析ELF文件，得到相关的入口地址，并自动执行init函数；规则卸载的时候，执行finit函数。

### Source Code

#ifndef EY\_EXPORT\_H

#define EY\_EXPORT\_H 1

#define EY\_EXPORT\_TYPE\_IDENT 1

#define EY\_EXPORT\_TYPE\_TYPE 2

#define EY\_EXPORT\_TYPE\_INIT 3

#define EY\_EXPORT\_TYPE\_FINIT 4

#define EY\_IDENT\_SECTION ".eyoung\_ident"

#define EY\_TYPE\_SECTION ".eyoung\_type"

#define EY\_INIT\_SECTION ".eyoung\_init"

#define EY\_FINIT\_SECTION ".eyoung\_finit"

**typedef** **struct** ey\_extern\_symbol

**{**

**char** **\***name**;**

**void** **\***value**;**

**char** **\***decl**;**

**char** **\***file**;**

**int** line**;**

**int** type**;**

**}**ey\_extern\_symbol\_t**;**

#define EY\_EXPORT\_IDENT(name, decl) \

static const ey\_extern\_symbol\_t \_\_eyoung\_ident\_##name \

\_\_attribute\_\_((section(EY\_IDENT\_SECTION), unused)) = \

{ \

#name, \

&name, \

decl, \

\_\_FILE\_\_, \

\_\_LINE\_\_, \

EY\_EXPORT\_TYPE\_IDENT \

};

#define EY\_EXPORT\_TYPE(name,decl) \

static const ey\_extern\_symbol\_t \_\_eyoung\_type\_##name \

\_\_attribute\_\_((section(EY\_TYPE\_SECTION), unused)) = \

{ \

#name, \

(void\*)0, \

decl, \

\_\_FILE\_\_, \

\_\_LINE\_\_, \

EY\_EXPORT\_TYPE\_IDENT \

};

#define EY\_EXPORT\_INIT(name) \

static const ey\_extern\_symbol\_t \_\_eyoung\_type\_##name \

\_\_attribute\_\_((section(EY\_INIT\_SECTION), unused)) = \

{ \

#name, \

name, \

(void\*)0, \

\_\_FILE\_\_, \

\_\_LINE\_\_, \

EY\_EXPORT\_TYPE\_INIT \

};

#define EY\_EXPORT\_FINIT(name) \

static const ey\_extern\_symbol\_t \_\_eyoung\_type\_##name \

\_\_attribute\_\_((section(EY\_FINIT\_SECTION), unused)) = \

{ \

#name, \

name, \

(void\*)0, \

\_\_FILE\_\_, \

\_\_LINE\_\_, \

EY\_EXPORT\_TYPE\_FINIT \

};

#endif

### 说明

1. EY\_EXPORT\_IDENT

这个宏可以将全局变量、全局函数的声明，以字符串的形式写入到动态链接库ELF文件的.eyoung\_ident段中。eyoung IPS引擎在加载IPS规则时，执行%import关键字时，会试图读取要加载的动态库中.eyoung\_ident段（如果存在）的内容，并将符号的声明直接写入到转换后的.eyc规则文件中。

1. EY\_EXPORT\_TYPE

这个宏将动态库源文件中的类型定义，写入到动态链接库ELF文件的.eyoung\_type段中。eyoung IPS引擎在加载IPS规则时，执行%import关键字时，会试图读取要加载的动态库中.eyoung\_type段（如果存在）的内容，并将符号的声明直接写入到转换后的.eyc规则文件中。

1. EY\_EXPORT\_INIT

这个宏可以将一个指定的函数入口，写入到动态链接库ELF文件的.eyoung\_init段中。eyoung IPS引擎在加载IPS规则时，执行%import关键字时，会试图读取要加载的动态库中.eyoung\_init段（如果存在）的内容，得到相关入口函数的地址后执行该入口函数，自动完成初始化功能。

1. EY\_EXPORT\_FINIT

这个宏可以将一个指定的函数入口，写入到动态链接库ELF文件的.eyoung\_finit段中。eyoung IPS引擎在加载IPS规则时，执行%import关键字时，会试图读取要加载的动态库中.eyoung\_finit段（如果存在）的内容，并记录finit函数的入口地址，待到规则卸载时执行该入口函数，自动完成终止化功能，以释放相关资源。

上述功能的示例，在本文2.5.3给出。

## 示例

### demo\_http\_xss.c

本示例主要展示***ey\_engine\_create、ey\_engine\_load***和***ey\_engine\_destroy***的使用。**libdecoder/http/testsuite/demo\_http\_xss.c**

