UNIVERSITÀ DEGLI STUDI DI PISA
DIPARTIMENTO DI INFORMATICA

Cron Research

Un Sistema di visualizzazione per l'archivio storico di ntop

David Lipparelli
Mat. 407651
david.lipparelli@icloud.com

Relazione di Progetto

Gestione di Rete
2012/13

Sommario

1. Introduzione

Nel monitoraggio di rete è frequente la produzione di una grande quantità di dati legata al bisogno di tenere un archivio storico (Database), da controllare in un secondo momento oppure su bisogno, dall'amministratore di rete, in caso di comportamenti anomali della rete.

Nel caso specifico del software di monitoraggio ntopng, nuova versione dell'ormai datato ntop, prodotto da Luca Deri professore del corso di Gestione di Rete presso l'Università di Pisa, la grande quantità di dati prodotta viene interrogata e rappresentata mediante un interfaccia web e una rappresentazione grafica.

2. Soluzione al problema

La soluzione pensata, risolve il problema dell'interrogazione dei dati con l'implementazione di un interfaccia web fornita di selettore di data e orario facendo così un interrogazione capillare al database con precisione al minuto, che restituirà i sessanta secondi associati al minuto cercato.

Il formato dei dati associati al minuto è un insieme di record (sessanta) composti dal secondo in formato epoch (UNIX time) e dalle n coppie [indirizzo IP:traffico generato] una per top host di quel secondo.
Queste informazioni sono elaborate e inserite in un grafico a barre impilate che ha sulle ascisse il tempo in secondi e sulle ordinate il traffico in byte, le barre sono colorate con colori diversi, e ogni colore corrisponde ad un diverso top host.

3. Implementazione del software e tecnologie utilizzate

L'architettura del sistema è mostrata in modo schematico in figura 1.

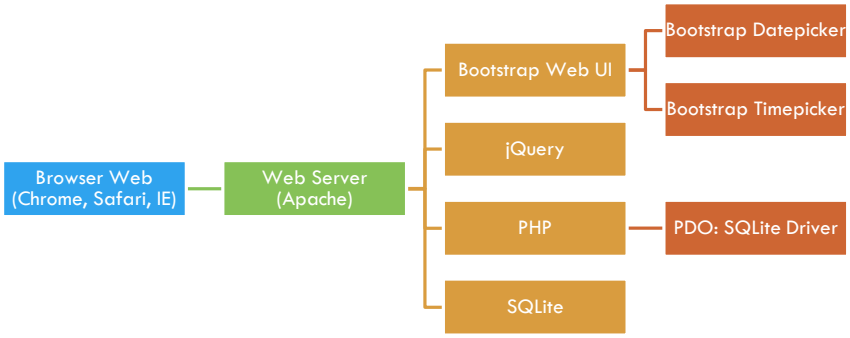

Figura 1 - Architettura del sistema

3.1. L'interfaccia Web

L'interfaccia web è basata sulla libreria *Twitter Bootstrap*, che mette a disposizione una serie di componenti di utilizzo comune nella costruzione dei siti internet, come ad esempio menu, barre di navigazione, pulsanti, modal ecc.... Per una lista completa delle funzionalità offerte si rimanda alla documentazione sul sito della libreria http://twitter.github.io/bootstrap/.
Sono presenti anche altre librerie di utilità come la *jquery*, la *Bootstarp Datepicker* e *Bootstrap Timepicker*, tutte queste librerie sono corredate dai rispettivi fogli di stile CSS.

Il layout dell'interfaccia è minimale e diviso in due parti, in alto alla pagina html è presente la barra di navigazione composta dal selettore di data da ora in poi chiamato datepicker, il selettore di orario da ora in poi chiamato timepicker, un pulsante per la ricerca, un pulsante che attiva la modalità di popolazione pseudo-casuale del database, un menu a tendina con le voci sulla documentazione e un

pulsante per le opzioni non ancora presenti, sotto alla barra di navigazione è presente la sezione dei risultati che mostra il grafico e la tabella che visualizzerà i valori selezionati sul grafico. Uno screenshot dell'interfaccia è in figura 2.

