

Проектирование сетевой фабрики на основе VxLAN EVPN для IPv6 потребителей с eBGP в Underlay на оборудовании Arista

Дизайн сетей ЦОД



**Меня хорошо видно
& слышно?**



Тема: Проектирование сетевой фабрики на основе VxLAN EVPN для IPv6 потребителей с eBGP в Underlay на оборудовании Arista



Ринат Гареев

Корпоративный архитектор
Архитектор прикладных решений
Технический руководитель в ИТ и Телеком



Содержание

Цель и задачи проекта

Какие технологии использовались

Что получилось

Выводы

Вопросы и рекомендации



Цель и задачи проекта

Цель проекта: изучить современные технологии при построении ЦОД на оборудовании Arista

- 1.** Спроектировать фабрику в соответствии с архитекторой Клоза в среде эмуляции EVE-NG;
- 2.** Предложить типовые решения по подключению потребителей в режиме assymmetric IRB L2VNI
- 3.** Предложить типовые решения по подключению потребителей в режиме assymmetric IRB L3VNI, MC-LAG (ESI LAG) и Anycast Gatway;
- 4.** Предложить решение для подключения фабрики к внешней сети;

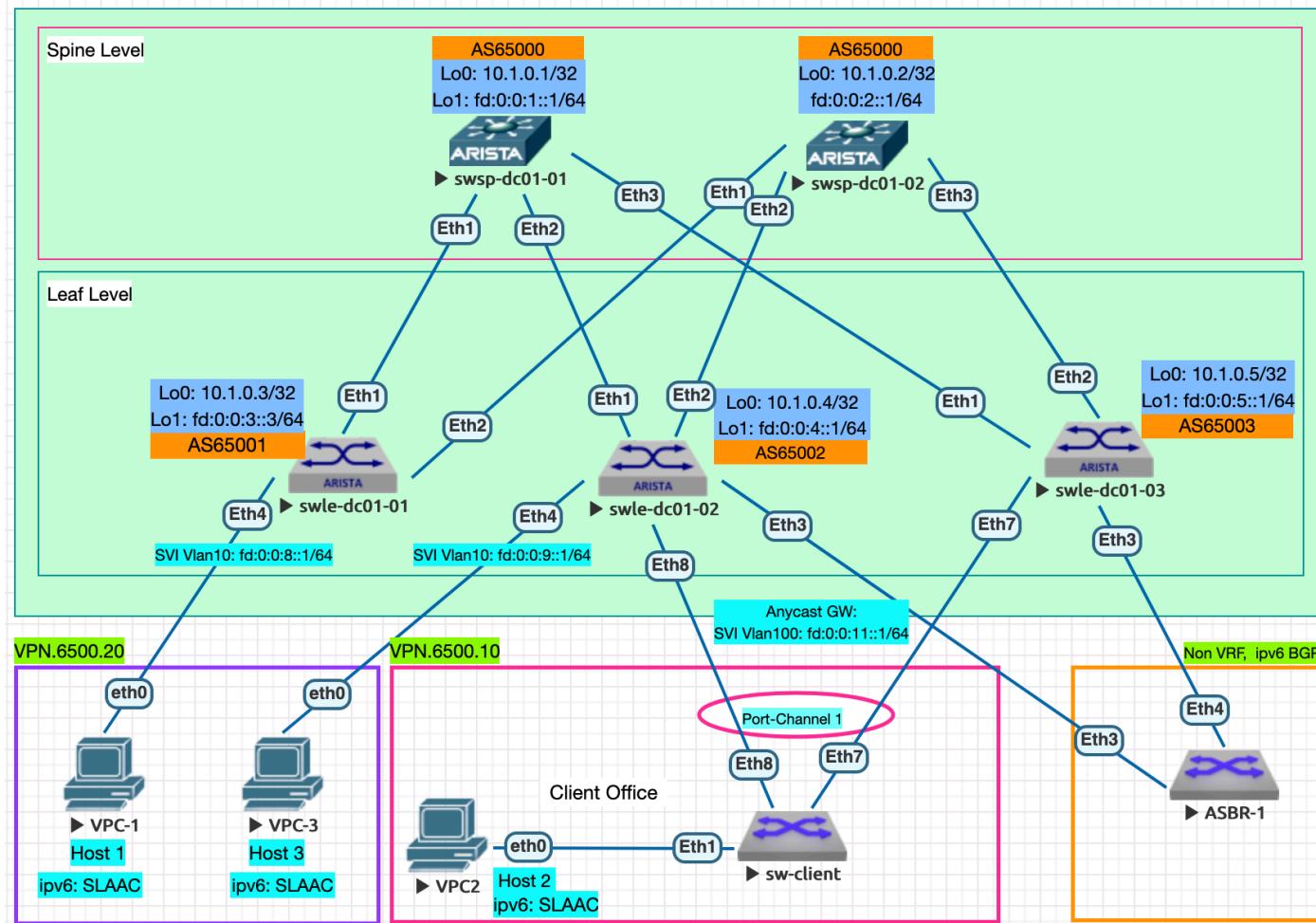


Какие технологии использовались

1. EVPN, VxLAN, ECMP
2. BGP в underlay
3. IPv6 в underlay и overlay
4. Anycast GW
5. ESI LAG
6. Bidirectional forwarding (оптимизация времени определения недоступности линка)



Что получилось. Сетевая схема фабрики с основными сущностями



Что получилось. Адресный план и прочие условия

Адресный план

- Адреса для Loopback интерфейсов ipv4: 10.00[DC num].0.0/23, 512 устройств на 1 DC, Таблица №1;
