A DNS Anomaly Detection and Analysis System

Hyo-Jeong Shin, KT

NANOG 40

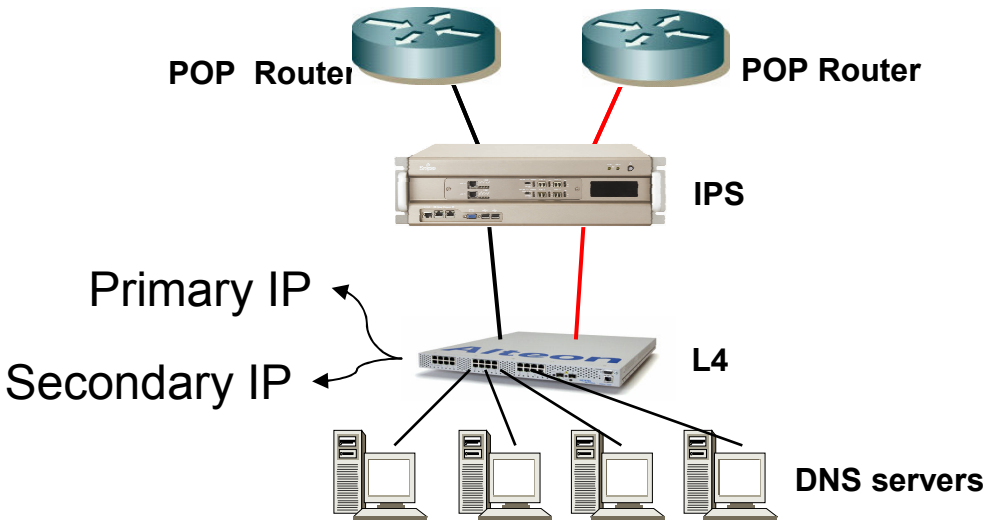
June 2007

KORNET DNS system

- More than 20 cache DNS server farms
- Each server farm advertises the *same* IP address using *anycast routing*
  - Each DNS server farm serves its local POP
- A central server farm (one of DNS server farms)
  - takes and serves queries for another server farm when the server farm fails

KORNET DNS system

- DNS server farms advertise two /32 addresses
  - One is primary DNS IP address
  - The other is secondary DNS IP address

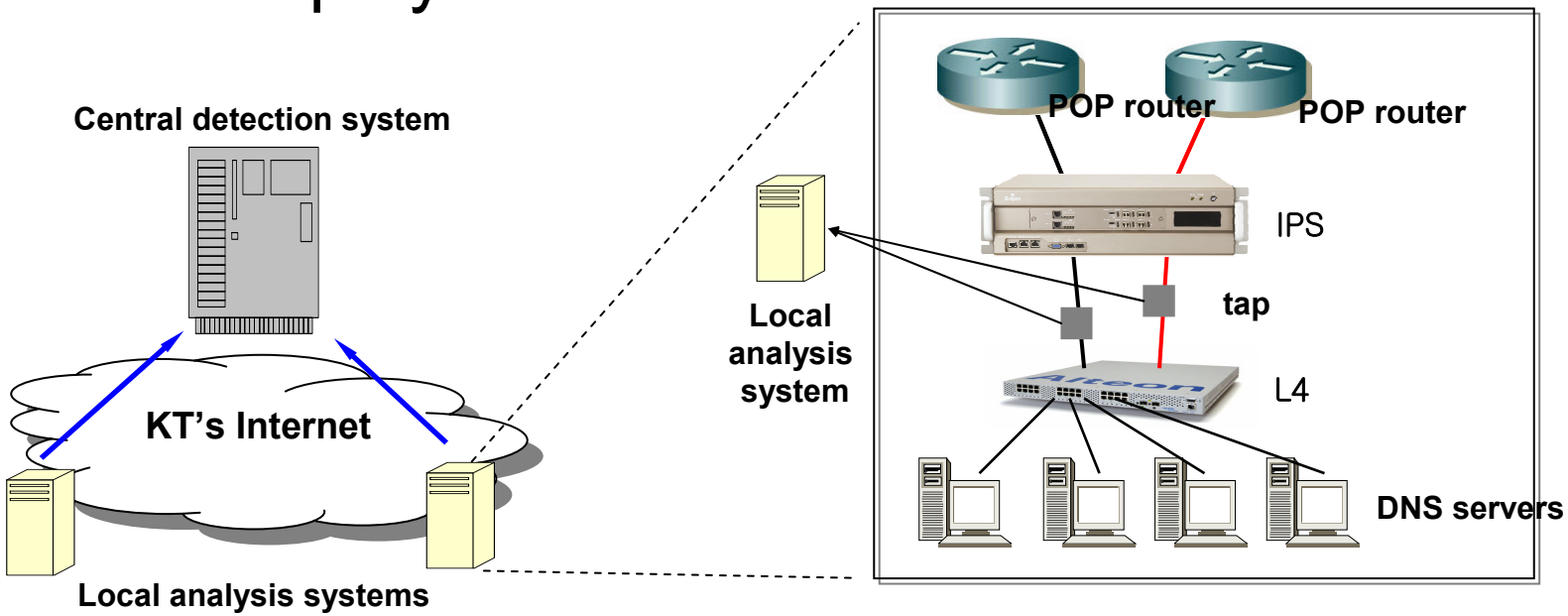


Purpose of our systematic approach

- Ever-increasing number of server farms and servers
  - Difficult to figure out what problems have happened on DNS
  - Manual dump and analysis ➔ time-consuming job
- Real-time systematic analysis would help
  - Investigate which server farm has problem
  - Investigate the problem instantly