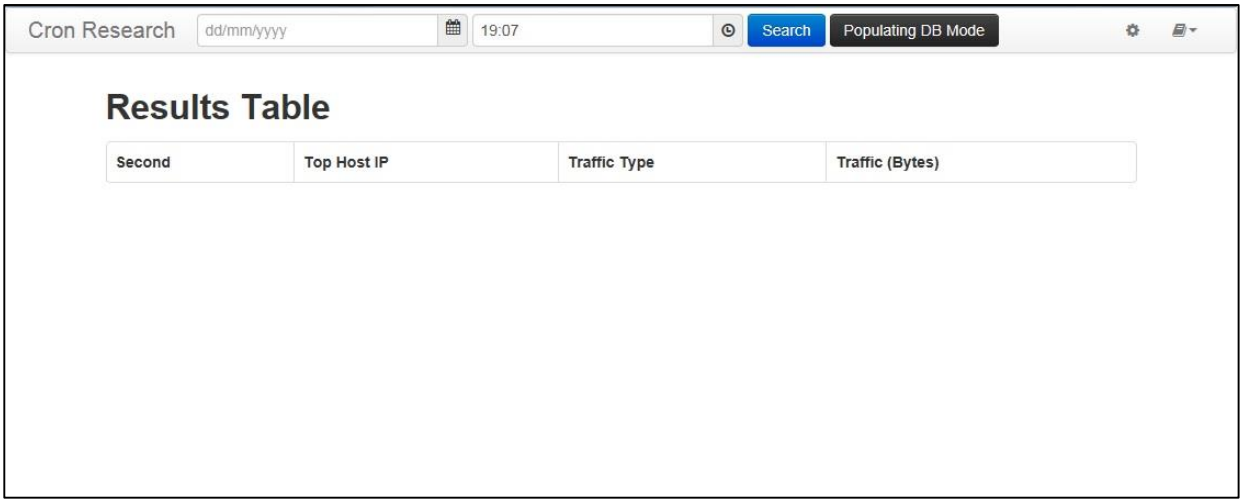

Figura 2 - Interfaccia web con barra di navigazione e sezione dei risulati

3.1.1. Il pulsante di ricerca

La query di ricerca viene composta leggendo i valori selezionati dall'utente dal datepicker e dal timepicker, viene poi creata una stringa formattata nel formato "*anno-mese-giorno-ora-minuti*" ed inviata tramite una richiesta Ajax al Web Server.

3.1.2. Il pulsante di popolamento

Il pulsante attiva la modalità popolamento che mostra un form composto da due datepicker e due timepicker per selezionare un intervallo temporale e un pulsante di invio richiesta denominato: *Populate DB*, che invia la richiesta al Web Server.
In figura 3 è mostrato lo screenshot della modalità popolamento.

