Jumbo Frames in AMS-IX

version 0.3

Maksym Tulyuk
maksym.tulyuk@ams-ix.net

RIPE 63. Vienna
03 November 2011

Survey: only 5 questions
https://www.surveymonkey.com/s/TGS5T2D

Maksym Tulyuk
maksym.tulyuk@ams-ix.net

RIPE 63. Vienna
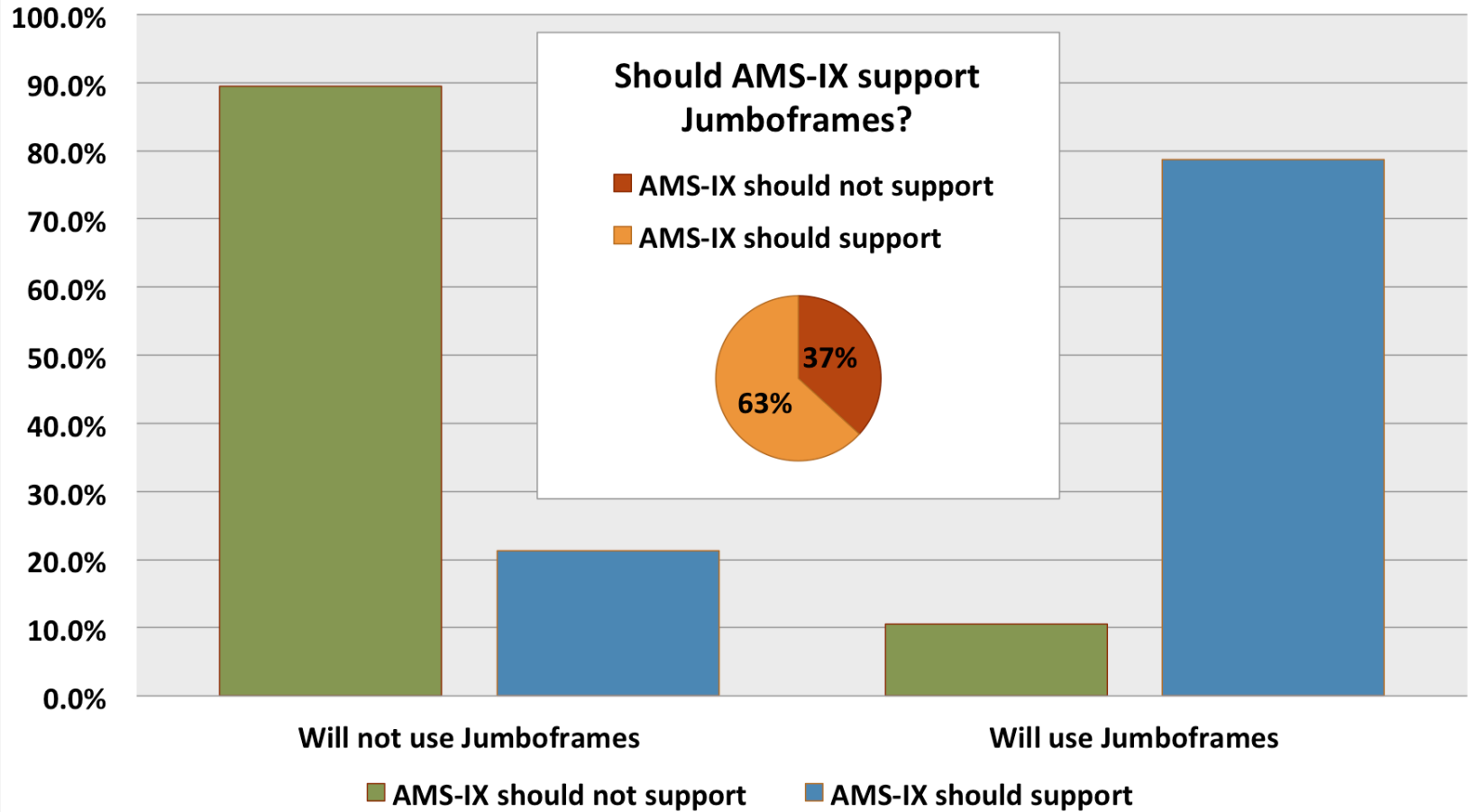03 November 2011

Content

1. AMS-IX survey and possible implementations

2. Advantages of Jumbo Frames

3. Disadvantages of Jumbo Frames

4. Summary and conclusion

1. AMS-IX survey and possible implementations

Maksym Tulyuk
maksym.tulyuk@ams-ix.net

RIPE 63. Vienna
03 November 2011

AMS-IX survey

Will you make use of Jumboframes if AMS-IX supports this?

