DNS-облака .RU/.РФ.

Anycast – как это устроено у нас

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

TLDCON-2017

Что в имени тебе моем?...

«Оно на памятном листке
Оставит мертвый след, подобный
Узору надписи надгробной
На непонятном языке.»

А.С.Пушкин

```
; <<>> DiG 9.8.3-P1 <<>> ru ns
;; global options: +cmd
;; Got answer:
;; ->>HEADER<<- opcode: QUERY, status: NOERROR, id:
28126
;; flags: qr rd ra; QUERY: 1, ANSWER: 5, AUTHORITY: 0,
ADDITIONAL: 0

;; QUESTION SECTION:
;ru.                            IN      NS

;; ANSWER SECTION:
ru.                     345600  IN      NS      d.dns.ripn.net.
ru.                     345600  IN      NS      e.dns.ripn.net.
ru.                     345600  IN      NS      f.dns.ripn.net.
ru.                     345600  IN      NS      a.dns.ripn.net.
ru.                     345600  IN      NS      b.dns.ripn.net.

;; Query time: 69 msec
;; SERVER: 192.168.1.1#53(192.168.1.1)
;; WHEN: Mon Sep  4 13:38:57 2017
;; MSG SIZE  rcvd: 112
```

Региональные облака

Anycast & DNS

Исконная цель Anycast для DNS – сокращение времени обслуживания резолверов клиентов сервиса DNS при ограничении размера UDP пакета.

Что мы имеем сейчас:

- Первоначальную цель (сокращение RTT)
- Решение вопроса устойчивости системы (DDoS)
- Балансировку нагрузки и прочее (подробно останавливаться на этом аспекте не будем)

Важное примечание: Anycast – это о маршрутизации, а не о DNS

Anycast

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].

Маршрут можно анонсировать:

- Для транзита; — **Глобальный DNS-узел**
- Только для пиринга. — **Локальный DNS-узел**

[1] RFC-4786

Локальный DNS-узел

Основное назначение локального узла:

- Обслуживание клиентов местного IX;
- Обслуживание клиентов местных локальных провайдеров.

Что произойдет если случиться «утечка» маршрута?

Скорее всего ничего страшного не случиться, т.к. Метрика такого маршрута будет хуже метрики маршрута для глобального узла. Во всяком случае об этом стоит позаботиться[1].

[1] Показателен случай Google&Verizon

Глобальный DNS-узел

Основное назначение глобального узла:

- Обслуживание клиентов региона;
- Обслуживание клиентов глобальных провайдеров.

