



MSK-IX

Ильин А.Ю.
kalend@ripn.net

Московский Internet Exchange



История

- Первая точка обмена трафиком в России была организована в Москве на ММТС-9 (M9) в 1995
 - 7 ISP пришли к соглашению о создании точки взаимного обмена трафиком
 - Координация M9-IX была поручена РосНИИРОС (RIPN)
 - Базировался M9-IX на Ethernet- коммутаторе CS5505, первые подключения выполнялись на скорости 10 Мбит/с



Технологическая площадка в РНЦ «Курчатовский Институт»



MSK
IX

Технологическая площадка в РНЦ «Курчатовский Институт»



Технологическая площадка в РНЦ «Курчатовский Институт»



MSK
IX

Технологическая площадка в РНЦ «Курчатовский Институт»



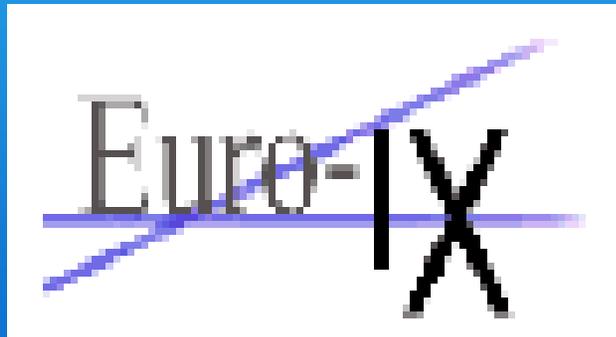
Участники

- В настоящее время 168 сетей имеют подключения к MSK-IX
- Большинство подключены к порту пропускной способностью 100 Мбит/с
- Крупные ISP подключены на 1 Гбит/с
- Подключения имеют Московские, региональные сети, сети ISP стран СНГ
- Около 85% участников используют Route Server



