UNIVERSITÀ DI PISA

Dipartimento di Informatica

Gestione di Rete

sFlow to Influx

Luca Pippi

Corso A

Matricola 533706

Introduzione

Il progetto ha lo scopo di realizzare un tool per la raccolta dei dati statistici dei counter samples di sFlow e memorizzarli direttamente su di un database per serie temporali, nel caso specifico in Influx.

Gli agent sFlow inviano i counter samples divisi per interfaccia e contengono i seguenti contatori:

- Indice di interfaccia (`ifIndex`)
- Tipo di interfaccia (`ifType`)
- Velocità dell'interfaccia (`ifSpeed`)
- Direzione di comunicazione (`ifDirection`)
- Stato operativo interfaccia (`ifStatus`)
- Bytes in ingresso (`ifInOctects`)
- Pacchetti unicast, multicast e broadcast in ingresso (`ifInUcastPkts,ifInMulticastcastPkts,ifInBroadcastPkts`)
- Errori e pacchetti scartati in ingresso (`ifInDiscards,ifInErrors`)
- Protocolli sconosciuti in ingresso (`ifInUnknownProtos`)
- Bytes in uscita (`ifOutOctects`)
- Pacchetti unicast, multicast e broadcast in uscita (`ifOutUcastPkts,ifOutMulticastcastPkts,ifOutBroadcastPkts`)
- Errori e pacchetti scartati in uscita (`ifOutDiscards,ifOutErrors`)
- Stato "promiscuous mode" (`ifPromiscuousMode`)

Il tool ignorerà i packet samples inviati dagli agent e prenderà in considerazione i soli contatori utili per le statistiche.

Prerequisiti

Per l'utilizzo del tool è necessario aver installato:

- InfluxDB (`https://www.influxdata.com/`)
- sflowtool (`https://github.com/sflow/sflowtool`)

È necessario inoltre un agent sFlow (`https://sflow.org/developers/tools.php`), che ho allegato al tool e verrà compilato assieme al tool per i test.

Implementazione

Il tool `sflowtoinflux` sfrutta `sflowtool` per raccogliere i dati dagli agent e poi li elabora per creare delle query corrette da inviare al database InfluxDB.

Per il corretto funzionamento di `sflowtoinflux` è necessario modificare il file di configurazione `conf.ini` con il percorso di installazione di `sflowtool` sulla propria macchina, l'indirizzo del database che si vuole utilizzare e la porta sul quale gli agent stanno comunicando.
`sflowtool` viene utilizzato in modo da raccogliere solamente i dati dei counter samples e non tutto il contenuto dei pacchetti sFlow dopodiché il tool identifica dall'ouput di `sflowtool` quali sono l'agent, l'interfaccia ed il timestamp del counter sample ricevuto e crea una query di inserimento per ogni contatore.

Formato file configurazione

Il file di configurazione conf.ini deve avere il seguente formato:

```
DATABASE=<indirizzo_influxDB/write?db=nome_db>
PATH=<posizione_sflowtool>
PORT=<porta_di_ascolto>
```

È necessario mantenere questo ordine preciso dei tag per la corretta lettura del file.

Compilazione e test

Per la compilazione di `sflowtoinflux` sarà sufficiente utilizzare il comando `make`.
Per eseguire dei test si può utilizzare il comando `sudo make test`, che decomprimerà, compilerà il programma agent e ne lancerà in esecuzione un'istanza sull'interfaccia attualmente attiva della propria macchina, che avrà anche la funzione di collector in quanto infine verrà eseguito sflowtoinflux.

Per eseguire individualmente un agent è sufficiente utilizzare il comando:

```
sudo sflsp -d <interfaccia> -P -s <sampling_rate>
                                    -A <indirizzo_agent>
                                        -C <indirizzo_collector>
                                            -c <porta_collector>
```

Per arrestare tutti gli agent a fine test ho incluso uno script `killall.sh`.

Analisi con Grafana

Il tool serve solamente a memorizzare i dati su un database a serie temporale, ma non crea nessun report statistico della rete né ne valuta le prestazioni.
Per utilizzare i dati memorizzati a fini statistici è necessario il supporto di Grafana che permette di manipolare i dati contenuti in time series database e creare grafici prestazionali quasi in tempo reale.

Per dare un esempio ho creato una semplice dashboard per visualizzare alcuni dati memorizzati tramite `sflowtoinflux`; la dashboard ha 3 panel che mostrano l'andamento nel tempo di:

- Traffico bytes: confronto fra il numero di byte in ingresso ed in uscita di ogni interfaccia di ogni agent.
- Volume traffico pacchetti: confronto fra numero di pacchetti in ingresso ed in uscita di ogni interfaccia di ogni client.
- Utilizzo banda: percentuale di occupazione della banda, confronto fra percentuale di utilizzo per input e per output.