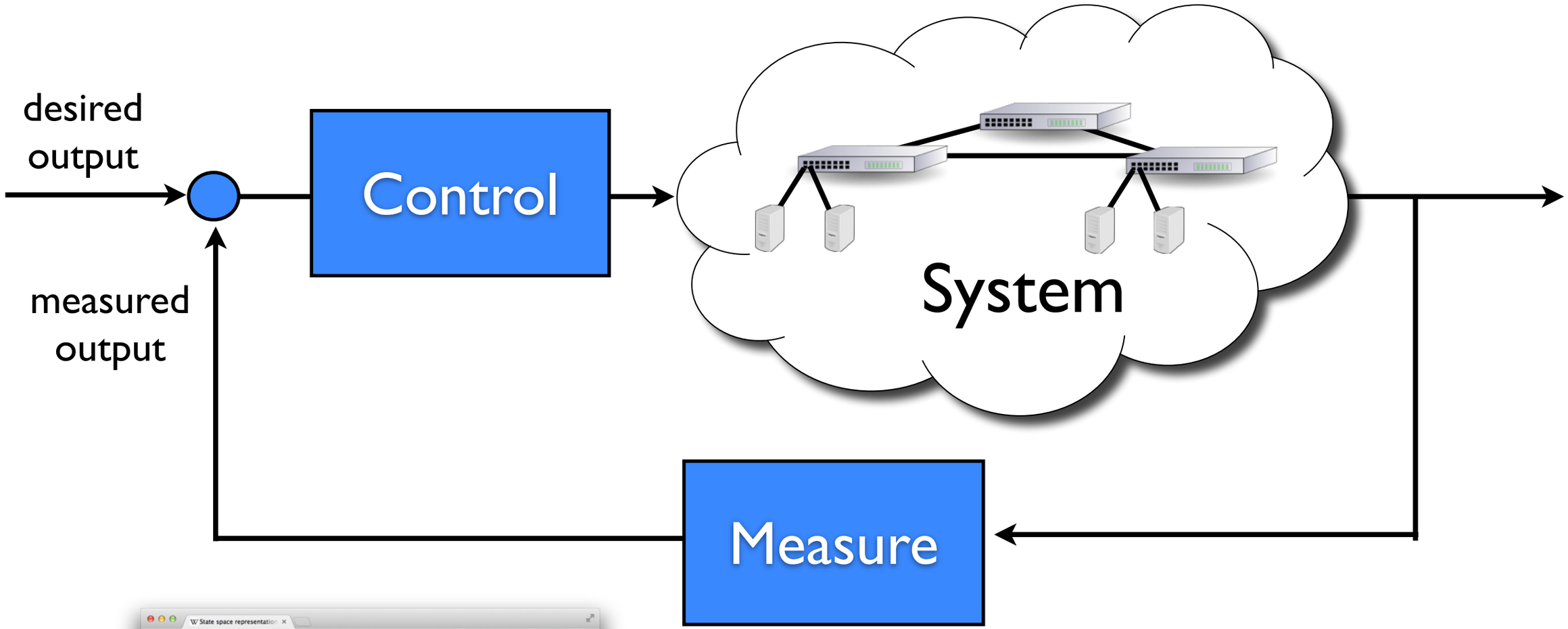
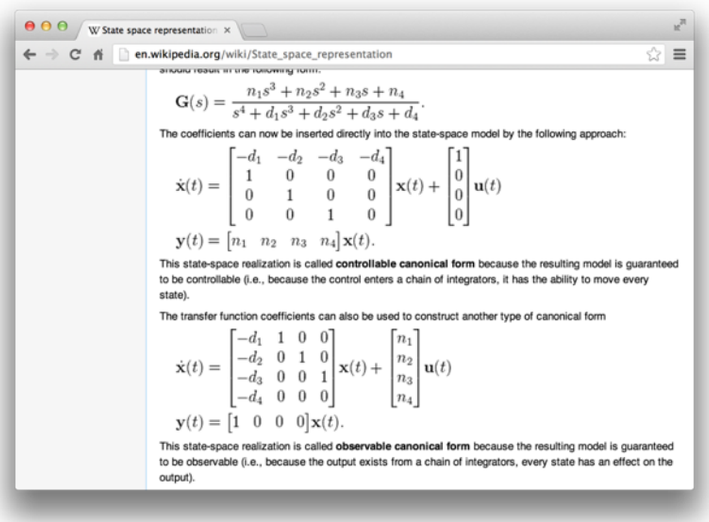
Real-time monitoring for adaptive control





Control Theory 101:
observability *required* for **controllability**

What is sFlow?

- Scalable Counter Push.
- Random 1:N sampling of network packet headers and forwarding actions.

Mature standard supported by 30+ vendors with exceptionally high levels of interoperability.

Lightweight - can run on all ports of all physical and virtual switches.