Overall system: overview

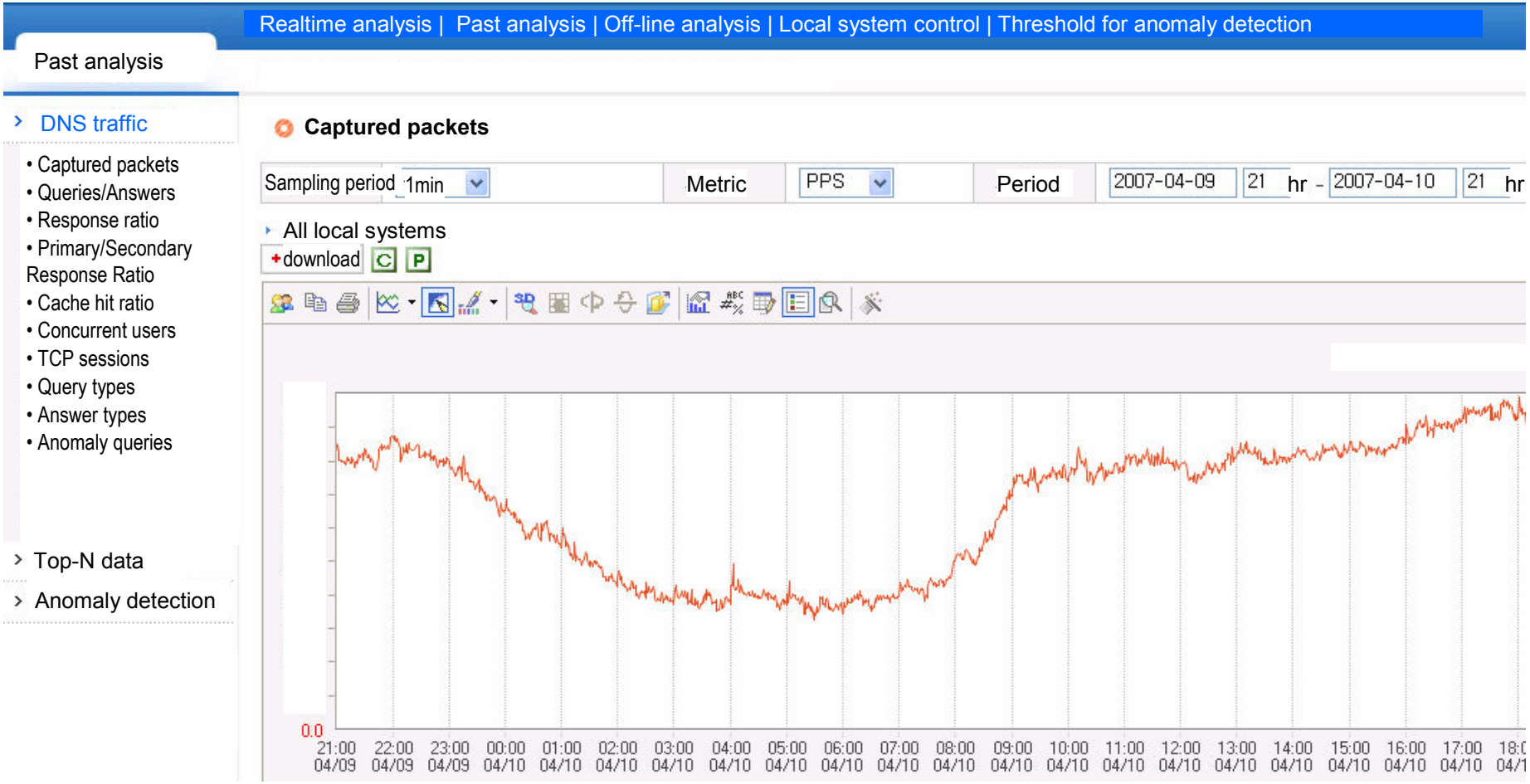
- A central detection system and multiple local analysis systems
  - Currently two local analysis systems have been deployed



Overall system: functions

- Local analysis system
  - Captures all packets to/from DNS server farms and inspects their contents with predefined rules
  - Sends analytical results to the central detection system
- Central detection system
  - Gathers the analytical results from all local analysis systems
  - Detects DNS anomalies

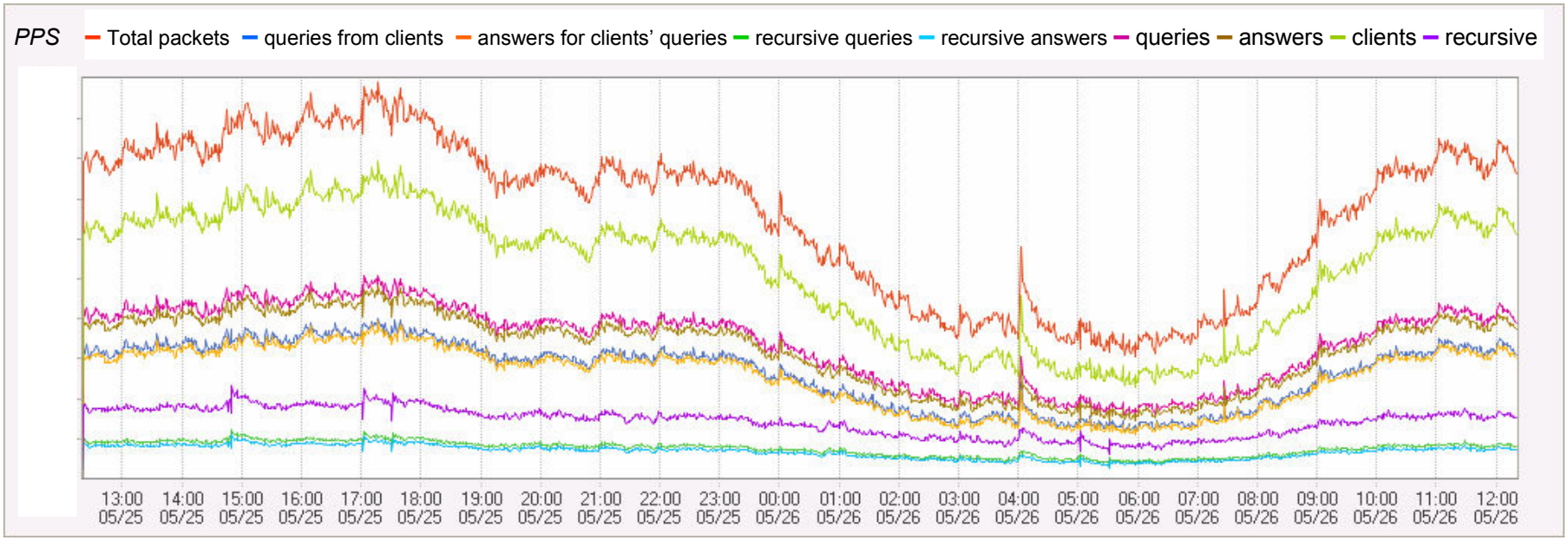
User Interface



The above chart shows PPS of captured DNS packet.

