Ай-ТИ Фреш

ТЕХНИЧЕСКИЙ РАЗБОР

Anycast DNS: отказоустойчивый резолв без единой точки отказа

Как мы строим DNS на PowerDNS 5.1/5.4 + FRR 10.6: BGP-anycast, BFD, health-check и зоны-как-код

Июль 2026

itfresh.ru · ИТ-аутсорсинг для юридических лиц

О ЧЁМ ЭТОТ РАЗБОР

Суть проблемы

Внутренний DNS в типовой сети — один-два Windows-DC в одном VLAN: при аварии свитча, петле или падении VPN до филиала резолв AD-имён исчезает, и встаёт всё сразу — 1С, почта, кассы, Wi-Fi с 802.1X. Мы убираем эту единую точку отказа: 2–3 узла PowerDNS анонсируют один anycast-IP по BGP, и отказ любого узла клиенты не замечают — трафик за секунды уходит на живой сервер.

Почему это важно бизнесу

- Час без DNS = час простоя всей компании: без резолва не работают 1С, почта, порталы и кассы, даже если сами серверы живы
- Филиалы перестают зависеть от VPN до головного офиса: локальный узел отвечает даже при обрыве канала
- Один anycast-IP на всех клиентах: замена или перенос DNS-сервера не требует перенастройки DHCP и станций
- Изменения зон идут через ревью в Git: ошибка в записи не улетает в прод мимо контроля и откатывается одним коммитом

ЧТО МЫ ЗАКЛАДЫВАЕМ

Ключевые параметры реализации

5.1

PowerDNS Authoritative: ветка с HTTP API v1 и LMDB-бэкендом — держит внутренние зоны на каждом…

по докам PowerDNS Auth 5.1

5.4

PowerDNS Recursor: кэширующий резолвер на anycast-IP, forward_zones на внутренние зоны

по докам PowerDNS Recursor 5.4

10.6

FRR на каждом DNS-узле: bgpd анонсирует /32 под route-map, bfdd контролирует пиринг с L3-ядром

FRR 10.6, docs.frrouting.org

300 мс

интервалы BFD rx/tx при detect-multiplier 3 — обрыв узла сеть фиксирует быстрее чем за секунду

по докам FRR bfdd

5 с

период health-check: dig SOA на anycast-IP узла (+time=2 +tries=1); при отказе — withdraw анон…

наш стандарт

x3

узла anycast в типовом проекте: по одному на площадку — переживаем отказ любого и плановый выв…

наш стандарт

РЕАЛИЗАЦИЯ 1 ИЗ 3

Маршрутизация: FRR и anycast-анонс на каждом узле

Что настраиваем

Три Linux-узла (Debian 12 / Astra Linux 1.7) + L3-ядро площадок (MikroTik CCR, Cisco, pfSense)

Как мы это делаем

1. Ставим FRR 10.6, в /etc/frr/daemons включаем bgpd и bfdd; anycast-IP 10.99.0.10/32 вешаем на lo и объявляем через network под route-map ANYCAST-DNS
2. iBGP с приватной AS 65100 до ядра площадки: neighbor <core> remote-as 65100 + bfd-профиль (rx/tx 300 мс, detect-multiplier 3)
3. Ближайший узел выигрывает по метрике IGP, остальные остаются запасными маршрутами — ECMP для DNS-UDP не обязателен
4. Плановый вывод узла — route-map с set as-path prepend 65100 65100: трафик стекает плавно, без обрыва запросов
5. Контроль: vtysh -c 'show bgp ipv4 unicast 10.99.0.10/32' на ядре — маршрут должен быть виден со всех узлов

РЕЗУЛЬТАТ

Отказ узла (питание, диск, канал) сеть отрабатывает сама: BFD валит сессию менее чем за секунду, маршрут исчезает, клиенты продолжают ходить на тот же IP через соседний узел. Ни одна станция и ни один сервер перенастройки не требуют.

КЛЮЧЕВОЙ НЮАНС

iBGP внутри компании не требует регистрации AS в RIPE — берём приватную AS из диапазона 64512–65534. На клонированных VM явно задаём bgp router-id: дубль тихо ломает пиринг.