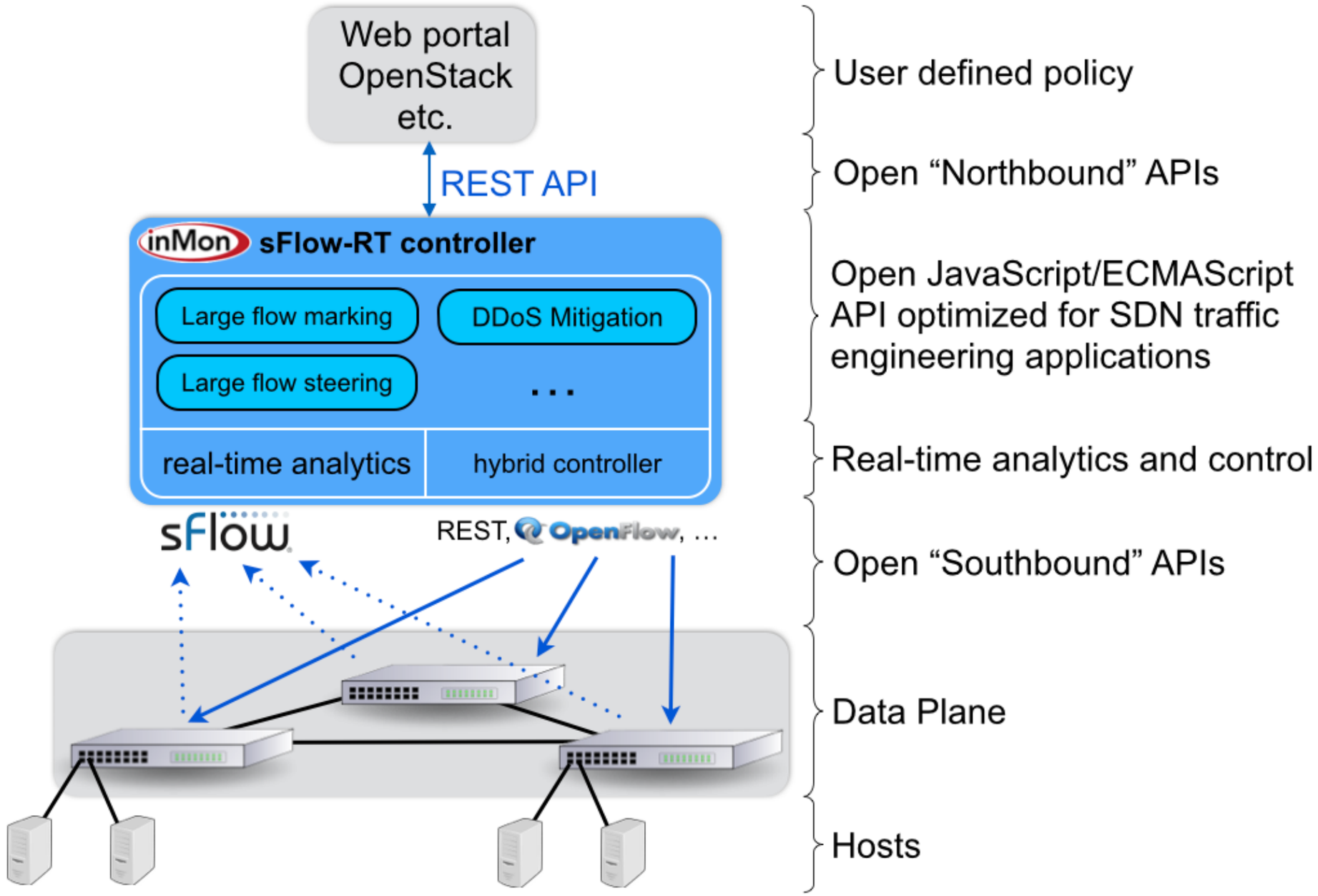
Asynchronous nature can be harnessed to provide real-time traffic visibility…