DNS traffic: Queries/Answers

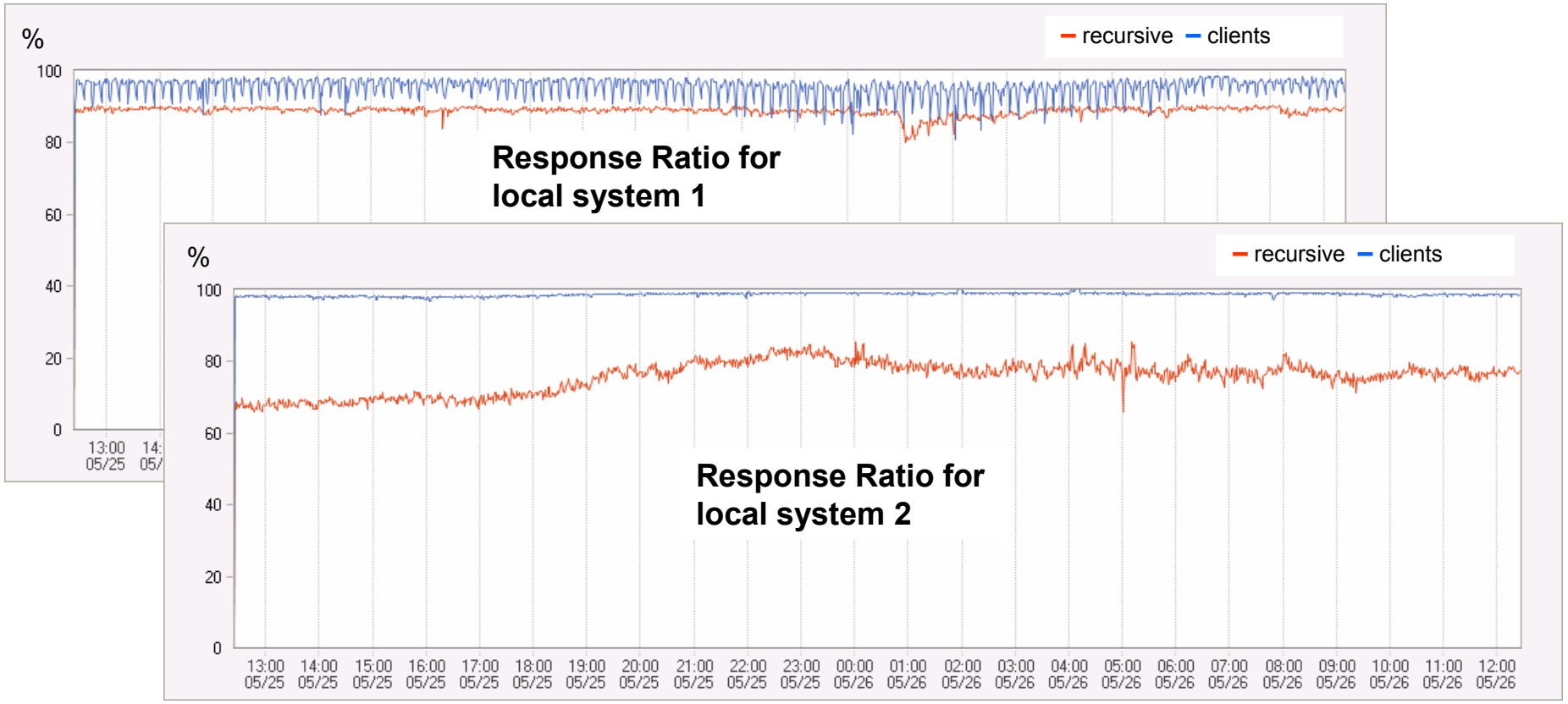
- *Total DNS packets*
- *Number of clients' queries/answers*
- *Number of recursive queries/answers*



