



BIRD Internet Routing Daemon

Route servery



Ondřej Filip • ondrej.filip@nic.cz
4. října 2014 • Linux Days • Praha

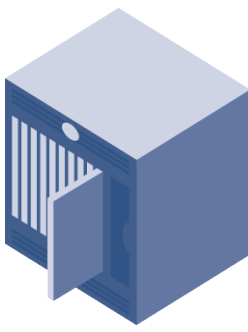


CZ.NIC, CZ.NIC Labs



- Přibližně 1,1M domén, 38% podepsáno DNSSEC
- Ne pouze registr domény .cz
- CSIRT.CZ
- Výzkumné a vývojové oddělení – CZ.NIC Labs
- Projekty pro dobro Internetu – Open Source
 - Knot DNS, Turris, Datovka, Tablexia...
- Sledujte na <http://labs.nic.cz>

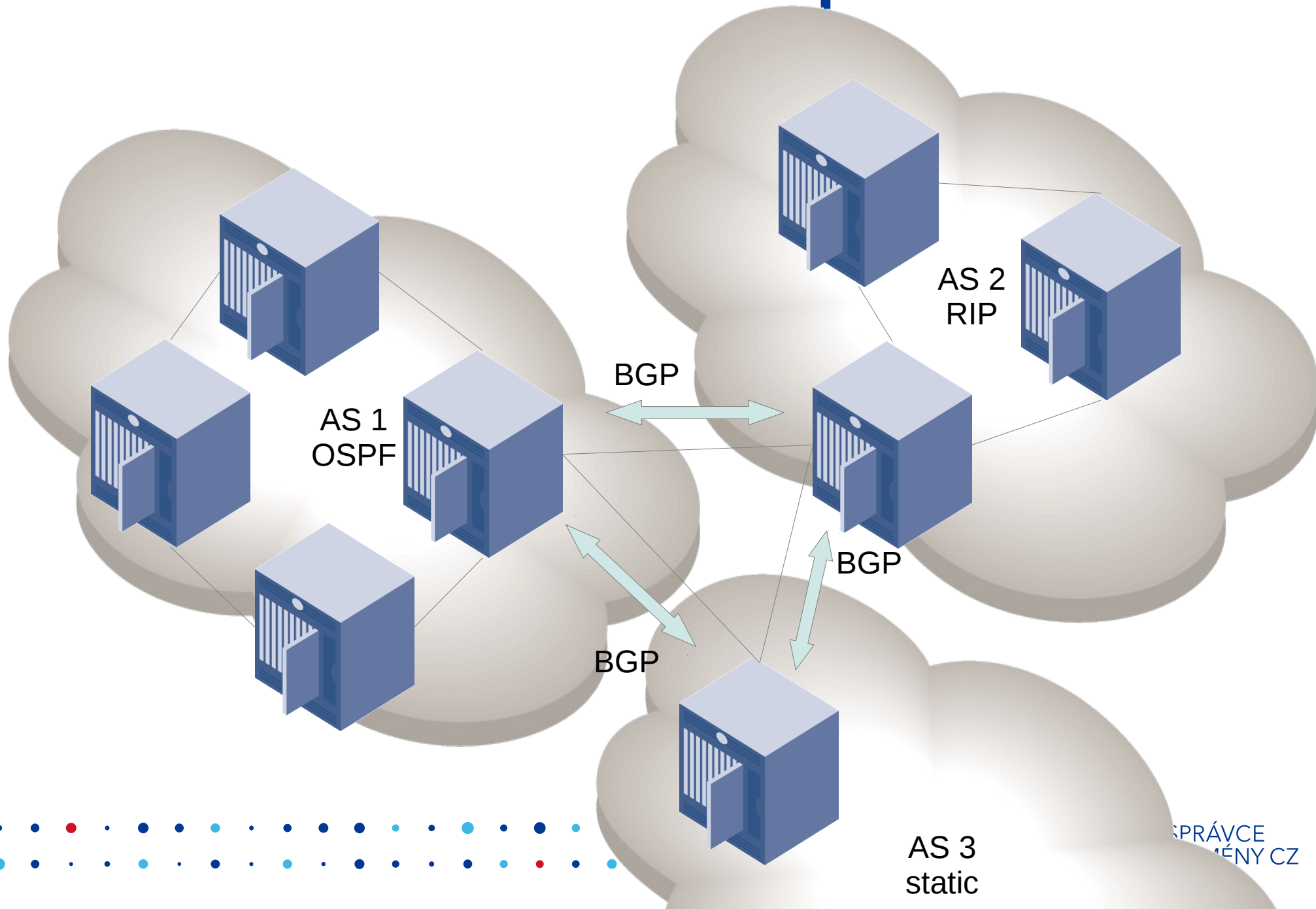




Směrování vs forwarding

- Router - zařízení připojené k více sítím
- Umí přeposlat „cizí“ zprávu - forwarding
- Cestu pozná podle směrovací (routovací) tabulky
- Sestavování routovací tabulky – routing
 - Statické
 - Dynamické
 - Interní - uvnitř AS rychlé, důvěřivé, přesné
 - RIP, OSPF, IS-IS
 - Externí (mezi AS, pomalé, filtrování, přibližné – pouze BGP)

Rozdělení směrovacích protokolů



Směrovací démon

- Na Linuxu (a ostatních UNIXech) – uživatelská aplikace mimo jádro, forwarding v jádře
- Obvykle implementuje více směrovacích protokolů