- Адреса для Loopback интерфейсов ipv6: fd::/61 в соответствии с Таблицей №1;
- Линковые адреса ipv4: 10.10[DC_num].16.0/20, для линковых сетей использовать /31, младший адрес на стороне Spine;
- Линковые адреса 'Spine - Leaf' ipv6: link-local
- Линковые интерфейсы 'Spine - Leaf' в 10 vlan (подинтерфейсы Eth [0..n].10, с тэгом 802.1q 10)

Таблица №1 Имена хостов и адреса Loopback, AS Number

Коммутатор	Hostname	IP Loopback 0	IP Loopback 1	BGP AS Number
Spine 1	swsp-dc01-01	10.1.0.1	fd:0:0:1::1/64	65000
Spine 2	swsp-dc01-02	10.1.0.2	fd:0:0:2:1/64	65000
Leaf 1	swle-dc01-01	10.1.0.3	fd:0:0:3::1/64	65001
Leaf 2	swle-dc01-02	10.1.0.4	fd:0:0:4::1/64	65002
Leaf 3	swle-dc01-03	10.1.0.5	fd:0:0:5::1/64	65003

<https://github.com/rinatkaa/DataCenterLabs/blob/main/Project/README.md>



Что получилось. Адресный план и прочие условия

Таблица №2 Настройки устройств подключенных к фабрике

Клиентское устройство	IPv4 адрес	IPv4 шлюза	IPv6 router	IPv6 интерфейса хоста
Host-1	-	-	fd:0:0:8::1/64	SLAAC
Host-3	-	-	fd:0:0:9::1/64	SLAAC
Host-2/LACP	-	-	fd:0:0:11::1/64	SLAAC
Host-2/LACP	-	-	fd:0:0:11::1/64	SLAAC

*** Host-2 подключен через клиентский коммутатор с включенной агрегацией каналов LACP в режиме Active в демонстративных целях функционала Multihoming***

Таблица №3 Настройки VRF

VRF	IPv6 network	RD	RT Both	VNI
VPN.6500.10	fd:0:0:8::1/64	65000:10	65000:10	10010
VPN.6500.20	fd:0:0:11::1/64	65000:20	65000:20	10020