DNS traffic : Response ratio

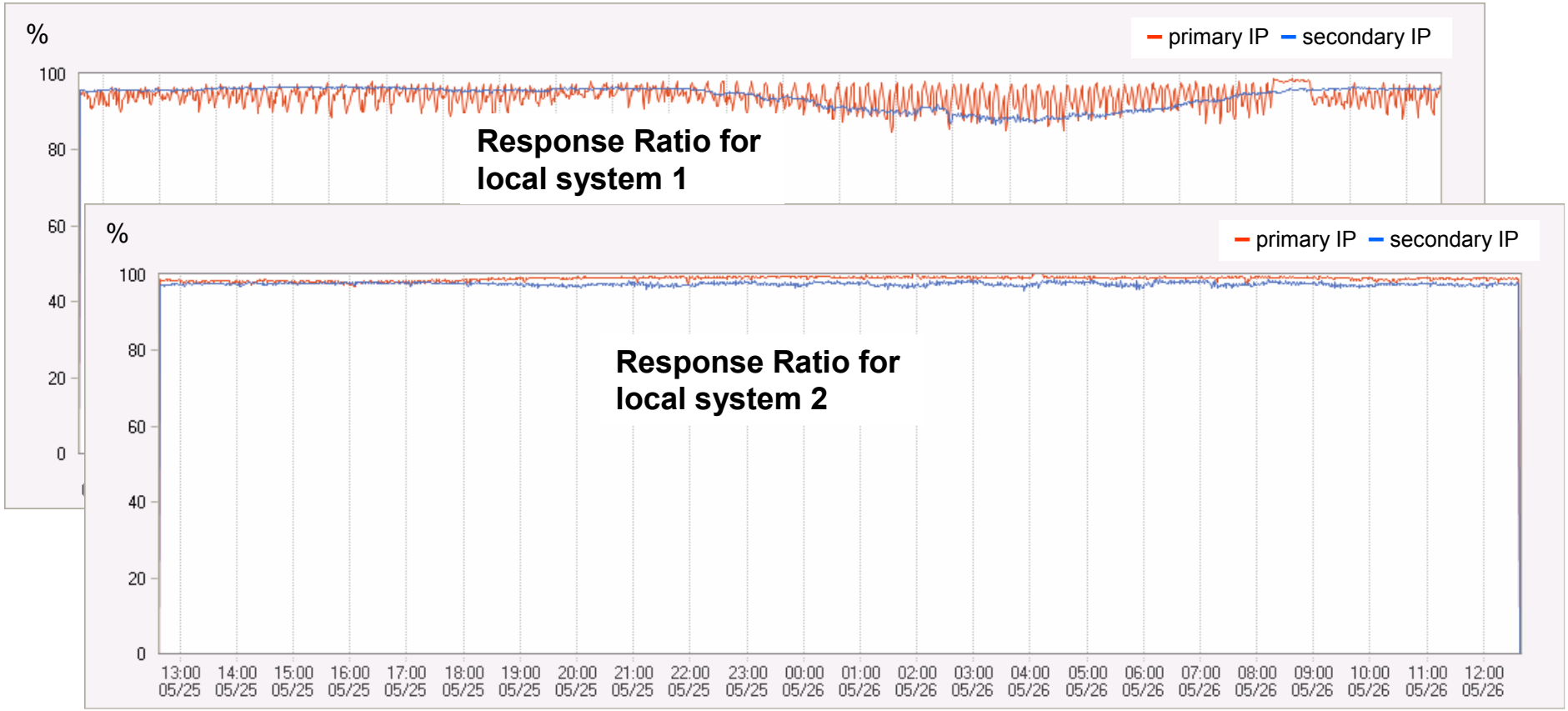
- *Ratio of resolved queries*

$$= \frac{\textit{number of answer packets}}{\textit{number of query packets}}$$

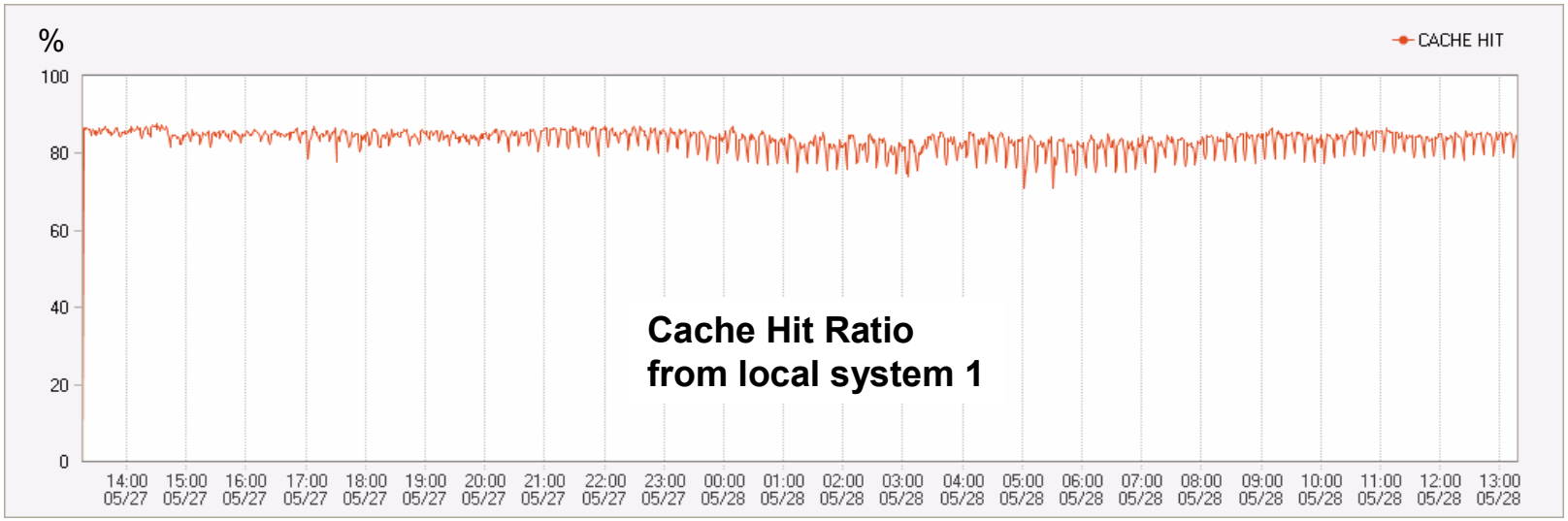


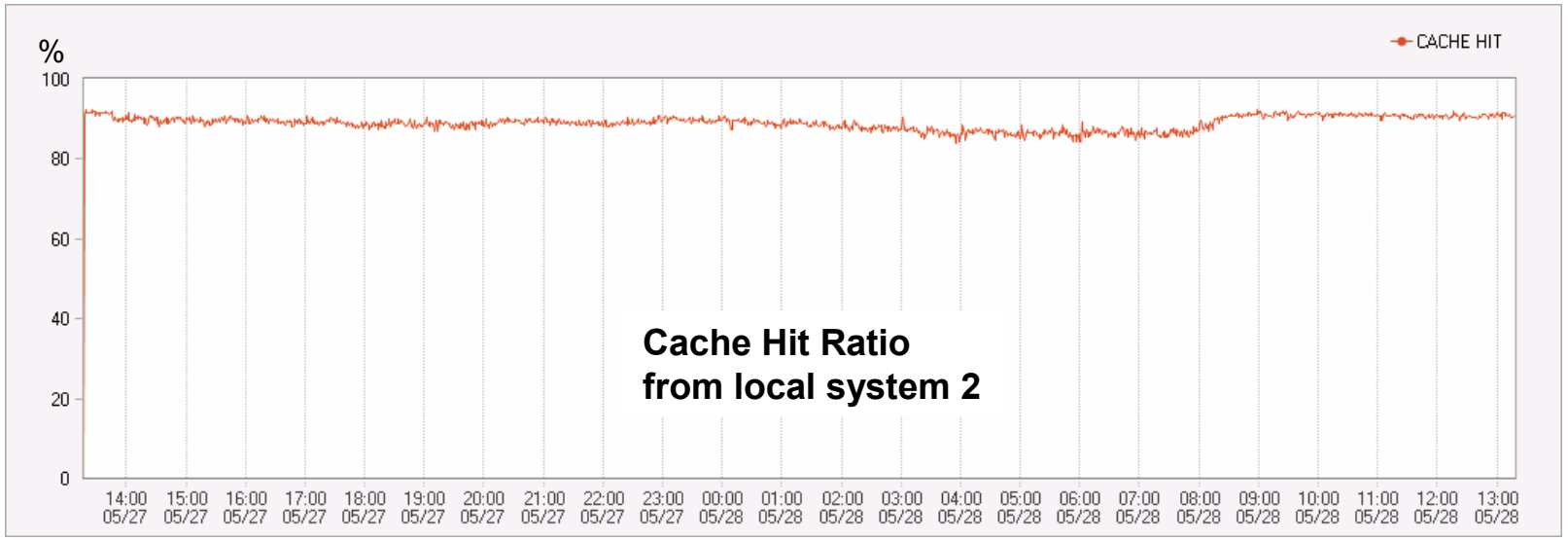
DNS traffic : Response ratio of primary/secondary DNS IP

