Internet Traffic Database

John Haverlack

IT Manager, School of Fisheries and Ocean Science, University of Alaska Fairbanks

jehaverlack@alaska.edu

RVTEC 2015-11-04
The Problem

- How much Internet capacity do we have?
- What is using our Internet connection?
- Are we using the capacity that we have efficiently?

Relevant to Captive Portal

- How much Internet have I used and how much do I have left?
- Am I being a responsible user?
Is there a way to measure Internet Usage?
Network Bastion Point Devices

Where is the best place to capture network traffic data?

- Cisco Switches (SNMP)
- Cisco Routers (SNMP)
- PaloAlto Firewall (XML API)
- Cyberoam Captive Portal (No API)
- PepLink Connection Aggregator (SNMP)
Network Bastion Point Devices

The best place is the **only place** we can get the data in which we are interested.

- Cisco Switches (SNMP)
- Cisco Routers (SNMP)
- **PaloAlto Firewall** (XML API)
- Cyberoam Captive Portal (No API)
- PepLink Connection Aggregator (SNMP)
The Solution

Clients → \textbf{Palo Alto (PA-500) Firewall} → Internet

\textbf{papoller.php}

A CLI PHP script that runs continuously polling the PA Firewall each minute with the PA XML-API. The Query filters only Internet bound IP traffic. Data is inserted into the \textbf{skqnetmon MySQL} database.

\textbf{R/V Sikuliaq}

School of Fisheries and Ocean Sciences

https://www.sikuliaq.alaska.edu
The skqnetmon Database

1,335,912 Internet Traffic Records in the last week, limit to 10 to get a sample of data.

Simplify the Sample Data Query Results

```
mysql> SELECT YYYY, MM, DD, hh, mm, ss, elapsed, src, dst, proto, dport, bytes_sent, bytes_received, pkts_sent, pkts_received FROM log_pa_internet_traffic_iweek LIMIT 10;
+-------+----+----+----+-----+-------------+-------------+------------+------------+-----------+-------------+-------------+-----------+-----------|
<table>
<thead>
<tr>
<th>YYYY</th>
<th>MM</th>
<th>DD</th>
<th>hh</th>
<th>mm</th>
<th>ss</th>
<th>elapsed</th>
<th>src</th>
<th>dst</th>
<th>proto</th>
<th>dport</th>
<th>bytes_sent</th>
<th>bytes_received</th>
<th>pkts_sent</th>
<th>pkts_received</th>
</tr>
</thead>
<tbody>
<tr>
<td>2015</td>
<td>9</td>
<td>30</td>
<td>23</td>
<td>9</td>
<td>6</td>
<td>16.1.0.165</td>
<td>66.58.255.34</td>
<td>tcp</td>
<td>443</td>
<td>1785</td>
<td>5617</td>
<td>17</td>
<td>18</td>
<td></td>
</tr>
<tr>
<td>2015</td>
<td>9</td>
<td>30</td>
<td>23</td>
<td>9</td>
<td>8</td>
<td>16.1.0.66</td>
<td>137.229.15.5</td>
<td>udp</td>
<td>53</td>
<td>127</td>
<td>131</td>
<td>1</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>2015</td>
<td>9</td>
<td>30</td>
<td>23</td>
<td>9</td>
<td>8</td>
<td>16.1.0.61</td>
<td>137.229.15.5</td>
<td>udp</td>
<td>53</td>
<td>78</td>
<td>82</td>
<td>1</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>2015</td>
<td>9</td>
<td>30</td>
<td>23</td>
<td>9</td>
<td>7</td>
<td>16.1.0.67</td>
<td>137.229.15.5</td>
<td>udp</td>
<td>53</td>
<td>78</td>
<td>82</td>
<td>1</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>2015</td>
<td>9</td>
<td>30</td>
<td>23</td>
<td>9</td>
<td>6</td>
<td>16.1.0.61</td>
<td>137.229.15.5</td>
<td>udp</td>
<td>53</td>
<td>78</td>
<td>82</td>
<td>1</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>2015</td>
<td>9</td>
<td>30</td>
<td>23</td>
<td>9</td>
<td>7</td>
<td>16.1.0.60</td>
<td>137.229.15.5</td>
<td>udp</td>
<td>53</td>
<td>78</td>
<td>82</td>
<td>1</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>2015</td>
<td>9</td>
<td>30</td>
<td>23</td>
<td>9</td>
<td>7</td>
<td>16.1.0.60</td>
<td>137.229.15.5</td>
<td>udp</td>
<td>53</td>
<td>78</td>
<td>82</td>
<td>1</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>2015</td>
<td>9</td>
<td>30</td>
<td>23</td>
<td>9</td>
<td>6</td>
<td>16.1.0.61</td>
<td>137.229.15.5</td>
<td>udp</td>
<td>53</td>
<td>78</td>
<td>82</td>
<td>1</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>2015</td>
<td>9</td>
<td>30</td>
<td>23</td>
<td>9</td>
<td>7</td>
<td>16.1.0.60</td>
<td>137.229.15.5</td>
<td>udp</td>
<td>53</td>
<td>78</td>
<td>82</td>
<td>1</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>2015</td>
<td>9</td>
<td>30</td>
<td>23</td>
<td>9</td>
<td>7</td>
<td>16.1.0.60</td>
<td>137.229.15.5</td>
<td>udp</td>
<td>53</td>
<td>78</td>
<td>82</td>
<td>1</td>
<td>1</td>
<td></td>
</tr>
</tbody>
</table>

10 rows in set (0.00 sec)
```
Database Tables

