

CERNET IPv6 multicast design and deployment

X. Li, C. Bao

2010-11-11

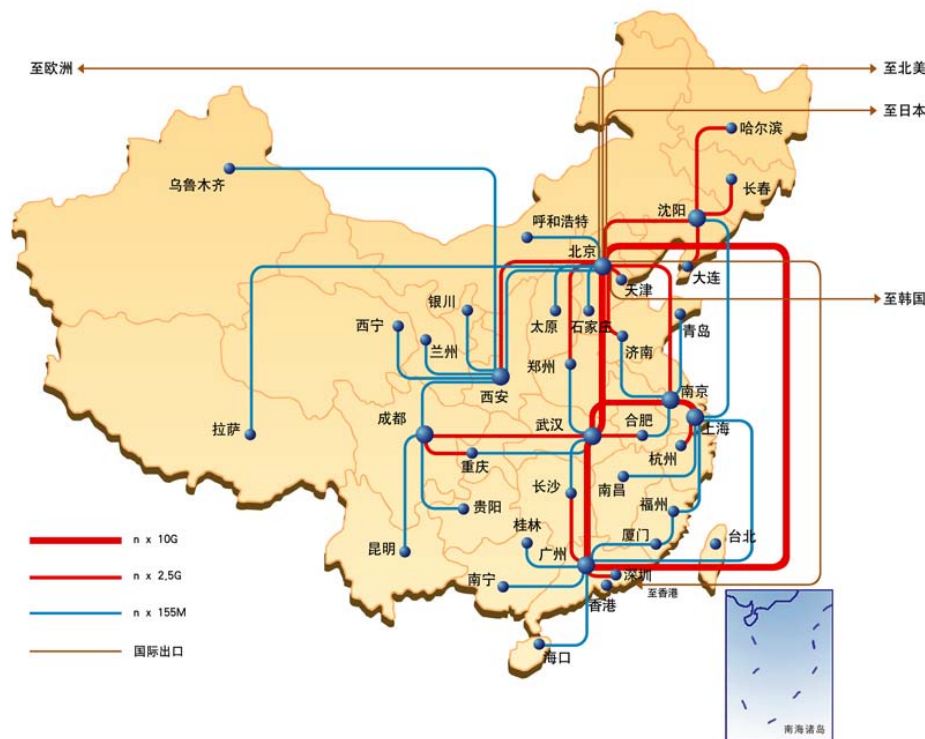
Outline



- Introduction
 - CERNET IPv4 multicast
 - CERNET2 IPv6 multicast
 - TEIN2/3 IPv4/IPv6 multicast
- CERNET2 IPv6 multicast
 - Design and implementation
- IPv4/IPv6 multicast translation
 - Prefix specific stateless translation (IVI)
 - Multicast IVI

The screenshot shows a Mozilla Firefox browser window with a blue title bar. The address bar displays 'http://lake.net.edu.cn:8033/map.html'. Below the address bar are navigation buttons (Back, Forward, Home) and a search bar. The main content area is divided into five columns, each representing a different network: CERNET, CNGI-CERNET2, CJI IPv6, NSFCNET, and TEIN2. Each column contains a map of China with network connections overlaid, and two line graphs showing network statistics over time (from 0 to 192). The graphs are labeled 'all-sum' and 'all-dnr'. The CERNET column has a 'main peers' link. The CNGI-CERNET2 column has a 'CNGI-6IX @' link. The bottom of the browser window shows a search bar with 'sam' and a '完成' (Complete) button. The Windows taskbar at the very bottom shows the '开始' (Start) button and several open applications, including Mozilla Firefox, Internet Explorer, and Microsoft Word. The system clock in the bottom right corner shows '22:58'.

CERNET (IPv4)



- CERNET is the first (1994) nation wide Internet backbone in China.
- CERNET ranks 30 in global CIDR report.
- Over 2,000 universities on CERNET with about 20M subscribers.

CERNET Multicast beacon

2010.11.10,22:04:49,0 (38/39)