- Směrovací politika - filtrování
- Quagga (Zebra) – Cisco syntax
<http://www.quagga.net>
- OpenBGPD - <http://www.openbgpd.org>
- GateD. – zastaralý, ne volná licence

• BIRD

Historie projektu

- Start projektu v roce 1999
- Seminární projekt – MFF UK Praha
- Projekt uspán
- Drobné probuzení v letech 2003 a 2006 (CESNET)
- Plně obnoveno na přelomu 2008/2009 v rámci Laboratoří CZ.NIC - <http://labs.nic.cz>



Cíle projektu (1999)

- Open source směrovací démon – alternativa k tehdejšímu démonu Quagga/Zebra (a GateD)
- Rychlý a efektivní
- Portabilní, modulární
- Podpora současných směrovacích protokolů
- IPv6 a IPv4 v jednom zdrojovém kódu – duální kompilace
- Jednoduchá konfigurace a rekonfigurace (!)



Vlastnosti

- Portabilní – Linux, FreeBSD, NetBSD, OpenBSD
- Podpora IPv4/IPv6, statické směrování, IPv6 router advertisements
- RIP, RIPv2, RIPng
- OSPFv2, OSPFv3
- MRTdump logging
- Bidirectional Forwarding Detection (BFD)
- BGP, BGP route server, BGP route reflector, add path, graceful restart

