

# Operation requirements

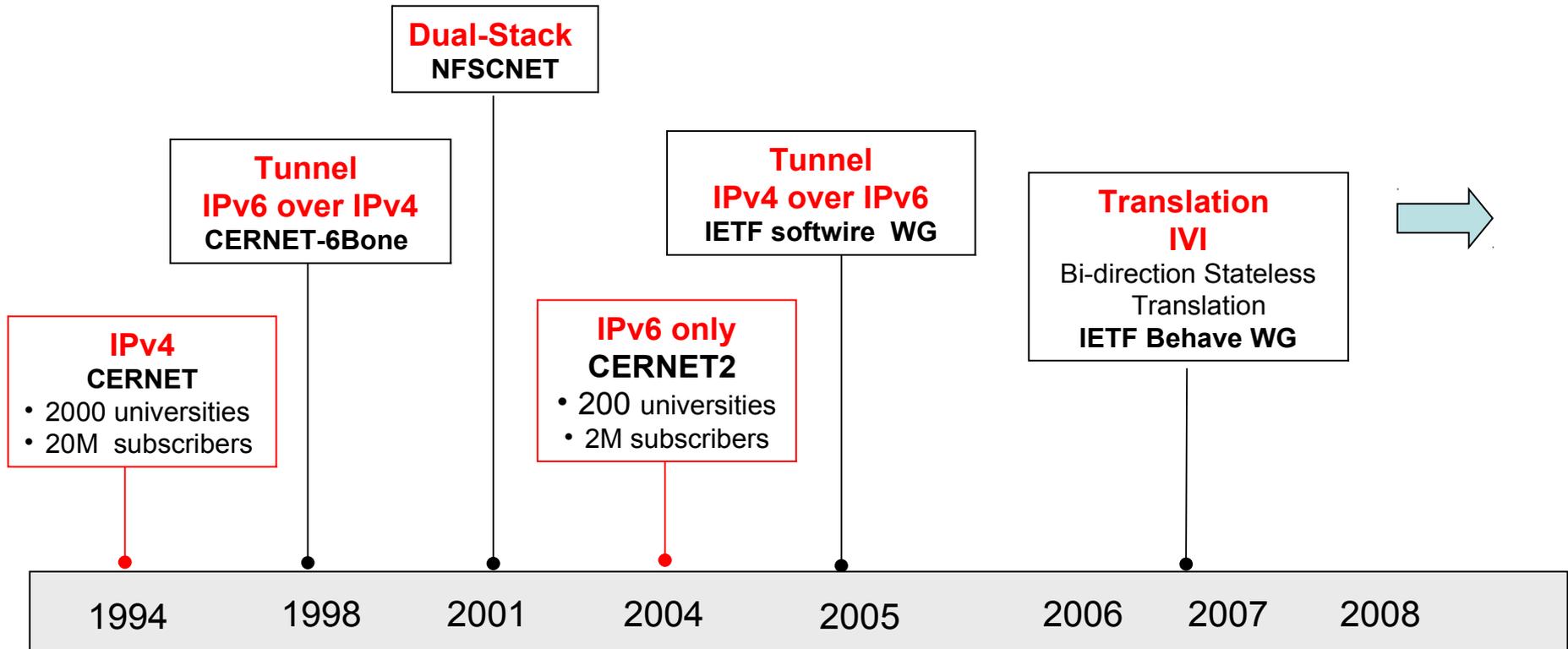
X. Li, C. Bao

2011-09-26

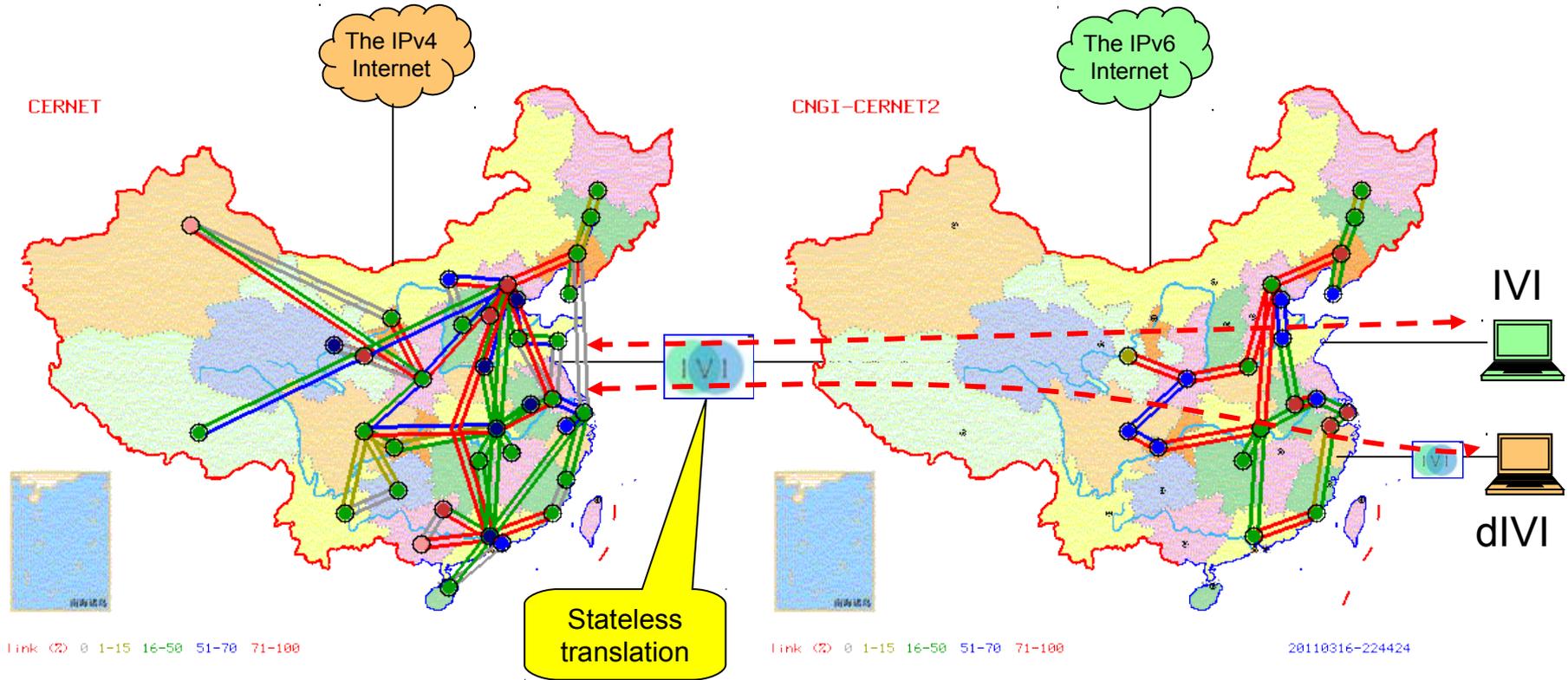
# Outline

- A brief history
- Networks
  - CERNET and CNGI-CERNET2
  - CERNET-NB
- Requirements
- Operation experience
- Summary