packet loss (%)	bj	pku	bpt	tjn	sjz	tyn	hht	xa	xnn	lzh	ych	wliq	cd	chq	gyn	knn	lva	gz	nan	qln	hak	szh	wh	chs	zhz	nj	huf	jnn	qda	sh	hzh	nch	fzh	xnn	sy	dln	chc	hrb	
bj 219.243.224.162	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
pku 219.243.224.170	2	0	1	0	1	0	1	0	0	0	5	2	0	4	2	3	0	3	4	3	1	1	0	0	2	1	0	1	1	0	1	3	0	2	NA	0	1	0	1
bpt 219.243.224.178	2	0	3	0	2	0	1	0	0	0	0	1	2	0	2	1	1	0	2	3	1	3	3	1	0	0	1	2	1	2	2	0	2	NA	0	2	1	0	
tjn 219.243.225.162	2	0	1	0	1	0	2	0	0	0	2	1	0	1	3	3	0	3	3	3	1	0	3	0	0	1	5	2	2	3	0	2	NA	0	1	1	0		
sjz 219.243.226.162	2	0	1	0	0	0	2	0	0	0	1	1	0	1	1	2	0	3	2	3	3	4	3	0	0	2	1	1	2	3	0	3	NA	0	1	2	2		
tyn 219.243.227.162	2	0	1	0	2	0	1	0	0	0	3	3	0	3	3	0	4	3	2	1	1	1	0	0	3	1	1	3	2	0	3	NA	0	4	1	2			
hht 219.243.228.162	2	0	3	0	1	0	0	4	0	0	1	2	0	1	1	4	0	3	2	3	2	3	2	0	0	2	2	1	2	3	0	3	NA	0	3	1	2		
xa 219.243.229.162	2	0	1	1	2	0	0	0	0	0	1	1	0	4	2	3	3	4	2	3	2	1	1	1	0	3	1	1	2	2	2	4	NA	0	2	0	1		
xnn 219.243.230.162	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0			
lzh 219.243.231.162	0	0	0	0	0	0	0	NA	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0		
ych 219.243.232.162	0	0	0	0	0	0	0	NA	NA	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0		
wliq 219.243.233.162	2	1	2	3	1	1	2	3	0	0	0	0	1	1	2	3	1	2	1	1	2	1	2	3	3	2	2	3	1	2	1	3	4	NA	0	1	1		
cd 219.243.234.162	1	2	2	3	3	2	0	2	0	0	0	2	0	0	2	2	2	2	2	2	2	2	2	2	2	2	0	0	1	1	4	3	2	2	NA	0	3	2	2
chq 219.243.235.162	4	0	3	3	5	0	0	2	0	0	0	2	5	3	2	4	3	4	4	5	3	3	3	0	5	3	3	4	5	3	4	NA	0	6	5	5			
gyn 219.243.237.162	2	0	2	2	1	0	0	1	0	0	0	2	2	2	0	0	1	2	3	2	2	2	2	2	1	0	1	1	1	2	2	4	4	NA	0	1	1	1	
knn 219.243.236.162	2	1	1	1	1	2	1	1	0	0	2	1	1	1	1	0	0	2	4	4	2	1	1	3	1	1	1	0	2	2	2	2	3	NA	1	1	1	1	
lva 219.243.238.162	3	0	3	0	1	0	0	3	NA	0	0	1	2	0	0	0	1	0	2	2	2	1	2	2	0	0	4	2	1	3	3	0	4	NA	0	2	1	2	
gz 219.243.239.162	0	0	5	0	5	0	5	0	0	5	3	0	3	5	5	0	2	4	2	4	3	3	0	0	3	5	4	3	4	0	4	NA	0	4	4	5			
nan 219.243.240.162	0	0	4	0	5	0	0	6	0	0	5	0	0	5	0	3	7	5	0	4	5	2	5	4	4	0	0	3	4	5	3	2	0	3	NA	0	5	4	5
qln 219.243.240.170	0	0	6	0	8	0	0	6	0	0	7	0	0	0	5	6	6	0	0	5	5	5	5	0	0	5	6	7	5	0	0	4	NA	0	7	6	7		
hak 219.243.241.162	0	0	4	0	6	0	0	5	0	0	5	3	0	2	5	5	0	4	4	0	0	5	5	0	0	3	3	5	4	3	0	2	NA	0	5	6	5		
szh 219.243.242.162	0	0	3	0	5	0	0	3	0	0	3	1	0	2	4	0	1	2	2	0	0	4	2	0	0	1	2	3	0	1	0	2	NA	0	5	4	4		
wh 219.243.243.162	0	0	4	5	3	0	0	2	0	0	0	3	3	2	3	4	3	3	6	4	5	0	3	3	0	3	4	3	4	4	3	3	NA	0	4	3	3		
chs 219.243.244.162	2	0	3	2	2	0	0	0	0	0	4	2	2	2	4	2	3	3	3	2	3	3	0	2	0	2	3	3	4	3	4	2	NA	0	3	2	2		
zhz 219.243.245.162	0	0	1	1	1	0	0	1	0	0	0	2	1	3	2	2	2	4	3	3	3	2	2	2	0	0	3	2	1	3	3	2	3	NA	0	1	1	3	
nj 219.243.246.162	0	0	0	0	0	0	2	0	0	0	2	1	0	1	2	3	0	2	2	3	1	1	1	0	0	1	0	0	2	0	0	2	NA	0	1	2	1		
huf 219.243.247.162	3	3	3	0	2	4	5	2	0	0	3	2	0	2	4	3	0	2	2	3	3	3	0	4	0	4	3	3	2	3	0	3	NA	3	3	0	0		
jnn 219.243.248.162	1	0	1	0	1	0	0	1	0	0	2	2	0	1	2	2	0	3	3	2	3	1	1	0	0	2	0	3	3	0	3	NA	0	2	1	1			
qda 219.243.248.170	2	0	2	0	1	0	0	1	0	0	1	2	0	1	1	0	0	3	1	2	0	2	2	0	0	1	2	0	3	2	0	1	NA	0	1	1	0		
sh 219.243.249.162	0	0	4	0	6	0	0	5	0	0	5	3	0	2	4	5	0	4	2	4	2	3	5	0	0	2	6	4	0	0	2	NA	0	6	6	7			
hzh 219.243.250.162	0	0	4	0	4	0	5	0	0	0	4	3	0	3	3	5	0	2	3	1	4	3	0	0	2	4	6	2	0	1	NA	0	5	5	4				
nch 219.243.251.162	1	0	1	0	1	0	0	1	0	0	1	1	0	1	1	1	0	2	3	2	1	1	0	0	1	1	1	1	2	0	1	NA	0	1	1	1			
fzh 219.243.252.162	7	0	4	0	4	0	0	6	0	0	5	2	0	3	4	5	0	4	2	3	2	3	2	0	0	1	4	4	1	2	0	0	NA	0	4	5	4		
xnn 219.243.252.170	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA		
sy 219.243.253.162	1	2	3	3	1	1	1	4	0	0	1	1	0	2	2	0	3	0	4	2	3	3	2	0	3	2	2	2	5	4	0	2	NA	0	3	2	2		
dln 219.243.253.170	2	2	3	0	2	2	5	2	0	0	1	1	0	1	2	2	0	0	3	3	2	5	4	0	1	2	1	1	1	2	1	3	NA	0	4	3	3		
chc 219.243.254.162	2	1	2	2	1	2	4	0	0	0	2	2	1	1	1	4	0	3	2	2	1	1	2	0	1	3	1	1	3	1	1	3	1	NA	0	3	2	1	
hrb 219.243.255.162	1	3	4	2	1	2	1	1	0	0	1	1	0	2	1	0	2	0	2	2	2	4	2	0	3	2	2	1	2	2	0	1	NA	0	2	2	0		

CERNET Multicast AccessGrid



CERNET2 (IPv6)



- Built in 2004, with national coverage
- CERNET2 is the largest IPv6 backbone in China.
- About **200** universities connected to CERNET2 with about **2M** subscribers.