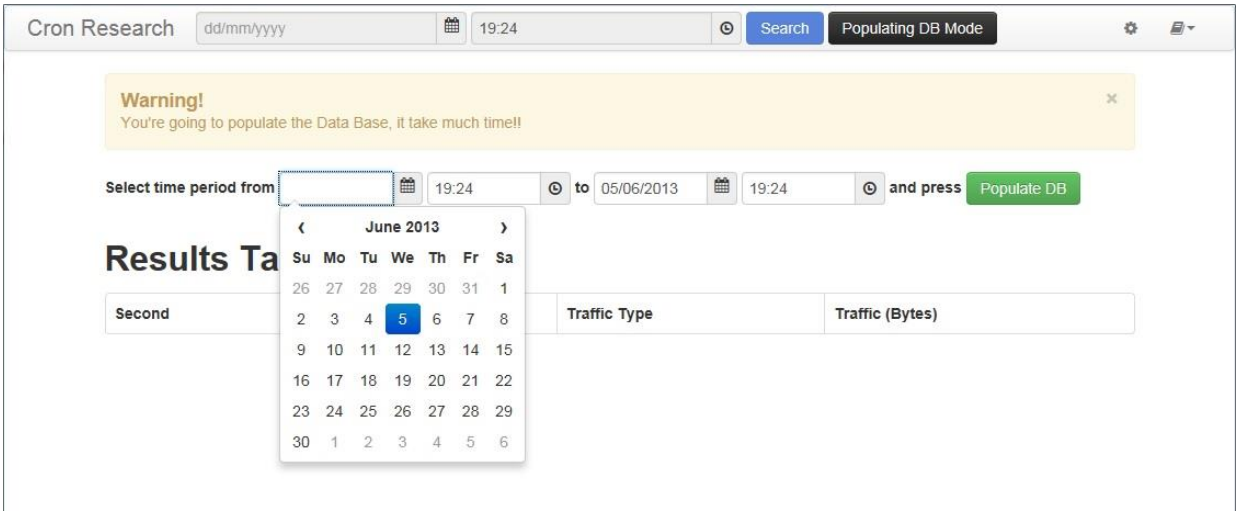

Figura 3 - Modalità popolamento del database

3.1.3. La visualizzazione del grafico e dei risultati

Il grafico è visualizzato sotto la barra di navigazione e subito sopra è presente la tabella dei risultati che si riempie man mano che sono selezionate le barre del grafico, fornendo informazioni aggiuntive quale: indirizzo IP, traffico generato dall'host, graduatoria dell'host e secondo temporale, in figura 4 è presente uno screenshot.