Wide Vendor Support

http://www.sflow.org/

Neil McKee neil.mckee@inmon.com SDN Workshop, LBNL July 2015

sFlow RT Real-time analytics engine

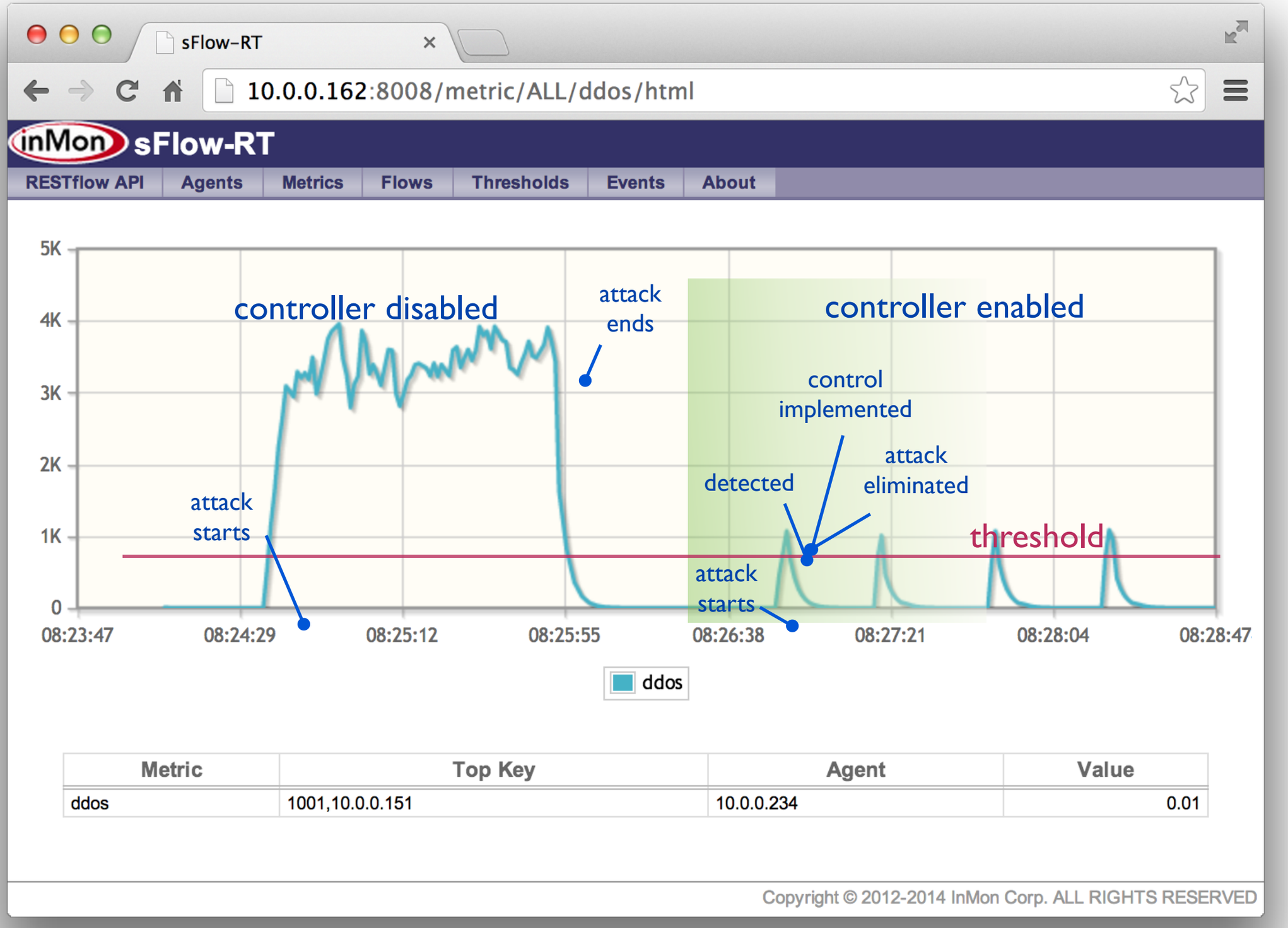


Free for non-production use: http://inmon.com/products/sFlow-RT.php

Use case examples

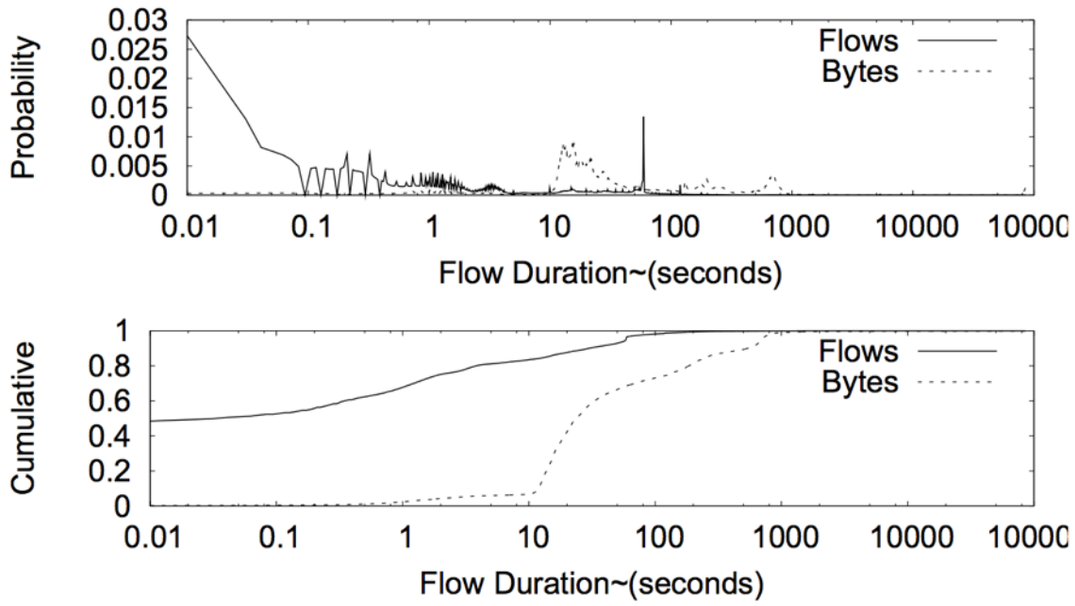
1. Sub-second DDoS Mitigation
2. Elephant Flow Marking & Steering
3. Scalable SDN Router
4. Security: Trigger Packet Capture

1. Sub-second DDoS Mitigation



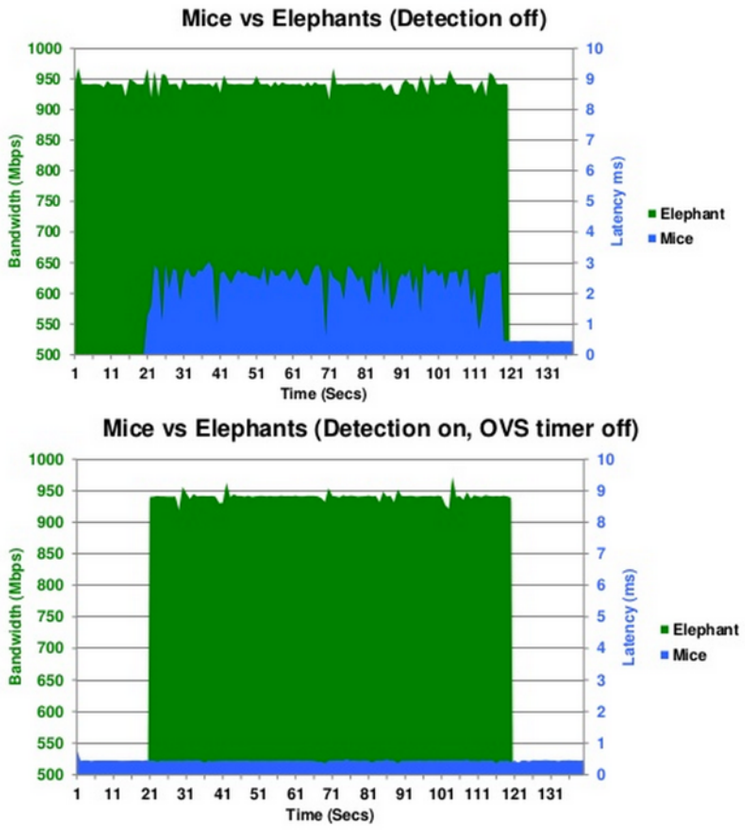
http://blog.sflow.com/2013/03/ddos.html

2. Elephant Flow marking & steering



Elephant flows are the small number of long lived large flows responsible for majority of bytes on network