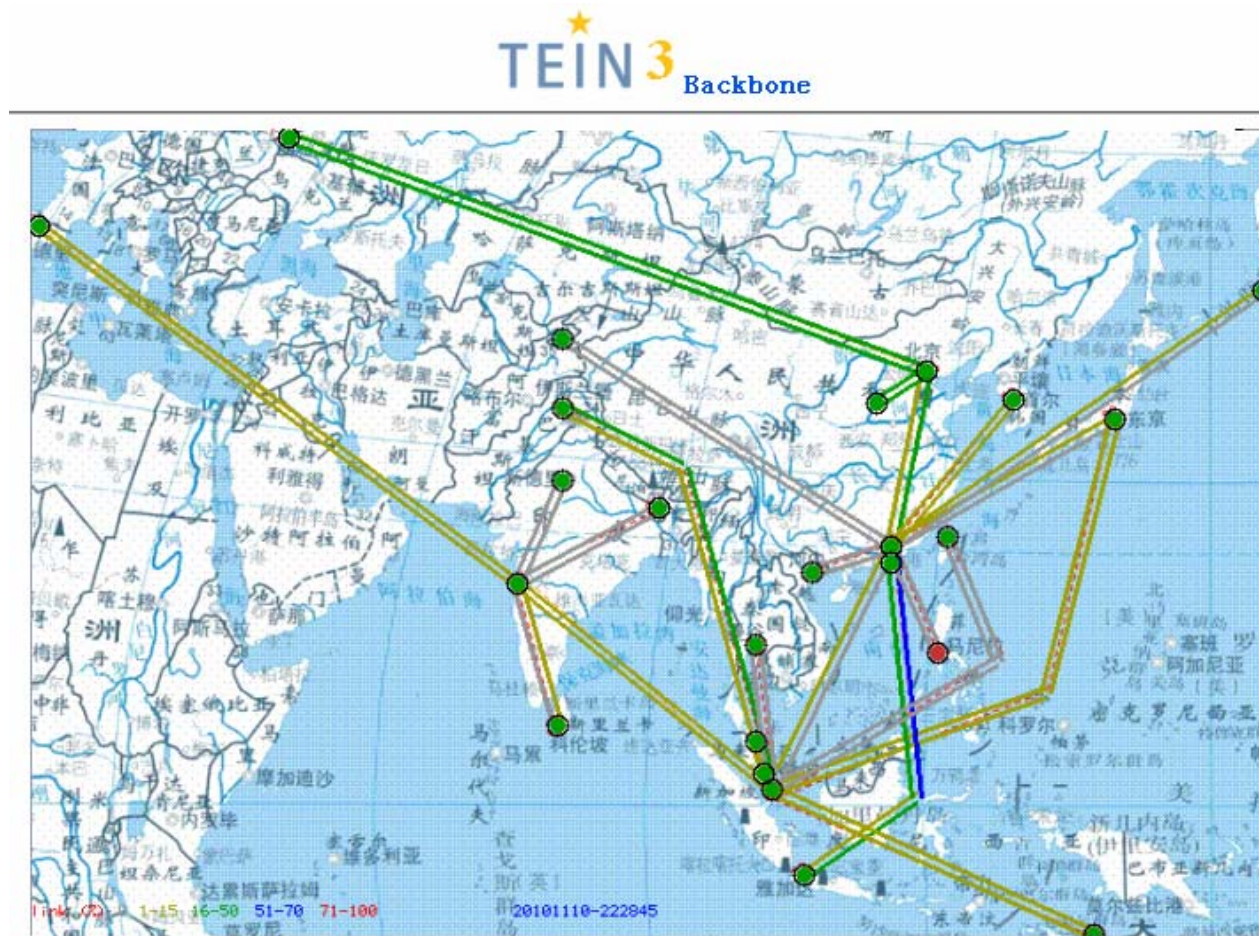
CERNET2 Multicast beacon

ssm

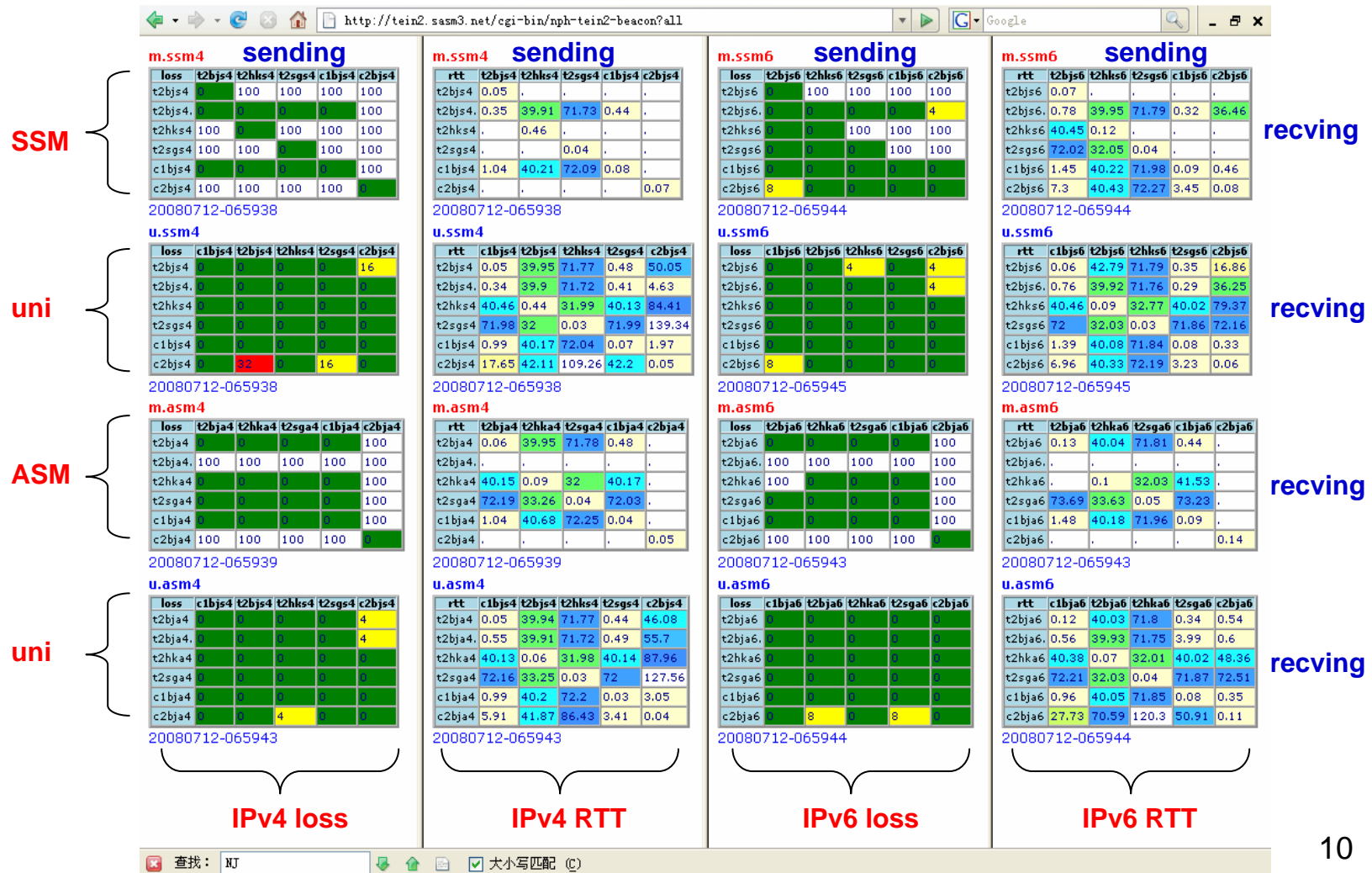
loss	bj	wh	gz	nj	sh	zhz	tjn	jnn	sy	dly	chc	hrb	pku	bupt	buaa	xa	cd	chq	lzh	hef	hzh	xmn	chs	fdv	tju
bj	0	0	0	20	100	0	0	100	100	0	0	100	100	100	100	0	0	0	0	25	0	0	100	25	0
wh	0	0	0	0	100	0	0	100	100	20	20	100	100	100	100	0	0	0	0	0	0	0	100	0	0
gz	0	0	0	0	100	0	0	100	100	0	0	100	100	100	100	0	0	0	0	0	0	0	100	0	0
nj	0	0	0	0	100	0	0	100	100	0	0	100	100	100	100	20	0	0	0	0	0	0	100	0	0
sh																									
zhz	0	20	0	0	100	0	0	100	100	0	40	100	100	100	100	0	0	0	0	0	0	0	100	0	0
tjn	0	0	0	0	100	0	0	100	100	0	0	100	100	100	100	0	0	0	0	0	0	0	100	0	0
jnn																									
sy	100	100	100	100	100	100	100	100	0	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100
dln	0	0	0	0	100	0	0	100	100	0	0	100	100	100	100	0	0	0	0	0	0	0	100	0	0
chc	0	0	0	20	100	0	0	100	100	0	0	100	100	100	100	0	0	0	0	0	0	0	100	0	0
hrb																									
pku																									
bupt																									
buaa	100	100	100	100	100	100	100	100	100	100	100	100	100	100	0	100	100	100	100	100	100	100	100	100	100
xa	0	0	0	0	100	0	0	100	100	0	0	100	100	100	100	0	0	0	0	0	0	0	100	0	0
cd	0	0	0	0	100	0	0	100	100	0	25	100	100	100	100	0	0	0	0	0	0	0	100	0	0
chq	0	0	0	0	100	0	0	100	100	0	0	100	100	100	100	0	0	0	0	0	0	0	100	0	0
lzh																									
hef	20	0	0	0	100	0	0	100	100	0	0	100	100	100	100	0	0	0	0	0	0	0	100	0	0
hzh	0	0	0	0	100	0	0	100	100	0	0	100	100	100	100	0	0	0	0	0	0	0	100	0	0
xmn	0	0	0	0	100	0	0	100	100	0	0	100	100	100	100	0	0	0	0	0	0	0	100	0	0
chs																									
fdv	0	0	0	0	100	0	0	100	100	0	0	100	100	100	100	0	0	0	0	0	0	0	100	0	0
tju	0	0	0	0	100	0	0	100	100	0	0	100	100	100	100	0	0	0	0	0	0	0	100	0	0