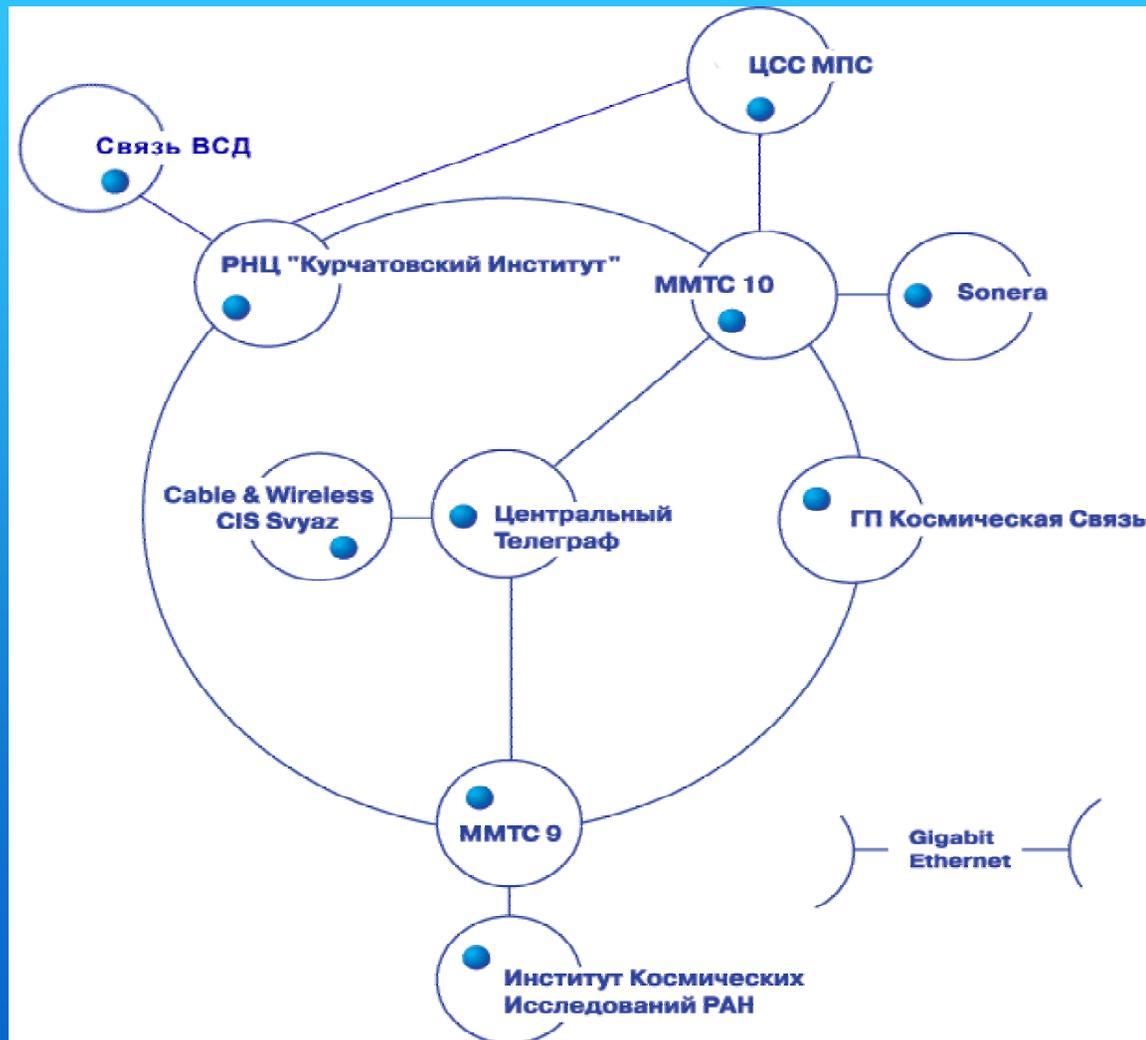
MSK-IX at EURO-IX

- MSK-IX член ассоциации европейских Internet Exchanges
- EURO-IX объединяет 35 IXP из 23 европейских стран