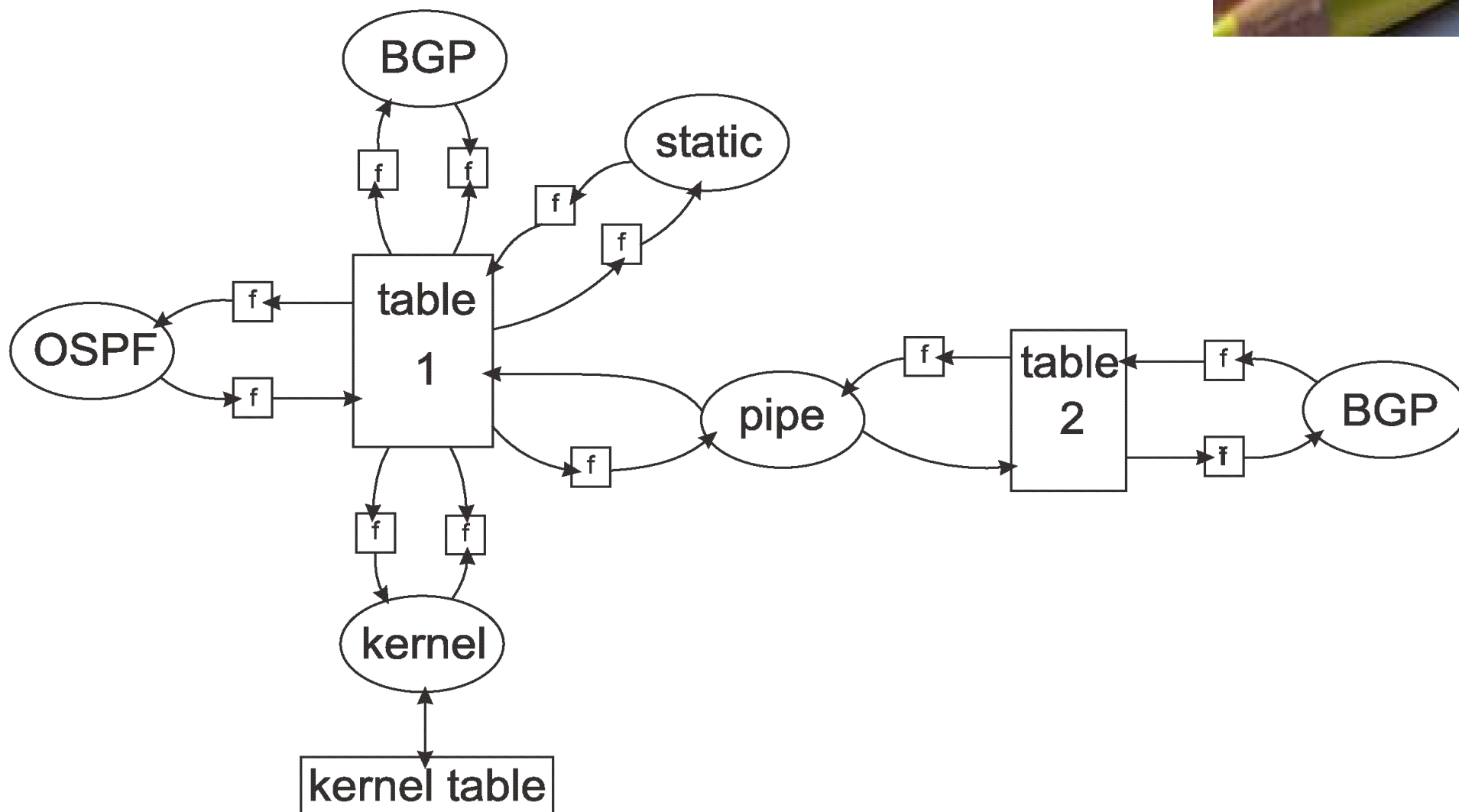


Vlastnosti

- Více směrovací tabulek - RIBs (interně a také synchronizace s OS)
- Více routerů, route reflektorů na jednom systému
- Protokol PIPE (!)
- Více směrovačů, route serverů a pod. na 1 systému
- Efektivní, silná konfigurace, Silný jazyk pro filtrování

• • • • • Příkazová řádka (show, restart, ...)

Design



Příklad konfigurace

```
log "/var/log/bird.log" all;
```

```
router id 193.51.100.238;
```

```
protocol static {
```

```
    route 10.0.0.0/8 drop;
```

```
    route 172.16.0.0/12 drop;
```

```
    route 192.168.0.0/16 drop;
```

```
}
```

```
filter bgp_out {
```

```
    if (net = 192.175.48.0/24 )
```

```
&&
```

```
    ::::: ( source ::: CZ.nic | SPRÁVCE  
    ::::: DOMÉNY CZ
```

```
BTS DEVICE) then accept;
```

Příkazová řádka



```
bird> show protocols
```

```
name          proto      table      state
```

```
since info
```

```
direct1      Direct    master     up
```

```
Apr11
```

```
kernel1     Kernel    master     up
```

```
Apr11
```

```
device1     Device    master     up
```

```
Apr11
```

```
static1     Static    master     up
```

```
Apr11
```

```
NIX_2       BGP       master     up
```

```
· Apr11 : Established :: :: :: ::
```

cz.nic

SPRÁVCE
DOMÉNY CZ

```
NIX_1       BGP       master     up
```

Příkazová řádka



```
bird> show route for 127.0.0.1
127.0.0.0/8          dev lo
[direct1 13:09] * (240)
```

```
bird> show route filter bgp_out
192.175.48.0/24     dev dummy0
[direct1 Apr1] * (240)
```

```
bird> show route count
1469 of 1469 routes for 849
networks
```



Odlehčená příkazová řádka

- Standardní příkazová řádka používá libncurses, libhistory, libreadline
- Lightweight CLI – žádné dodatečné knihovny
- Zejména pro embedded systémy (OpenWRT)