Data is stored in **1 table per day**.

MySQL Merge tables are used to concatenate recent data.

- Easy to clean up old data, just archive and delete tables
- Efficient to query recent data
- Able to perform comprehensive data queries
CREATE TABLE log_pa_internet_traffic_YYYY_MM_DD (  
itid SERIAL PRIMARY KEY,  
created TIMESTAMP DEFAULT CURRENT_TIMESTAMP ON UPDATE CURRENT_TIMESTAMP,  
YYYY INT UNSIGNED NOT NULL,  
MM TINYINT UNSIGNED NOT NULL,  
DD TINYINT UNSIGNED NOT NULL,  
hh TINYINT UNSIGNED NOT NULL,  
min TINYINT UNSIGNED NOT NULL,  
ss TINYINT UNSIGNED NOT NULL,  
sessionid MEDIUMINT UNSIGNED NOT NULL,  
start TIMESTAMP NOT NULL,  
elapsed MEDIUMINT UNSIGNED NOT NULL,  
src VARCHAR(15) NOT NULL,  
src_fqdn VARCHAR(255),  
dst VARCHAR(15) NOT NULL,  
dst_fqdn VARCHAR(255),  
proto VARCHAR(15) NOT NULL,  
dport SMALLINT UNSIGNED NOT NULL,  
bytes_sent INT UNSIGNED NOT NULL,  
bytes_received INT UNSIGNED NOT NULL,  
pkts_sent SMALLINT UNSIGNED NOT NULL,  
pkts_received SMALLINT UNSIGNED NOT NULL,  
UNIQUE KEY (sessionid, start, elapsed)  
) ENGINE=MYISAM;
CREATE TABLE log_pa_internet_traffic_3days ( 
    litid SERIAL PRIMARY KEY, 
    created TIMESTAMP DEFAULT CURRENT_TIMESTAMP ON UPDATE CURRENT_TIMESTAMP, 
    YYYY INT UNSIGNED NOT NULL, 
    MM TINYINT UNSIGNED NOT NULL, 
    DD TINYINT UNSIGNED NOT NULL, 
    hh TINYINT UNSIGNED NOT NULL, 
    min TINYINT UNSIGNED NOT NULL, 
    ss TINYINT UNSIGNED NOT NULL, 
    sessionid MEDIUMINT UNSIGNED NOT NULL, 
    start TIMESTAMP NOT NULL, 
    elapsed MEDIUMINT UNSIGNED NOT NULL, 
    src VARCHAR(15) NOT NULL, 
    src_fqdn VARCHAR(255), 
    dst VARCHAR(15) NOT NULL, 
    dst_fqdn VARCHAR(255), 
    proto VARCHAR(15) NOT NULL, 
    dport SMALLINT UNSIGNED NOT NULL, 
    bytes_sent INT UNSIGNED NOT NULL, 
    bytes_received INT UNSIGNED NOT NULL, 
    pkts_sent SMALLINT UNSIGNED NOT NULL, 
    pkts_received SMALLINT UNSIGNED NOT NULL, 
    UNIQUE KEY (sessionid, start, elapsed) 
) ENGINE=MERGE UNION=(log_pa_internet_traffic_2015_03_30, log_pa_internet_traffic_2015_03_31, log_pa_internet_traffic_2015_04_01) INSERT_METHOD=NO;
### Internet Usage per Day

```sql
mysql> SELECT CONCAT(YYYY, '-', MM, '-', DD) AS Date, SUM(bytes_sent)/(1024*1024) AS Sent, SUM(bytes_received)/(1024 *1024) AS Received, SUM(Elapsed)/3600 AS Elapsed FROM log_pa_internet_traffic_skq201505s GROUP BY YYYY, MM, DD ORDER BY start;
```