115 participants

Possible implementations

Two ways to implement Jumbo Frames:

1. Change MTU for the existing VLAN

2. Make a new VLAN with Jumbo Frame support

Support in existing VLAN

YES

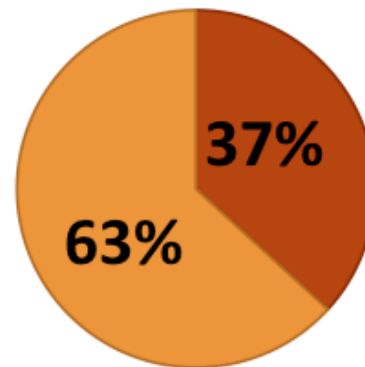1. Technical possibility: our equipment supports it

NO

1. Customers don't like changes in existing VLAN

2. No official standards

3. No MTU negotiation protocol

4. Path MTU discovery (PMTUD) protocol doesn't work

AMS-IX survey

Should AMS-IX support Jumboframes?

- ■ AMS-IX should not support
- ■ AMS-IX should support

37%

63%

115 participants

~~Support in existing VLAN~~

YES

1. Technical possibility: our equipment supports it

NO

1. **Customers don't like changes in existing VLAN**
2. No official standards
3. **No MTU negotiation protocol**
4. Path MTU discovery (PMTUD) protocol doesn't work

Make a new VLAN

YES

1. Technical possibility: our equipment supports it

NO

1. New port for each customer: will anyone pay for it?

2. No official standards: what size of Jumbo Frames?

3. Path MTU discovery (PMTUD) protocol doesn't work
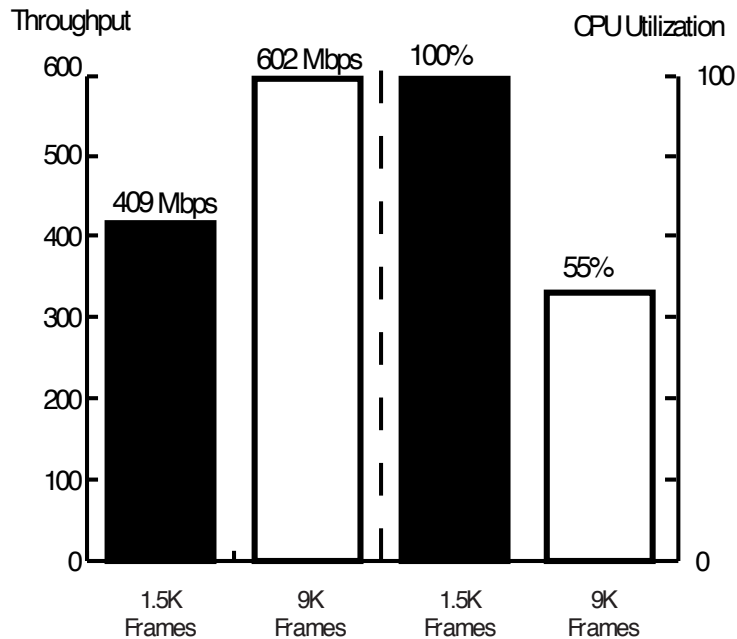
Make a new VLAN

Research is needed

NO

1. New port for each customer: **will anyone pay for it**?

2. No official standards: what size of Jumbo Frames?

3. Path MTU discovery (PMTUD) protocol doesn't work

2. Advantages of Jumbo Frames

Maksym Tulyuk
maksym.tulyuk@ams-ix.net

RIPE 63. Vienna
03 November 2011

2.1. Less CPU load

Figure 1. Extended Ethernet Frames vs. Standard Ethernet Frames*

Throughput — CPU Utilization

- 409 Mbps — 1.5K Frames
- 602 Mbps — 9K Frames
- 100% — 1.5K Frames
- 55% — 9K Frames

* Using Gigabit Ethernet. Throughput on tests was limited to SBus capacity. TCP tests used dual 300 Mhz Sun servers running Solaris 2.5.1
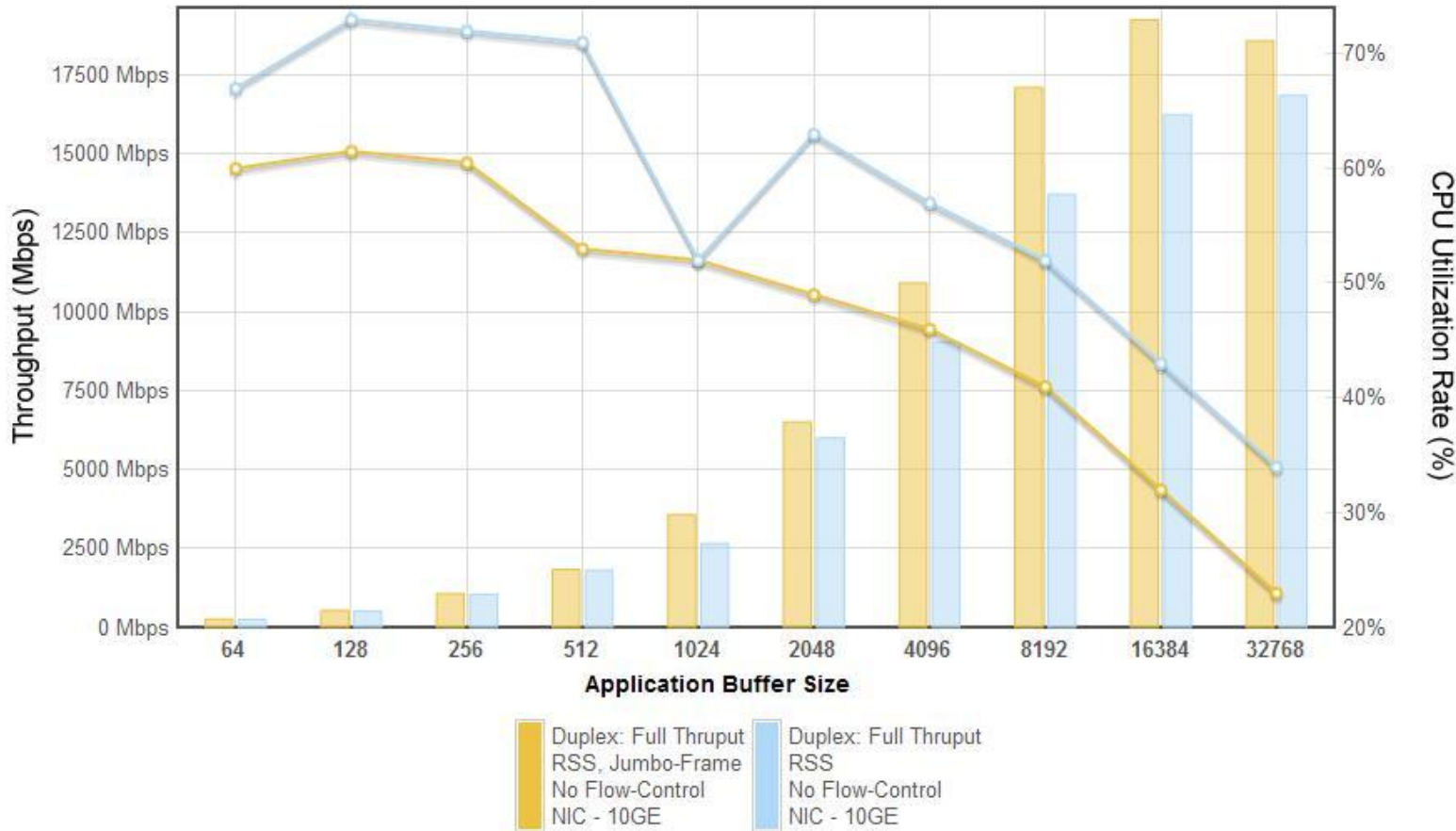
Alteon Networks. Extended Frame sizes for Next Generations Ethernets