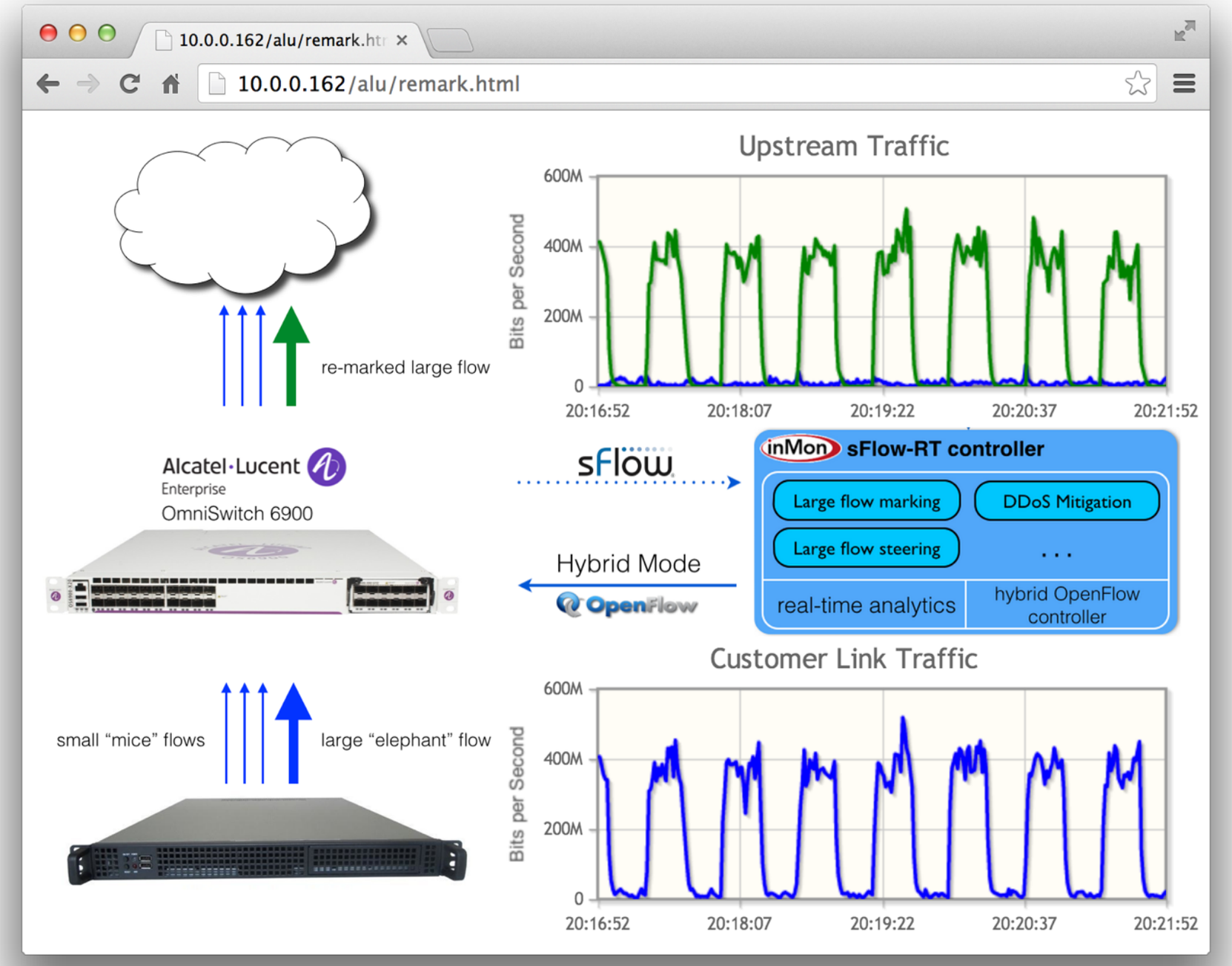
http://research.microsoft.com/en-us/UM/people/srikanth/data/imc09_dcTraffic.pdf



Marking and priority queuing, minimal impact on large flow throughput and greatly improved latency for small flows

http://www.slideshare.net/martin_casado/elephants-and-mice-elephant-detection-in-the-vswitch-with-hardware-handling

Elephant Flow marking



http://blog.sflow.com/2014/03/performance-optimizing-hybrid-openflow.html

sFlow-RT mark.js

```
// Define large flow as greater than 1Mbits/sec for 1 second or longer
var bytes_per_second = 1000000/8, duration_seconds = 1;

// get topology from Mininet
setTopology(JSON.parse(http("http://192.168.56.101/topology.json")));

// define TCP flow cache
setFlow('tcp',
 {keys:'ipsource,ipdestination,tcpsourceport,tcpdestinationport', filter:'direction=ingress',
  value:'bytes', t:duration_seconds}
);

// set threshold identifying Elephant flows
setThreshold('elephant', {metric:'tcp', value:bytes_per_second, byFlow:true, timeout:4});
```