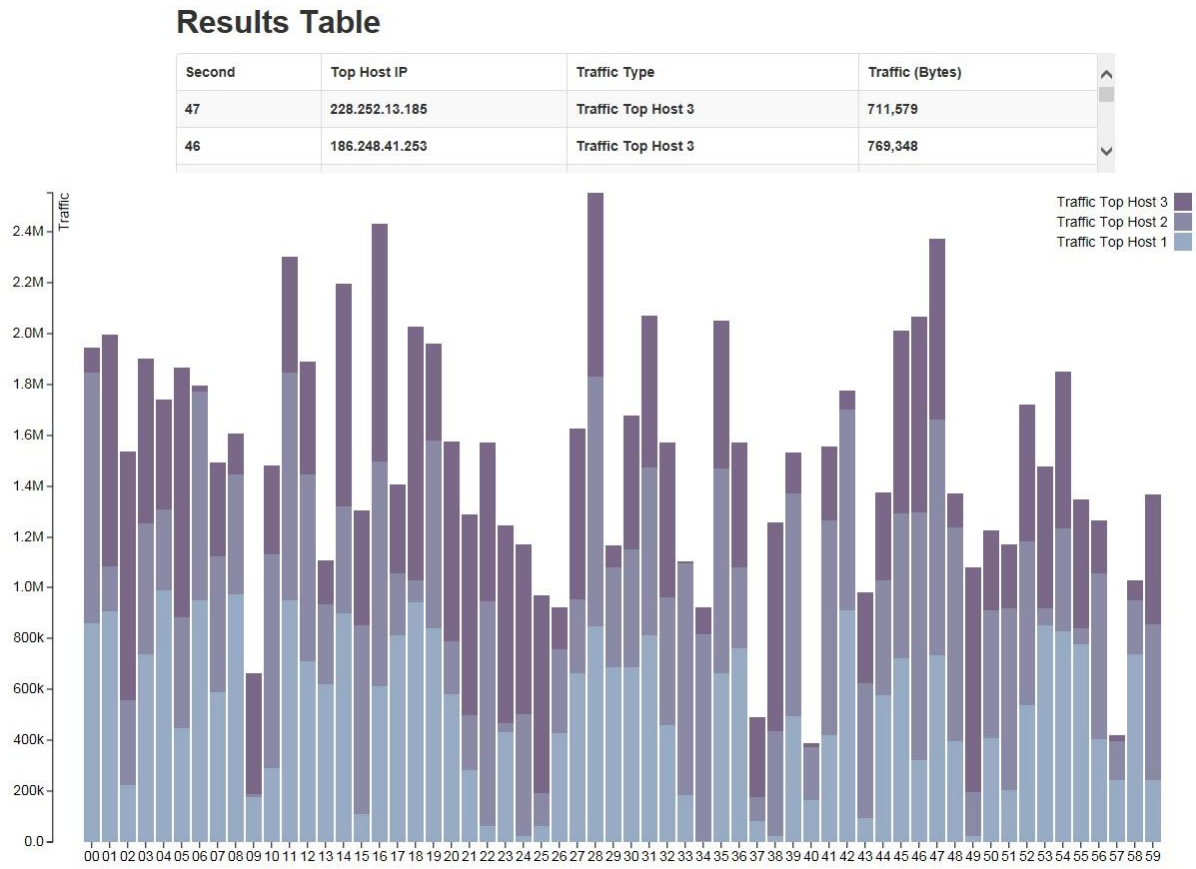

Figura 4 - Grafico con tabella dei risultati

3.2. Il Web Server e il Database

Il web server è Apache2 con il modulo per gli script php attivato, il server carica di default la pagina iniziale dell'interfaccia web *index.html* e le varie richieste Ajax di ricerca e popolamento sono servite da due script php: *searchRequest.php* e *popDB.php*, che interagiscono con il database SQLite che risiede in un unico file denominato: *cronDB.sqlite*.

3.2.1. Lo script searchRequest.php

Questo script elabora le ricerche nel database pervenute tramite richieste Ajax, la stringa precedentemente formattata nel formato "*anno-mese-giorno-ora-minuti*" è utilizzata per la creazione dell'epoch e per individuare il record corrispondente nel database, se presente vengono estratti dal database i sessanta record associati al minuto cercato.

I sessanta record sono inseriti in un file di formato *csv* (comma separated value) ed elaborati dal codice javascript che genera il grafico a barre impilate.

Il codice viene inviato come risposta della richiesta ed eseguito dal browser che mostra il grafico nell'apposita sezione.

3.2.2. Lo script popDB.php

Questo script genera in modo pseudo-casuale i sessanta record da inserire nel database. Ogni record è composto da l'epoch time del secondo e le n coppie corrispondenti a gli n top hosts che hanno generato più traffico in quel determinato secondo. L'indirizzo IP e il traffico associato al top host sono generate in modo pseudo-casuale, in particolare del traffico si può dare un limite superiore di default è impostato a 1.000.000 di byte.

3.2.3. Il database SQLite

Il database è basato su tecnologia SQLite, un formato di memorizzazione compresso che mantiene tutte le tabelle di un database all'interno di un unico file, indipendente dall'architettura e dal sistema operativo.

Le query sul database sono eseguite grazie ad una libreria sviluppata per php chiamata **PDO: PHP Data Objects**, che astrae dal database e quindi permette di accedere con le funzioni messe a disposizione da questa, su qualsiasi database esempio: **MySQL**, **PGSQL**, **Oracle**, **SQLite** ecc.… basta implementare le funzioni dell'interfaccia messa a disposizione dalla PDO.

4. Il grafico e la libreria D3js

Il grafico è costruito tramite la libreria D3js, una innovativa libreria per la creazione di grafici in formato svg di tutti i tipi, anche molto complessi, programmata completamente in codice javascript e quindi compatibile con la maggior parte dei browser esistenti, per maggiori informazioni visitare il sito internet dello sviluppatore http://d3js.org/.