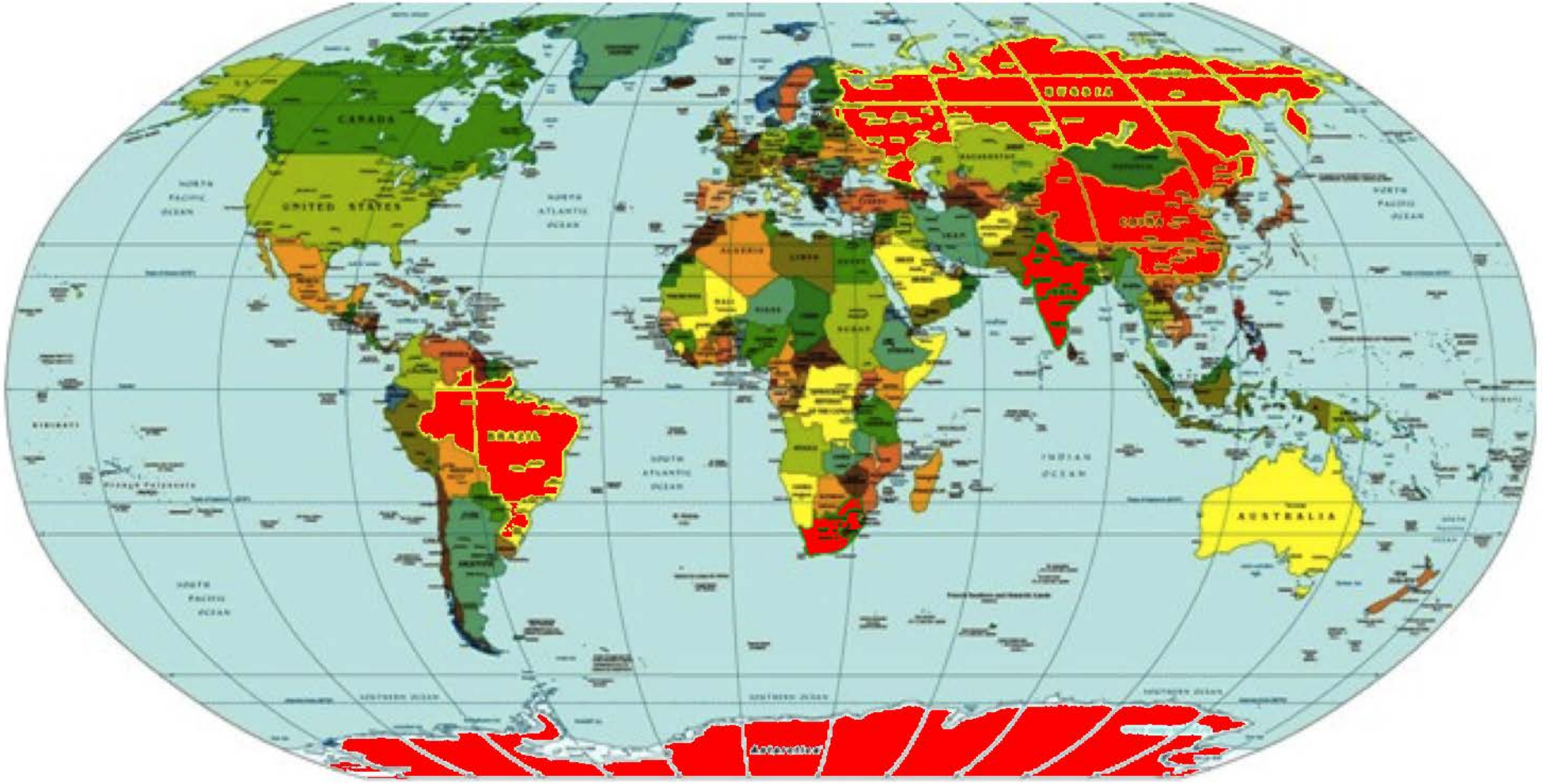
Что может подстерегать здесь?

Возможно размещение узла, на который не придет трафик в силу особенностей маршрутизации между провайдерами региона. Азия – как раз такой регион.

Принципы размещения узлов

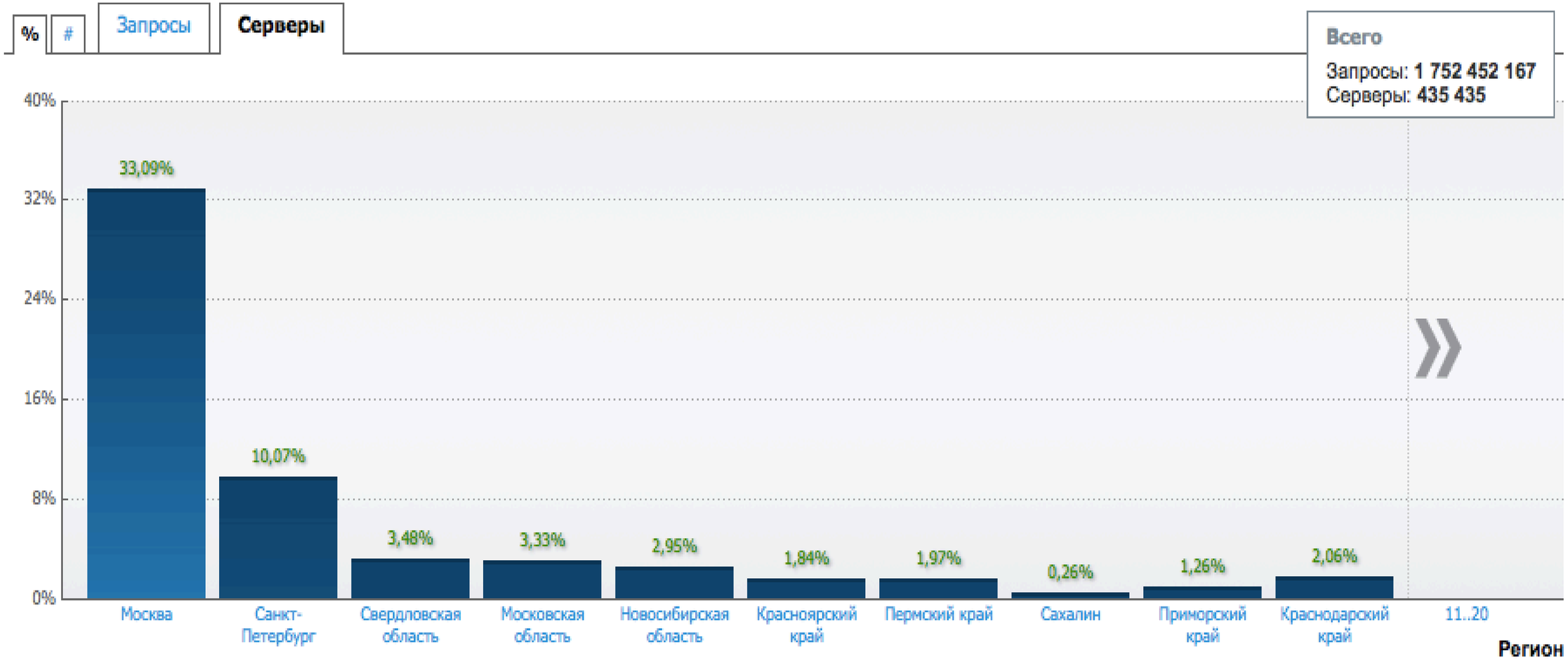
- Политическая география;
- Плотность резолверов-клиентов/запросов с резолверов клиентов;
- Топология сети;
- Соблюдение SLA&