#define \_GNU\_SOURCE

#include <stdio.h>

#include <stdlib.h>

#include <string.h>

#include <ctype.h>

#include "libengine.h"

#include "http.h"

#include "html.h"

**int** parse\_http\_file**(**http\_handler\_t decoder**,**

**const** **char** **\***filename**)**

**{**

**char** **\***line **=** NULL**;**

size\_t len **=** 0**;**

ssize\_t read **=** 0**;**

**int** ret **=** **-**1**;**

http\_work\_t **\***work **=** NULL**;**

**int** lines **=** 0**;**

FILE **\***fp **=** fopen**(**filename**,** **"r");**

**if(!**fp**)**

**{**

fprintf**(**stderr**,**

**"failed to open file %s\n",** filename**);**

**goto** failed**;**

**}**

work **=** **http\_work\_create(**decoder**,** 0**);**

**if(!**work**)**

**{**

fprintf**(**stderr**,**

**"failed to alloc http private data\n");**

**goto** failed**;**

**}**

**while** **((**read **=** getline**(&**line**,** **&**len**,** fp**))** **!=** **-**1**)**

**{**

lines**++;**

**if(**read **<=**1**)**

**{**

fprintf**(**stderr**,**

**"invalid line(%d): %s\n",** lines**,** line**);**

**goto** failed**;**

**}**

**if(**toupper**(**line**[**0**])==**'C' **&&** line**[**1**]==**':'**)**

**{**

**if(**http\_decode\_data**(**work**,** line**+**2**,** read**-**2**,** 1**,** 0**))**

**{**

fprintf**(**stderr**,**

**"parse client failed, line(%d): %s\n",**

lines**,** line**);**

**goto** failed**;**

**}**

**}**

**else** **if** **(**toupper**(**line**[**0**])==**'S' **&&** line**[**1**]==**':'**)**

**{**

**if(**http\_decode\_data**(**work**,** line**+**2**,** read**-**2**,** 0**,** 0**))**

**{**

fprintf**(**stderr**,**

**"parse server failed, line(%d): %s\n",**

lines**,** line**);**

**goto** failed**;**

**}**

**}**

**else**

**{**

fprintf**(**stderr**,**

**"invalid line(%d): %s\n",** lines**,** line**);**

**goto** failed**;**

**}**

**}**

*/\*give end flag to parser\*/*

**if(**http\_decode\_data**(**work**,** **"",** 0**,** 1**,** 1**))**

**{**

fprintf**(**stderr**,**

**"parse client end flag failed");**

**goto** failed**;**

**}**

**if(**http\_decode\_data**(**work**,** **"",** 0**,** 0**,** 1**))**

**{**

fprintf**(**stderr**,**

**"parse server end flag failed");**

**goto** failed**;**

**}**

ret **=** 0**;**

*/\*pass through\*/*

fprintf**(**stderr**,** **"parser OK!\n");**

**failed:**

**if(**work**)**

**http\_work\_destroy(**work**);**

**if(**line**)**

free**(**line**);**

**if(**fp**)**

fclose**(**fp**);**

**return** ret**;**

**}**

**int** main**(int** argc**,** **char** **\***argv**[])**

**{**

**int** ret **=** 0**;**

http\_handler\_t decoder **=** NULL**;**

engine\_t engine **=** NULL**;**

**if(**argc **!=** 3**)**

**{**

fprintf**(**stderr**,**

**"Usage: http\_parser <sig\_file> <msg\_file>\n");**

**return** **-**1**;**

**}**

debug\_http\_server\_lexer **=** 0**;**

debug\_http\_server\_parser **=** 0**;**

debug\_http\_client\_lexer **=** 0**;**

debug\_http\_client\_parser **=** 0**;**

debug\_http\_mem **=** 0**;**

debug\_http\_detect **=** 0**;**

debug\_html\_lexer **=** 0**;**

debug\_html\_parser **=** 0**;**

debug\_html\_mem **=** 0**;**

debug\_html\_detect **=** 0**;**

debug\_engine\_parser **=** 0**;**

debug\_engine\_lexier **=** 0**;**

debug\_engine\_init **=** 0**;**

debug\_engine\_compiler **=** 0**;**

debug\_engine\_runtime **=** 0**;**

engine **=** **ey\_engine\_create("http");**

**if(!**engine**)**

**{**

fprintf**(**stderr**,** **"create http engine failed\n");**

ret **=** **-**1**;**

**goto** failed**;**

**}**

**if(ey\_engine\_load(**engine**,** **&**argv**[**1**],** 1**))**

**{**

fprintf**(**stderr**,** **"load http signature failed\n");**

ret **=** **-**1**;**

**goto** failed**;**

**}**

decoder **=** http\_decoder\_init**(**engine**,** NULL**);**

**if(!**decoder**)**

**{**

fprintf**(**stderr**,** **"create http decoder failed\n");**

ret **=** **-**1**;**

**goto** failed**;**

**}**

ret **=** parse\_http\_file**(**decoder**,** argv**[**2**]);**

**failed:**

**if(**decoder**)**

http\_decoder\_finit**(**decoder**);**

**if(**engine**)**

**ey\_engine\_destroy(**engine**);**

**return** ret**;**

**}**

### http\_decode.c

本示例主要展示

***ey\_engine\_work\_create、ey\_engine\_work\_detect\_data、ey\_engine\_work\_create\_event、ey\_engine\_work\_detect\_event、ey\_engine\_work\_destroy\_event和ey\_engine\_work\_destroy***的用法。**libdecoder/http/decode/http\_decode.c**

#include <stdlib.h>

#include <stdio.h>

#include <string.h>

#include <assert.h>

#include "http.h"

#include "http\_private.h"

#include "libengine.h"

http\_work\_t http\_work\_create**(**http\_handler\_t handler**,** **int** greedy**)**