- Calculate response ratio of primary/secondary DNS IP separately



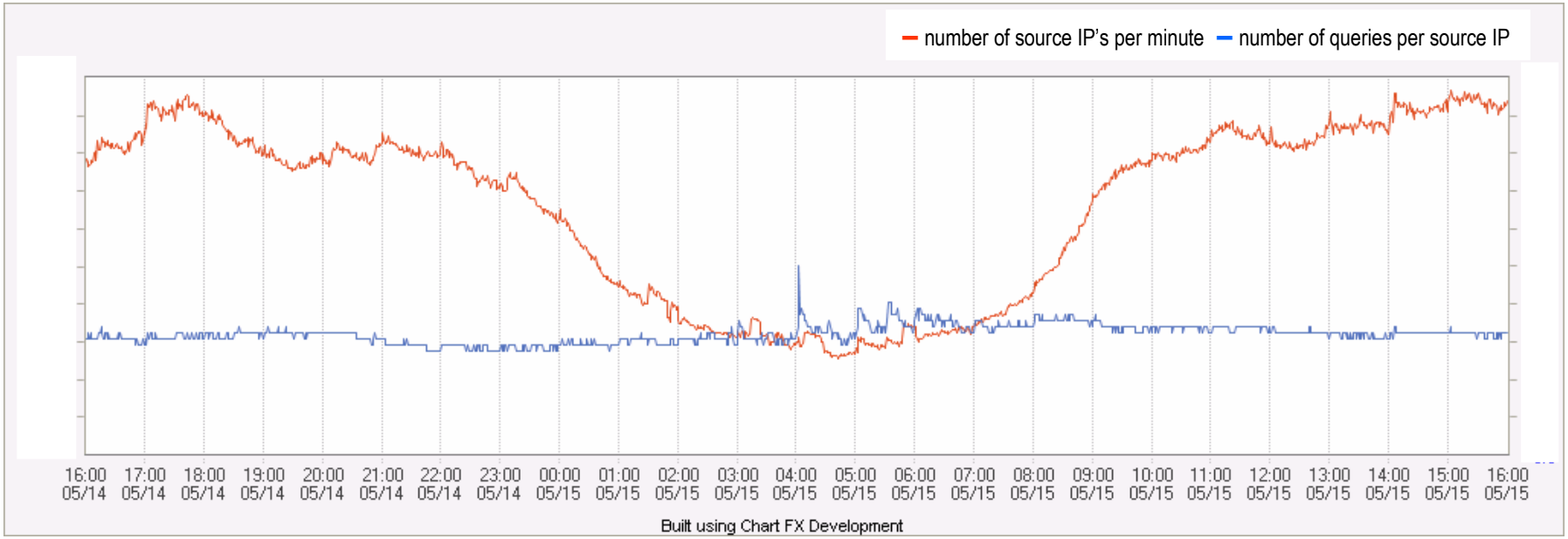
DNS traffic : Cache hit ratio





DNS traffic : Concurrent users

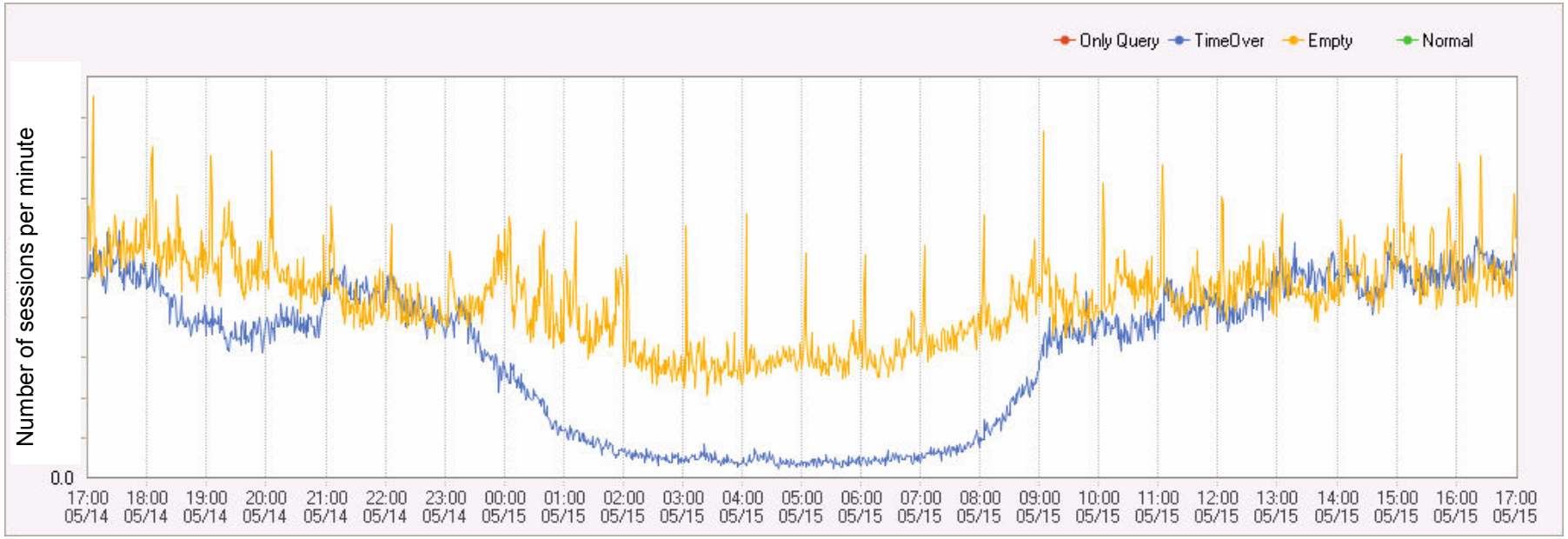
- Count number of source IP's per minute
- Count number of queries per source IP