<https://github.com/rinatkaa/DataCenterLabs/blob/main/Project/README.md>



Что получилось. Базовые настройки underlay, eBGP

1. Настроено eBGP соседство между Leaf - Spine устройствами (ipv6 link local)
2. Включен bfd для быстрого определения ошибок состояния линков
3. Маршруты Loopback интерфейсов распространяются по фабрике

```
swle-dc01-01#sh bfd peers
VRF name: default
-----
DstAddr          MyDisc   YourDisc  Interface/Transport  Type
-----
fe80::5200:ff:fe03:3766  766824962  636697255    Ethernet2.10(30)  normal
fe80::5200:ff:fed5:5dc0  4293376334  3116120061   Ethernet1.10(31)  normal
-----
LastUp           LastDown   LastDiag   State
-----
05/19/25 17:51   05/19/25 17:51   No Diagnostic     Up
05/19/25 17:51   05/19/25 17:51   No Diagnostic     Up
```

```
swle-dc01-01#sh bgp ipv6 unicast
BGP routing table information for VRF default
Router identifier 10.1.0.3, local AS number 65001
Route status codes: s - suppressed contributor, * - valid, > - active, E - ECMP head, e - ECMP
S - Stale, c - Contributing to ECMP, b - backup, L - labeled-unicast
% - Pending BGP convergence
Origin codes: i - IGP, e - EGP, ? - incomplete
RPKI Origin Validation codes: V - valid, I - invalid, U - unknown
AS Path Attributes: Or-ID - Originator ID, C-LST - Cluster List, LL Nexthop - Link Local Nexthop
```

	Network	Next Hop	Metric	AIGP	LocPref	Weight	Path
* >	fd:0:0:1::/64	fe80::5200:ff:fed5:5dc0%Et1.10	0	-	100	0	65000 i
* >	fd:0:0:2::/64	fe80::5200:ff:fe03:3766%Et2.10	0	-	100	0	65000 i
* >	fd:0:0:3::/64	-	-	-	0	i	
* >Ec	fd:0:0:4::/64	fe80::5200:ff:fed5:5dc0%Et1.10	0	-	100	0	65000 65002 i
* ec	fd:0:0:4::/64	fe80::5200:ff:fe03:3766%Et2.10	0	-	100	0	65000 65002 i
* >Ec	fd:0:0:5::/64	fe80::5200:ff:fed5:5dc0%Et1.10	0	-	100	0	65000 65003 i
* ec	fd:0:0:5::/64	fe80::5200:ff:fe03:3766%Et2.10	0	-	100	0	65000 65003 i



Что получилось. Базовые настройки Overlay фабрики и L2VNI

1. Настроено EVPN BGP соседство между Leaf - Spine устройствами (ipv6 link local)
2. Настроен VNI 20200 на Vxlan интерфейсе
3. Настроен MAC-VRF

```
vlan 20
  rd auto
  route-target both 20:20
  redistribute learned
```

```
swle-dc01-01#sh run int vx1
interface Vxlan1
  vxlan source-interface Loopback1
  vxlan udp-port 4789
  vxlan encapsulation ipv6
  vxlan vlan 10 vni 10101
  vxlan vlan 20 vni 20200
```

```
swle-dc01-01#sh bgp evpn vni 20200 next-hop fd:0:0:4::1
BGP routing table information for VRF default
Router identifier 10.1.0.3, local AS number 65001
Route status codes: * - valid, > - active, S - Stale, E - ECMP head, e - ECMP
                    c - Contributing to ECMP, % - Pending BGP convergence
Origin codes: i - IGP, e - EGP, ? - incomplete
AS Path Attributes: Or-ID - Originator ID, C-LST - Cluster List, LL Nexthop - Link Local Nexthop

      Network          Next Hop           Metric LocPref Weight Path
* >Ec    RD: 10.1.0.4:20 mac-ip 0050.7966.680a
                  fd:0:0:4::1        -     100     0   65000 65002 i
*  ec    RD: 10.1.0.4:20 mac-ip 0050.7966.680a
                  fd:0:0:4::1        -     100     0   65000 65002 i
* >Ec    RD: 10.1.0.4:20 imet fd:0:0:4::1
                  fd:0:0:4::1        -     100     0   65000 65002 i
*  ec    RD: 10.1.0.4:20 imet fd:0:0:4::1
                  fd:0:0:4::1        -     100     0   65000 65002 i
```