**{**

**if(!**handler**)**

**{**

http\_debug**(**debug\_http\_mem**,**

**"null http decoder handler\n");**

**return** NULL**;**

**}**

http\_data\_t **\***priv\_data **=** NULL**;**

http\_decoder\_t **\***decoder **=** **(**http\_decoder\_t**\*)**handler**;**

engine\_t engine **=** decoder**->**engine**;**

priv\_data **=** http\_alloc\_priv\_data**(**decoder**,** greedy**);**

**if(!**priv\_data**)**

**{**

http\_debug**(**debug\_http\_mem**,**

**"failed to alloc http private data\n");**

**return** NULL**;**

**}**

priv\_data**->**decoder **=** handler**;**

**if(**engine**)**

**{**

**engine\_work\_t \*engine\_work =**

**ey\_engine\_work\_create(engine);**

**if(!**engine\_work**)**

**{**

http\_debug**(**debug\_http\_mem**,**

**"failed to alloc engine work\n");**

http\_free\_priv\_data**(**decoder**,** priv\_data**);**

**return** NULL**;**

**}**

priv\_data**->**engine\_work **=** engine\_work**;**

engine\_work**->**predefined **=** **(void\*)**priv\_data**;**

**}**

**return** priv\_data**;**

**}**

**void** http\_work\_destroy**(**http\_work\_t work**)**

**{**

**if(**work**)**

**{**

http\_data\_t **\***priv\_data **=** **(**http\_data\_t**\*)**work**;**

http\_decoder\_t **\***decoder **=**

**(**http\_decoder\_t**\*)(**priv\_data**->**decoder**);**

**if(**priv\_data**->**engine\_work**)**

**ey\_engine\_work\_destroy(priv\_data->engine\_work);**

http\_free\_priv\_data**(**decoder**,** priv\_data**);**

**}**

**}**

**int** http\_decode\_data**(**http\_work\_t work**,**

**const** **char** **\***data**,** size\_t data\_len**,**

**int** from\_client**,** **int** last\_frag**)**

**{**

assert**(**work **!=** NULL**);**

http\_data\_t **\***http\_data **=** **(**http\_data\_t**\*)**work**;**

**if(**http\_data**->**engine\_work**)**

**ey\_engine\_work\_detect\_data(http\_data->engine\_work,**

**data, data\_len, from\_client);**

**if(**from\_client**)**

**return** parse\_http\_client\_stream**(**http\_data**,**

data**,** data\_len**,** last\_frag**);**

**else**

**return** parse\_http\_server\_stream**(**http\_data**,**

data**,** data\_len**,** last\_frag**);**

**}**

http\_handler\_t http\_decoder\_init**(**engine\_t engine**,**

**void** **\***html\_decoder**)**

**{**

http\_decoder\_t **\***decoder **=**

**(**http\_decoder\_t**\*)**http\_malloc**(sizeof (**http\_decoder\_t**));**

**if(!**decoder**)**

**{**

http\_debug**(**debug\_http\_mem**,**

**"failed to alloc http decoder\n");**

**goto** failed**;**

**}**

memset**(**decoder**,** 0**,** **sizeof(\***decoder**));**

**if(**http\_mem\_init**(**decoder**))**

**{**

http\_debug**(**debug\_http\_mem**,**

**"failed to init http decoder mem\n");**

**goto** failed**;**

**}**

decoder**->**html\_decoder **=** html\_decoder**;**

**if(**engine**)**

**{**

decoder**->**engine **=** engine**;**

http\_server\_register**(**decoder**);**

http\_client\_register**(**decoder**);**

**}**

**return** **(**http\_handler\_t**)**decoder**;**

failed**:**

**if(**decoder**)**

http\_decoder\_finit**((**http\_handler\_t**)**decoder**);**