The above chart shows: num of source IP's and num of queries

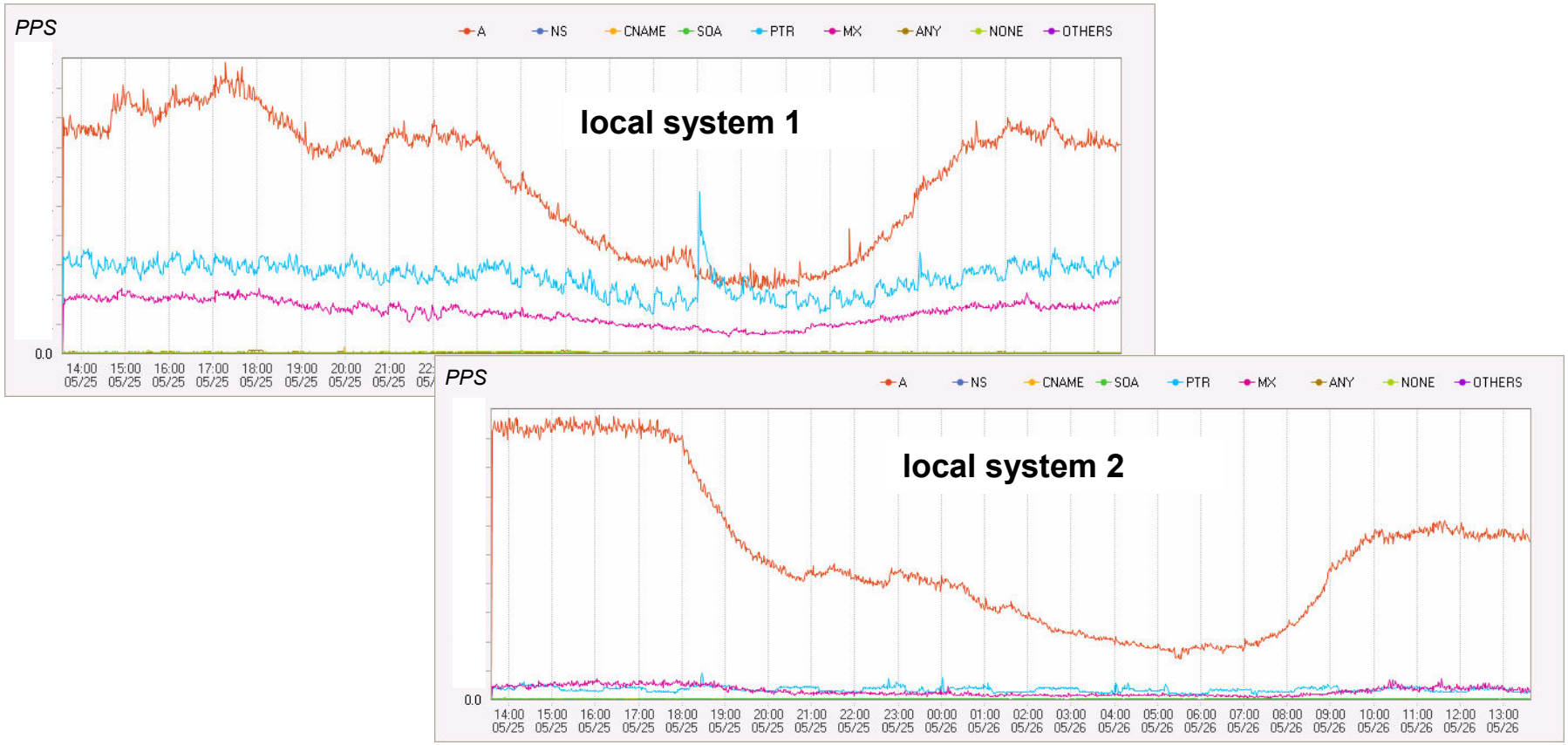
DNS traffic : TCP sessions

- Number of TCP sessions per minute
  - Only Query, Time Over, Empty, Normal



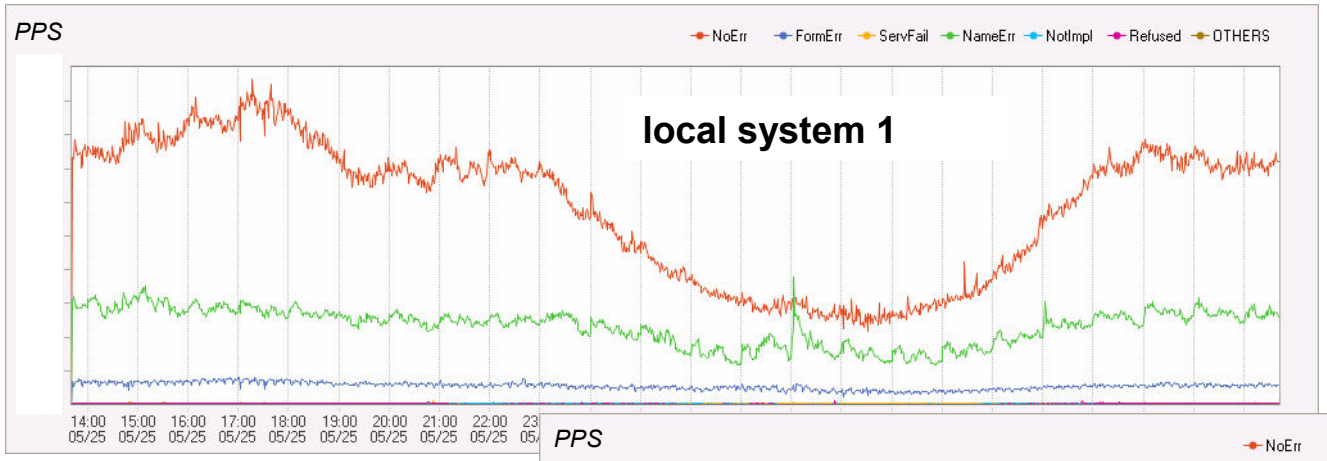
DNS traffic: Query types

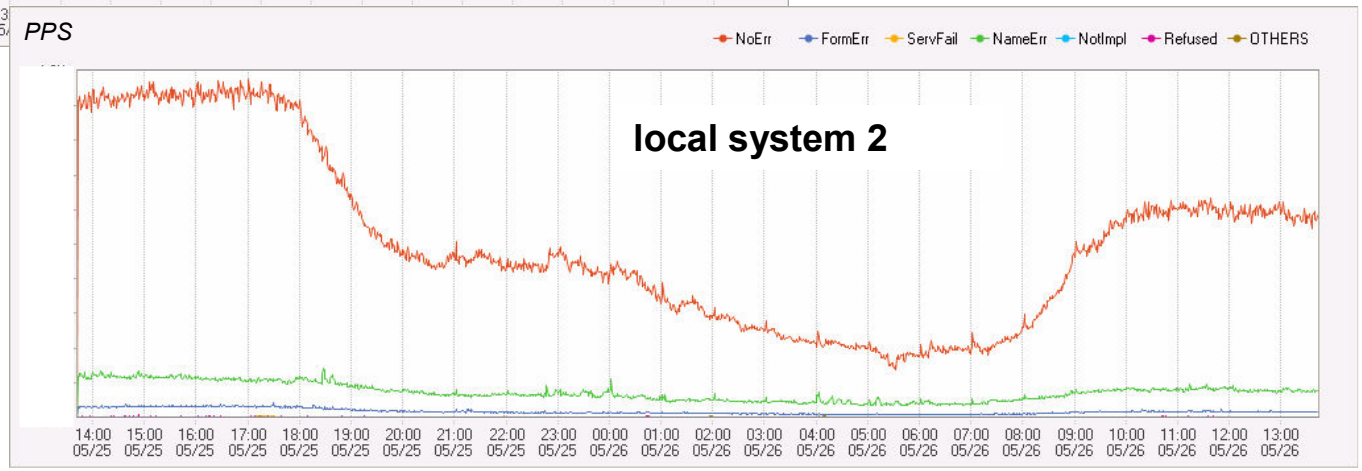
- *Distribution of query types*
  - A, MX, PTR, CNAME, …