Наблюдаем наличие imet и mac-ip префиксов по EVPN BGP дважды т.к. Работает ECMP



Что получилось. Базовые настройки Overlay фабрики. Symmetric IRB (L3VNI)

- Настроен vrf VPN.6500.10 (Host-1, Host-3), в нем включена маршрутизация ipv6 и redistribute connected
- Host'ы подключены к Ethernet 4 коммутаторов Leaf-1, Leaf-2
- Настроен общий L2VNI (MAC VPN) для 10 и 20 van
- Настроен L3VNI на обоих коммутаторах
- Настроены соответствующие L2VNI на Vxlan интерфейсе:

```
vrf VPN.6500.10
rd 65000:10
route-target import evpn 65000:10
route-target export evpn 65000:10
redistribute connected
!
```

```
vlan-aware-bundle l3vnilab
rd 10.1.0.3:101
route-target both 101:101
redistribute learned
vlan 10,20,100
.
```

```
interface Vxlan1
vxlan source-interface Loopback1
vxlan udp-port 4789
vxlan encapsulation ipv6
vxlan vlan 10 vni 10101
vxlan vlan 20 vni 20202
vxlan vlan 100 vni 10100
vxlan vrf VPN.6500.10 vni 10010
```

```
swle-dc01-01#sh bgp evpn vni 20202
BGP routing table information for VRF default
Router identifier 10.1.0.3, local AS number 65001
Route status codes: * - valid, > - active, S - Stale, E - ECMP head, e - ECMP
                  c - Contributing to ECMP, % - Pending BGP convergence
Origin codes: i - IGP, e - EGP, ? - incomplete
AS Path Attributes: Or-ID - Originator ID, C-LST - Cluster List, LL Nexthop - Link Local Nexthop

      Network          Next Hop           Metric LocPref Weight Path
* >   RD: 10.1.0.3:101 mac-ip 20202 0050.7966.6806
      -                 -               -       -       0       i
* >   RD: 10.1.0.3:101 mac-ip 20202 0050.7966.6806 fd::8:2050:79ff:fe66:6806
      -                 -               -       -       0       i
* >Ec  RD: 10.1.0.4:101 mac-ip 20202 0050.7966.680a
      fd:0:0:4::1        -               -       100     0       65000 65002 i
* ec   RD: 10.1.0.4:101 mac-ip 20202 0050.7966.680a
      fd:0:0:4::1        -               -       100     0       65000 65002 i
* >Ec  RD: 10.1.0.4:101 mac-ip 20202 0050.7966.680a fd::9:2050:79ff:fe66:680a
      fd:0:0:4::1        -               -       100     0       65000 65002 i
* ec   RD: 10.1.0.4:101 mac-ip 20202 0050.7966.680a fd::9:2050:79ff:fe66:680a
      fd:0:0:4::1        -               -       100     0       65000 65002 i
* >   RD: 10.1.0.3:101 imet 20202 fd::0:0:3::1
      -                 -               -       -       0       i
* >Ec  RD: 10.1.0.4:101 imet 20202 fd::0:0:4::1
      fd:0:0:4::1        -               -       100     0       65000 65002 i
* ec   RD: 10.1.0.4:101 imet 20202 fd::0:0:4::1
      fd:0:0:4::1        -               -       100     0       65000 65002 i
swle-dc01-01#sh ip6c 2
```

Наблюдаем связность между хостами в разных Vlan и разных ipv6 сетях, обеспеченную наличием двух записей type-2 (imet): MAC и MAC - IPv6



