Università di Pisa
Dipartimento di Informatica
Corso di Laurea in Informatica

Syn flood detection

Corso di Gestione di Rete
Anno accademico 2016-2017

Prof. Luca Deri

Marco Zizi
(matr. 424213)

1 L'attacco SYN flood.

Il SYN flood è un attacco di tipo *denial of service (DOS)*, mirato cioè all'esaurimento delle risorse di un sistema destinate ad un servizio, in modo tale da rendere impossibile l'erogazione di quel servizio.
Esso sfrutta il meccanismo utilizzato dal protocollo TCP per stabilire una connessione.
Quando due host utilizzano una connessione TCP per comunicare tra di loro, avviene uno scambio di alcuni messaggi, chiamati di *handshaking*, al fine di sincronizzare i valori necessari al corretto funzionamento della comunicazione.
Consideriamo ad esempio che un client A voglia stabilire una connessione con un Server B che è in ascolto sulla porta 80. I passi che devono essere compiuti sono i seguenti:

1. Il client A, per comunicare che vuole stabilire una connessione sulla porta 80, invia un pacchetto al Server B, che contiene al suo interno un SYN (flag SYN);

2. Dopo aver ricevuto il pacchetto sulla porta in ascolto, il Server B risponde al client A con un messaggio, che ha al suo interno i flag SYN e ACK;

3. Il client A, ricevendo il messaggio di risposta dal Server B, risponde a sua volta con un terzo pacchetto di conferma (ACK).

Questa procedura è chiamata *three way handshake* ed è utilizzata per stabilire ogni connessione TCP.
Se per qualche motivo il client non è in grado di inviare il riscontro finale (ACK) -o questo riscontro viene perso-, il Server si trova in uno stato intermedio (SYN-RECEIVED), in cui non può far altro che attendere il messaggio di conferma e rinviare il SYNACK alla scadenza di un time-out (*Retransmission Time Out*), per richiedere nuovamente conferma al client.
Lo scopo dell'attacco SYN flood è quello di riempire la coda delle connessioni attive (*backlog queue*) su una determinata porta del Server in modo da rendere il servizio inaccessibile per le nuove connessioni in arrivo.
Un utente malevolo che effettua questo tipo di attacco invia numerosi pacchetti SYN verso un altro host e fa in modo di non inviare il riscontro finale (ACK).

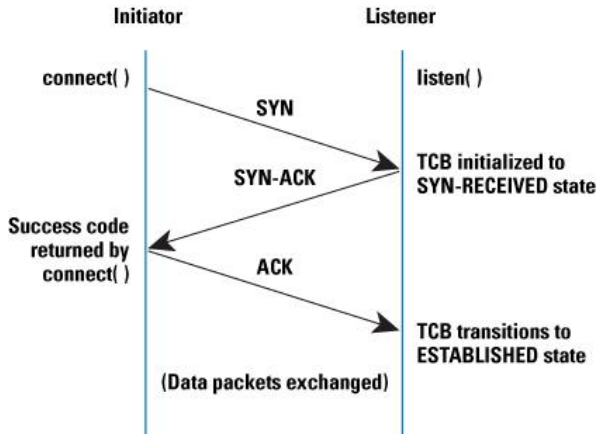

fig. 1 (a). Three-way handshake.

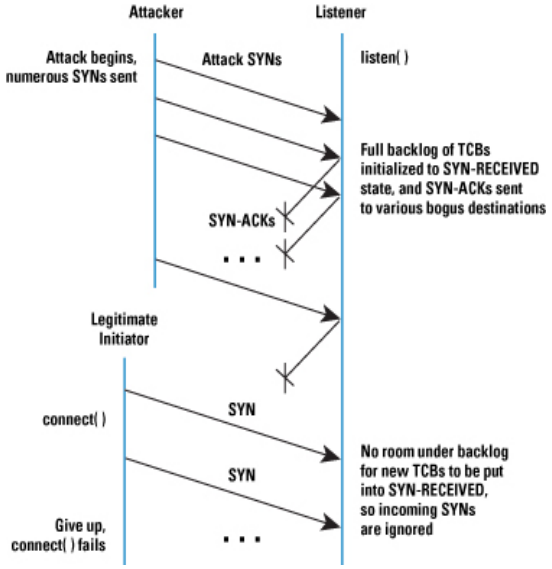

fig. 1 (b). SYN flood con IP spoofing.

Una variante consiste nell'inviare i pacchetti SYN con indirizzi IP diversi dal proprio (IP *spoofing*), in modo che il SYNACK dell'host attaccato sia inviato verso destinatari che non hanno fatto alcuna richiesta e che quindi non trasmetteranno alcun ACK di risposta.
In entrambi i casi il bersaglio dell'attacco non vedendo l'ACK aspetta lo scadere di un time-out (RTO) prima di rinviare il SYNACK. Il meccanismo di ritrasmissione si ripete per un numero prefissato di volte prima di considerare la connessione non più attiva e liberare la *backlog queue*.

2 Una possibile soluzione.

Un modo semplice per identificare l'attacco è quello di considerare la differenza tra il numero di pacchetti SYNACK e SYN in un intervallo non troppo lungo. Se durante questo spazio di tempo la differenza supera una soglia prestabilita allora si rileva l'attacco.
Analizzando il valore SYNACK − SYN si presentare diverse situazioni:

1. Se SYNACK − SYN= 0 si verifica il caso ideale (e anche irrealistico) in cui tutte le connessioni sono stabilite senza problemi;

2. Se SYNACK ≫ SYN probabilmente si sta verificando un attacco SYN flood. Il numero di pacchetti SYNACK cresce velocemente a causa dalle ritrasmissioni;

3. Se SYNACK ≪ SYN allora si possono presentare due sotto casi:

   3.1. È in atto uno scanning delle porte;

   3.2. Si tratta di un attacco SYN flood, il cui numero di pacchetti SYN è così alto da superare il numero di SYNACK. Questo può verificarsi anche quando la *backlog queue* è piena e non potendo aggiungere nuove connessioni scarta i pacchetti SYN.

Per rilevare un attacco di SYN flood si è scelto di considerare quanto la differenza SYNACK−SYN cresce/decresce in un secondo.

2.1 Strumenti di sviluppo.

Nello svolgimento del progetto sono adoperati i due strumenti presentati durante il corso: la libreria pcap e RRDtool. RRDtool è uno strumento che permette di lavorare su un insieme di dati ordinati rispetto al tempo (*time series*) e di salvarli in un database chiamato RRD. Un RRD è un *Time Series DataBase* (TSDB) che divide la *timeline* in intervalli di tempo chiamati *step* e si aspetta di ricevere dei dati per ognuno di questi.
Al momento della creazione del database, oltre a definire la durata degli step, viene richiesto il tipo dei dati (COUNTER, GAUGE, DERIVE, etc.), quale intervallo di dati considerare (numero di step) e come processarli (AVERAGE, MAX, MIN, etc.). Infine è possibile visualizzare dei grafici, che rappresentano lo scambio dei pacchetti durante un periodo circoscritto, permettendo un'analisi rapida e intuitiva del dato.

2.2 Sviluppo del programma.

Sulla base della soluzione proposta è stato implementato un programma che si occupa di catturare i pacchetti SYN e SYNACK, di calcolarne la differenza e mostrare in tempo reale un grafico del comportamento assunto.
Il software, scritto in linguaggio C, è costituito dal singolo file sorgente `sfd.c`.
Le prime operazioni all'interno del *main* riguardano l'acquisizione di alcuni parametri (*cfr. 2.3*) che possono essere personalizzati dall'utente prima dell'esecuzione.
All'avvio viene inizializzato il contatore della differenza `diff` e creato un RRD i cui step hanno la durata di un secondo. Ad ogni step è immesso il valore corrente della differenza nel *Data Source* (DS) `Rate`, definito con il tipo DERIVE. I dati non necessitano di essere raggruppati per l'analisi e vengono considerati validi per un arco di 60 secondi, dopo il quale sono eliminati.
L'esecuzione equivalente da linea di comando è la seguente:

```
rrdtool create rrd_name --start=now --step=1  DS:Rate:DERIVE:10:U:U  RRA:LAST:0.999:1:60
```

Nella cattura dei pacchetti sono impiegate le funzioni della libreria pcap.
Un filtro creato allo scopo identifica i segmenti TCP che hanno i flag SYN e SYNACK. Viene quindi definito un *thread* ed avviata la cattura vera e propria. Ogni volta che viene catturato un SYNACK il contatore della differenza riporta un incremento, mentre viene decrementato alla cattura di un SYN.
Il *thread* invece si occupa di aggiornare ad ogni secondo il database RRD e di generare un allarme se viene superata la soglia. Se la soglia è superata per 3 intervalli consecutivi il *thread* genera un ulteriore allarme.
Il valore della soglia è 40, che è puramente convenzionale, stabilito in base al numero delle connessioni possibili.
Infine ogni 5 secondi viene generato il grafico che permette all'utente di visualizzare lo stato corrente.

2.3 Utilizzo del programma.

Il tool è utilizzabile per individuare un possibile attacco SYN Flood su un interfaccia ethernet che utilizzi IPv4. Una volta compilato il codice sorgente è possibile avviare l'eseguibile con alcune semplici opzioni i cui argomenti sono obbligatori:

Opzione	**Descrizione**
`-i interface name`	Di default il programma avvia la cattura dei pacchetti sull'interfaccia di loopback. Tramite questa opzione è possibile scegliere l'interfaccia desiderata
`-n RRD name`	Per indicare il nome del database che verrà creato tramite RRDtool. Se questa opzione non viene utilizzata il programma creerà un RRD con il nome `myrrd.rrd`
`-g image name`	Se si vuole indicare il nome e il formato che deve avere il file immagine contenente il grafico. Per conoscere i formati disponibili vedere la documentazione di RRDtool. Il valore di default in questo caso è `grafico.png`
`-t value`	Per modificare il valore della soglia. Il valore deve essere un interno positivo e maggiore di 0. Il valore di default del programma è 40.

Tab. 1. Le opzioni di configurazione del programma.