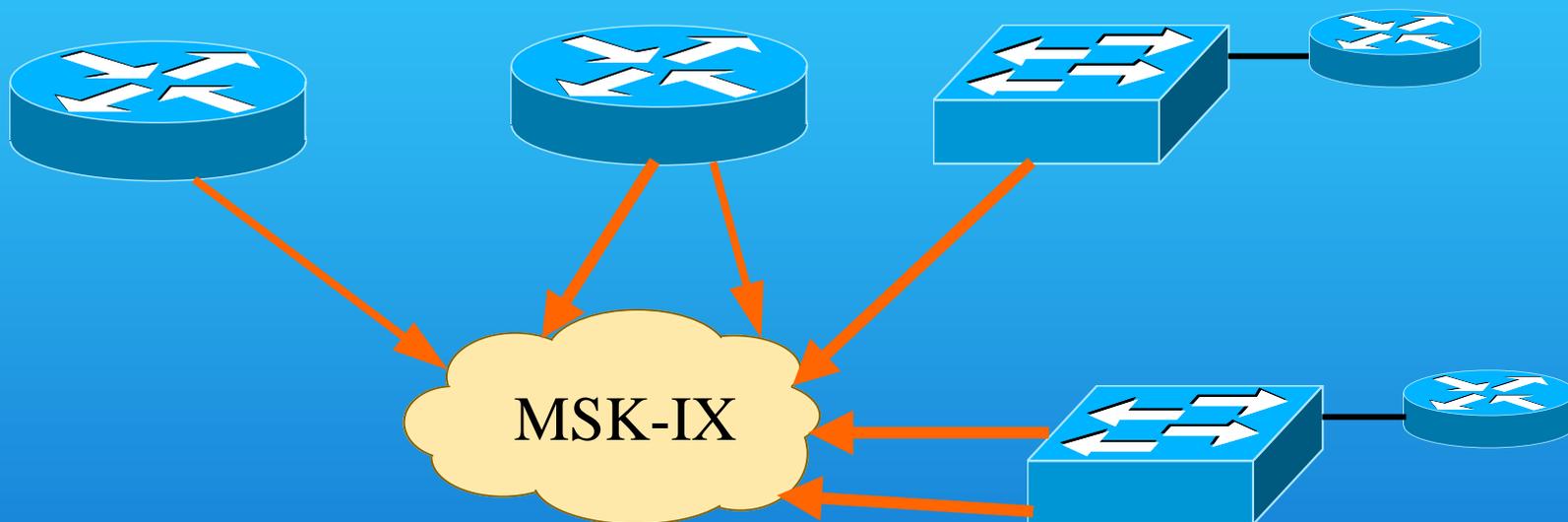


Структура MSK-IX

- 10 центральных узлов соединенных по Gigabit Ethernet
- Резервирование по питанию и оптике
- CS 6506 в основных точках
- 10 Gigabit между центральными узлами на M9
- 24x7 мониторинг MSK-IX



Варианты подключений к IX



- Предпочтительнее разные точки подключения к IX для резервирования
- Обязательно использовать разные номера VLAN в случае свитча
- Недопускается подключение нескольких маршрутизаторов через один порт MSK-IX

Адресация MSK-IX

- Клиенту выделяется два адреса из двух сетей /23 с целью резервирования.
- В зависимости от схемы подключения клиент должен использовать два адреса на одном устройстве или подключаться к двум портам MSK-IX.
- Клиент обязан использовать два адреса на MSK-IX и устанавливать BGP сессии с другими участниками с использованием двух адресов

Услуги MSK-IX

- Подключение к портам оборудования MSK-IX (Ethernet 10/100/1000)
- Полное резервирование оборудования
- Два подключения к оборудованию MSK-IX, размещенному на разных технологических площадках



Услуги MSK-IX

- **Приватные Ethernet-соединения между технологическими площадками MSK-IX для Участников подключения (peer2peer connection)**
- **Приватные соединения между участниками MSK-IX (private peering group)**
- **Размещение оборудования Участников подключения на технологических площадках MSK-IX (colocation)**



Услуги MSK-IX

- **Multicast Internet Exchange (Multicast-IX)**
- **14 Участников имеют подключения к Multicast-IX**
- **IPv6 Internet Exchange (Pilot Project)**
- **Доступ к ISP в Санкт-Петербурге непосредственно с VLAN MSK-IX (MSK-SPB-IX)**



Услуги MSK-IX

Информационные услуги

- **Route Server**
- **F.Root name server**
- **IP transit**
- **RIPE RIS**



Route server

- **RS предоставляет возможность уменьшить количество "парных взаимодействий" с переходом к обмену роутинговой информацией через одну точку**
- **Интернет-трафик передается между подключенными к MSK-IX сетями напрямую, минуя RS.**
- **Входящие фильтры RS участников строятся в соответствии с RIPv6db раз в сутки.**
- **2 сервера RS расположены на двух различных площадках**



F.Root name server

- Под управлением ISC (<http://www.isc.org>)
- Размещен в KIAE-HOUSE
- дублированная схема подключения к IX
- Подключен к RS (доступ всех к F.Root)



Routing Information Service

- **Statistics**

- • [RISreport](#) generates graphs by using data from the RIS database. This tool was designed by Thomas Franchetti for his masters thesis.
- • [IPv6 routing activities](#)
- • [Weekly Statistics Report](#)
- • [AS Statistics](#) (per RRC)
- • [Prefix Statistics](#) (per RRC/entire RIS).
- • [Martians](#), daily lists of prefixes within "special use" IPv4 address ranges as documented by Bill Manning in his [Internet Draft](#).
- • [BGP Traffic Hot Spots](#) (per RRC/entire RIS). Daily lists of prefixes for which high BGP announcement or withdrawal activity has been observed by our Route Collectors.
- • [MOAS Report](#). Prefixes being announced by multiple origin ASes.
- [Agilent Labs' RIPE NCC data analysis project](#) is to visualize and characterize Border Gateway Protocol (BGP4) behavior.