Что получилось. Базовые настройки Overlay фабрики для Multihoming

1. Настроить Port-channel на коммутаторах Leaf-2 и Leaf-3, задать ESI
2. Настройка SVI на leaf-2, leaf-3
3. Убедиться в наличии префиксов с 'route-type 1,4' по evpn/bgp для соседа Leaf-2, Leaf-3 (вывод на примере Leaf-2):

```
swle-dc01-03#sh run | b Port-
interface Port-Channel1
description - Client SW
switchport trunk allowed vlan 100
switchport mode trunk
!
evpn ethernet-segment
    identifier 0000:8888:8888:8888:8888
    route-target import 88:88:88:88:88:88
    lacp system-id 0001.0002.0003
!
```

```
swle-dc01-03#sh run int vl100
interface Vlan100
vrf VPN.6500.10
ip address unnumbered Loopback0
ipv6 enable
ipv6 address virtual fd:0:0:11::1/64
!
```

```
swle-dc01-02#sh bgp evpn next-hop fd:0:0:5::1
BGP routing table information for VRF default
Router identifier 10.1.0.4, local AS number 65002
Route status codes: * - valid, > - active, S - Stale, E - ECMP head, e - ECMP
                                         c - Contributing to ECMP, % - Pending BGP convergence
Origin codes: i - IGP, e - EGP, ? - incomplete
AS Path Attributes: Or-ID - Originator ID, C-LST - Cluster List, LL Nexthop - Link Local Nexthop

      Network          Next Hop           Metric LocPref Weight  Path
* >Ec   RD: 10.1.0.5:101 auto-discovery 10100 0000:8888:8888:8888
                                         fd:0:0:5::1   -     100   0  65000 65003 i
*   ec   RD: 10.1.0.5:101 auto-discovery 10100 0000:8888:8888:8888
                                         fd:0:0:5::1   -     100   0  65000 65003 i
* >Ec   RD: 10.1.0.5:1 auto-discovery 0000:8888:8888:8888
                                         fd:0:0:5::1   -     100   0  65000 65003 i
*   ec   RD: 10.1.0.5:1 auto-discovery 0000:8888:8888:8888
                                         fd:0:0:5::1   -     100   0  65000 65003 i
* >Ec   RD: 10.1.0.5:101 auto-discovery 10100 fd:0:0:5::1
                                         fd:0:0:5::1   -     100   0  65000 65003 i
*   ec   RD: 10.1.0.5:101 imet 10100 fd:0:0:5::1
                                         fd:0:0:5::1   -     100   0  65000 65003 i
* >Ec   RD: 10.1.0.5:101 imet 10100 fd:0:0:5::1
                                         fd:0:0:5::1   -     100   0  65000 65003 i
*   ec   RD: 10.1.0.5:101 imet 10101 fd:0:0:5::1
                                         fd:0:0:5::1   -     100   0  65000 65003 i
* >Ec   RD: 10.1.0.5:101 imet 10101 fd:0:0:5::1
                                         fd:0:0:5::1   -     100   0  65000 65003 i
*   ec   RD: 10.1.0.5:101 imet 10101 fd:0:0:5::1
                                         fd:0:0:5::1   -     100   0  65000 65003 i
* >Ec   RD: 10.1.0.5:1 ethernet-segment 0000:8888:8888:8888 fd:0:0:5::1
                                         fd:0:0:5::1   -     100   0  65000 65003 i
*   ec   RD: 10.1.0.5:1 ethernet-segment 0000:8888:8888:8888 fd:0:0:5::1
                                         fd:0:0:5::1   -     100   0  65000 65003 i
* >Ec   RD: 65000:10 ip-prefix fd:0:0:13::/64
                                         fd:0:0:5::1   -     100   0  65000 65003 ?
*   ec   RD: 65000:10 ip-prefix fd:0:0:13::/64
                                         fd:0:0:5::1   -     100   0  65000 65003 ?
```