# A brief history



# CERNET/CNGI-CERNET2

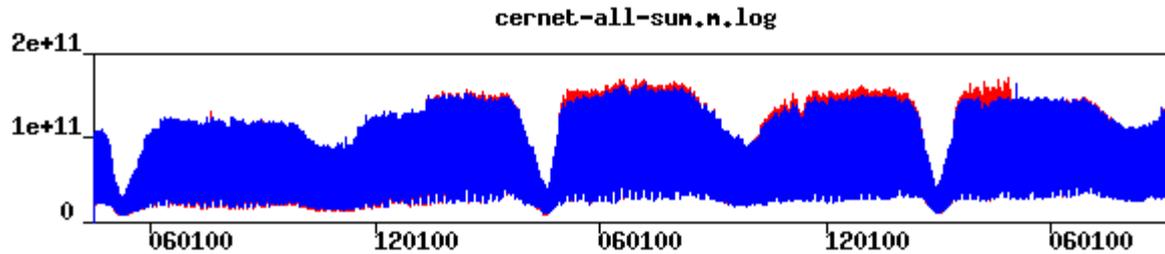


IPv4

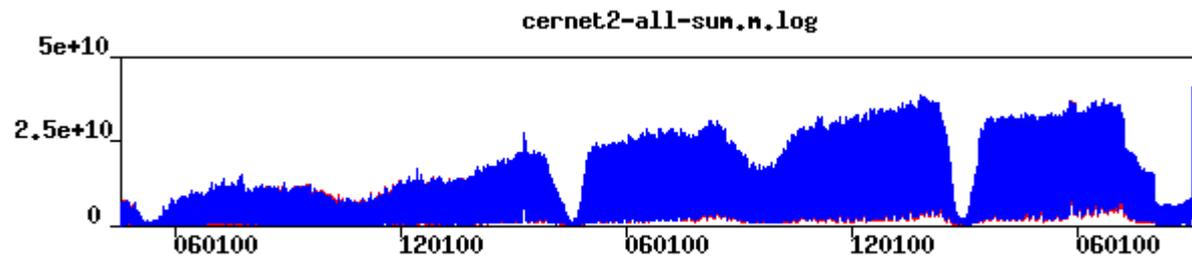
IPv6

# Traffic comparisons

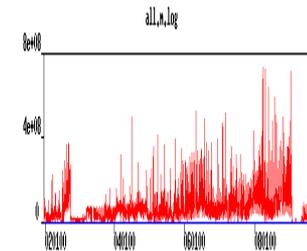
IPv4



IPv6



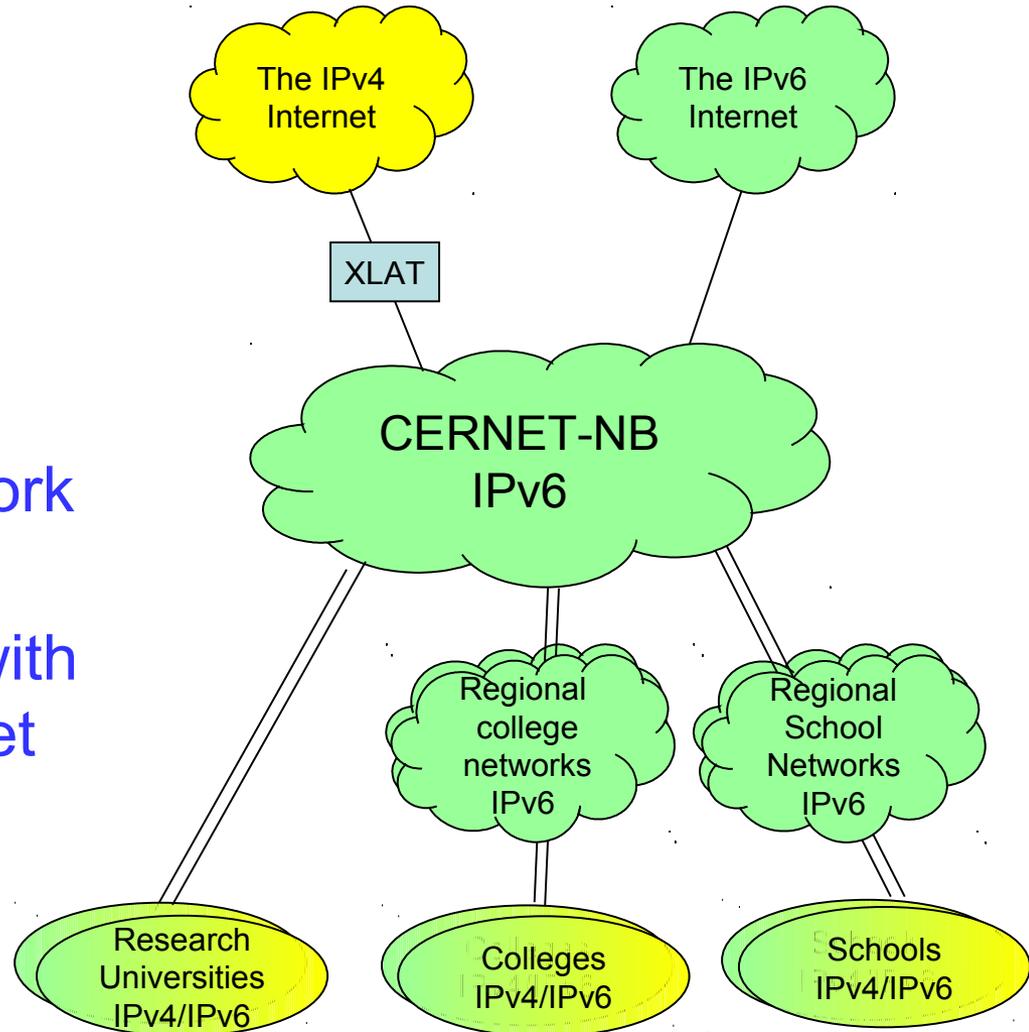
IVI



- **CERNET IPv6' traffic is about 20% of IPv4**

# CERNET-NB

- **Scale**
  - 400K schools
  - 320M students
- **Requirements**
  - IPv6-only network
  - Be able to communicate with the IPv4 Internet



## Requirements

- Stateless translators
- Communicate with IPv4

# Requirements

Requirement	Stateless tunneling	Stateless dual translation
<b>Minimum impact to business</b>		
Communicate with IPv4	yes	no
Solve ALG problem	yes	yes
Bi-directional initiated communications	yes	yes
<b>O&amp;M considerations</b>		
Being scalable	yes	yes
Prefer stateless in the core	yes	yes
Use familiar O&M Tools	yes	no
User log for trace back	yes	yes
Renumbering and address migration	yes	yes
<b>Promoting IPv6 transition</b>		
Unified with 3GPP, DOCSIS if possible	yes	no
IPv6-only device communicates with IPv4	yes	no

dIVI is not harmful

X. Li, C. Bao

2011-09-26

# dIVI is not harmful

- IPv4
  - Options
    - Dropped (no harm)
  - Type of service
    - Copied to traffic class
  - MTU and fragmentation
    - Processed according to RFC6145/RFC6146 (same as tunnel)
  - ICMP/ICMPv6
    - Processed according to RFC6145 (no harm)
  - Transport layer (TCP, UDP)
    - Processed according to RFC6145

# IPv4 Options

- The IPv4 option will not be translated to IPv6.

	packets	ratio	%
total	40702507685		1
option	1255	3.08335E-08	0.000003
ICMP	176	4.32406E-09	
TCP	1079	2.65094E-08	