In una prima fase vengono creati gli assi x e y e il container svg che conterrà gli elementi del grafico, poi i dati vengono processati dalle funzioni di mapping che associano i dati contenuti nel file csv (vedi figura 5) con le varie componenti del grafico.

```
State,Traffic Top Host 1,Traffic Top Host 2,Traffic Top Host 3
00,16.185.180.33:859558,185.165.180.201:987397,84.246.111.59:98022
01,95.231.173.202:908722,201.11.194.9:173889,140.215.185.169:911591
02,97.112.87.89:223877,39.206.149.182:335846,39.212.177.243:978455
03,207.129.121.94:740082,255.222.236.87:511841,154.187.212.227:651642
04,99.237.122.122:992005,217.30.13.33:316620,39.215.230.140:430664
05,209.229.66.74:449860,53.173.148.180:433289,130.163.12.170:984955
06,80.116.126.208:949097,100.4.57.98:824250,212.236.248.57:21057
07,36.139.143.47:588837,115.76.108.124:535889,105.81.142.36:368561
08,248.3.29.34:974061,148.203.45.107:472015,7.120.231.104:159790
09,123.132.155.7:176544,218.202.252.6:11047,44.221.255.148:478241
10,5.144.167.235:289551,176.212.56.4:841950,50.152.12.254:350769
11,196.185.156.98:951264,228.170.229.222:893921,160.249.188.146:458527
12,55.70.70.161:711975,179.214.29.14:734833,157.97.88.152:441650
13,100.44.0.170:620026,161.43.33.8:313415,209.45.26.80:172699
14,167.203.57.241:901490,147.37.91.195:420196,169.31.149.203:873841
```

Figura 5 - File csv generato dalla libreria D3js

Quindi i dati vengono prelevati dal file csv e i secondi vengono associati all'asse delle ordinate il traffico all'asse delle ascisse, e poi viene creato un array di pile dove ogni pila è un array contenente tutte le informazioni: indirizzo IP del i-esimo host, traffico in byte, secondo temporale e graduatoria dell'host esempio: top host 1 , top host 2 ecc... degli n top hosts di quel secondo e la gamma di colori a loro associata.

5. Test e conclusioni

Mostro adesso l'esecuzione del software in tutte le sue fasi:

5.1. Popolamento del database

Come si vede figura 6 e 7 il database è stato popolato con successo con due minuti di traffico.