**return** NULL**;**

**}**

**void** http\_decoder\_finit**(**http\_handler\_t handler**)**

**{**

**if(**handler**)**

**{**

http\_decoder\_t **\***decoder **=** **(**http\_decoder\_t**\*)**handler**;**

http\_mem\_finit**(**decoder**);**

http\_free**(**decoder**);**

**}**

**}**

**int** http\_element\_detect**(**http\_data\_t **\***http\_data**,**

**const** **char** **\***event\_name**,** **int** event\_id**,** **void** **\***event**,**

**char** **\***cluster\_buffer**,** size\_t cluster\_buffer\_len**)**

**{**

**if(!**http\_data**->**engine\_work**)**

**{**

http\_debug**(**debug\_http\_detect**,**

**"engine work is not created, skip scan\n");**

**return** 0**;**

**}**

**if(!**event\_name**)**

**{**

http\_debug**(**debug\_http\_detect**,** **"event name is null\n");**

**return** 0**;**

**}**

**if(**event\_id **<** 0**)**

**{**

http\_debug**(**debug\_http\_detect**,**

**"event id %d for event %s is illegal\n",**

event\_id**,** event\_name**);**

**return** 0**;**

**}**

**if(!**http\_data **||** **!**event**)**

**{**

http\_debug**(**debug\_http\_detect**,**

**"bad parameter for event %s\n",** event\_name**);**

**return** 0**;**

**}**

engine\_action\_t action **=** **{**ENGINE\_ACTION\_PASS**};**

engine\_work\_t **\***work **=** http\_data**->**engine\_work**;**

**engine\_work\_event\_t \*work\_event =**

**ey\_engine\_work\_create\_event(work, event\_id, &action);**

**if(!**work\_event**)**

**{**

http\_debug**(**debug\_http\_detect**,**

**"create event for %s failed\n",** event\_name**);**

**return** 0**;**

**}**

**ey\_engine\_work\_detect\_event(work\_event, event);**

**ey\_engine\_work\_destroy\_event(work\_event);**

http\_debug**(**debug\_http\_detect**,**

**"detect %s[%d], get actoin %s\n",**

event\_name**,** event\_id**,**

ey\_engine\_action\_name**(**action**.**action**));**

**if(**action**.**action**==**ENGINE\_ACTION\_PASS**)**

**return** 0**;**

**return** **-**1**;**

**}**

### export\_test.c

本示例主要展示**EY\_EXPORT\_IDENT、EY\_EXPORT\_INIT、EY\_EXPORT\_TYPE和EY\_EXPORT\_TYPE**的用法。

**test/export\_test.c**

#include <stdio.h>

#include "ey\_export.h"

**int** a**=**1**;**

**int** foo**(void** **\***link**,** **void** **\***event**)**

**{**

printf**("call foo, a=%d\n",** **++**a**);**

**return** 1**;**

**}**

**int** bar**(void** **\***link**,** **void** **\***event**)**

**{**

printf**("call bar, a=%d\n",** **++**a**);**

**return** 0**;**

**}**

**int** test\_init**(void** **\***eng**)**

**{**

printf**("call init, a=%d\n",** a**++);**

**}**

**int** test\_exit**(void** **\***eng**)**

**{**

printf**("call finit, a=%d\n",** a**--);**

**}**

**struct** s

**{**

**int** a**;**

**int** b**;**

**};**

**EY\_EXPORT\_IDENT(**a**,** **"extern int a;");**

**EY\_EXPORT\_IDENT(**foo**,** **"int foo(void \*link, void \*event);");**

**EY\_EXPORT\_IDENT(**bar**,** **"int bar(void \*link, void \*event);");**

**EY\_EXPORT\_TYPE(**s**,** **"struct s{int a; int b;};");**

**EY\_EXPORT\_INIT(**test\_init**);**

**EY\_EXPORT\_FINIT(**test\_exit**);**

# 协议分析器

协议分析器对Layer7协议报文做协议分析，并且以层次化、结构化的方式将分析之后的数据以协议事件（event）的形式提交给eyoung IPS引擎，用以在规则的指引下进行分析和检测。为了达到这个目的，eyoung协议分析器从结构上大致分为协议词法分析器Lexer和协议语法分析器Parser。这一点同编译器的实现有很大的相同之处。