Наблюдаем наличие префиксов route type 1,4 с leaf-1, работает ECMP

<https://github.com/rinatkaa/DataCenterLabs/tree/main/lab07>



Что получилось. Базовые настройки Overlay фабрики для внешнего подключения

1. Выполняем подключение граничного маршрутизатора ASBR по L3 согласно схемы, настраиваем на нем loopback 1 с fd:0:0:13::1/64;
2. Поднимаем BGP сессии с Leaf-2, Leaf-3 в VRF VPN.6500.10
3. На в сторону ASBR, настраиваем redistribute connected;
4. Убеждаемся на Leaf-1,2 наличие маршрута type-5 (ip-prefix)

```
vrf VPN.6500.10
rd 65000:10
route-target import evpn 65000:10
route-target export evpn 65000:10
neighbor fd:0:0:100::2 peer group asbr
neighbor fd:0:0:100::2 remote-as 5555
```

```
swle-dc01-02#sh run int e3
interface Ethernet3
  description - ASBR-1 1/13
  no switchport
  vrf VPN.6500.10
  ip address unnumbered Loopback0
  ipv6 enable
  ipv6 address fd:0:0:100::1/64
```

```
swsp-dc01-02#sh bgp evpn route-type ip-prefix fd:0:0:13::/64
BGP routing table information for VRF default
Router identifier 10.1.0.2, local AS number 65000
BGP routing table entry for ip-prefix fd:0:0:13::/64, Route Distinguisher: 65000:10
Paths: 2 available
 65003 5555
    fd:0:0:5::1 from fe80::5200:ff:fe15:f4e8%Et3.10 (10.1.0.5)
      Origin IGP, metric -, localpref 100, weight 0, tag 0, valid, external, ECMP head, ECMP, best, ECMP contributor
      Extended Community: Route-Target-AS:65000:10 TunnelEncap:tunnelTypeVxlan EvpnRouterMac:50:00:00:15:f4:e8
      VNI: 10010
  65002 5555
    fd:0:0:4::1 from fe80::5200:ff:fecb:38c2%Et2.10 (10.1.0.4)
      Origin IGP, metric -, localpref 100, weight 0, tag 0, valid, external, ECMP, ECMP contributor
      Extended Community: Route-Target-AS:65000:10 TunnelEncap:tunnelTypeVxlan EvpnRouterMac:50:00:00:cb:38:c2
      VNI: 10010
```

```
router bgp 5555
  router-id 10.5.5.5
  graceful-restart restart-time 300
  neighbor leaves peer group
  neighbor leaves bfd
  neighbor leaves timers 20 60
  neighbor leaves graceful-restart
  neighbor fd:0:0:100::1 peer group leaves
  neighbor fd:0:0:100::1 remote-as 65002
  neighbor fd:0:0:101::1 peer group leaves
  neighbor fd:0:0:101::1 remote-as 65003
!
address-family ipv6
  neighbor leaves activate
  redistribute connected
  redistribute static
```

Наблюдаем наличие префикса от внешнего маршрутизатора ASBR с коммутатора leaf-1, работает ECMP



Выводы

1. Подтверждена работы с eBGP в underlay и overlay в стеке ipv6
2. Не смотря на поддержку ipv6, для работы BGP необходим BGP RID в ipv4
3. Не смотря на поддержку ipv6 на SVI Leaf для корректной работы необходим ipv4 (использовался unnumbered)
4. Допустимо использовать link local ipv6 адреса в фабрике, не беспокоясь о назначении линковых IPv6 адресов
5. Успешно сэмбулирована работа наиболее востребованных сценариев в виртуальный среде



Ответьте на вопросы
одногруппников и
преподавателей и получите
обратную связь на свою работу