20101110-214721

TEIN3 (IPv4/IPv6)

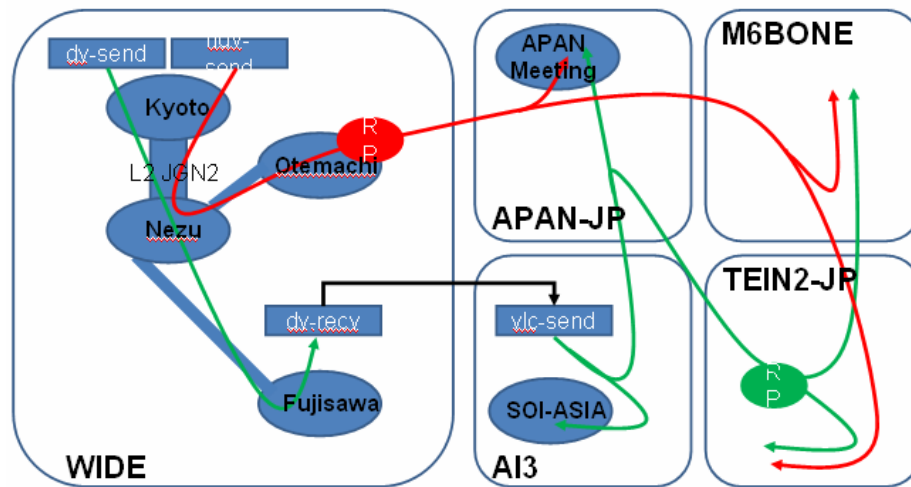


TEIN2/TEIN3 multicast beacon

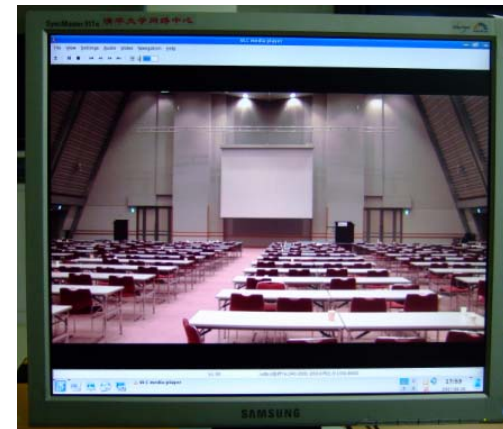


SIGCOMM 2007 via TEIN2

Multicast Topology @ SIGCOMM 2007



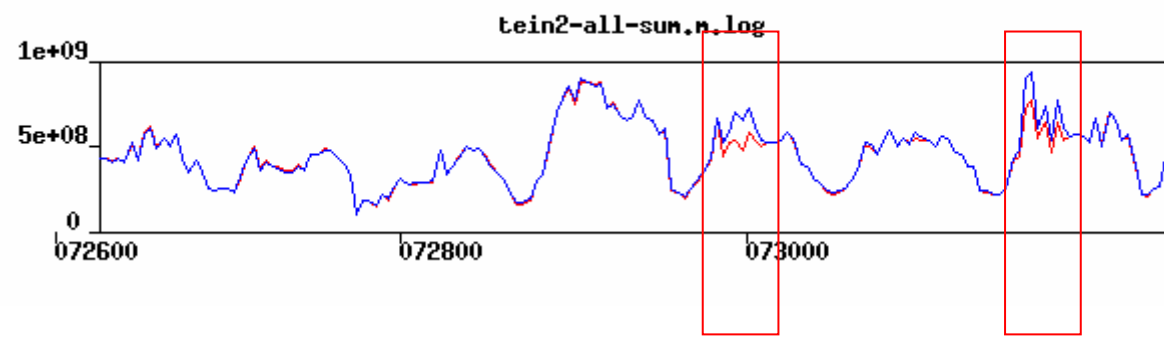
- Streaming through TEIN2 IPv6 multicasts;
 - Audience:
 - INHERENT universities (Indonesia)
 - MYREN universities (Malaysia)
 - UNINET universities (Thailand)
 - Client application : VideoLAN
 - Streams: 1Mbps and 5Mbps



```
smjoin6 ff7e:240:2001:200:0:ff01:0:100 -p 8000 -c -1 -k 3000
```

mode	source addr:port	group addr:port	pack size	counter
ASM:	[2001:200:0:ff20:216:d3ff:fe3b:6803]:32769	[ff7e:240:2001:200:0:ff01:0:100]:8000	16 bytes MSG	(13452000)
ASM:	[2001:200:0:ff20:216:d3ff:fe3b:6803]:32769	[ff7e:240:2001:200:0:ff01:0:100]:8000	16 bytes MSG	(13456000)
ASM:	[2001:200:0:ff20:216:d3ff:fe3b:6803]:32769	[ff7e:240:2001:200:0:ff01:0:100]:8000	16 bytes MSG	(13458000)
ASM:	[2001:200:0:ff20:216:d3ff:fe3b:6803]:32769	[ff7e:240:2001:200:0:ff01:0:100]:8000	103 bytes MSG	(13461000)
ASM:	[2001:200:0:ff20:216:d3ff:fe3b:6803]:32769	[ff7e:240:2001:200:0:ff01:0:100]:8000	18 bytes MSG	(13464000)
ASM:	[2001:200:0:ff20:216:d3ff:fe3b:6803]:32769	[ff7e:240:2001:200:0:ff01:0:100]:8000	103 bytes MSG	(13467000)