## Stream-mode协议词法分析器



### 设计

应用层协议词法分析器（Lexer）在协议分析器中更贴近底层，即更多地与网络数据直接交互，主要的作用是对获取的网络数据，按照分析规则，将原始报文拆解成一系列的“词元”。这一点同编译器中的词法分析器的作用是一致的。不同的是，编译器是对本地磁盘上的文件进行操作，多采用阻塞式的文件IO功能（例如fread、fscanf等等），而用于协议分析的词法分析器不可能采用阻塞的方式，必须在“流”上进行词法分析。为此，eyoung改造了flex的用法，解决了以下问题：

1. 输入缓冲区自动管理。对于上一个输入缓冲区未被匹配的部分，例如当前的输入缓冲区的结尾部分仅部分匹配一个模式，此时eyoung lexer不按照default模式报告，而是将结尾部分未完全匹配的部分”copy and save”。待到下一次送入输入缓冲区后，再次拼接并重新匹配。
2. greedy模式。如果在贪婪模式下，如果上一个缓冲区的结尾恰好匹配了某一个模式，此时将eyoung lexer直接返回匹配的模式；反之，eyoung lexer按照1)的逻辑保存最后匹配的内容，并在下一个输入缓冲区到来时拼接缓冲区并重新匹配。这是因为，要匹配的模式可能存在包含关系，例如“b”和“ba”，如果当前缓冲区以“b”结尾，此时有可能匹配“ba”（只需要下一个输入缓冲区以“a”开头即可）。在greedy模式下，eyoung lexer会返回第一个模式“b”；反之则保存结尾的“b”，待到下一个输入缓冲区到来之后再做判断。注意：eyoung lexer使用的是flex工具，而flex使用的是greedy模式且不能修改，但是flex并没有在stream上考虑这个问题。eyoung的非greedy模式仅仅是针对当前输入报文结尾部分的分析，并不会导致输入报文非结尾的词法分析使用非greedy模式。

### 示例

**%{**

#include <stdlib.h>

#include <stdio.h>

#include <string.h>

#include <assert.h>

#include "http.h"

#include "http\_server\_parser.h"

#include "http\_private.h"

#define **YY\_USER\_ACTION** \

if(yyg->yy\_hold\_char==YY\_END\_OF\_BUFFER\_CHAR && \

**save\_server\_stream\_context**(yyg,0)) \

return TOKEN\_SERVER\_CONTINUE;

#ifdef **YY\_RESTORE\_YY\_MORE\_OFFSET**

#undef YY\_RESTORE\_YY\_MORE\_OFFSET

#define YY\_RESTORE\_YY\_MORE\_OFFSET \

{ \

if(**save\_server\_stream\_context**(yyg,1)) \

return TOKEN\_SERVER\_CONTINUE; \

}

#endif

**%}**

**%**option header**-**file**="http\_server\_lex.h"**

**%**option outfile**="http\_server\_lex.c"**

**%**option prefix**="http\_server\_"**

**%**option bison**-**bridge

**%**option noyywrap

**%**option reentrant

**%**option case**-**insensitive

**%**option ansi**-**definitions

**%**option noinput

**%**option nounput

**%**option noyyalloc

**%**option noyyfree

**%**option noyyrealloc

**…………**

**%%**

**…………**

**%%**

**…………**

**struct** yy\_buffer\_state**\*** http\_server\_scan\_stream**(**

**const** **char** **\***new\_buf**,** size\_t new\_buf\_len**,** http\_data\_t **\***priv**)**

**{**

YY\_BUFFER\_STATE b**;**

**char** **\***buf**;**

yy\_size\_t n**,** \_yybytes\_len**;**

**char** **\***last\_buf **=** priv**->**response\_parser**.**saved**;**

size\_t last\_buf\_len **=** priv**->**response\_parser**.**saved\_len**;**

yyscan\_t scanner **=** **(**yyscan\_t**)**priv**->**response\_parser**.**lexier**;**

\_yybytes\_len **=** new\_buf\_len **+** last\_buf\_len**;**

n **=** \_yybytes\_len **+** 2**;**

buf **=** **(char** **\*)**http\_server\_alloc**(**n**,** scanner**);**

**if** **(!**buf**)**

**{**

http\_debug**(**debug\_http\_server\_lexer**,**

**"out of dynamic memory in http\_server\_scan\_stream()\n");**

**return** NULL**;**

**}**

**if(**last\_buf**)**

memcpy**(**buf**,** last\_buf**,** last\_buf\_len**);**

**if(**new\_buf**)**

memcpy**(**buf**+**last\_buf\_len**,** new\_buf**,** new\_buf\_len**);**

buf**[**\_yybytes\_len**]** **=** buf**[**\_yybytes\_len**+**1**]** **=** YY\_END\_OF\_BUFFER\_CHAR**;**

http\_debug**(**debug\_http\_server\_lexer**,** **"[HTTP SERVER]: %s\n",** buf**);**

*//b = http\_server\_scan\_buffer(buf, n, scanner);*

b **=** **(**YY\_BUFFER\_STATE**)**http\_server\_alloc**(sizeof(struct**

yy\_buffer\_state**),** scanner**);**

**if** **(** **!** b **)**

**{**

http\_debug**(**debug\_http\_server\_lexer**,**

**"failed to alloc server buffer state\n");**

http\_server\_free**(**buf**,** scanner**);**

**return** NULL**;**

**}**

b**->**yy\_buf\_size **=** n **-** 2**;** */\* "- 2" to take care of EOB's \*/*

b**->**yy\_buf\_pos **=** b**->**yy\_ch\_buf **=** buf**;**

b**->**yy\_is\_our\_buffer **=** 0**;**

b**->**yy\_input\_file **=** 0**;**

b**->**yy\_n\_chars **=** b**->**yy\_buf\_size**;**

b**->**yy\_is\_interactive **=** 0**;**

b**->**yy\_at\_bol **=** 1**;**

b**->**yy\_fill\_buffer **=** 0**;**

b**->**yy\_buffer\_status **=** YY\_BUFFER\_NEW**;**

http\_server\_switch\_to\_buffer**(**b**,** scanner**);**

b**->**yy\_is\_our\_buffer **=** 1**;**

**if(**priv**->**response\_parser**.**saved**)**

**{**

http\_server\_free**(**priv**->**response\_parser**.**saved**,** scanner**);**

priv**->**response\_parser**.**saved **=** NULL**;**

priv**->**response\_parser**.**saved\_len **=** 0**;**

**}**

**return** b**;**

**}**

**static** **int** save\_server\_stream\_context**(**

yyscan\_t yyscanner**,** **int** from\_default**)**

**{**

**struct** yyguts\_t **\*** yyg **=** **(struct** yyguts\_t**\*)**yyscanner**;**

http\_data\_t **\***priv **=** **(**http\_data\_t**\*)**http\_server\_get\_extra**(**yyg**);**

**int** len **=** 0**;**

**if(**priv**->**response\_parser**.**saved**)**

**{**

http\_server\_free**(**priv**->**response\_parser**.**saved**,** yyg**);**

priv**->**response\_parser**.**saved **=** NULL**;**

priv**->**response\_parser**.**saved\_len **=** 0**;**

**}**

**if(!**priv **||** priv**->**response\_parser**.**last\_frag **||**

**(!**from\_default **&&** **!**priv**->**response\_parser**.**greedy**))**

**{**

http\_debug**(**debug\_http\_server\_lexer**,**

**"No need to save stream context\n”);**

**return** 0**;**

**}**

len **=** from\_default**?**yyleng**-**1**:**yyleng**;**

**if(!**len**)**

**{**

http\_debug**(**debug\_http\_server\_lexer**,**

**"Exit save stream context for ZERO length yytext\n");**

**return** 1**;**

**}**

priv**->**response\_parser**.**saved **=** http\_server\_alloc**(**len**,** yyg**);**

**if(!**priv**->**response\_parser**.**saved**)**

**{**

http\_debug**(**debug\_http\_server\_lexer**,**

**"out of memory while saving context\n");**

**return** 0**;**

**}**

memcpy**(**priv**->**response\_parser**.**saved**,** yytext**,** len**);**

priv**->**response\_parser**.**saved\_len **=** len**;**

http\_debug**(**debug\_http\_server\_lexer**,**

**"Save stream context, string: %s, len: %d\n",** yytext**,** len**);**

**return** 1**;**

**}**

其中：

http\_server\_scan\_stream函数用来组装当前的扫描缓冲区，如果eyoung lexer之前没有保存待确定的内容，则直接用当前的输入缓冲区做本次扫描缓冲区；否则，则分配一块缓冲区，其长度为保存的待确定内容长度+当前输入缓冲区长度，并把两段缓冲区copy到新分配的缓冲区中。这里存在优化空间，可以不保存第一次匹配的结尾数据，而只保存结尾的匹配状态，这样可以避免一次数据拷贝！