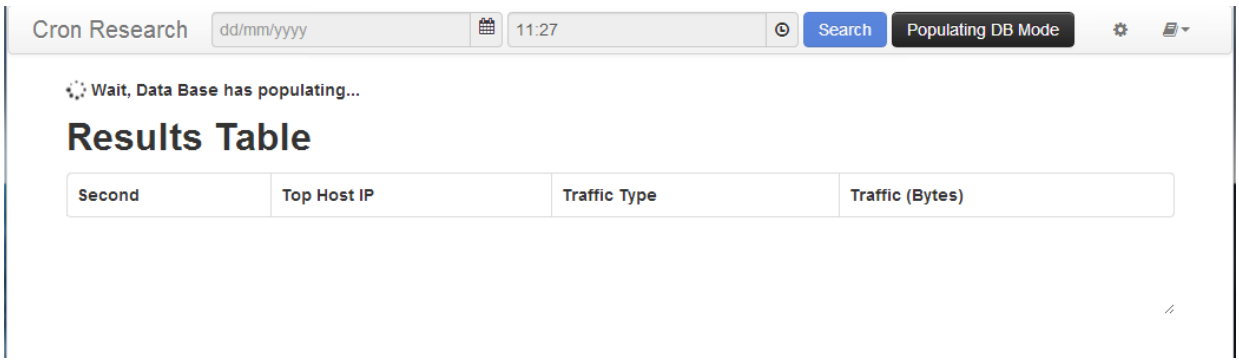

Figura 6 - Fase di popolamento del database

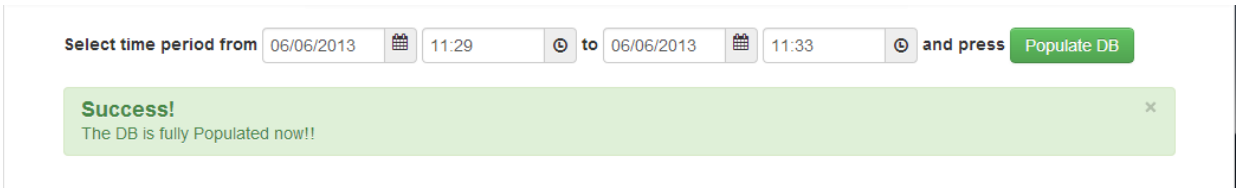

Figura 7 - Database popolato con successo

5.2. Ricerca nel database

Come mostrato in figura 8 e 9 si seleziona giorno e orario da ricercare all'interno del database.