TEIN2 traffic



multicast

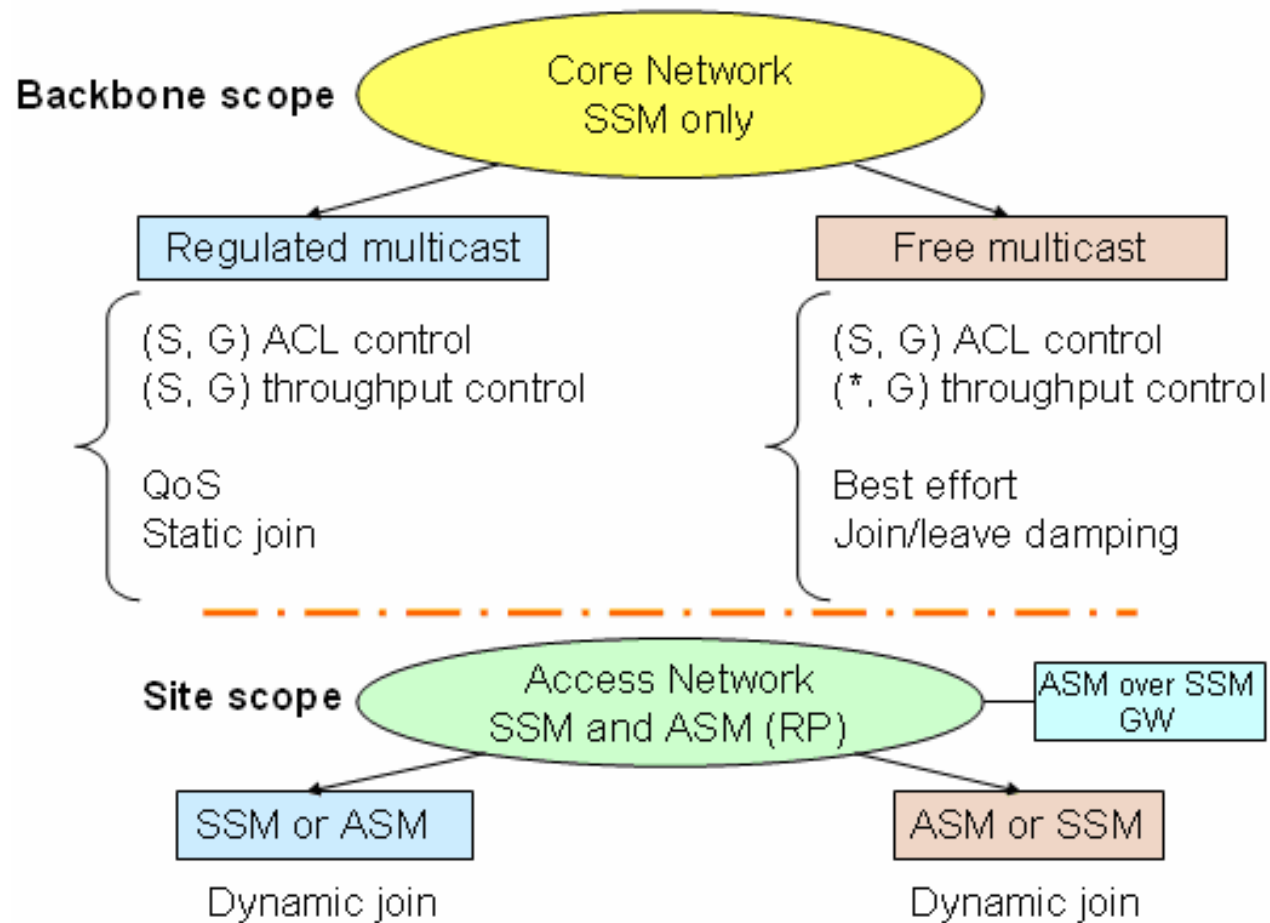
CERNET2 multicast design

- CERNET2 is an IPv6-only network
- We try to design a scalable and controllable IPv6 multicast network
- The challenges are
 - The individual (non-aggregate) state in the multicast forwarding path has to be maintained to map the specific group address to a set of destination unicast address in a specific time.
 - The state is dynamic, triggered by the joins/leaves of the receivers.
 - The source addresses which are sending packets to the specific group address have to be found in order to form the distribution tree.

CERNET2's approach

- Provide IPv6 SSM only service
- Only a special IPv6 block can serve as multicast source
- Embed bandwidth parameter into the IPv6 group address
- Static join in PEs
- Provide ASM to SSM translation service

SSM



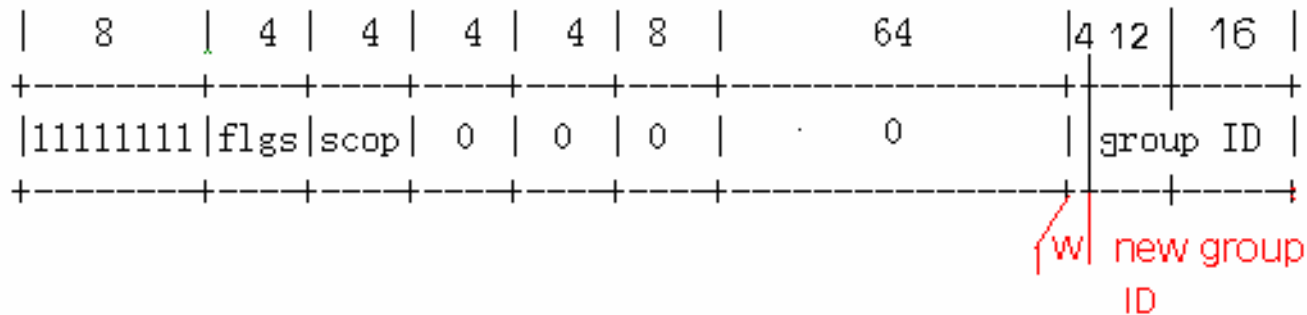
Multicast source



- CERNET2 address plan 2001:da8::/32
 - Each campus network will get
 - A /48 for unicast
 - 2001:da8:200::/48
 - A /64 for multicast source
 - 2001:da8:3ffe:200::/64

Group address

(a) SSM



The (W) represents the service throughput

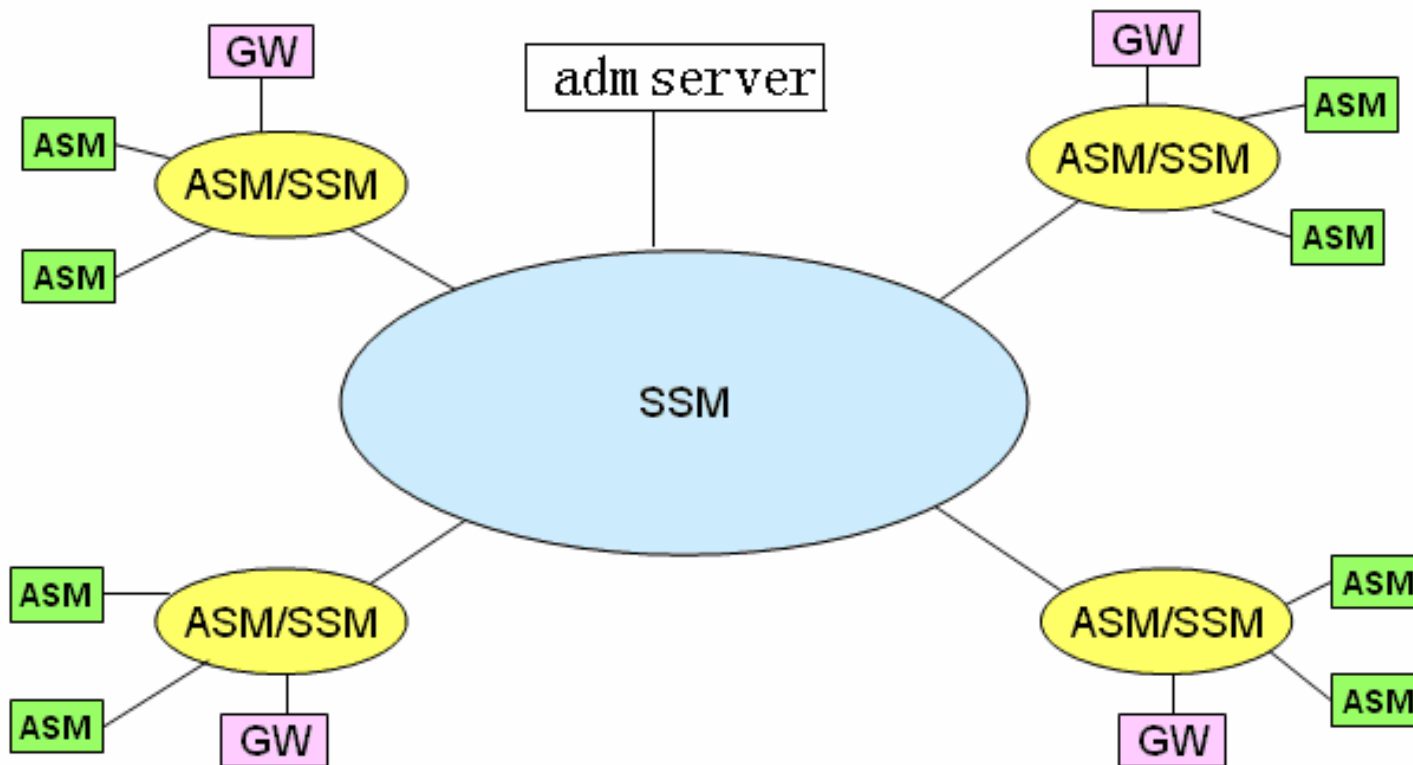
- W= 0x8: 100 Kbps,
- W= 0xC: 1 Mbps,
- W= 0xE: 10 Mbps,
- W= 0xF: 100 Mbps.