DNS traffic: Answer types

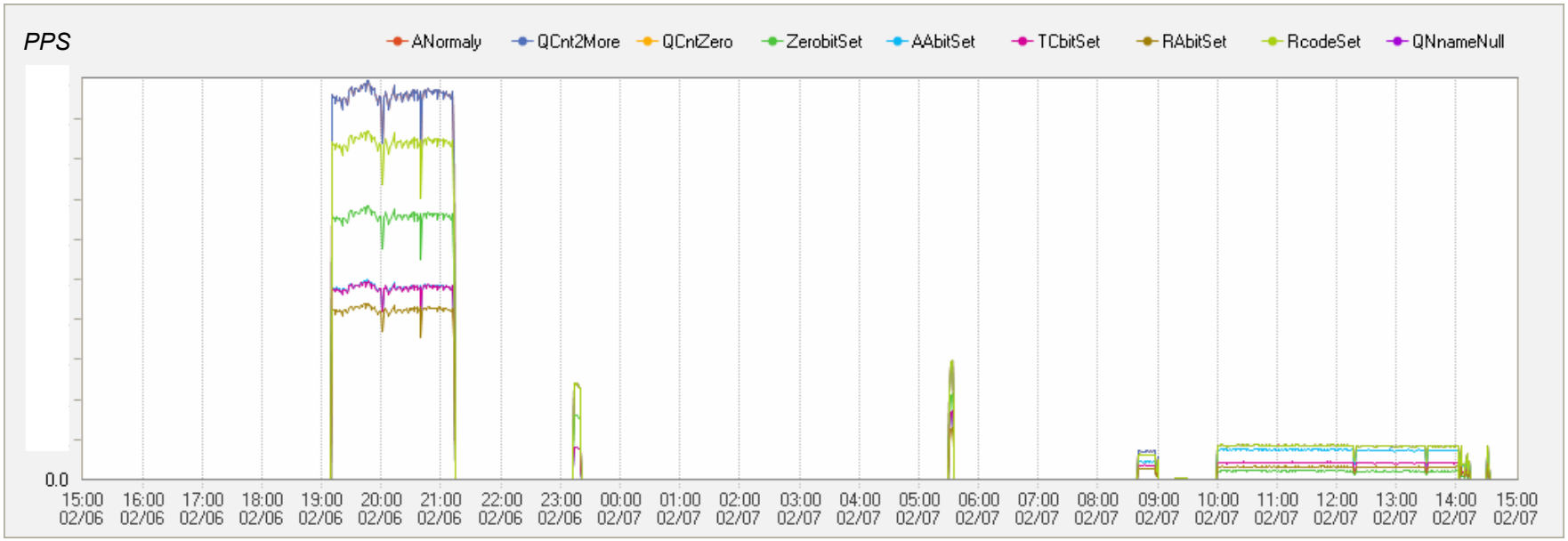
- *Distribution of answer types*
  - No Error, Serv Fail, Name Error,…





DNS traffic: Anomaly queries

- *Number of packets* which violate the format constraints of DNS header
  - These packets have been generated by malicious users



Top-N data: Source IP

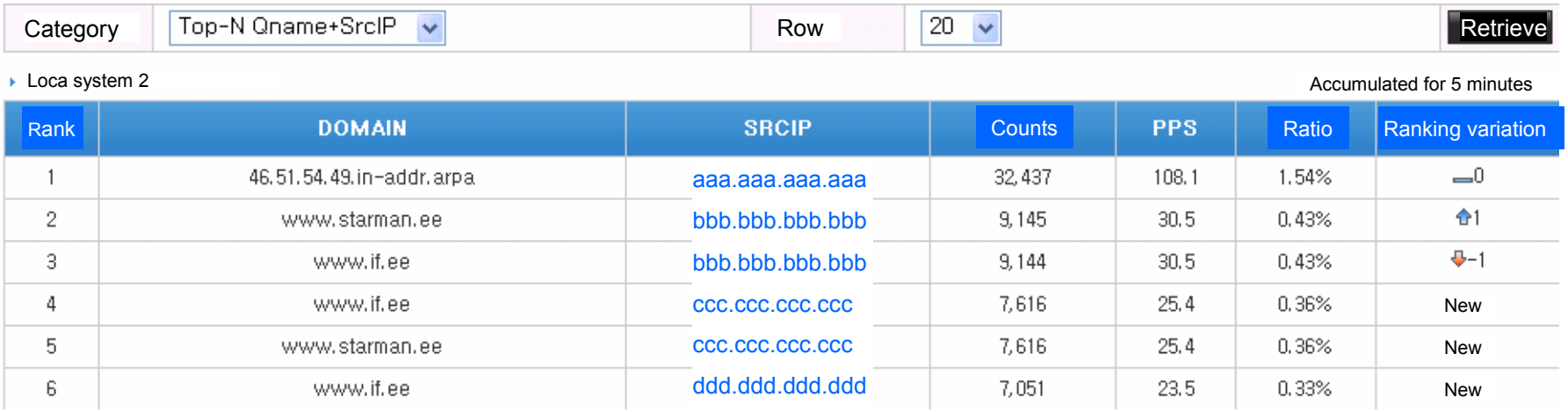
Category	Top-N Source IP's	Sampling period	5min	
Time	2007-3-8 10 hr 0 min	Row	20	Retrieve

accumulated for 5 minutes

Rank	SRCIP	Counts	QCNT ZERO	QCNT2 MORE	ZEROBIT SET	AABIT SET	TCBIT SET	RABIT SET	RCODE SET	QNAME NULL	PPS	Ratio	Ranking variation
1	aa.aa.aa.aa	54642	0	0	0	0	0	0	0	0	182.1	0.0	New
2	bb.bb.bb.bb	34376	0	0	0	0	0	0	0	0	114.6	0.0	-1
3	cc.cc.cc.cc	25965	0	0	0	0	0	0	0	0	86.6	0.0	3
4	dd.dd.dd.dd	22444	0	0	0	0	0	0	0	0	74.8	0.0	-1
5	ee.ee.ee.ee	19662	0	0	0	0	0	0	0	0	65.5	0.0	0
6	ff.ff.ff.ff	19407	0	1	0	1	1	0	1	1	64.7	0.0	1