Routing Information Service

- Peering information for all RRCs

	AS number
• RRC00, RIPE NCC	12654
• RRC01, LINX	12654
• RRC02, SFINX	12654
• RRC03, AMS-IX, NL-IX, GN-IX	12654
• RRC04, CIXP	12654
• RRC05, VIX	12654
• RRC06, NSPIXP2	12654
• RRC07, Netnod	12654
• RRC10, MIX	12654
• RRC11, NYIIX	12654
• RRC12, DE-CIX	12654
• RRC13, MSK-IX	12654
• RRC14, PAIX	12654



УСЛОВИЯ ПОДКЛЮЧЕНИЯ К MSK-IX

- Участник обязан иметь собственную AS, и зарегистрировать роутинговую политику этой AS в RIPv6db
- Участник обязан использовать один MAC-адрес для каждого выделенного для подключения порта
- Участник обязан отключить на интерфейсе подключения к IX все виды broadcast за исключением ARP
- Полный текст требований приведен по адресу <http://www.msk-ix.ru/rus/order/required.shtml>



Основные проблемы клиентов на IX

- Несоблюдения правил подключения к IX
- Подключение свитчем на порт MSK-IX с включенным STP
- Подключение двумя портами в одном VLANe (!)
- Паразитный трафик (OSPF, DHCP, TFTP...)
- Анонсирование сетей IX в мир
- Невыключенный проху-arp
- Отсутствует BGP next-hop-self



Ошибки использования RS на IX

- Выдача upstream провайдеру полученных анонсов
- На MSK-IX **2 (!)** Route Server на 2-х технологических площадках – нужно поднимать BGP сессии с каждым из них
- Участник должен принимать **все (!)** анонсы от RS
- Корректная настройка BGP с RS: (пример)
bgp always-compare-med
set local-preference (одинаковое для всех пиров на MSK-IX)
set metric (меньшее значение для анонсов с RS)



Пример конфигурации Cisco Router

```
no service dhcp / no ip bootp server
```

```
no cdp run
```

```
!
```

```
Interface FastEthernet0
```

```
no ip redirects
```

```
no ip proxy-arp
```

```
no cdp enable
```

```
no ip directed-broadcasts
```

```
no mop enable
```

```
duplex full (no autonegotiation)
```

```
no keepalive
```



Пример конфигурации свитча (IOS)

```
vtp mode transparent
!  
No spanning-tree vlan 100
!  
Vlan 100
  name MSK-IX
!  
Interface FastEthernet0/0
  description MSK-IX connection
  switchport mode access
  switchport nonegotiate
  switchport access vlan 100
  no keepalive
  speed nonegotiate
  no cdp enable
  no udld enable
  spanning-tree bpdupfilter enable
!
```

Обязательно отключать
посылку BPDU фреймов в
сторону MSK-IX(!)

- bpdupfilter enable
- no spanning tree vlan xxx

Защита IX от проблем

- На подключениях клиентов `bpdu-guard` с `errdisable timeout` равным 5 минутам
- Увеличено `span aging time`
- `Port-security`
- `Block unknown unicast/multicast`
- Мониторинг IX

Мониторинг MSK-IX

- 24 x 7
- статус подключений
- загрузка каналов
- ошибки/потери
- отсутствие трафика
- перегрузка
- журнал событий
- приоритет обслуживания



Полный вид на MPS-S1 - Microsoft Internet Explorer

File Edit View Favorites Tools Help

Address

Tue Sep 13 22:06:02 2005 1622[TOTAL]