РЕАЛИЗАЦИЯ 2 ИЗ 3

DNS-плоскость: Authoritative + Recursor на одном узле

Что настраиваем

Каждый anycast-узел: PowerDNS Authoritative 5.1 (зоны) + PowerDNS Recursor 5.4 (кэш и внешний резолв)

Как мы это делаем

1. Authoritative 5.1 слушает 127.0.0.1:5300 (local-address=127.0.0.1:5300), бэкенд LMDB или gsqlite3; включаем webserver и api-key для OctoDNS
2. Recursor 5.4 слушает anycast-IP:53; в recursor.yml — forward_zones: corp.example.ru → 127.0.0.1:5300: внутренние зоны идут в Authoritative, остальное резолвится нар...
3. incoming.allow_from ограничиваем корпоративными подсетями (RFC1918) — узел никогда не станет открытым резолвером
4. AD-integrated зоны: PowerDNS как secondary (autosecondary) с AXFR/NOTIFY от контроллеров домена — миграция без big bang
5. Клиентам через DHCP option 6 раздаём два anycast-IP (10.99.0.10 / 10.99.0.11) как primary/secondary

РЕЗУЛЬТАТ

Один узел закрывает и авторитативные ответы по внутренним зонам, и кэширующий резолв наружу; AD продолжает управлять своими зонами, а клиенты получают отказоустойчивый резолв на неизменных IP-адресах.

КЛЮЧЕВОЙ НЮАНС

Auth и Recursor не делят порт 53: разносим по адресам/портам и отключаем stub systemd-resolved (DNSStubListener=no), иначе порт занят и сервис молча не поднимается.

РЕАЛИЗАЦИЯ 3 ИЗ 3

Health-check, зоны-как-код и мониторинг

Что настраиваем

Все узлы: dns-health.sh в systemd-timer, GitLab CI + OctoDNS, Zabbix 7.0 LTS

Как мы это делаем

1. dns-health.sh по systemd-timer каждые 5 с: dig @10.99.0.10 SOA corp.example.ru +time=2 +tries=1 — проверяет цепочку Recursor→Authoritative; при отказе vtysh снимает…
2. Возврат анонса — только после 3 подряд успешных проверок: гистерезис против флаппинга маршрута на полуживом узле
3. Зоны — YAML в Git; в MR джоб octodns-sync без --doit показывает diff (dry-run), после merge джоб с --doit раскатывает изменения через API PowerDNS на все узлы
4. Zabbix: qps и cache-hit из rec_control get-all и API Authoritative, статус BGP/BFD из vtysh (json), алерт при peer != Established
5. Раз в квартал — учения: гасим узел, замеряем время переключения клиентов (цель — до 5 с) и фиксируем в регламенте

РЕЗУЛЬТАТ

Сценарий «анонсирует, но не отвечает» исключён как класс: маршрут живёт ровно столько, сколько жив сервис. Любое изменение зоны проходит с ревью и аудитом, откат — один revert-коммит.

КЛЮЧЕВОЙ НЮАНС

Порог отказа держим агрессивным (таймаут 2 с, одна попытка), а восстановление — консервативным: быстрее убрать больной узел важнее, чем быстрее вернуть его в строй.

ЧТО МЫ ВЫУЧИЛИ НА ПРАКТИКЕ

Подводные камни

✕ BGP жив — DNS мёртв

BGP не видит состояние сервиса: узел с упавшим PowerDNS продолжает анонсировать IP. Лечим health-check'ом с withdraw через vtysh каждые 5 с.

✕ Порт 53 занят systemd-resolved

Stub-листенер 127.0.0.53 конфликтует с Recursor. Ставим DNSStubListener=no в resolved.conf и статический resolv.conf на узле.

✕ Дубль router-id на клонах VM

Узлы разворачивают клонированием — bgp router-id совпадает, iBGP-сессии флапают. Прописываем router-id явно из адреса управления.

✕ Открытый резолвер наружу

Recursor на anycast-IP без ограничений отвечает всем и участвует в амплификации. Жёстко ограничиваем incoming.allow_from внутренними подсетями.

✕ Флаппинг анонса

Полуживой узел то проходит, то валит проверку — маршрут дёргается. Гистерезис: withdraw после 1 фейла, возврат после 3 успешных подряд.

✕ AD-зоны не отдаются по AXFR

DC по умолчанию запрещает zone transfer — PowerDNS-secondary остаётся пустым. На DC разрешаем Zone Transfers на IP узлов и включаем Notify.

✕ Все узлы в одном VLAN

Anycast не спасёт от L2-аварии, если узлы за одним свитчем. Разносим по площадкам/стойкам с независимыми аплинками и питанием.

✕ Жёсткий вывод узла в обслуживание

Мгновенный shutdown рвёт запросы в полёте. Сначала as-path prepend или метрика — трафик стекает на соседей, затем гасим сервис.