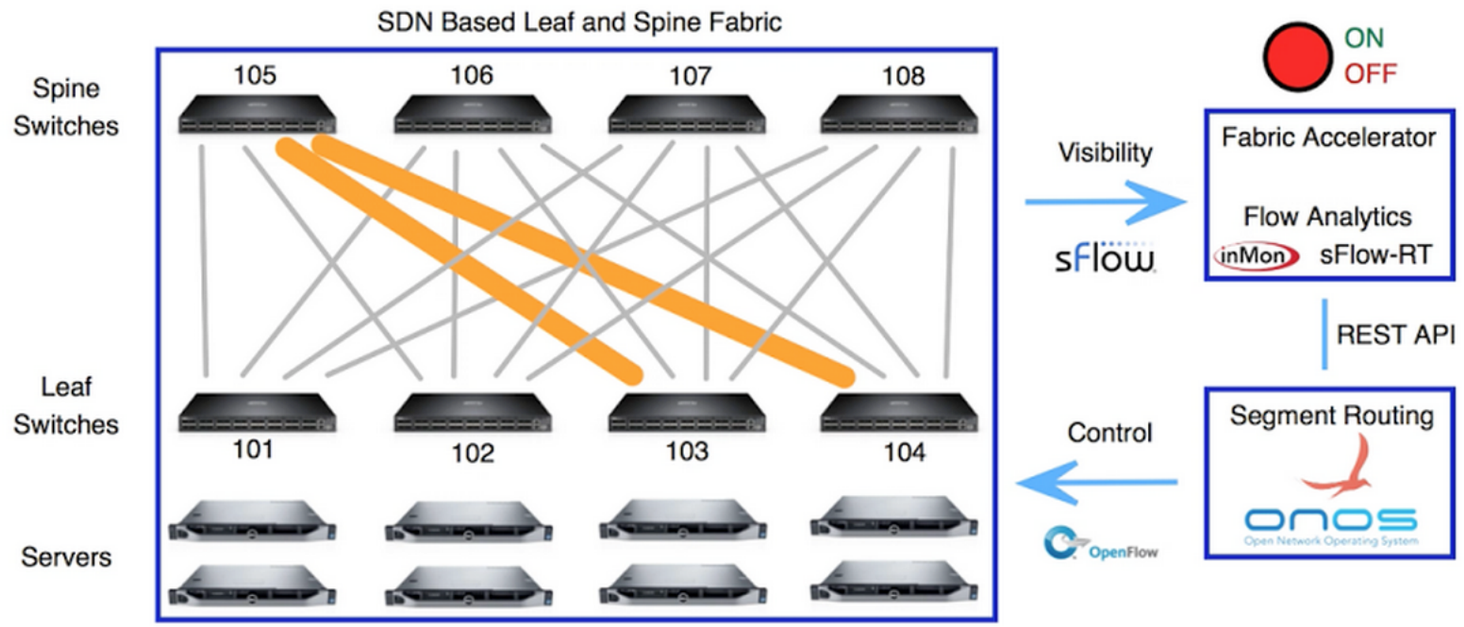
Define “Flow”

set Threshold

```
// set OpenFlow marking rule when Elephant is detected
var idx = 0;
setEventHandler(function(evt) {
 if(topologyInterfaceToLink(evt.agent,evt.dataSource)) return;
 var port = ofInterfaceToPort(evt.agent,evt.dataSource);
 if(port) {
  var dpid = port.dpid;
  var id = "mark" + idx++;
  var k = evt.flowKey.split(',');
  var rule= {
    match:{in_port: port.port, dl_type:2048, ip_proto:6, nw_src:k[0], nw_dst:k[1], tcp_src:k[2], tcp_dst:k[3]},
    actions:["set_ip_dscp=8","output=normal"], priority:1000, idleTimeout:2
  };
  logInfo(JSON.stringify(rule,null,1));
  setOfRule(dpid,id,rule);
 }
},['elephant']);
```