Устройство: MPS-S1

статус	журнал	график	увел	сводка	карточка	YP-NEW	архив	войти	стат
Имя	Приоритет	Причина	Стат	Время	Загрузка CPU	Свободная память	Общая CPU	i/o	
MPS-S1	3	27d3h	U	27d3h	0%	OK	K	K	
int name	prior	in/out	ST	input		output			
GE2/49 m10-s2-mps-s1	3	=====	U	0.3% 1398.3 p/s	0.0% err	0.6%	1320.6 p/s	0.0% drps	
GE2/50 mps-s1-kiac-s11	3	=====	U	8.9% 16190.9 p/s	0.0% err	5.2%	15648.3 p/s	0.0% drps	
FE2/1 mps	2	=====	U	4.1% 1998.0 p/s	0.0% err	4.3%	936.4 p/s	0.0% drps	
FE2/2 tsinet	2	=====	U	7.5% 104.8 p/s	0.0% err	1.0%	71.7 p/s	0.0% drps	
FE2/3 pcsnet	3	=====	U	0.9% 784.4 p/s	0.0% err	4.2%	806.6 p/s	0.0% drps	
FE2/4 okbte1	3	=====	U	1.1% 39.8 p/s	0.0% err	0.6%	34.0 p/s	0.0% drps	
FE2/5 rootelecom	2	=====	U	0.0% 0.1 p/s	0.0% err	0.0%	5.7 p/s	0.0% drps	
FE2/6 ctk	3	=====	U	11.8% 1513.5 p/s	0.0% err	12.6%	733.0 p/s	0.0% drps	
FE2/7 ltech	3	=====	U	0.8% 85.4 p/s	0.0% err	12.4%	652.7 p/s	0.0% drps	
FE2/8 sovintel-3.2	1	=====	U	11.0% 945.1 p/s	0.0% err	96.5%	1509.5 p/s	0.0% drps	
FE2/9 progivk	3	=====	U	14.2% 702.3 p/s	0.0% err	54.8%	756.8 p/s	0.0% drps	
FE2/10 runnet-3.2	1	=====	U	15.6% 4265.7 p/s	0.0% err	36.0%	5193.6 p/s	0.0% drps	
FE2/11 pcsnet-2.2	1	=====	U	6.0% 1310.6 p/s	0.0% err	3.1%	1253.3 p/s	0.0% drps	
FE2/12 373centr	1	=====	U	0.0% 0.5 p/s	0.0% err	0.0%	0.7 p/s	0.0% drps	
FE2/13 progivk-2.2	1	=====	U	25.8% 5229.7 p/s	0.0% err	24.8%	5126.5 p/s	0.0% drps	

Мониторинг MSK-IX

утилита IXP Watch + Ethereal

- «слушает» трафик на свободном порту IX
- ARP, Excessive ARPS
- NON-IP traffic
- NON Unicast traffic (OSPF, DHCP, ...)
- Испорченные фреймы
- Странные ICMP



Мониторинг MSK-IX

```
- PuTTY
-----
MCK-IX Traffic Summary Report - 2005-09-13
-----
Analysis based on a sample of 10 minutes.
Started at 2005-09-13 23:00, ended at 2005-09-13 23:10
-----
ARP Queries.....: 2978
ARPs per minute..: 297
IP Packets.....: 0
ICMP Packets.....: 0
NON-IP Packets...: 2
Incorrect ARPs...: 27
-----

::::: NON IP TRAFFIC
-----
  1 00:11:92:1a:1a:82 -> ab:00:00:02:00:00 0x6002 DEC DNA Remote Console
  1 00:11:92:1a:1a:80 -> ab:00:00:02:00:00 0x6002 DEC DNA Remote Console
-----

::::: NON UNICAST IP TRAFFIC
-----

::::: ICMP MESSAGES
-----

::::: ARPs for/from non-MCK-IX IP addresses
-----
  8 00:02:17:43:68:01 -> ff:ff:ff:ff:ff:ff ARP Who has 10.0.7.19? Tell 193.232.244.221
  2 00:05:5d:75:6f:ed -> ff:ff:ff:ff:ff:ff ARP Who has 193.232.244.151? Tell 212.113.1
14.3
```

Контакты

Административные вопросы:

E-mail: msk-ix-adm@ripn.net

Технические вопросы:

E-mail: msk-ix-noc@ripn.net