2.1. Less overhead

Dell Inc. Internal Report

2.2. Network overhead

Theoretical maximum throughout on 1Gbit

9000 bytes

TCP: 990.042 Mbps

UDP: 992.697 Mbps

1500 bytes

TCP: 941.482 Mbps

UDP: 957.087 Mbps

TCP: 48.56 Mbps, UDP: 35.61 Mbps

TCP: 4.8%, UDP: 3.5%

http://sd.wareonearth.com/~phil/net/overhead/

2.3. TCP performance

Double MTU size - Double TCP throughout

$$Throughput = \frac{(0.7 * MSS)}{RTT * \sqrt{(Packet\_Loss\_Rate)}}$$

MSS = MTU – 40 RTT = 40ms Loss = 0.01%

Frame size = 9000 bytes   Frame size = 1500 bytes

Throughput = 40Mbit/s   **Throughput = 6.5Mbit/s**

http://www2.rad.com/networks/2003/largemtu/tcperf.htm

2.1. Advantages

1. Less CPU load

2. Less network overhead

3. Better TCP performance

3. Disadvantages of Jumbo Frames

Maksym Tulyuk
maksym.tulyuk@ams-ix.net

RIPE 63. Vienna
03 November 2011

3.1. No standard: header

1. IEEE 802.3 specification – 1518 bytes

2. 802.1Q (VLANs) – 1522 bytes

3. 802.1ad (Provider Bridge) – 1526 bytes

4. 802.1AS – 2000 bytes

5. 802.3AE – 1582 bytes

6. MPLS – 1518 bytes + N * 4 bytes

3.1. No standard: header

1. IEEE 802.3 specification – 1518 bytes

2. 802.1Q (VLANs) – 1522 bytes

3. 802.1ad (Provider Bridge) – 1526 bytes

4. 802.1AS – 2000 bytes

5. 802.3AE – 1582 bytes

6. MPLS – 1518 bytes + N * 4 bytes

But they say about Ethernet **header** only i.e. payload still **1500** bytes

3.1. No standard: payload

1. Ethernet (IEEE 802.3 specification) – 1518 bytes

2. FCoE (T11 specification) – 2166 bytes

3. iSCSI (VMWare de-facto ?) – 9000 bytes

3.1. No standard: terminology

1. Baby Giant – MPLS, 802.1Q, 802.1ad, 802.3AE

2. Mini Jumbo – FCoE

3. Giant Jumbo – payload more that 1500 bytes

4. Payload MTU – size of payload

5. Link MTU – size headers plus payload

3.2. Increase delay and jitter

Double MTU size - Double delay

Transmission Time per Frame in Microseconds		
Link Speed, Gigabits per second (Gbps)	1500 byte MTU frame	9000 byte MTU frame
1 Gbps Ethernet	12.00	72.00
10 Gbps Ethernet	1.20	7.20
40 Gbps Ethernet	0.30	1.80
100 Gbps Ethernet	0.12	0.72