Static join

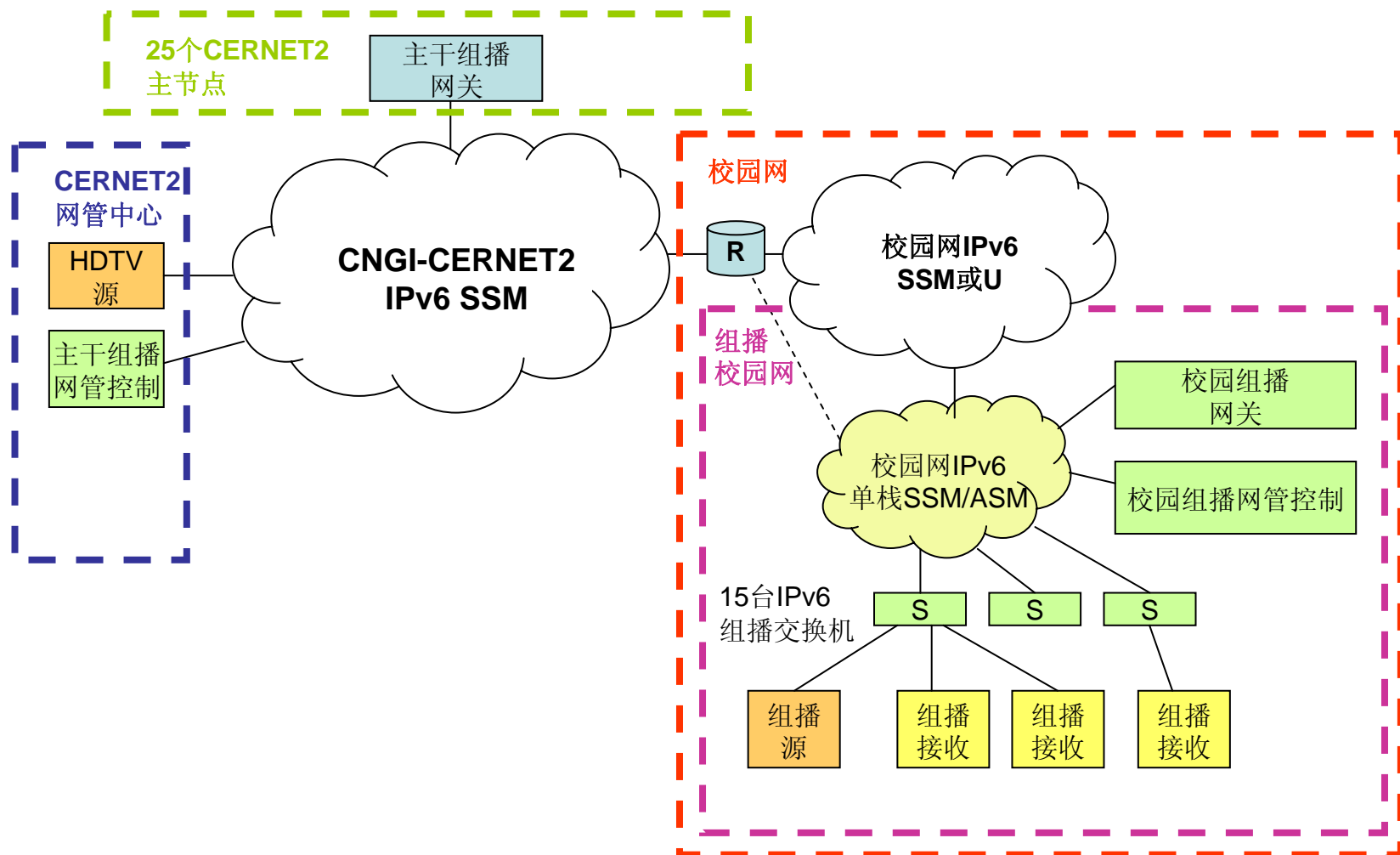


- There will be no dynamic join and state changes in the core network and this will make the core more stable;
- The join convergence time will be eliminated;
- It is possible to do the source and group aggregation in the future.