- Category
  - Top-N source IP's
  - Top-N source IP's with type A, MX, PTR…
  - Top-N source IP's generating anomaly queries

Top-N data: Qname + Source IP



Category: Top-N Qname+SrcIP | Row: 20 | Retrieve

Loca system 2

Accumulated for 5 minutes

Rank	DOMAIN	SRCIP	Counts	PPS	Ratio	Ranking variation
1	46.51.54.49.in-addr.arpa	aaa.aaa.aaa.aaa	32,437	108.1	1.54%	0
2	www.starman.ee	bbb.bbb.bbb.bbb	9,145	30.5	0.43%	1
3	www.if.ee	bbb.bbb.bbb.bbb	9,144	30.5	0.43%	-1
4	www.if.ee	ccc.ccc.ccc.ccc	7,616	25.4	0.36%	New
5	www.starman.ee	ccc.ccc.ccc.ccc	7,616	25.4	0.36%	New
6	www.if.ee	ddd.ddd.ddd.ddd	7,051	23.5	0.33%	New

- Category
  - Top-N Qname + source IP's
  - Top-N Qname + source IP’s with type A, NS, PTR, ServFail …
  - Top-N Black domain + source IP's
  - Top-N Qname + source IP's who generate queries to secondary DNS IP

Top-N data: Qname

Category: Top-N Qname | Row: 20 | Retrieve

Loca system 1 — Accumulated for 5 minutes

Rank	DOMAIN	Counts	PPS	Ratio	Ranking variation
1	time.nist.gov	350,150	1,167.2	4.73%	0
2	213.43.122.222.in-addr.arpa	80,495	268.3	1.09%	46
3	vip.mk.co.kr	65,825	219.4	0.89%	-1
4	cmd.msnplus.co.kr	61,860	206.2	0.84%	-1
5	rev1.kornet.net	55,192	184	0.75%	-1
6	clock.iptime.co.kr	51,264	170.9	0.69%	-1
7	rev2.kornet.net	46,328	154.4	0.63%	-1

- Category
  - Top-N Qname
  - Top-N Qname with type A, MX, PTR, ServFail, NamErr…
  - Top-N Recursive Qname

Top-N DNS server

- Top-N DNS server
  - Count queries to authoritative DNS servers which serve recursive queries

Accumulated for 5 minutes

Category		Counts	PPS	%	root only
Root server	J-root	19,304	64.3	1.71%	50%
	F-root	8,014	26.7	0.71%	21%
	M-root	2,799	9.3	0.25%	7%
	Other root	8,650	28.8	0.77%	22%
gTLD server		116,425	388.1	10.32%	
kr server		13,237	44.1	1.17%	
others		959,586	3,198.60	85.07%	
Total		1,128,015	3,760.10	100%	

Top-N TLD, SLD/3LD

- Top-N TLD

Accumulated for 5 minutes

rank	DOMAIN	counts	PPS	ratio	ranking variation
1	com	2,732,130	9,107.10	44.3	0
2	kr	1,330,693	4,435.60	21.58	0
3	net	900,962	3,003.20	14.61	0
4	gov	335,277	1,117.60	5.44	0
5	org	178,693	595.6	2.9	0

- Top-N SLD/3LD

Accumulated for 5 minutes

rank	DOMAIN	counts	PPS	ratio	ranking variation
1	naver.com	459,799	1,532.70	7.48	0
2	nate.com	351,624	1,172.10	5.72	0
3	nist.gov	333,805	1,112.70	5.43	0
4	kornet.net	179,113	597	2.92	0
5	daum.net	150,804	502.7	2.45	0
6	mk.co.kr	125,271	417.6	2.04	0
7	nasads.com	82,771	275.9	1.35	1
8	hanmail.net	76,573	255.2	1.25	-1
9	iveconmiga.info	65,856	219.5	1.07	0
10	msnplus.co.kr	65,174	217.2	1.06	0

Anomaly detection logic

- Based upon the variation ratio of “DNS traffic”
- Basic ideas on detection formula
  - Use the collected parameters of “DNS traffic”
  - Keep the limited history of values on each parameter
  - Put more weights on recent data
  - Rely on the threshold value to make decision
- When anomaly detected
  - Check the details on who or what incurred problems with “Top-N data” page

Indicators of anomaly

- Increase in number of DNS packets
- Increase in number of queries to DNS server farm's secondary IP address
- Decrease in cache hit ratio
- Increase in average DNS queries of individual source IP addresses
- Increase in number of recursive queries
- Increase in number of TCP sessions
- Increase in number of source IP addresses within a limited time slot
- Decrease in ratio of resolved queries
- Variation of query types *cf*. decrease in ratio of type A
- Variation of answer types *cf*. decrease in type "No Error"

An example of “Anomaly detection”

• Local system 2

Time	packet counts	secondary/primary	cache hit ratio	clients queries	recursive queries	TCP sessions	concurrent users	Qtype A	Rcode Noerror
Threshold (%)	30 50 80	50 70 90	20 50 100	30 50 80	20 50 100	50 70 90	20 50 100	20 50 100	20 50 100
2007-05-27 00:36:00.0									
2007-05-27 00:35:00.0									
2007-05-27 00:34:00.0									
2007-05-27 00:33:00.0									
2007-05-27 00:32:00.0									
2007-05-27 00:31:00.0									
2007-05-27 00:30:00.0									
2007-05-27 00:29:00.0									

- Warning level: green-yellow-orange-red
  - Red means the situation is critical
  - Second row shows threshold values between warning levels

Case of utilizing the system(1)

- Detected unexpected increase on MX queries
  - In Feb, 2007
  - Response ratio was just 10% or higher
    - Most of clients' queries are dropped
  - Then we need to know
    - what queries they are
    - who generates
    - what query types
  - Multiple source IP addresses kept generating MX queries

➔ Blocked the traffic from the KORNET to the infecting system

Case of utilizing the system(2)

- Detected tons of queries on “time.nist.gov”
  - In March, 2007
  - The amount of DNS queries had been doubled or tripled at all DNS server farms
  - Then we need to know
    - What queries they are
      - More than 60% of queries were on “nist.time.gov”
    - Who generates
      - They were generated from some home gateways

➔ Investigated why the home gateways had generated the queries and advised the company to patch their firmware

On-going plan of this year

- Install another local analysis systems to cover all DNS server farms of KORNET
- Improve the anomaly detection logic