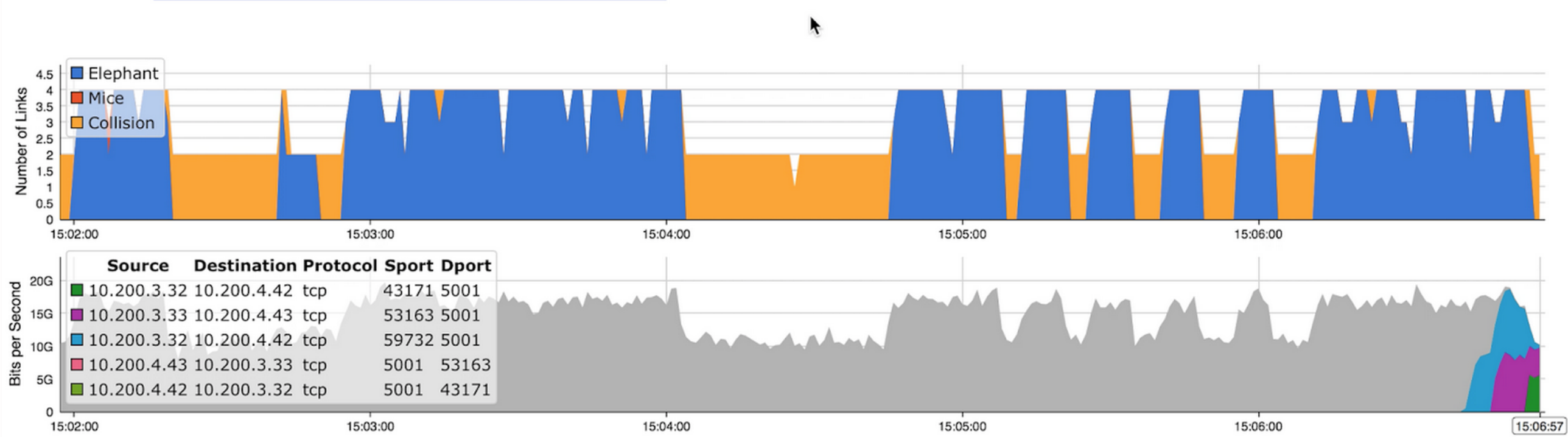
Event Handler: issue control

Elephant Flow Steering
(a) Controller Off - hash collisions impact throughput



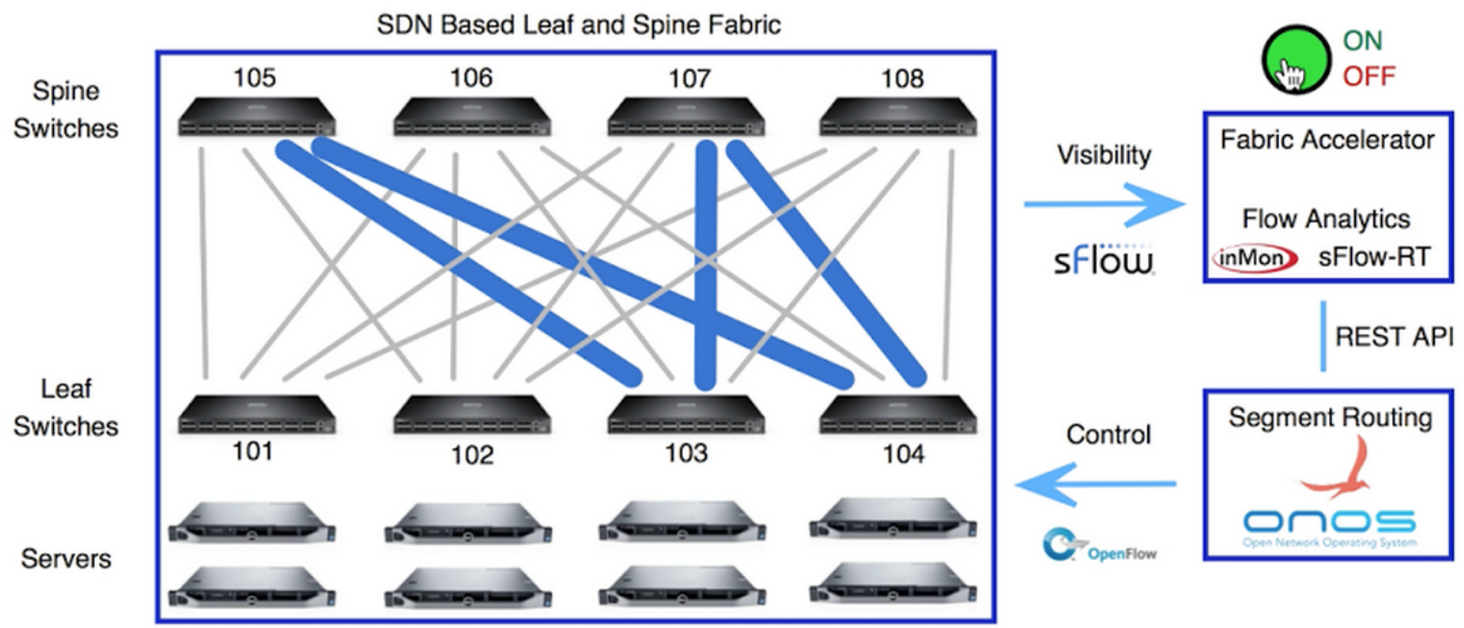
Segment Routing ECMP Override Policies

Source Destination Prot Sport Dport Tunnel



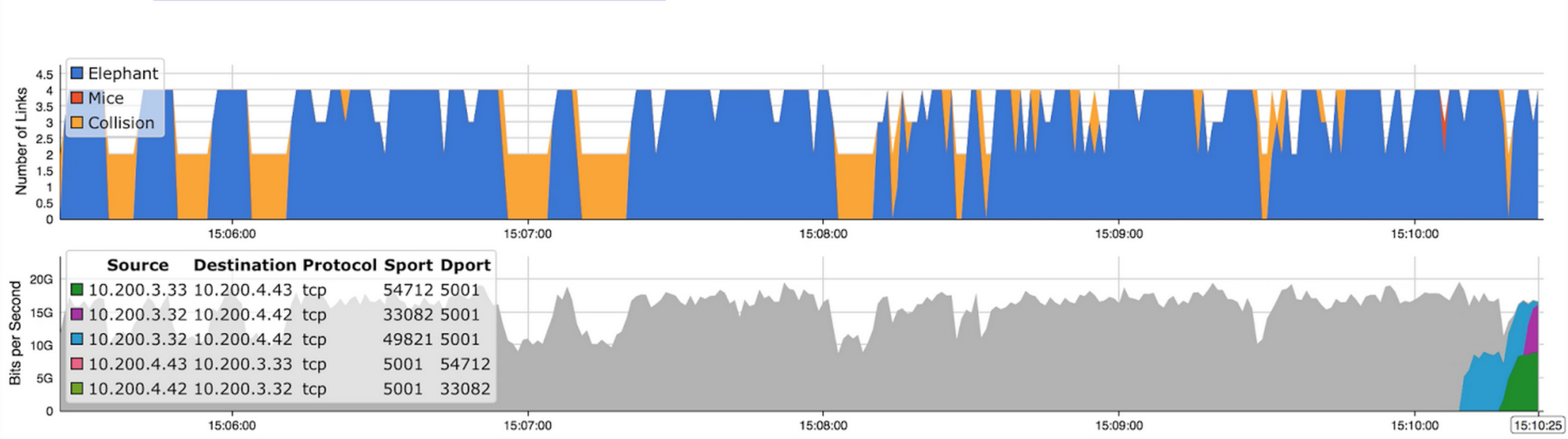
http://blog.sflow.com/2015/06/leaf-and-spine-traffic-engineering.html

Elephant Flow Steering
(b) Controller On, full throughput restored



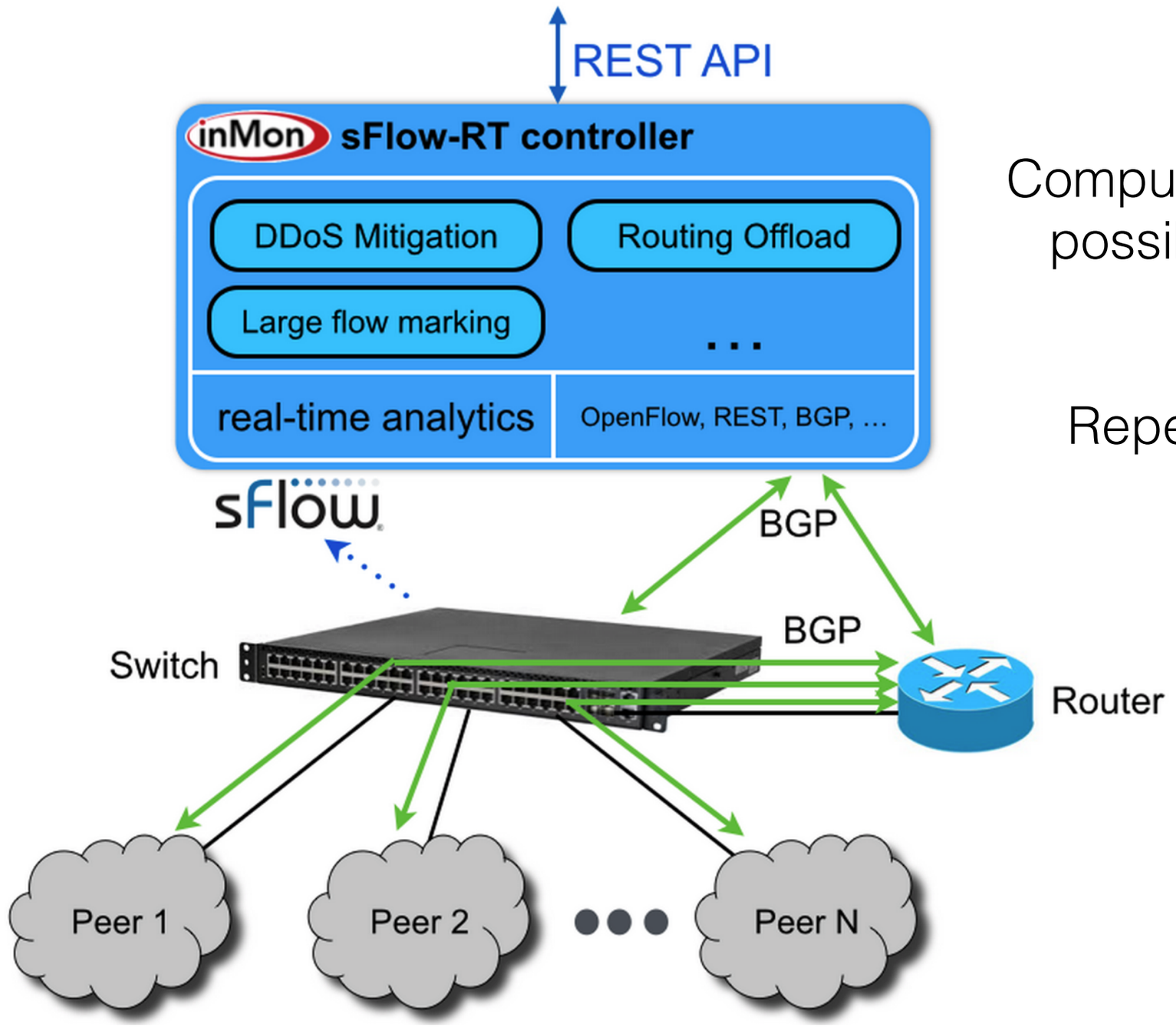
Segment Routing ECMP Override Policies

Source	Destination	Prot	Sport	Dport	Tunnel
10.200.3.33	10.200.4.43	tcp	54712	5001	103,107,104



http://blog.sflow.com/2015/06/leaf-and-spine-traffic-engineering.html

3. Scaleable SDN Router



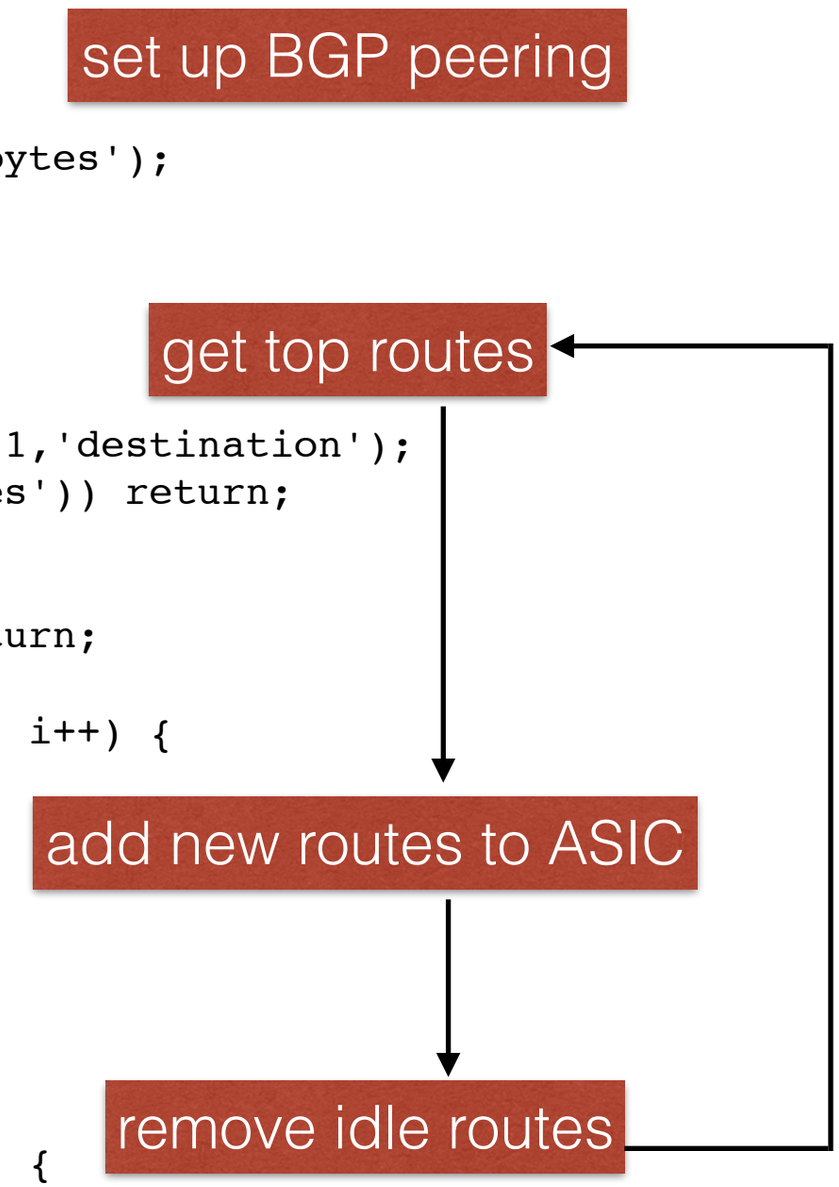
Compute top 30K routes (from possible 500K+) and push into ASIC…