ICMP: 176

IP option 7, Record Route 22

Unreachable 154

type=0, code=0

type=3, code=2, option=0 or 1

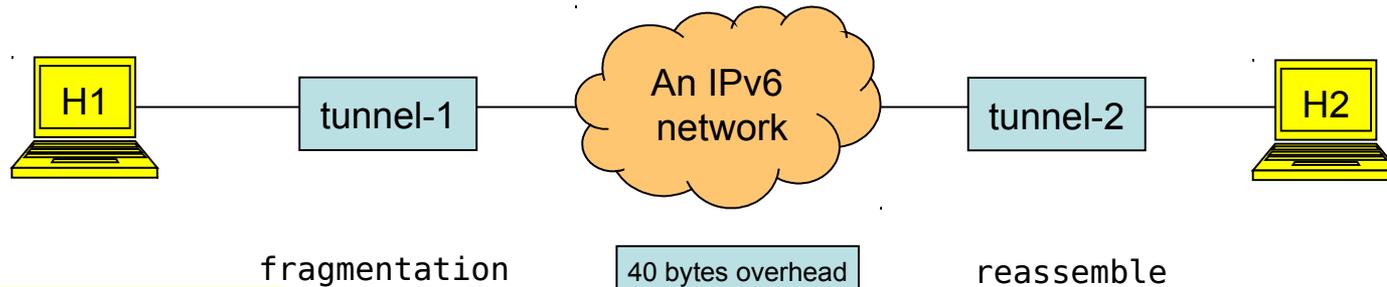
TCP: 1079

1 port=80, option=0 or

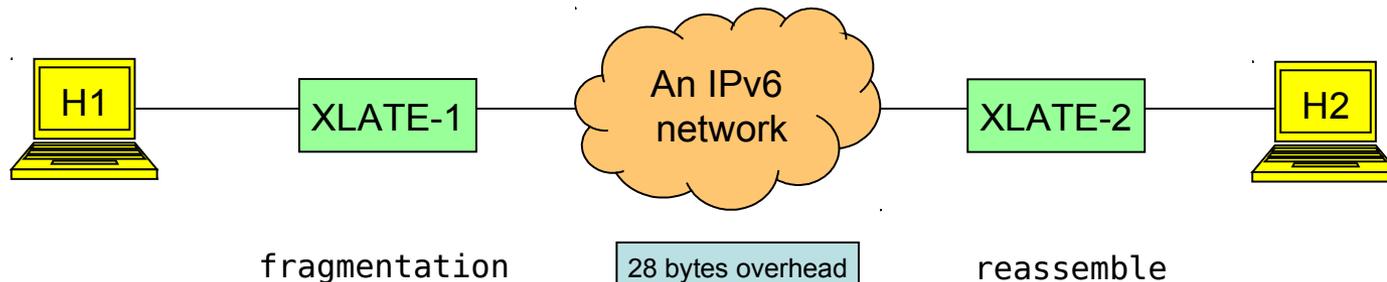
Conclusion: Lost of option field results in no harm.

# MTU and fragmentation (1)

- Comparisons



RFC6333  
draft-murakami-software-4rd



draft-xli-behave-dIVI  
draft-xli-behave-dIVI-pd



Conclusion: dual stateless translation are about the same, except for 12-20 bytes savings.

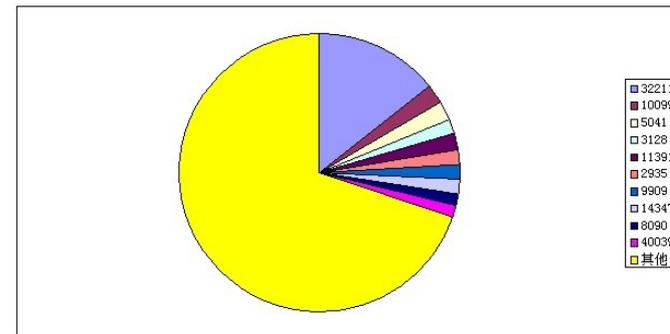
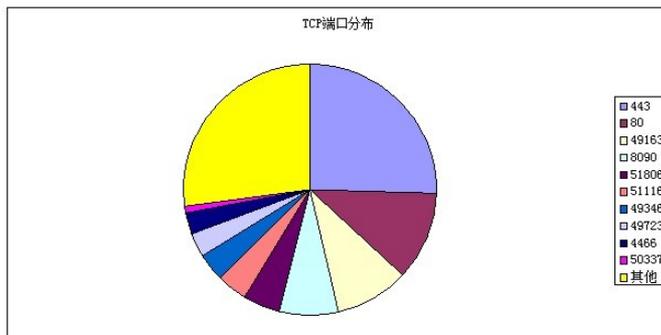
# MTU and fragmentation (2)

	packets	ratio	%
total	40702507685	1	
frag	41990319	0.00103164	0.103163961
TCP	5843	1.43554E-07	1.43554E-05
ICMP	18414278	0.000452411	0.045241139
UDP	22786760	0.000559837	0.055983676
GRE	783259	1.92435E-05	0.001924351

TCP mainly 80 (HTTP) or 443 (HTTPS)

UDP random (p2p?)

