DNS–clouds
Anycast: .RU/РФ experience

Pavel Khramtsov, ccTLD
p.khramtsov@msk-ix.ru

APTLD–2017
Tbilisi
What do you see in my name?

Upon the memo sheet, in grief,
Its imprint in the stillborn gloom,
Much like the writing on the tomb,
In foreign language it will leave.

Alexander Pushkin
(Translated by M. Kneller)
The initial goal: «The primary motivation for the development and deployment of these practices is to increase the distribution of Domain Name System (DNS) servers to previously under-served areas of the network topology and to reduce the latency for DNS query responses in those areas.» RFC-3258.

Nowadays:
- Initial goal
- DDoS mitigation
- Load balancing etc

Important remark: Anycast – it is about Routing but not about DNS
To distribute a service using anycast, the service is first associated with a stable set of IP addresses, and reachability to those addresses is advertised in a routing system from multiple, independent service nodes.¹

There are different ways to announce routes:

- The prefix can be advertised within the neighbor’s AS, and not to all other external AS’s;
- The prefix can be advertised to all external AS’s via neighbor’s AS.

¹ RFC-4786
In case of extension of our DNS-network we apply the following general principles:

- DNS-nodes should be installed in the most geographically remote from each other locations (focused on the most active users).
- The maximum concentration of local Russian providers in a point of DNS-node deployment (it is applied to nodes in the territory of the Russian Federation).
- Existence of unoccupied capacities and possibility of interaction with node via the control channel (it is applied to local nodes).
- Placement of nodes in networks of providers which can provide the protection against DDOS and flood.
Map of the Globe and DNS-nodes locations traps
Concentration of the DNS-queries

Павел Храмцов, MSK-IX, p.khramtsov@msk-ix.ru

Источник: https://tcinet.ru/dnsstat/all/reports/
Network topology

Источник: https://tcinet.ru/dnsstat/all/reports/

Павел Храмцов, MSK-IX, p.khramtsov@msk-ix.ru
Service Level Agreement Matrix

<table>
<thead>
<tr>
<th>Parameter</th>
<th>SLR (monthly basis)</th>
</tr>
</thead>
<tbody>
<tr>
<td>DNS service availability</td>
<td>0 min downtime = 100% availability</td>
</tr>
<tr>
<td>DNS name server availability</td>
<td>≤ 432 min of downtime (≈ 99%)</td>
</tr>
<tr>
<td>TCP DNS resolution RTT</td>
<td>≤ 1500 ms, for at least 95% of the queries</td>
</tr>
<tr>
<td>UDP DNS resolution RTT</td>
<td>≤ 500 ms, for at least 95% of the queries</td>
</tr>
<tr>
<td>DNS update time</td>
<td>≤ 60 min, for at least 95% of the probes</td>
</tr>
<tr>
<td>RDDS availability</td>
<td>≤ 864 min of downtime (≈ 98%)</td>
</tr>
</tbody>
</table>

It is basic deviation from the principle of the minimum RTT

1 ICANN SLA for new gTLD
Model of threats: Mirai example:

ISP responsibility zone

ISP

DNS Provider

UDP Flood
ACK Flood
GRE Flood

DNS water torture
Amplification (a bite in Mirai case)

Internet

incoming traffic

invalid

valid

// define ATK_VEC_PROXY 0 /* Back connection */
#define ATK_VEC_UDP_PLAIN 9 /* Plain UDP flood optimized for speed */
#define ATK_VEC_HTTP 10 /* HTTP layer 7 flood */
Real time mitigation:

In case of DSN water torture (DNS valid traffic):
- Distribute traffic across DNS clouds & nodes;
- Switch on ACL and RRL on the servers.

In case of flood (Invalid traffic):
- Switch off announces of local nodes;
- Automatically switch on ISP protection service.

The cloud is pulled together in one protected node.
The node has to be highly productive.
If to look more widely:

- The increase of number of nodes increases probability of an error of a configuration;
- Complicates synchronization of files of zones;
- Demands the protected file transfer of zones via a public network of data transmission;
- Demands protection against flood;
- Demands protection against «DNS water torture»;
- Determination of procedures of interaction with ISP;
- SLA definition of «The special period»;
- Definitions of "incident".
Anycast DNS-node architecture

• Connectivity: two channels (management and service, synchronization of zones via a control channel);
• "Hardware" vs "Virtual";
• Structure (router, server, statistics server, monitoring);
• Choice of hardware and software;
• DNS-node tuning.

The personnel skills are cornerstone of your choice of hardware and software.

The choice of architecture depends on the node productivity which is defined by a location selection.
Structure of the current DNS-network

Five anycast DNS-clouds:

Cloud «Russia» - 17 servers (9 nodes),
Cloud «Moscow» - 4 servers (2 nodes),
Cloud «Europe» - 6 servers (3 nodes).
Cloud «Asia» - 4 servers (3 nodes).
Cloud «America» - 4 servers (3 nodes).

Totally - 35 authoritative DNS-servers in 20 nodes.
We use channels of the following ISPs:

- MSK-IX (regional IXs),
- RIPN (local nodes),
- RETN (Global transit),
- CW (Global ISP),
- LEVEL3 (Global ISP),
- TATA (Global ISP),
- PCCWG (Global ISP),
- IPTP (Global ISP),
- DENIC (IX),
- NICBR (IX),
- KAZNIC (Local ISP).
The current indicators of loading on nodes

Moscow (all nodes) – 1,5 Gqpd  
Санкт-Петербург – 0,9 Gqpd  
Нью-Йорк – 0,8 Gqpd  
Амстердам – 0,7 Gqpd  
Гонконг – 0,6 Gqpd  
Лос-Анджелес – 0,5 Gqpd  
Франкфурт – 0,5 Gqpd  
Сингапур – 0,2 Gqpd  
Сан-Паулу – 0,07 Gqpd  
Ростов-на-Дону – Gqpd  
Самара – 0,07 млрд Gqpd  
Новосибирск – 0,03 Gqpd  
Прага – 0,013 млрд Gqpd  
Астана – 0,014 млрд Gqpd  
Екатеринбург – 0,03 Глрд Bqpd  
Владивосток – 0,003 Глрд Bqpd  
Ставрополь – 0,002 Gqpd  
Казань – 0,001 млрд Gqpd

In case of attack the volume of a traffic increases in tens times. But the system allows to maintain SLA conditions.
Resume

• Anycast resolves some problems

• Anycast produces some problems

• Anycast does not an ultimate decision for DNS-service

• Anycast needs additional efforts, skills and experience for organization of the reliable DNS-service

• We have got it and we are ready to cooperation
Questions?