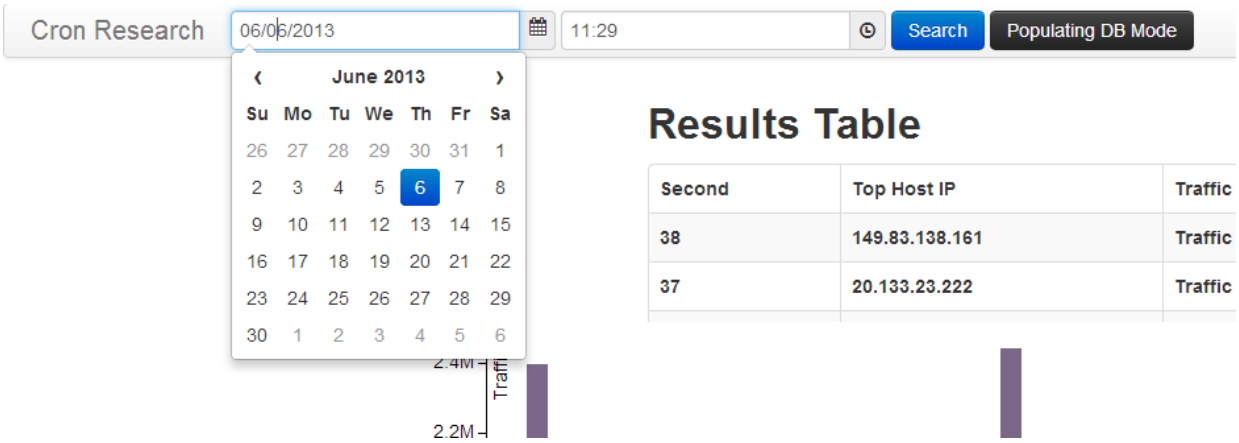

Figura 8 – Selezione della data

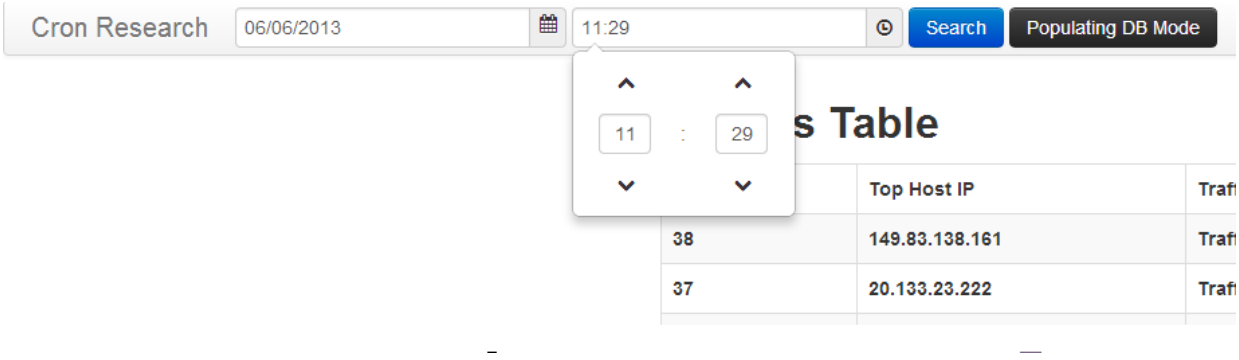

Figura 9 - Selezione dell'orario

5.3. Visualizzazione dei risultati

Come si vede in figura 10 il server restituisce il grafico corrispondente al minuto cercato, selezionando le barre con il mouse è possibile vedere i valori associati nella tabella dei risultati.

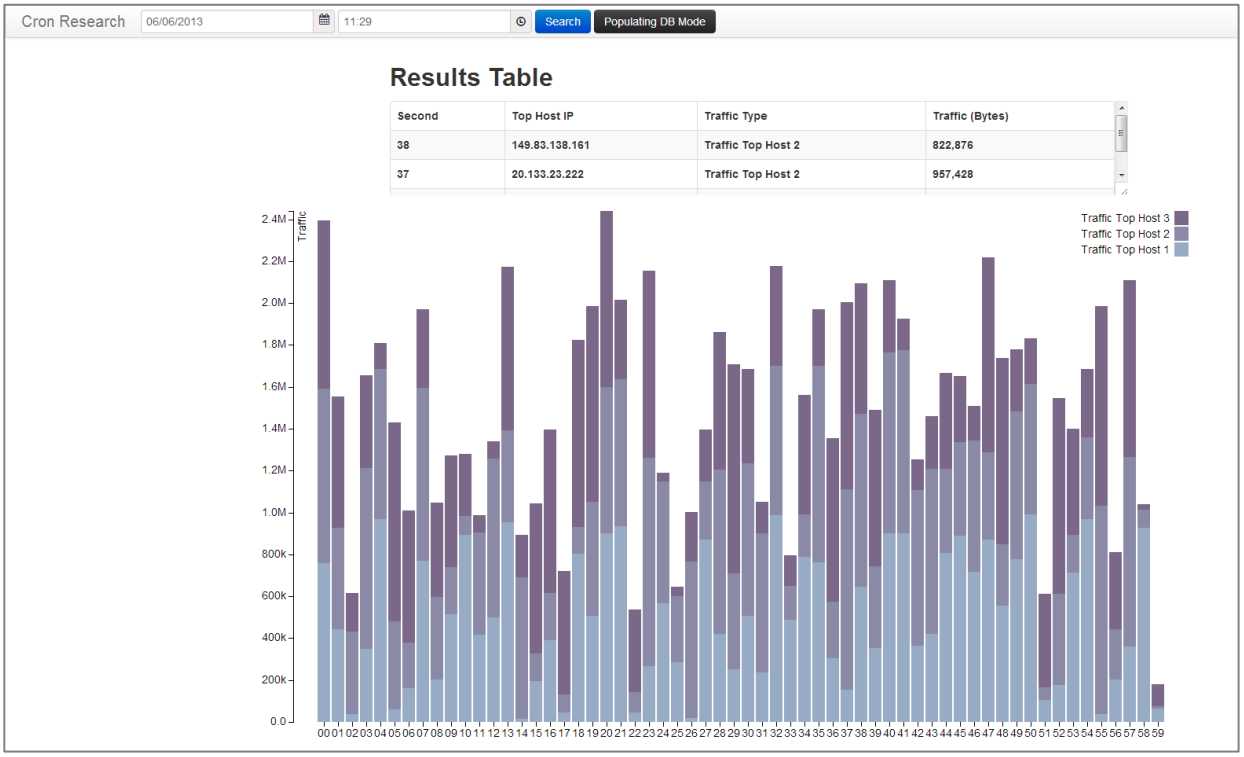

Figura 10 - Grafico e visualizzazione dei valori nella tabella