ICMP ICMP type 8, Echo request message or ICMP type 3, Destination unreachable



# ICMP/ICMPv6 (1)

- RFC6145

- 4.2. Translating ICMPv4 Headers into ICMPv6 Headers

- ICMPv4 query messages
    - ICMPv4 error messages

- 4.3. Translating ICMPv4 Error Messages into ICMPv6

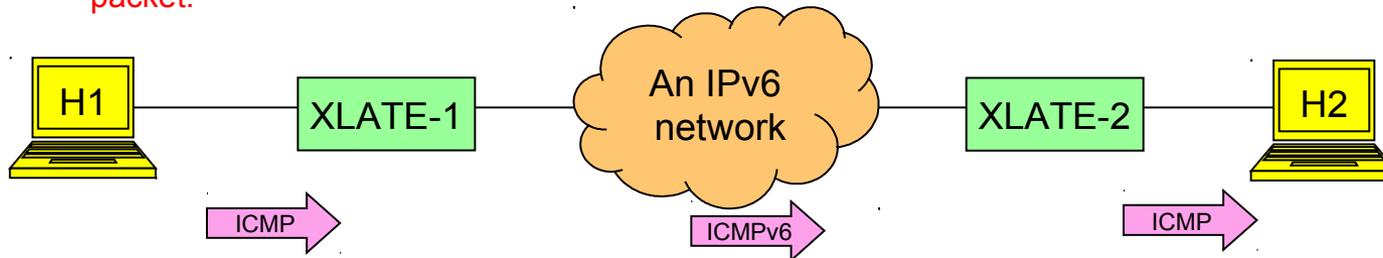
- The ICMP error messages containing the packet in error MUST be translated just like a normal IP packet.

- 5.2. Translating ICMPv6 Headers into ICMPv4 Headers

- ICMPv6 informational messages
    - ICMPv6 error messages

- 5.3. Translating ICMPv6 Error Messages into ICMPv4

- The ICMP error messages containing the packet in error MUST be translated just like a normal IP packet.



Conclusion: most of the ICMP and ICMP error messages can be carried end-to-end.

# ICMP/ICMPv6 (2)

RFC6145				packets	ratio
v4type	v4code	v6type	v6code	11332799	
0		129		1103832	0.097402
3	0	1	0	22294	0.001967
3	1	1	0	241089	0.021274
3	2	4	1	505006	0.044561
3	3	1	4	3225497	0.284616
3	4	2	0	80760	0.007126
3	5	1	0	0	0
3	6	1	0	0	0
3	7	1	0	0	0
3	8	1	0	0	0
3	9	1	1	1306	0.000115
3	10	1	1	246199	0.021724
3	11	1	0	0	0
3	12	1	0	0	0
3	13	1	1	62716	0.005534
3	15	1	1	0	0
3	others			0	0
8		128		5491161	0.484537
11		3		325596	0.02873
12	0	4	0	23	2.03E-06
12	2	4	0	0	0
				<b>11305479</b>	<b>0.99759</b>

RFC6145				packets	ratio	notes
v4type	v4code	v6type	v6code	11332799		
3	14	1	无	0	0	DROP
4		n/a		644	5.683E-05	DROP
5		n/a		26438	0.0023329	DROP
6		n/a		0	0	DROP
9		n/a		0	0	DROP
10		n/a		0	0	DROP
12	1		4	0	0	DROP
12	others			0	0	DROP
13		n/a		10	8.824E-07	DROP
14		n/a		1	8.824E-08	DROP
15		n/a		0	0	DROP
16		n/a		1	8.824E-08	DROP
17		n/a		2	1.765E-07	DROP
18		n/a		0	0	DROP
				<b>27096</b>	<b>0.002391</b>	

- RFC6145
  - **Source Quench (Type 4): Obsolete in ICMPv6. Silently drop.**
  - **Redirect (Type 5): Single-hop message. Silently drop.**
- IANA
  - # Type 13 — Timestamp
  - # Type 14 — Timestamp Reply
  - # Type 15 — Information Request
  - # Type 16 — Information Reply
  - # Type 17 — Address Mask Request
  - # Type 18 — Address Mask Reply