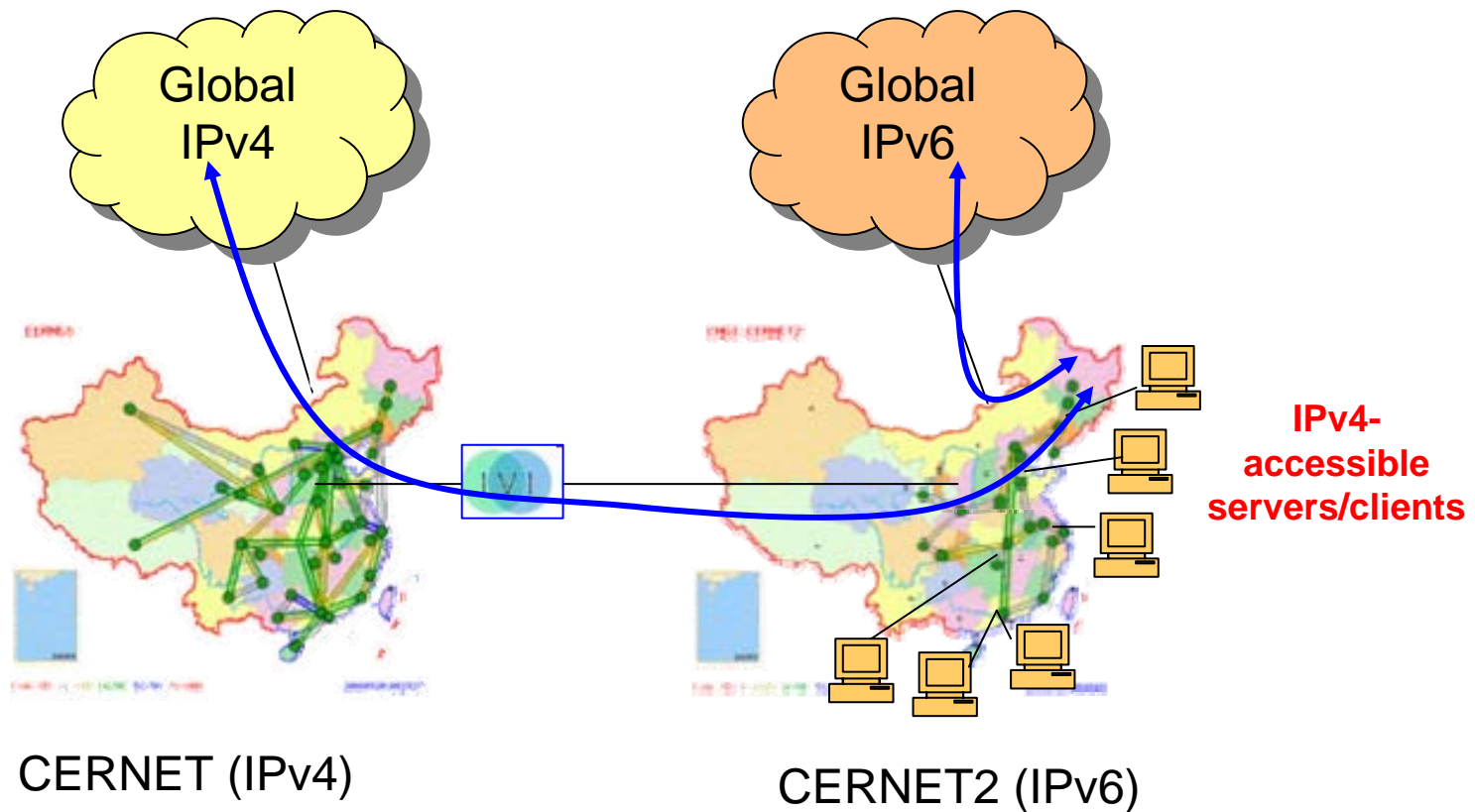
ASM to SSM translation



CERNET2: 100 campus project



Prefix specific and stateless Translation



IETF standards

Internet Engineering Task Force (IETF) C. Bao
Request for Comments: 6052 CERNET Center/Tsinghua University
Updates: 4291 C. Huitema
Category: Standards Track Microsoft Corporation
ISSN: 2070-1721 M. Bagnulo
UC3M
M. Boucadair
France Telecom
X. Li
CERNET Center/Tsinghua University
October 2010

IPv6 Addressing of IPv4/IPv6 Translators

behave F. Baker
Internet-Draft Cisco Systems
Intended status: Informational X. Li
Expires: November 19, 2010 C. Bao
CERNET Center/Tsinghua University
K. Yin
Cisco Systems
May 18, 2010

Framework for IPv4/IPv6 Translation
draft-ietf-behave-v6v4-framework-09

Network Working Group X. Li
Internet-Draft C. Bao
Intended status: Informational M. Chen
Expires: July 10, 2010 H. Zhang
J. Wu
CERNET Center/Tsinghua University
January 6, 2010

The CERNET IVI Translation Design and Deployment for the IPv4/IPv6
Coexistence and Transition
draft-xli-behave-ivi-07

behave X. Li
Internet-Draft C. Bao
Obsoletes: 2765 (if approved) CERNET Center/Tsinghua University
Intended status: Standards Track F. Baker
Expires: November 19, 2010 Cisco Systems
May 18, 2010

IP/ICMP Translation Algorithm
draft-ietf-behave-v6v4-xlate-20

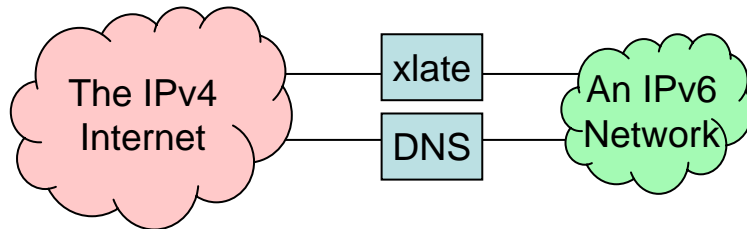
BEHAVE WG M. Bagnulo
Internet-Draft UC3M
Intended status: Standards Track P. Matthews
Expires: January 11, 2011 Alcatel-Lucent
I. van Beijnum
IMDEA Networks
July 10, 2010

Stateful NAT64: Network Address and Protocol Translation from IPv6
Clients to IPv4 Servers
draft-ietf-behave-v6v4-xlate-stateful-12

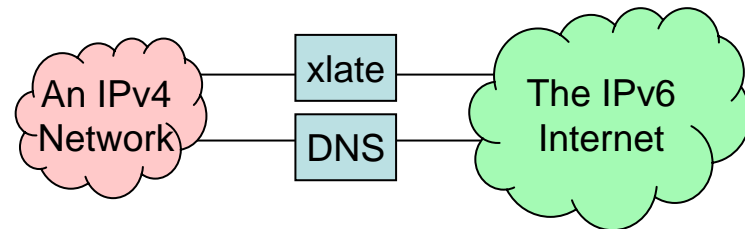
BEHAVE WG M. Bagnulo
Internet-Draft UC3M
Intended status: Standards Track A. Sullivan
Expires: April 4, 2011 Shinkuro
P. Matthews
Alcatel-Lucent
I. van Beijnum
IMDEA Networks
October 1, 2010

DNS64: DNS extensions for Network Address Translation from IPv6 Clients
to IPv4 Servers
draft-ietf-behave-dns64-11

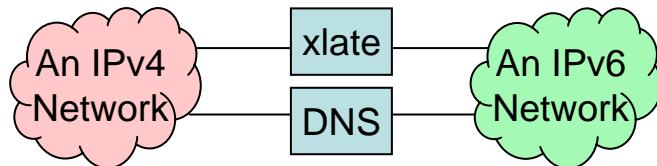
Translation scenarios



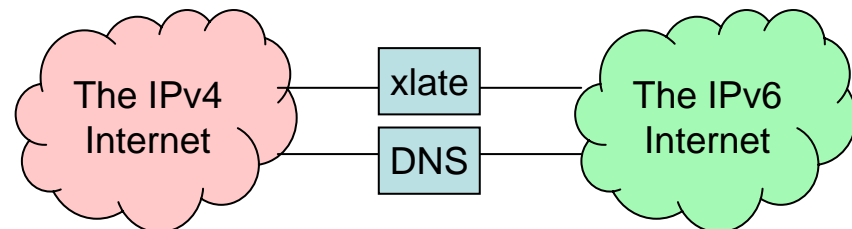
IVI { Scenario 1 “an IPv6 network to the IPv4 Internet” < **NAT64**
 Scenario 2 “the IPv4 Internet to an IPv6 network”



Scenario 3 “an IPv4 network to the IPv6 Internet” < **NAT64**
 Scenario 4 “the IPv6 Internet to an IPv4 network”