save\_server\_stream\_context函数被调用的地方有两处：YY\_USER\_ACTION宏和YY\_RESTORE\_YY\_MORE\_OFFSET宏。YY\_USER\_ACTION是当某一个模式被匹配时调用，按照2.1.1中2)的介绍，此时需要判断当前匹配位置是否到了输入缓冲区的结尾同时判断是否开启了greedy模式，如果到了结束位置且开启了greedy模式，则保存结尾的匹配部分，并返回TOKEN\_CONTINUE通知parser；而对于另一处调用点YY\_RESTORE\_YY\_MORE\_OFFSET，则是在lexer到达当前输入缓冲区结尾后，且并没有找到任何可匹配模式，实际上进入default处理时被调用，由于eyoung lexer是stream-based，这种情况可能不代表不匹配任何模式，此时如果当前输入缓冲区不是stream上的最后一个缓冲区，则保存最后不匹配任何模式的部分，并返回TOKEN\_CONTINUE通知parser等待下一个输入缓冲区。

## Push-Mode协议语法分析器

### 设计

eyoung协议分析器的核心是一个协议语法分析器（Parser），这个分析器在协议语法规则（通常是一个上下文无关文法）的指引下，进行语法分析。语法分析器一方面检查协议数据在语法和语义上是否符合协议规范，另一方面自动地把分析后的相关数据传递到eyoung IPS引擎进行攻击检测。

传统的语法分析器，例如bison生成的语法分析器，其工作模式通常采纳“拉”（PULL）的方式，即语法分析器内部通过默认的接口调用词法分析器获取下一个词元，词法分析器再进行文件和缓冲区操作。这种方式对于阻塞式I/O的词法分析器是合适的，但是对于网络数据检测，这种方式的阻塞式特性会对性能造成非常严重的影响。为了解决这个问题，eyoung利用了bison的新特性，那就是Push-Mode Parser。这种模式下，词法分析器的调用实际是先于语法分析器被调用，然后将词法分析器得到的词元，以参数的形式“压”（PUSH）入语法分析器。

在eyoung中，由于词法分析器当遇到数据流结尾位置时，可能会进行缓冲区copy and save，并返回TOKEN\_CONTINUE。因此，当控制流得到TOKEN\_CONTINUE之后，将不调用语法分析器的功能。

### 示例

libdecoder/http/decode/http\_server.y

**int** parse\_http\_server\_stream**(**http\_data\_t **\***priv**,**

**const** **char** **\***buf**,** size\_t buf\_len**,** **int** last\_frag**)**

**{**

http\_server\_pstate **\***parser **=**

**(**http\_server\_pstate**\*)**priv**->**response\_parser**.**parser**;**

yyscan\_t lexier **=** **(**yyscan\_t**)**priv**->**response\_parser**.**lexier**;**