# Operational perspective

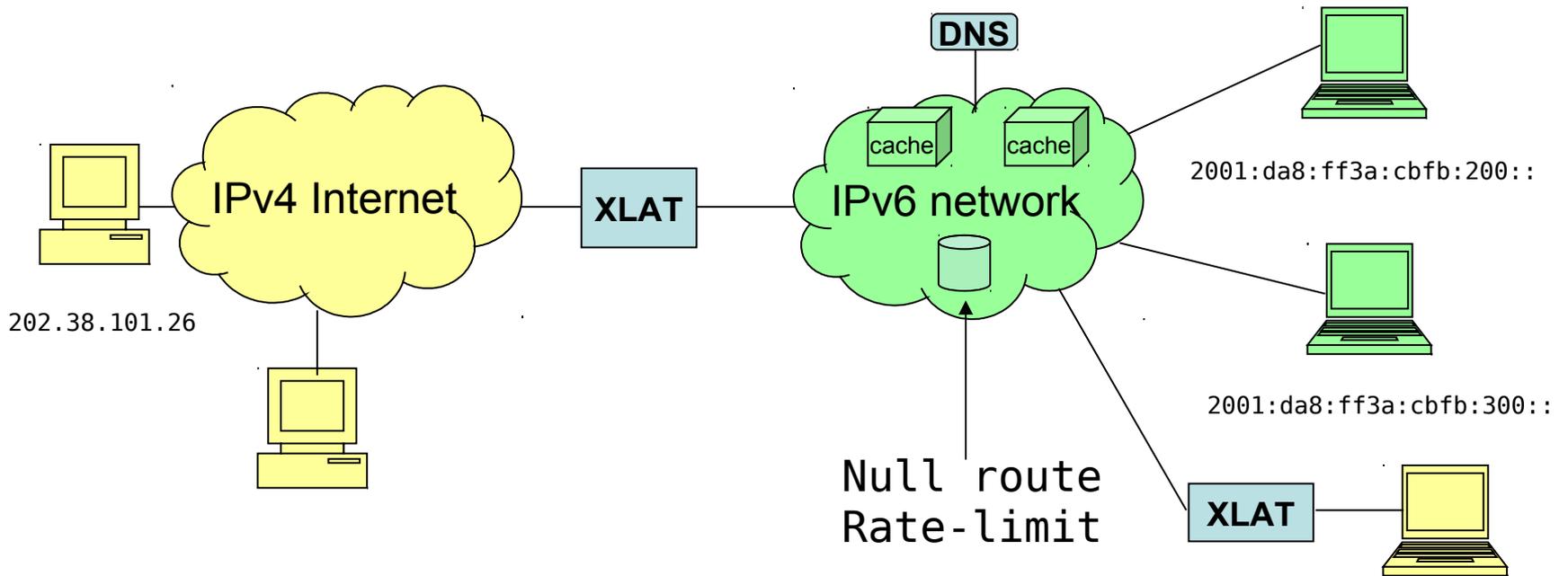
X. Li, C. Bao

2011-09-26

# O&M requirements

- Tools
  - IPv6 null route to discard IPv4 traffic
  - IPv6 ACL to filter IPv4 traffic
  - IPv6 traffic shaping to rate limit IPv4 traffic
- Caching
  - CDN
- We cannot do this with tunnelling using existing tools

# Topology



# dIVI: config on r-bj6 (null route)

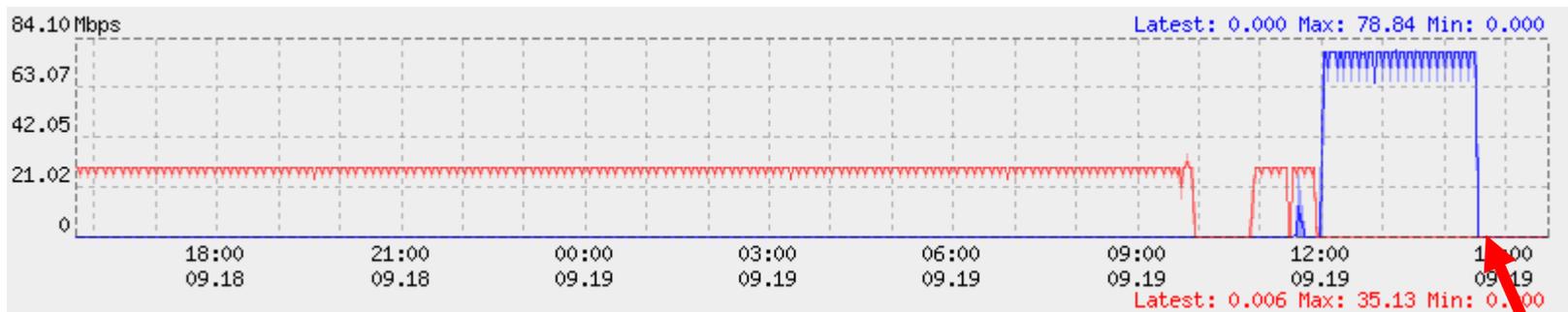
```
show configuration routing-options
rib inet6.0 {
  static {
    route 2001:DA8:FF3A:C8FB:300::/128 discard;
  }
}
```

```
ge-2/2/0 {
  description 1G_to_Dragon-Lab;
  unit 0 {
    family inet {
      address 202.38.120.189/29;
    }
    family inet6 {
      address 2001:DA8:1:30::1/64;
    }
  }
}
```