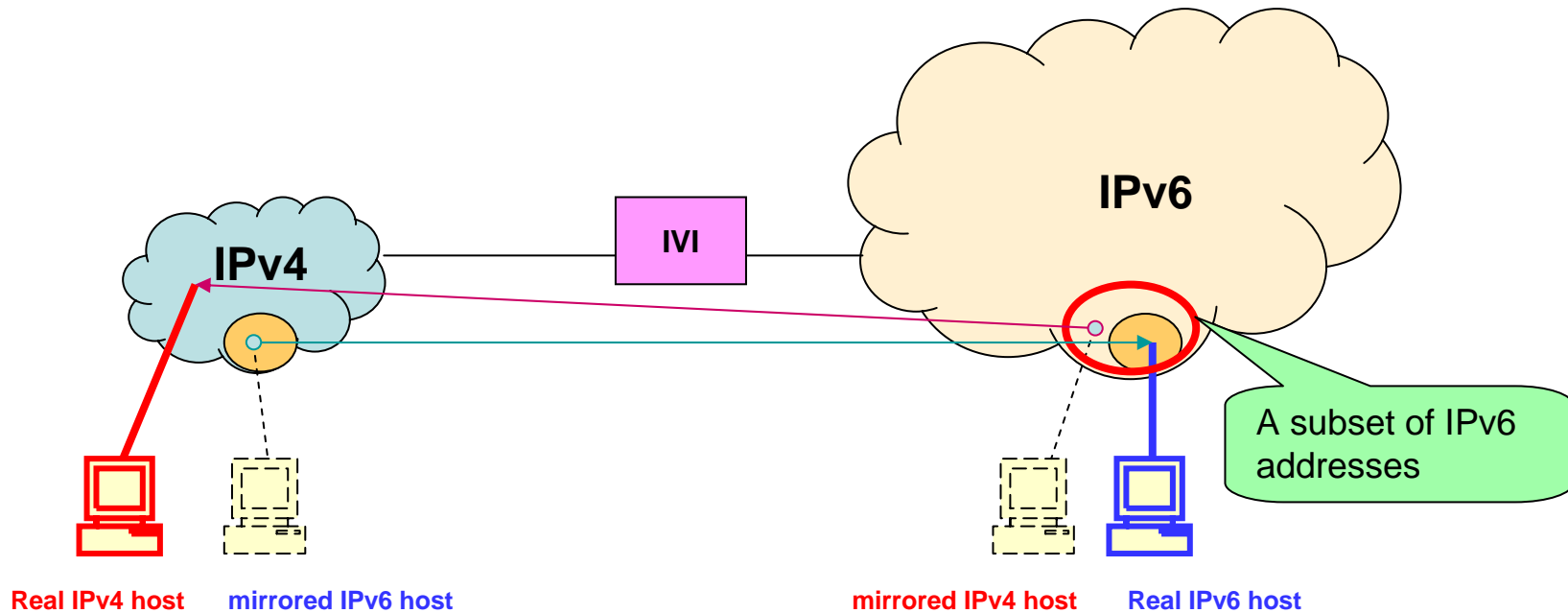


IVI { Scenario 5 “an IPv6 network to an IPv4network” < **NAT64**
 Scenario 6 “an IPv4 network to an IPv6 network”



Scenario 7 “the IPv6 Internet to the IPv4 Internet”
 Scenario 8 “the IPv4 Internet to the IPv6Internet”

Stateless translation concepts

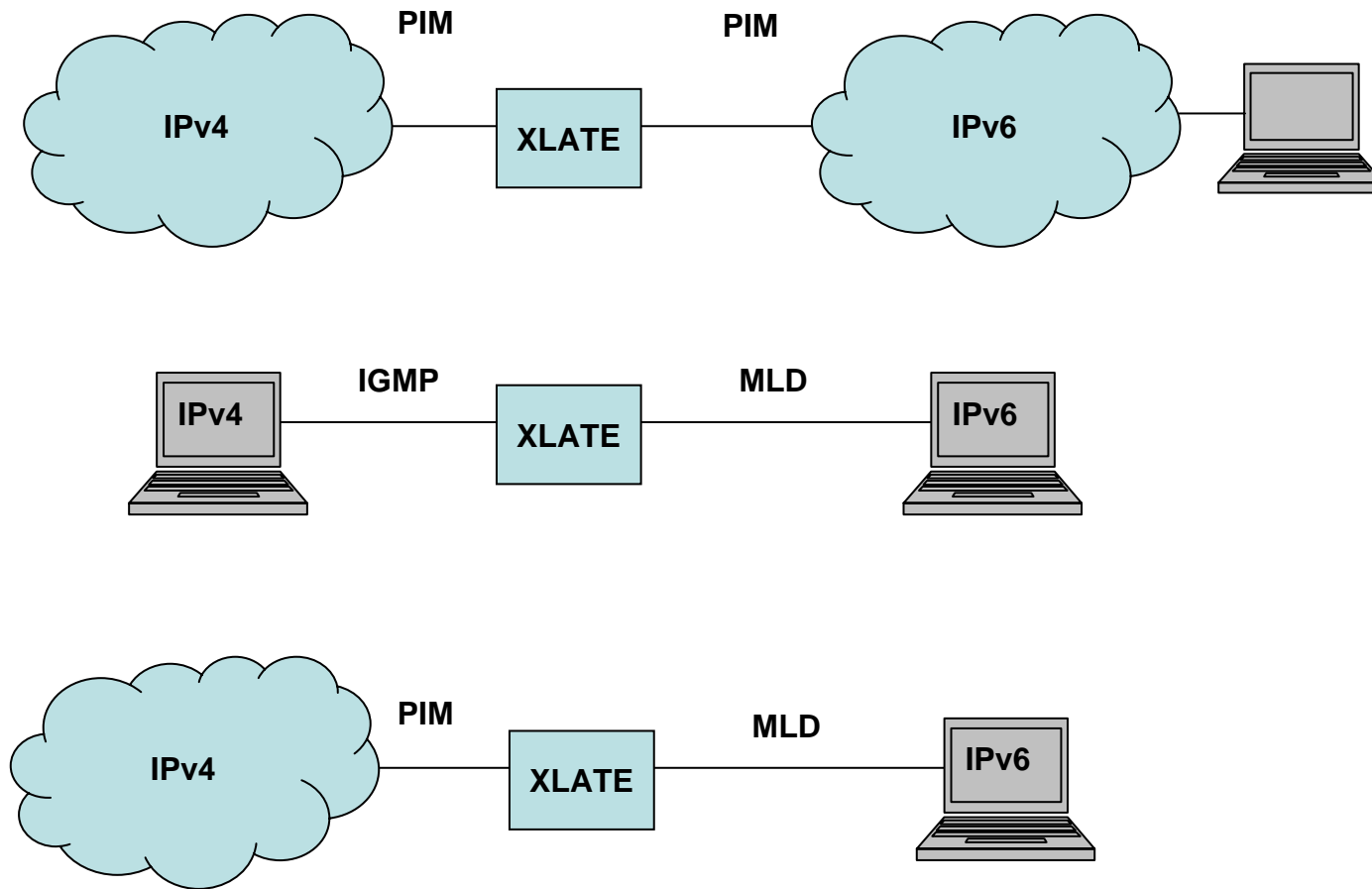


A subset of IPv6 addresses

Operation scenarios

- Stateless mode (IVI) supports
 - Scenario 1
 - an IPv6 network to the IPv4 Internet
 - IPv4 sending and IPv6 joining
 - Scenario 2
 - the IPv4 Internet to an IPv6 network
 - IPv6 sending and IPv4 joining
 - Scenario 5 and scenario 6
 - Same as scenario 1 and scenario 2
- Stateful mode (NAT64) supports
 - Scenario 1
 - an IPv6 network to the IPv4 Internet
 - IPv4 sending and IPv6 joining
 - Scenario 3
 - the IPv6 Internet to an IPv4 network
 - IPv4 sending and IPv6 joining
 - Scenario 5
 - Same as scenario 1

Routing



IVI Multicast support

- SSM is supported for the IVI
 - no MSDP in IPv6
 - no embedded RP in IPv4
- Group address mapping rule (there will be 2^{24} group ID available)
 - 232.0.0.0/8 → ff3e:0:0:0:0:0:f000:0000/96
 - 232.255.255.255/8 → ff3e:0:0:0:0:0:f0ff:ffff/96
- For the cross address family SSM
 - the source address in IPv6 has to be IVI6 for the RPF scheme
- The inter operation of PIM-SM in IPv4 and IPv6
 - Application layer gateway
 - Static join using IGMPv3 and MLDv2

Remarks

- Multicast is very useful for academic applications
- IPv6 multicast has more flexibility
- Tools and router access are both important.
- IPv6 multicast configuration and debugging process is an enjoyable process (-: