William Guglielmo, matricola 516326

Progetto GR: rilevazione traffico TINC su nDPI

Codice

Cliccare qui.

Introduzione

Il progetto consiste in un' estensione di nDPI, in cui è stata aggiunta la rilevazione del protocollo usato da TINC.

nDPI è una libreria per effettuare deep-packet inspection.

TINC è un demone VPN (Virtual Private Network) che utilizza tunneling e crittografia per creare una rete privata protetta tra gli host su Internet.

Protocollo di TINC

TINC prevede due flussi, uno usa TCP e l'altro UDP. Il protocollo è P2P, entrambi i nodi hanno un server ***sulla stessa porta*** (di default la 655) sia per il flusso su TCP che per quello su UDP.

- Flusso su TCP

  Questo flusso è usato per stabilire la connessione. ***Prevede inizialmente l' autenticazione dei due nodi***. Solo dopo la fase di autenticazione tutto il traffico di questo flusso viene criptato e vengono scambiati meta-dati e varie informazioni.

- Flusso su UDP

Tutti i datagrammi scambiati su questo flusso sono criptati fin da subito. Questo flusso viene avviato a seguito della fase di autenticazione del flusso su TCP e qui vengono scambiati i dati che viaggiano all' interno della VPN. I datagrammi hanno come porta sorgente la porta del server UDP del nodo che sta inviando il datagramma e come porta destinazione la porta del server UDP a cui si sta inviando il datagramma.

Strategia di rilevazione

- Flusso su TCP

  Per la rilevazione del flusso su TCP è stata sfruttata la procedura di autenticazione descritta qui. In particolare il flusso viene considerato TINC quando vengono rilevati i primi quattro messaggi.

  - All'avvio della connessione TCP, entrambi i nodi inviano un pacchetto con il seguente payload:

    0 HOSTNAME 17

    dove

    - 0 è il numero di sequenza del messaggio;
    - HOSTNAME è una stringa contenente il nome dell' host associato al demone che inizia la connessione, in particolare è una stringa priva di spazi e ritorni a capo, differente per ogni host della VPN;
    - 17 è la versione del protocollo di TINC;
    - termina con un ritorno a capo.

  - In seguito ai primi due messaggi, entrambi i nodi inviano un pacchetto con il seguente payload:

    1 NUM1 NUM2 NUM3 NUM4 STRING

dove:

- 1 è il numero di sequenza del messaggio;
- NUM1, NUM2, NUM3 sono tre numeri contenenti informazioni sulla chiave;
- STRING è una stringa formata solo da lettere maiuscole e numeri;
- termina con un ritorno a capo.

Attenzione

La connessione è P2P, quando uno dei due nodi si connette a un altro entrambi inviano il primo messaggio, e quando lo hanno ricevuto entrambi inviano il secondo.

L' unico modo per identificare il nodo che avvia la connessione è tramite il pacchetto contenente il solo flag SYN settato, è quindi necessario catturare anche i pacchetti privi di payload.

- Flusso su UDP

  Non è possibile effettuare la rilevazione di questo flusso tramite l' analisi dei datagrammi UDP, in quanto completamente criptati.

  Per rilevarlo ho sfruttato il fatto che la porta del server del flusso UDP e del server del flusso TCP è la stessa porta. Un flusso su UDP viene considerato TINC quando viene rilevato il corrispettivo flusso su TCP e ricevuto il primo datagramma del flusso su UDP.

  Viene aspettata la ricezione del primo datagramma del flusso UDP, perché al momento della rilevazione del flusso TCP non si hanno ancora tutte le infomazioni che identificare il flusso su UDP.

  Considerando come:

- nodo1: il nodo che avvia la connessione;
- nodo2: il nodo a cui il nodo1 sta cercando di connettersi.

Al momento della rilevazione del flusso TCP si conosce:

1. l' indirizzo IP del nodo1;
2. l' indirizzo IP del nodo2;
3. la porta del server UDP del nodo2;
4. il protocollo usato (si presume TINC).

Non si conosce ancora la porta del server UDP del nodo1, questo perché i pacchetti TCP con cui abbiamo rilevato il flusso su TCP avevano come porta associata al nodo1 una porta fittizia.

Appena viene catturato il primo datagramma avente per sorgente o destinazione la porta e l' indirizzo IP del nodo2, e per indirizzo IP rimanente quello del nodo1, si considera la porta del nodo1 del datagramma come la porta mancante del flusso. Il flusso viene quindi identificato come TINC.

Problemi riscontrati

Per rilevare il flusso UDP, alla ricezione di un datagramma è necessario controllare se è stato rilevato il flusso su TCP corrispondente. Per far questo ho dovuto introdurre dello stato globale, inserendo nella struttura ndpi_detection_module_struct una semplice cache LRU di dimensione limitata per evitare un eccessivo assorbimento di memoria.

Test

Ho aggiunto nella test-suite di nDPI un semplice test-case in cui vengono riconosciute due connessioni TINC parallele.