2001:DA8:FF3A:C8FB:300:: ↔ 58.200.251.3

# dIVI: IPv6 null route to discard IPv4 traffic

2001:da8:ff3a:c8fb:300:: -> 202.38.101.26 70M (blue)



IPv6 null route

# dIVI: config on r-bj6 (shaping)

## Interface to Dragon-lab

```
interfaces {
  ge-2/2/0 {
    description 1G_to_Dragon-
    Lab;
    unit 0 {
      family inet6 {
        filter {
          output ivi-qos;
        }
        address
        2001:DA8:1:30::1/64;
      }
    }
  }
}
```

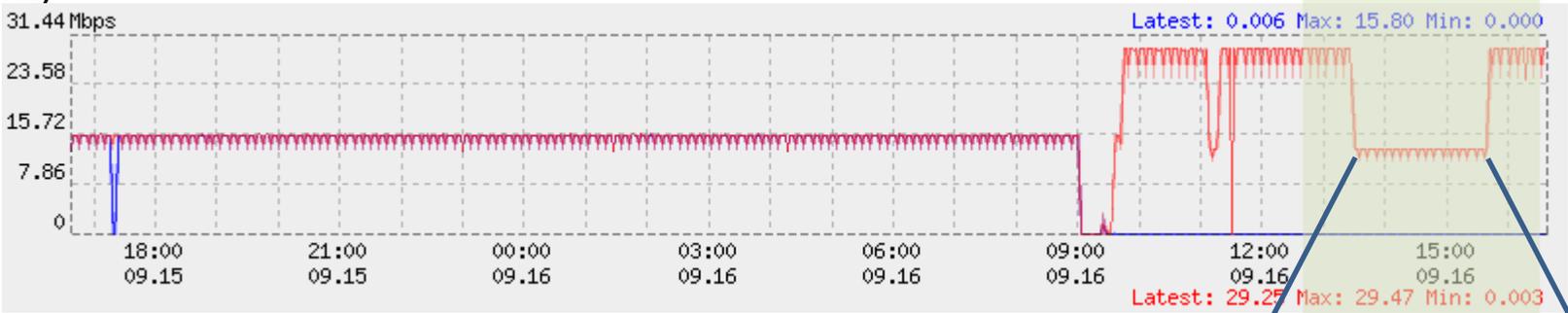
## Filter & Policer

```
firewall {
  family inet6 {
    filter ivi-qos {
      term blue {
        from {
          destination-address {
            2001:DA8:FF3A:C8FB:200::/128;
          }
        }
        then policer ivi-qos;
      }
      term default {
        then accept;
      }
    }
  }
  policer ivi-qos {
    if-exceeding {
      bandwidth-limit 15m;
      burst-size-limit 1500;
    }
    then discard;
  }
}
```

2001:DA8:FF3A:C8FB:200:: ↔ 58.200.251.2

# dIVI: IPv6 traffic shaping for IPv4 traffic

202.38.101.26 -> 2001:da8:ff3a:cbfb:200:: (w/ rate-limit)



202.38.101.26 -> 2001:da8:ff3a:cbfb:300:: (w/o rate-limit)

Beginning of rate-limit

End of rate-limit

# dIVI: Squid Server Deployment

- Squid Server is used to cache the content of IPv4 Web Server, through core IVI translation
- Squid Server is working in the Reverse Proxy Mode, listening to 80 HTTP port
- Dynamic Configuration: propagate the A && AAAA of Web Servers' records pointing to Squid Server into DNS system

# Summary

- dIVI/dIVI-PD can use existing tools for O&M
  - Null route
  - ACL
  - eACL
  - PBR
  - QoS
  - Caching
- Tunneling
  - No tools