<table>
<thead>
<tr>
<th>Date</th>
<th>Sent</th>
<th>Received</th>
<th>Elapsed</th>
</tr>
</thead>
<tbody>
<tr>
<td>2015-3-26</td>
<td>0.0315</td>
<td>0.4121</td>
<td>0.2836</td>
</tr>
<tr>
<td>2015-3-27</td>
<td>245.6690</td>
<td>1002.3526</td>
<td>1225.7939</td>
</tr>
<tr>
<td>2015-3-28</td>
<td>486.2423</td>
<td>1672.8752</td>
<td>2324.5017</td>
</tr>
<tr>
<td>2015-3-29</td>
<td>497.1199</td>
<td>1740.7115</td>
<td>2348.3669</td>
</tr>
<tr>
<td>2015-3-30</td>
<td>500.1793</td>
<td>1566.1931</td>
<td>2673.9889</td>
</tr>
<tr>
<td>2015-3-31</td>
<td>536.1364</td>
<td>1773.3232</td>
<td>2811.7033</td>
</tr>
<tr>
<td>2015-4-1</td>
<td>602.7032</td>
<td>2071.5144</td>
<td>2277.5997</td>
</tr>
<tr>
<td>2015-4-2</td>
<td>790.8056</td>
<td>4195.2148</td>
<td>2557.1444</td>
</tr>
<tr>
<td>2015-4-3</td>
<td>1054.1565</td>
<td>2887.2841</td>
<td>2833.2742</td>
</tr>
<tr>
<td>2015-4-4</td>
<td>694.1730</td>
<td>3119.8923</td>
<td>2767.1881</td>
</tr>
<tr>
<td>2015-4-5</td>
<td>616.1213</td>
<td>2884.4760</td>
<td>2766.2283</td>
</tr>
<tr>
<td>2015-4-6</td>
<td>587.5304</td>
<td>2833.9949</td>
<td>2404.4144</td>
</tr>
<tr>
<td>2015-4-7</td>
<td>668.2687</td>
<td>2940.2123</td>
<td>2702.0506</td>
</tr>
<tr>
<td>2015-4-8</td>
<td>556.5085</td>
<td>2381.1070</td>
<td>2451.0817</td>
</tr>
<tr>
<td>2015-4-9</td>
<td>450.3210</td>
<td>2159.7080</td>
<td>1855.4378</td>
</tr>
</tbody>
</table>

15 rows in set (5.58 sec)
Internet Usage per Hour of Day
Internet Usage per Dest Protocol/Port

SIKULIAQ Internet Traffic 2015-03-27 to 2015-04-09

99% of traffic Bytes and Elapsed Time (Seconds) per Protocol and Port

R/V Sikuliaq
School of Fisheries and Ocean Sciences
https://www.sikuliaq.alaska.edu
Internet Usage per Source IP
Internet Usage per Source IP Subnet

SIKULIAQ Internet Traffic Mar 27 to Apr 20 2015

Traffic for SHPNET and SHRNET

Data (UTC)

SHIPWIDE Received  SHPNET Received  SHIPWIDE Sent  SHPNET Sent

R/V Sikuliaq
School of Fisheries and Ocean Sciences

https://www.sikuliaq.alaska.edu
Next Steps

1. Automated Webpage to help users see their usage.
2. Automate Usage Reporting back to Shore
The PAPoller Source Code and link for this presentation can be found at:

https://www.sikuliaq.alaska.edu/ops/?q=node/206