Политическая география



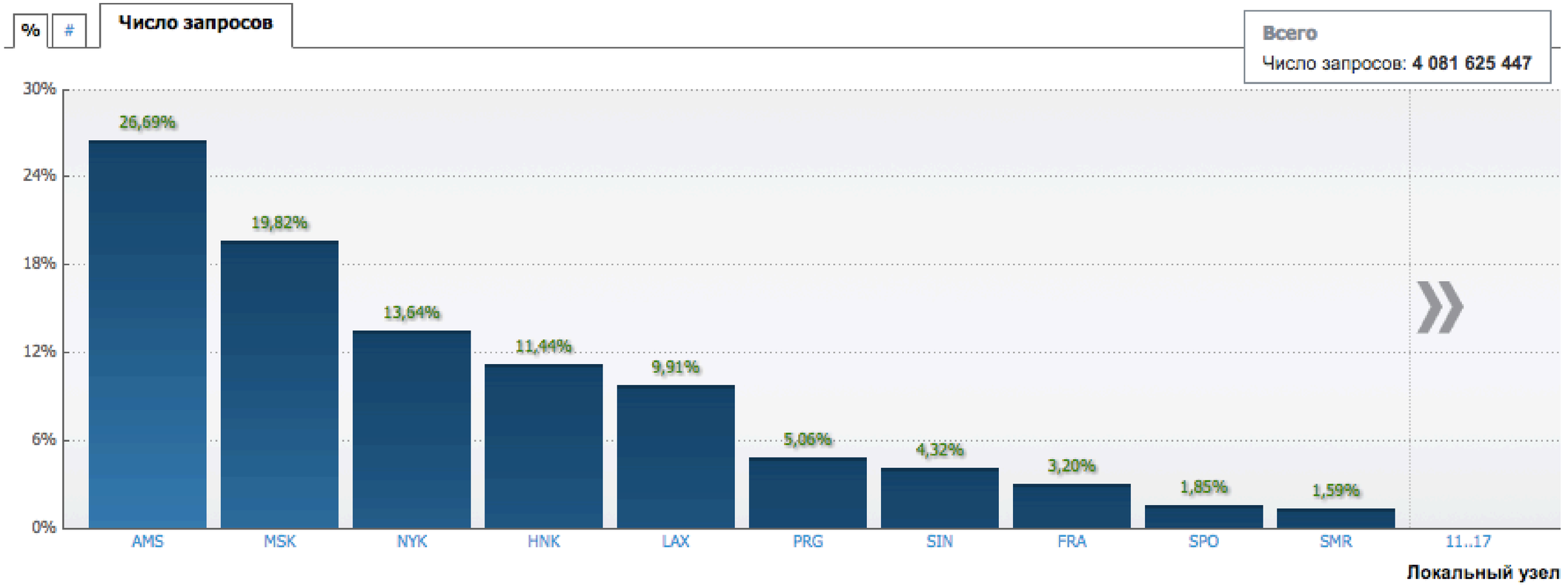
Павел Храмцов, MSK-IX, p.khramtsov@msk-ix.ru

Плотность резолверов-клиентов/запросов



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

Топология сети



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

Service Level Agreement (SLA)

Показатель[1]	**Значение показателя (за месяц)**
Доступность сервиса DNS	100%
Доступность сервера DNS	99%
Время отклика TCP DNS	1500 мс, для как минимум 95% запросов
Время отклика UDP DNS	500 мс, для как минимум 95% запросов
Время обновления зоны на DNS-узлах	60 минут, для как минимум 95% запросов

Принципиальное отступление от принципа минимального RTT

[1] ICANN SLA для new gTLD

Архитектура DNS-узла

- Связность: два канала (управление и сервис, синхронизация зон по каналу управления);
- «Железный» vs «Виртуальный»;
- Состав (маршрутизатор, сервер, сервер статистики, мониторинг);
- Выбор «железа» и ПО;
- Тюнинг DNS-узла.

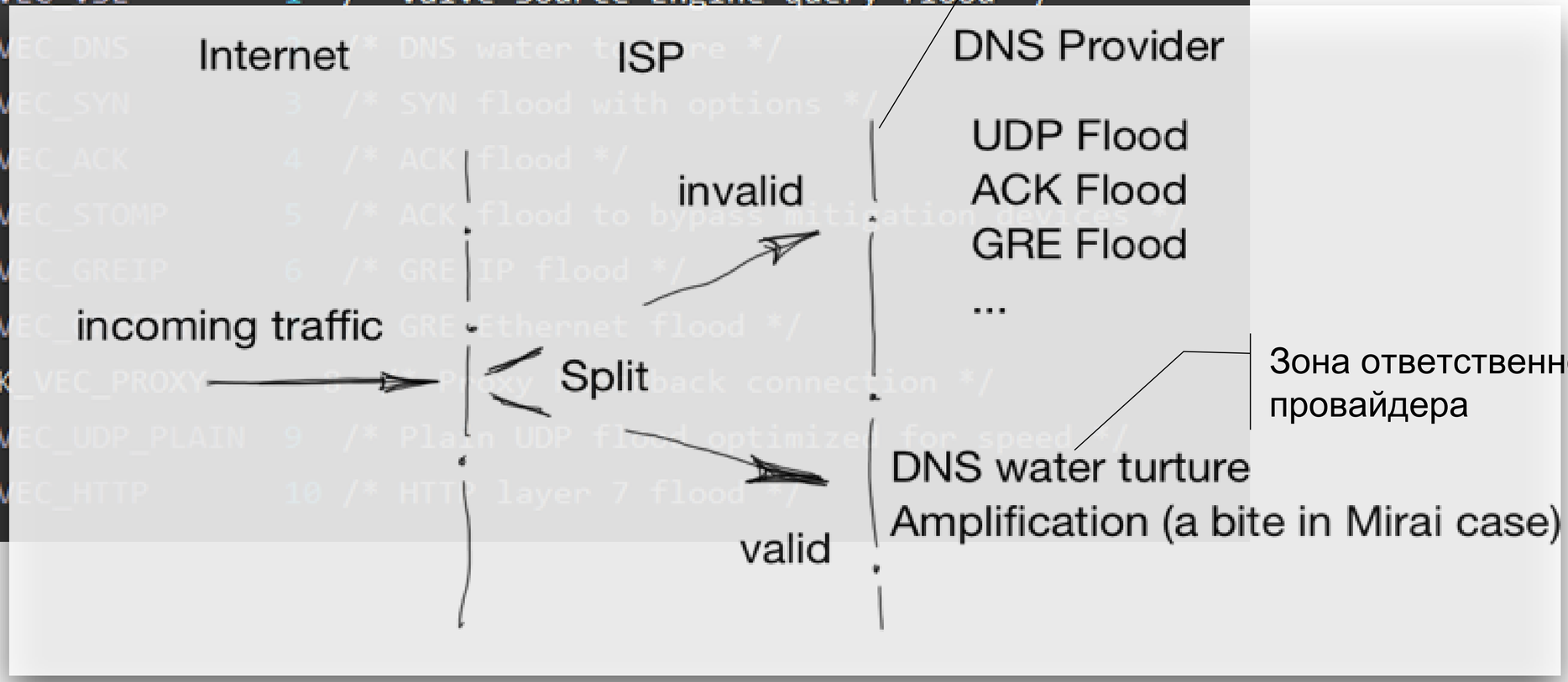
Принципиальное значение имеет компетенция персонала

Выбор архитектуры зависит от расчетной производительности, которая определяется местом размещения

Стабильность/надежность/безопасность

Модель угроз на примере кода Mirai[1]:

```
#define ATK_VEC_UDP         0  /* Straight up UDP flood */
#define ATK_VEC_VSE         1  /* Valve Source Engine query flood */
#define ATK_VEC_DNS         2  /* DNS water torture */
#define ATK_VEC_SYN         3  /* SYN flood with options */
#define ATK_VEC_ACK         4  /* ACK flood */
#define ATK_VEC_STOMP       5  /* ACK flood to bypass mitigation devices */
#define ATK_VEC_GREIP       6  /* GRE IP flood */
#define ATK_VEC_GRETH       7  /* GRE Ethernet flood */
//#define ATK_VEC_PROXY     8  /* Proxy knockback connection */
#define ATK_VEC_UDP_PLAIN   9  /* Plain UDP flood optimized for speed */
#define ATK_VEC_HTTP        10 /* HTTP layer 7 flood */
```



ISP должен отфильтровать флуд

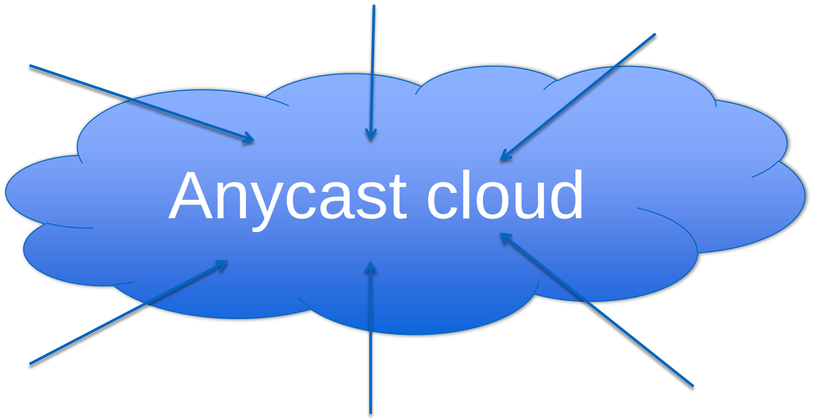
Зона ответственности DNS-провайдера

[1] https://blog.radware.com/security/2016/11/insight-into-mirais-source-code/

Противодействие атаке в реальном времени:

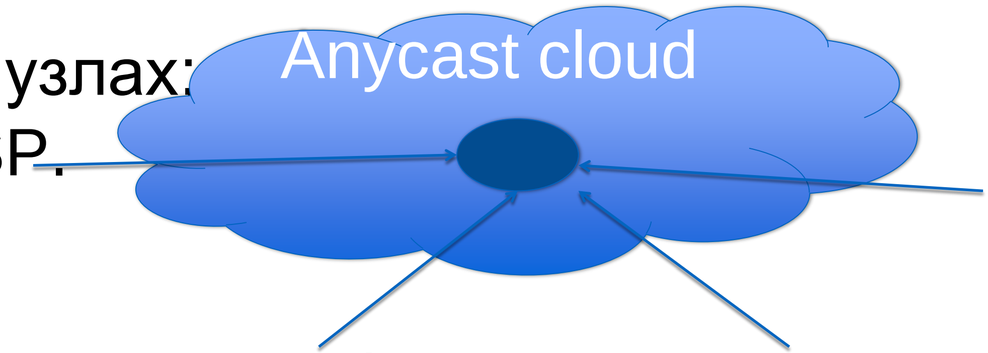
В случае валидного DNS-трафика:

- Распределение трафика по узлам ;
- Включение ACL и RTT серверах.