http://www.ethernetalliance.org/files/static_page_files/DCB whitepaper.pdf

3.3. Increase buffers

Double MTU size - Double buffers

Port_Buffers = 2 * MTU + link_delay * link_speed

Switch_Buffers = Num_Ports * Num_Queues * Port_Buffer

link_delay = 0, Num_Ports = 24

MTU = 1518   MTU = 9000

Num_Queues = 8   Num_Queues = 8

Buffers = 582 912   Buffers = 3 456 000

http://www.ethernetalliance.org/files/static_page_files/DCB whitepaper.pdf

3.4. PMTUD doesn't work

1. RFC 1191 – standard

2. Easy to break, difficult to debug (RFC 2923)

3. TCP timeouts if it doesn't work

4. Solution (de-facto): set up the lowest MTU on servers/ CPE to exclude any issues

3.5. Low traffic with max. size

Frame Size Distribution - weekly

	Current	Average	Maximum	Minimum
0-63 bytes	0.0%	0.0%	0.0%	0.0%
64-127 bytes	41.2%	41.1%	45.7%	38.7%
128-255 bytes	3.5%	3.4%	4.9%	2.8%
256-511 bytes	2.1%	1.9%	2.2%	1.6%
512-1023 bytes	2.7%	2.5%	2.8%	2.1%
1024-1513 bytes	28.8%	27.8%	29.4%	24.8%
1514 bytes	21.8%	23.3%	26.1%	21.5%
> 1514 bytes	0.0%	0.0%	0.0%	0.0%
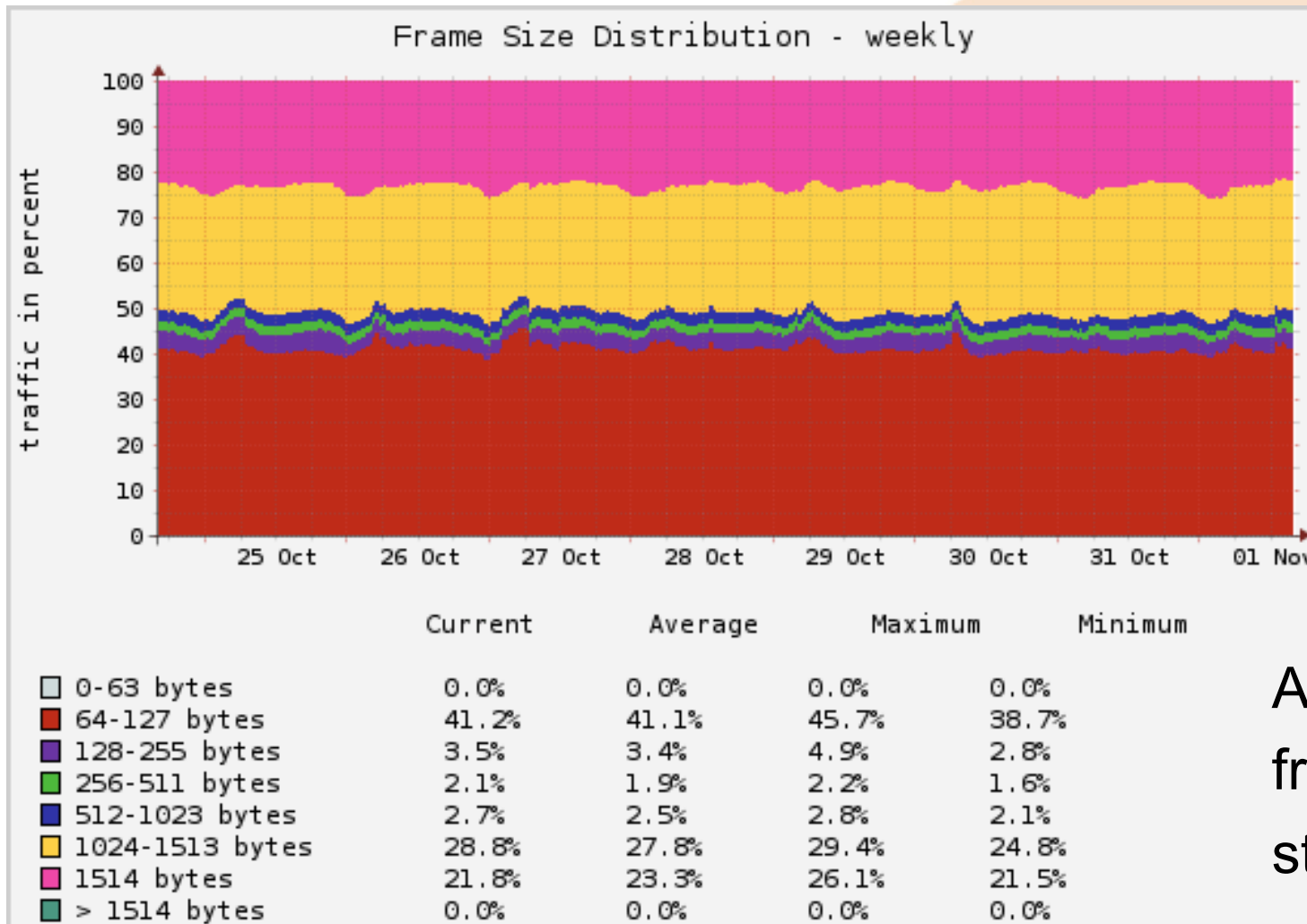
AMS-IX frame size statistics