УРОВНИ РЕАЛИЗАЦИИ

Как правильно

МИНИМУМ

- 2 узла PowerDNS на разных площадках, общий anycast-IP через iBGP (FRR)
- Health-check dig SOA каждые 5 с с автоматическим withdraw через vtysh
- Клиенты и DHCP option 6 переведены на anycast-адреса

НОРМАЛЬНО

- 3 узла, BFD 300 мс / x3 на пирингах, приватная AS, route-map на анонсы
- Зоны в Git + OctoDNS: dry-run в MR, apply с --doit из CI через API PowerDNS
- PowerDNS как secondary для AD-зон (AXFR/NOTIFY с контроллеров)
- Zabbix: qps, latency, cache-hit, статус BGP/BFD по каждому узлу

ХОРОШО

- Узлы на 3 независимых площадках/ЦОД, drain через as-path prepend
- dnsdist 2.1 перед Recursor: rate-limit, правила и телеметрия на порту 53
- DNSSEC-подпись внешних зон в Authoritative, RPZ-фильтрация на Recursor
- Квартальные failover-учения с замером времени переключения

ПРОВЕРЬТЕ У СЕБЯ

Чек-лист самопроверки

- [ ] Переживёт ли резолв внутренних имён отказ любого одного DNS-узла без действий инженера?
- [ ] Снимается ли BGP-анонс автоматически, когда PowerDNS перестал отвечать, а не только при падении сервера?
- [ ] Разнесены ли DNS-узлы по разным L2-сегментам и площадкам, а не по одному VLAN?
- [ ] Хранятся ли зоны в Git и проходит ли каждое изменение через dry-run octodns-sync и ревью коллеги?
- [ ] Есть ли secondary-копия AD-зон на PowerDNS и настроены ли AXFR/NOTIFY с контроллеров домена?
- [ ] Закрыта ли рекурсия списком allow_from только для внутренних подсетей?
- [ ] Мониторятся ли qps, cache hit ratio, задержка резолва и состояние BGP/BFD-сессий по каждому узлу?
- [ ] Есть ли регламент плановых работ: drain узла через prepend вместо жёсткого выключения?
- [ ] Проводились ли учения: выключение узла с замером фактического времени переключения клиентов?

Если хотя бы на два вопроса ответ «нет» или «не знаю» — тема требует внимания.

ЧТО МЫ ДЕЛАЕМ

Как поможет ITFresh

ITFresh — ИТ-аутсорсинг для юридических лиц до 50 рабочих мест в Москве и области. 15+ лет практики, собственная инфраструктура в дата-центре МТС (8 серверов Dell Xeon Platinum).

- Аудит текущей DNS-инфраструктуры: точки отказа, зависимость филиалов от VPN, тайминги резолва
- Разворачиваем 2–3 anycast-узла под ключ: PowerDNS 5.1/5.4 + FRR 10.6, пиринг с вашим L3-ядром
- Миграция AD-зон без простоя: secondary-режим и поэтапный перевод клиентов на anycast-IP
- Зоны-как-код: репозиторий, OctoDNS, CI-пайплайн dry-run/apply и регламент изменений
- Мониторинг в Zabbix и эксплуатация: health-check, drain узлов, failover-учения

15+
лет в ИТ-поддержке

50
рабочих мест — наш профиль

МТС
дата-центр, Москва

КОНТАКТЫ

Обсудить вашу задачу

Сайт	**itfresh.ru**
Телефон	**+7 903 729-62-41**
Telegram	**@ITfresh_Boss**

Бесплатно посмотрим вашу инфраструктуру по этому чек-листу и скажем, где тонко — без обязательств.

itfresh.ru

ДОКУМЕНТАЦИЯ И НАШ СТЕК

Техническая база

01 PowerDNS Authoritative 5.1: Settings, Backends (LMDB), HTTP API (doc.powerdns.com — 5.1)

02 PowerDNS Recursor 5.4: YAML Settings (forward_zones, incoming.allow_from),… (doc.powerdns.com — 5.4)

03 FRR User Guide: BGPd (network, route-map, timers), BFD (bfdd) (docs.frrouting.org — 10.6)

04 OctoDNS: Getting Started (dry-run/--doit), провайдер octodns-powerdns (octodns.readthedocs.io — 1.21)

05 RFC 4786 — Operation of Anycast Services (BCP 126) (rfc-editor.org — 2006)

06 Шаблоны ITfresh: dns-health.sh, unit/timer systemd, Zabbix-шаблон DNS Anyc… (itfresh.ru — 2026)

Основано на официальной документации продуктов и нашей практике внедрения.