В случае флуда:

- Выключение анонсов на незащищенных узлах:
  - Автоматическое включение защиты ISP.



Облако сжимается, что требует высокой производительности защищенных узлов.

Стабильность/надежность/безопасность

Если взглянуть шире:

- Увеличение количества узлов повышает вероятность ошибки конфигурации;
- Усложняет синхронизацию файлов зон;
- Требует защищенной передачи файлов зон по публичной сети передачи данных;
- Защита от флуда;
- Защита от нагрузок валидным трафиком;
- Налаживание процедур взаимодействия с ISP;
- Определения SLA «Особого периода»;
- Определения ситуации «инцидента».

От теории к практике

Национальные доменные зоны RU/РФ поддерживает 5 DNS-облаков:

a.dns.ripn.net. – регион Москва;
b.dns.ripn.net. – регион Россия;
d.dns.ripn.net. – регион Азия;
e.dns.ripn.net. – регион Европа;
f.dns.ripn.net. – регион Северная, Центральная и Южная Америки

Каждому из облаков назначены свои адреса IPv4 и IPv6.

От теории к практике

При размещении узлов системы DNS доменов RU/РФ применяются следующие общие принципы:

- Вынесение узлов в географически разнесенные локации (максимально приближенные к наиболее активным пользователям, применяется для глобальных узлов).
- Максимальная концентрация локальных российских провайдеров в точке размещения узла (применяется для узлов на территории РФ).
- Наличие свободных емкостей и удобство взаимодействие с площадкой (применяется для локальных узлов).
- Размещение узлов в сетях провайдеров, которые могут обеспечить защиту от DDOS атак «очистку» трафика на своих каналах.

От теории к практике

Система поддержки национальных доменов RU/РФ включает в себя:

Облако «Россия» - 17 серверов (9 узлов),
Облако «Москва» - 4 сервера (2 узла) ,
Облако «Европа» - 6 серверов (3 узла).
Облако «Азия» - 4 сервера (3 узла).
Облако «Америки» - 4 сервера (3 узла).

Всего по всем облакам - 35 авторитетных серверов DNS.

От теории к практике

В настоящее время узлы системы DNS подключены через каналы следующих провайдеров:

- MSK-IX (региональные IX-ы),
- RIPN (локальные узлы),
- RETN (Глобальный транзит),
- CW (Глобальный провайдеров),
- LEVEL3 (Глобальный провайдер),
- TATA (Глобальный провайдер),
- PCCWG (Глобальный провайдер),
- IPTP (Глобальный провайдер),
- DENIC (Немецкий IX),
- NICBR (Бразильский IX),
- KAZNIC (Региональный провайдер).

От теории к практике

По узлам в среднем при стационарных условиях наблюдаются следующие показатели :

Москва (все узлы) - 1,5 млрд запросов/сутки
Санкт-Петербург - 0,9 млрд запросов/сутки
Нью-Йорк - 0,8 млрд запросов/сутки
Амстердам - 0,7 млрд запросов/сутки
Гонконг - 0,6 млрд запросов/сутки
Лос-Анжелес - 0,5 млрд запросов/сутки
Франкфурт - 0,5 млрд запросов/сутки
Сингапур - 0,2 млрд запросов/сутки
Сан-Паулу - 0,07 млрд запросов/сутки
Ростов-на-Дону - 0,02 млрд запросов/сутки
Самара - 0,07 млрд запросов/сутки
Новосибирск - 0,03 млрд запросов/сутки
Прага - 0,013 млрд запросов/сутки
Астана - 0,014 млрд запросов/сутки
Екатеринбург - 0,03 млрд запросов/сутки
Владивосток - 0,003 млрд запросов/сутки
Ставрополь - 0,002 млрд запросов/сутки
Казань - 0,001 млрд запросов/сутки

При атаках трафик увеличивается кратно, но система позволяет выдерживать SLA

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

Резюме

- DNS- облака RU/РФ в настоящее время организованы по принципу приближения к резолверам клиентов с учетом региональных особенностей и топологии сети;
- Управление облаками автоматизировано;
- Программное обеспечение узлов настроено на достижение максимальной производительности с учетом особенностей аппаратной части и ПО;
- Параллельно с системой RU/РФ, с учетом опыта ее эксплуатации развернута еще одна система для поддержки sTLD и new gTLD.

Вопросы?