2.4 Lettura e interpretazione del grafico.

Nel grafico sono rappresentate, tramite aree colorate, le oscillazioni del valore prestabilito SYNACK – SYN durante l'arco di 60 secondi. Nell'asse X sono indicati gli intervalli temporali, mentre l'asse delle ordinate individua la variazione di SYNACK – SYN in un secondo.
I colori delle aree forniscono un'informazione immediata sullo stato delle connessioni nel periodo in esame. Il verde indica valori nella norma ("Good"), il giallo l'avvicinarsi al valore della soglia prestabilita ("Warning") e il rosso che questa soglia è stata superata ("Bad").
Altre considerazioni permettono di comprendere l'intensità e la durata dell'attacco. Se ad esempio il diagramma presenta un'unica area continua ed omogena di colore rosso al di sotto dell'asse della X per un intervallo notevole (Fig. 2), questo significa che il numero di pacchetti SYN inviato supera ampiamente la soglia (*SYN flood* o *Port Scanning*).
Quando invece al di sotto dell'asse X le aree rosse variano per brevi intervalli (Fig. 3), allora il numero di pacchetti SYN è moderato e supera di poco la soglia. Le oscillazioni sono pertanto dovute alle ritrasmissioni dei SYNACK.
Infine se le aree "anomale" sono isolate e collocate al di sopra dell'asse X (Fig. 4), allora è stato inviato un numero di pacchetti SYN elevato per un brevissimo intervallo di tempo. I picchi rossi e gialli sono dovuti alle ritrasmissioni.

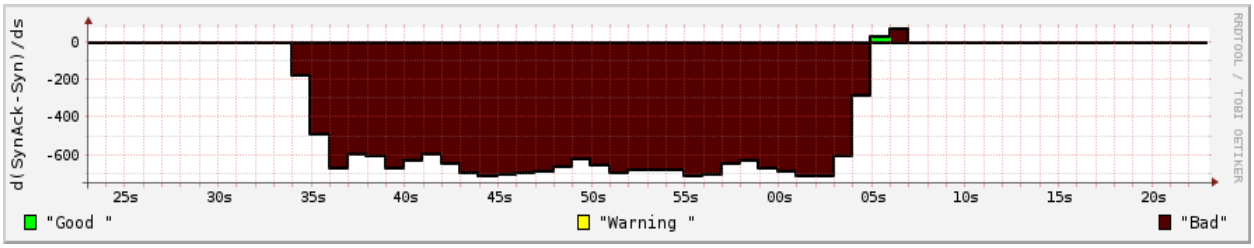

fig. 2. SYN flood effettuato con 1000 pacchetti al secondo. Soglia 40.

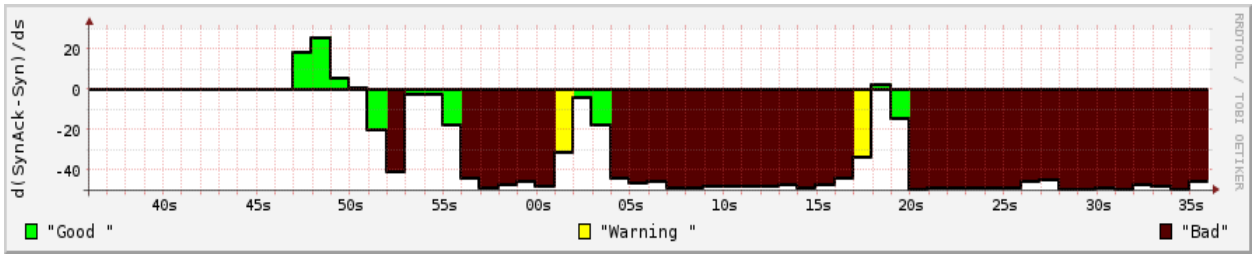

fig. 3. SYN flood effettuato con 50 SYN al secondo. Soglia 40.

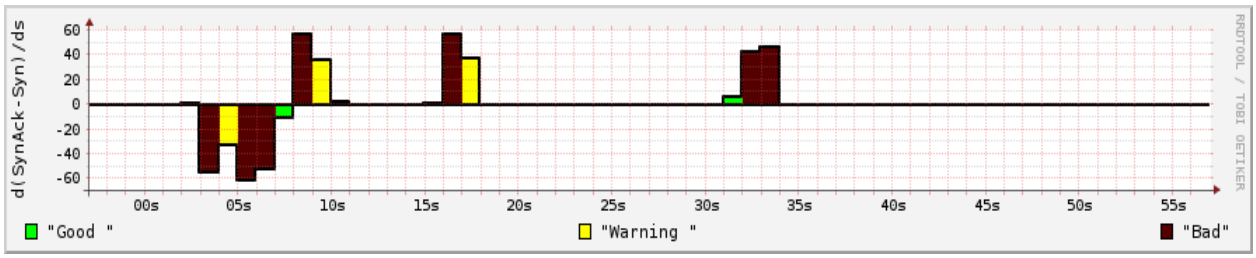

fig. 4. SYN flood effettuato inviando 100 SYN al secondo per 5 secondi. Soglia 40.