YY\_BUFFER\_STATE input **=** NULL**;**

**int** token **=** 0**,** parser\_ret **=** 0**;**

HTTP\_SERVER\_STYPE value**;**

yydebug **=** debug\_http\_server\_parser**;**

priv**->**response\_parser**.**last\_frag **=** last\_frag**;**

input **=** **http\_server\_scan\_stream(**buf**,** buf\_len**,** priv**);**

**if(!**input**)**

**{**

http\_debug**(**debug\_http\_server\_parser**,**

**"create http server stream buffer failed\n");**

**return** 1**;**

**}**

**while(**1**)**

**{**

memset**(&**value**,** 0**,** **sizeof(**value**));**

**if(http\_server\_lex\_body\_mode(**lexier**))**

token **=** **http\_server\_body\_lex(&**value**,** lexier**);**

**else**

token **=** **http\_server\_lex(&**value**,** lexier**);**

**if(**token **==** TOKEN\_SERVER\_CONTINUE**)**

**break;**

parser\_ret **=** **http\_server\_push\_parse(**parser**,** token**,**

**&**value**,** **(void\*)**priv**);**

**if(**parser\_ret **!=** YYPUSH\_MORE**)**

**break;**

**}**

**http\_server\_delete\_buffer(**input**,** lexier**);**

**if(**parser\_ret **!=** YYPUSH\_MORE **&&** parser\_ret **!=** 0**)**

**{**

http\_debug**(**debug\_http\_server\_parser**,**

**"find error while parsing http server stream\n");**

**return** 2**;**

**}**

**return** 0**;**

**}**

这个函数在底层I/O收到收到一段网络报文<buf, buf\_len>之后，以PUSH的方式调用协议分析器。其中：

* http\_server\_scan\_stream负责组装词法分析器使用的buffer结构，细节在3.1.2中已经做了介绍。
* http\_server\_delete\_buffer负责销毁创建的buffer结构，这个是flex提供的API函数，详情参见flex manual。
* http\_server\_lex\_body\_mode是一个返回布尔值的函数，用来判断当前词法分析器是否分析到HTTP报文的BODY部分。由于HTTP BODY不管是使用CONTENT-LENGTH定义还是使用CHUNK TRANSFER-ENCODING，BODY内容的解析都是依赖长度信息的。这种计数的解析并不适合使用基于正则表达式的FLEX完成。所以，eyoung在解析此类数据时，将BODY部分和非BODY部分分别解析。http\_server\_lex\_body\_mode返回TRUE时，eyoung HTTP协议分析其使用http\_server\_body\_lex对BODY部分进行词法分析；反之使用http\_server\_lex进行非BODY部分的解析，http\_server\_lex是由FLEX生成的标准函数。
* http\_server\_body\_lex或http\_server\_lex返回SYM\_CONTINUE时，意味着当前缓冲区已经全部读取完毕，并可能把当前缓冲区的最后一段copy and save。此时，整个解析过程应该跳出while循环。
* http\_server\_push\_parse函数是Push-Mode Parser的入口，该函数由bison自动生成，函数的参数是词法分析器的结果。http\_server\_push\_parse返回0时，表示输入的词元序列已经完全被语法分析器所接受；返回YYMORE\_PUSH，表示当前输入词元符合语法，但是需要后续词元序列推进语法分析器的状态，直到返回0或发现错误；其他返回值意味着发现语法错误。

## 事件

### 事件的类型

事件的定义，是使用规则文件完成的，详细的语法通过《Writing Signatures for eyoung》描述。另外，在本文档的2.2.2节的ey\_engine\_find\_event的介绍中，已经讲述了如何在eyoung协议分析器中获取已经创建事件的索引值的方法。本节讲述的是事件的类型定义机制以及事件的发射机制。

在上下文无关语言的语法分析中，通常采用“自底向上”的LR分析法，即“移动——规约”分析法。在这种分析法中，由词法分析器所得到的词元被称为“终结符”，而在规约过程中产生的符号被称为“非终结符”。**eyoung的事件本质上就是由LR分析中产生的“非终结符”**。伴随着协议语法的分析过程，语法分析器会自底向上地产生一系列“非终结符”，这些“非终结符”从概念上也必然是“自底向上”的，即从微观的概念到宏观的概念。这样IPS规则就获得了在不同概念层面描述协议内容细节的能力。

**事件的名称就是“非终结符”的名称，事件的类型就是“非终结符”的类型**。在bison语法定义中，可以通过使用***%union***和***%type***关键字对“非终结符”定义类型。例如，http\_server.y中如下一段定义：

%union

{

……

**http\_response\_header\_t \*header;**

……

}

……

**%type <header> response\_header\_host**

……

这段代码意味着“非终结符”response\_header\_host在语法分析器中的类型是http\_response\_header\_t的指针。如果要让eyoung IPS引擎知道事件response\_header\_host的类型，必须通过规则文件。所以上述例子在HTTP规则文件中有如下的描述：

Libdecoder/http/testsuite/http\_xss.ey

……

**%event “response\_header\_host” “http\_response\_header\_t \*”**

……

### 事件的提交

协议语法分析器在“自底向上”的语法分析过程中，每规约出来一个“非终结符”，就将此“非终结符”以对应的事件格式提交给eyoung IPS引擎进行扫描和检查。为了达到这个目的，eyoung修改了bison生成代码的模板yacc.c，patch文件是tool/yacc.c.diff。同时，在协议分析语法文件中需要做出配合修改：

Libdecoder/http/decode/http\_server.y文件片段：

#ifdef YY\_REDUCTION\_CALLBACK

#undef YY\_REDUCTION\_CALLBACK

#endif

#define YY\_REDUCTION\_CALLBACK(data,name,id,val) \

do \

{ \

if(http\_element\_detect(data,name,id,val)<0) \

{ \

http\_debug(debug\_http\_detect, "find attack!\n"); \

return -1; \

} \

}while(0)

YY\_REDUCTION\_CALLBACK这个宏是eyoung对bison的扩展，该宏是在LR语法分析器在规约动作发生时被自动调用。http\_element\_detect函数已经在2.5.2中给出了实现。经过Patch之后的bison模板文件yacc.c会显示YY\_REDUCTION\_CALLBACK是如何被调用的，读者可以自己查看。