Repeat every 5 seconds.

http://blog.sflow.com/2015/07/sdn-router-using-merchant-silicon-top.html

sFlow-RT router.js

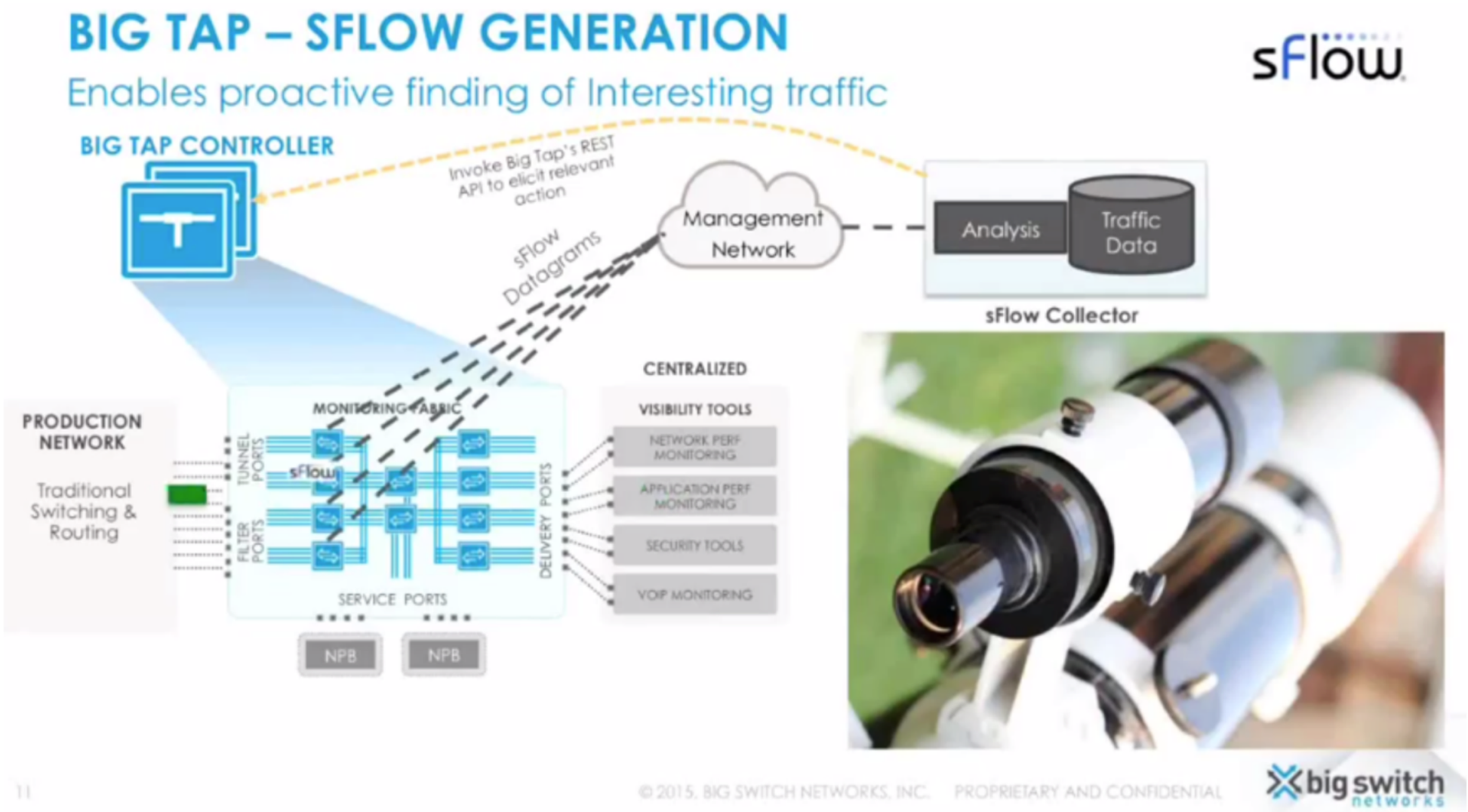
```
bgpAddNeighbor('10.0.0.254',65000);
bgpAddSource('10.0.0.253','10.0.0.254',10,'bytes');
bgpAddNeighbor('10.0.0.253',65000);

var installed = {};
setIntervalHandler(function() {
  let now = Date.now();
  let top = bgpTopPrefixes('10.0.0.254',100,1,'destination');
  if(!top || !top.hasOwnProperty('topPrefixes')) return;

  let tgt = bgpTopPrefixes('10.0.0.253',0);
  if(!tgt || 'established' != tgt.state) return;

  for(let i = 0; i < top.topPrefixes.length; i++) {
     let entry = top.topPrefixes[i];
     if(bgpAddRoute('10.0.0.253',entry)) {
       installed[entry.prefix] = now;
     }
  }
  for(let prefix in installed) {
     let time = installed[prefix];
     if(time === now) continue;
     if(bgpRemoveRoute('10.0.0.253',prefix)) {
        delete installed[prefix];
     }
  }
}, 5);
```



http://blog.sflow.com/2015/07/sdn-router-using-merchant-silicon-top.html

4. sFlow RT and BigSwitch Big Tap

Low-latency trigger (on any fields) for security/troubleshooting packet-capture.
e.g. match against list of 1,000,000+ “bad actor” IP addresses, even if talking over tunnel.



http://blog.sflow.com/2015/04/big-tap-sflow-enabling-pervasive-flow.html

Control plane experiences

- BGP
  - Good interoperability, but destination-CIDR only
- Hybrid OpenFlow (1.0, 1.3)
  - needs “super-normal” (http://blog.sflow.com/2014/10/super-normal.html)
- MPLS Segment Routing (ONOS)
  - Powerful mechanism for WAN controls - with control at ingress
- Cumulus Linux ACL Server (REST API)
  - Roll your own! Or use ours: https://github.com/pphaal/acl_server/