Вопросы и рекомендации



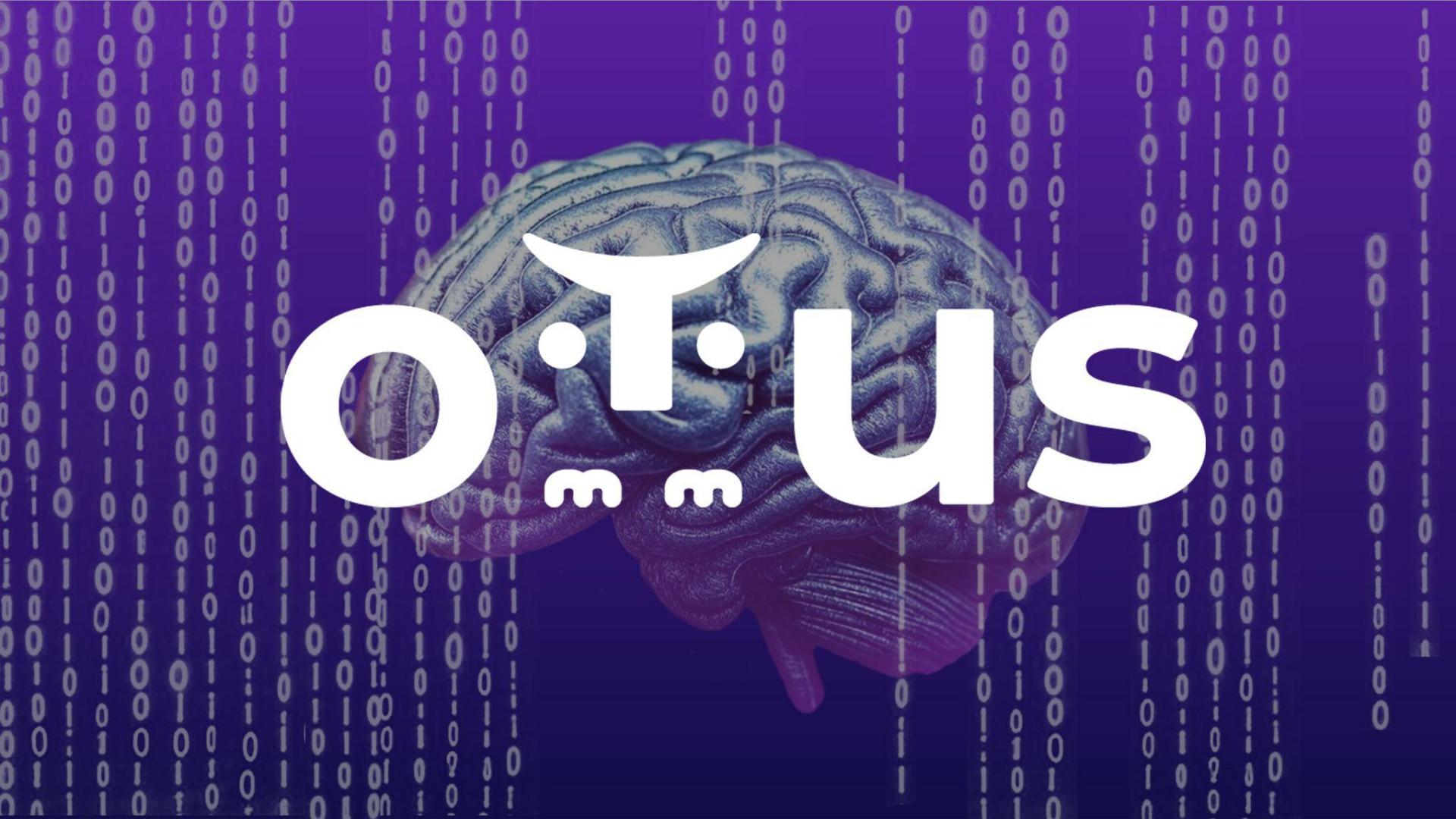
если есть вопросы



если вопросов нет

Спасибо за внимание!





opus