https://www.ams-ix.net/sflow-stats/size/

26

3. Disadvantages

1. No standard/agreement for size of Jumbo Frames

2. Increased transmission time, packet delay, jitter, etc

3. Require bigger buffers on equipment

4. Path MTU Discovery (PMTUD) doesn't work

5. Low traffic with the current maximum size

4. Summary and conclusion

Maksym Tulyuk
maksym.tulyuk@ams-ix.net

RIPE 63. Vienna
03 November 2011

4. Pros and Cons: summary

Pros

1. Less CPU load

2. Less network packet overhead

3. Better TCP performance

Cons

1. No standard/agreement for size of Jumbo Frames

2. Increase transmission time, packet delay, jitter, etc.

3. Require bigger buffers on equipment

4. PMTUD doesn't work

5. Low traffic with the current maximum size

4. Pros and Cons: applications

Pros

1. Data transfer (Backups/Clusters/NFS/NNTP)

2. VPNs with payload 1500 bytes
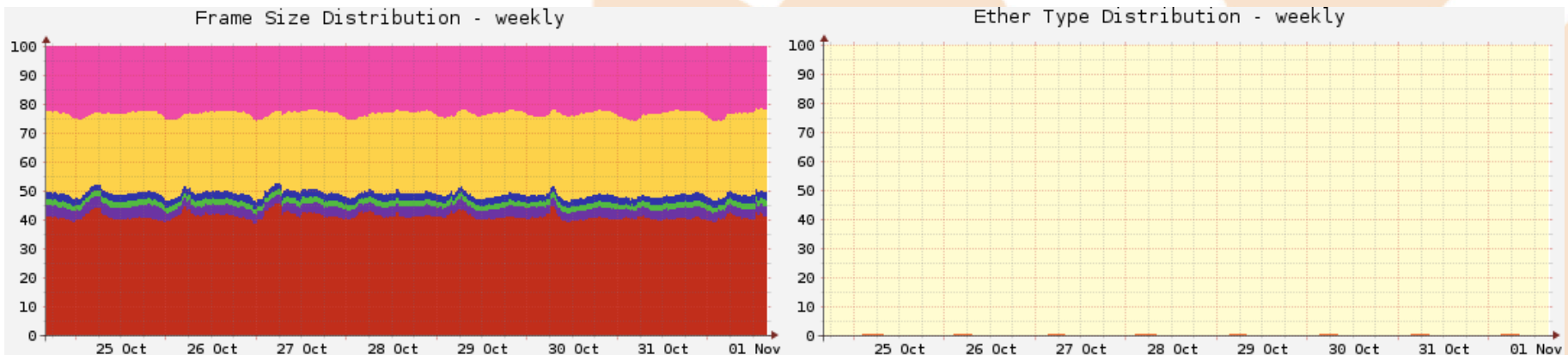
3. SAN (FCoE/iSCSI)

Cons

1. Inter-process communication (IPC)

2. Protocols using small packets (DNS, VoIP, etc)

3. Interoperability (no standards, broken PMTUD)

4. Conclusion

Personal

1. Nature of Internet traffic: small packets

2. All talks about Jumbo Frames are similar to IPv6 talks: started in 90x but IPv4 addresses are over and Ethernet with 1500 bytes still works fine

4. Conclusion

Official

1. Postpone for now

2. Ask our customers

3. Discuss with community

4. Make another survey ☺

Take part in survey!

https://www.surveymonkey.com/s/TGS5T2D

Comments?

Maksym Tulyuk
maksym.tulyuk@ams-ix.net

RIPE 63. Vienna
03 November 2011