Filtrování



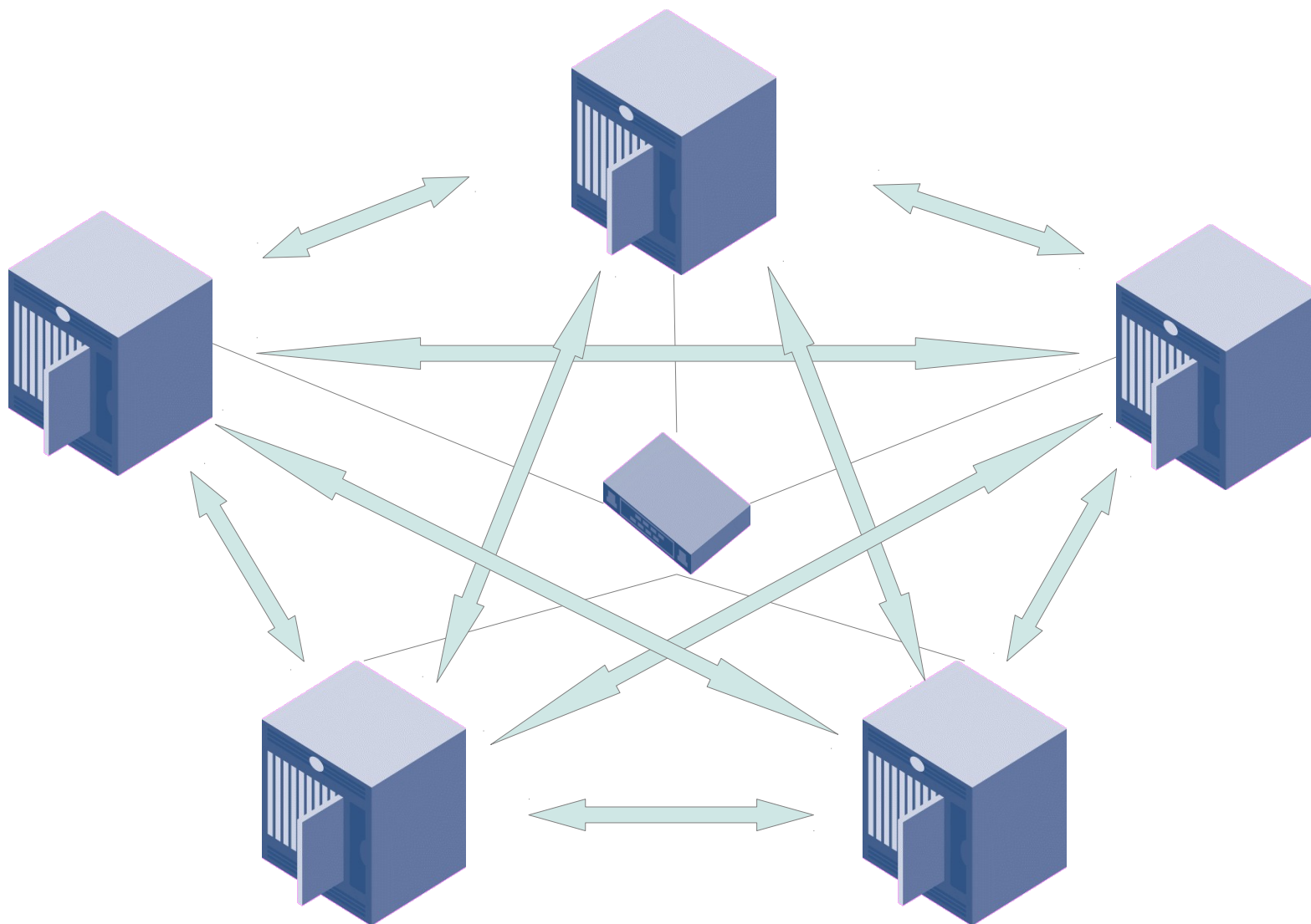
```
function avoid_martians()  
prefix set martians;  
{  
    martians =  
        [ 169.254.0.0/16+, 172.16.0.0  
/12+,          192.168.0.0/16+,  
10.0.0.0/8+, 224.0.0.0/4+,  
    240.0.0.0/4+, 0.0.0.0/32-,  
0.0.0.0/0{25,32},  
    0.0.0.0/0{0,7}  ];
```

Filtry

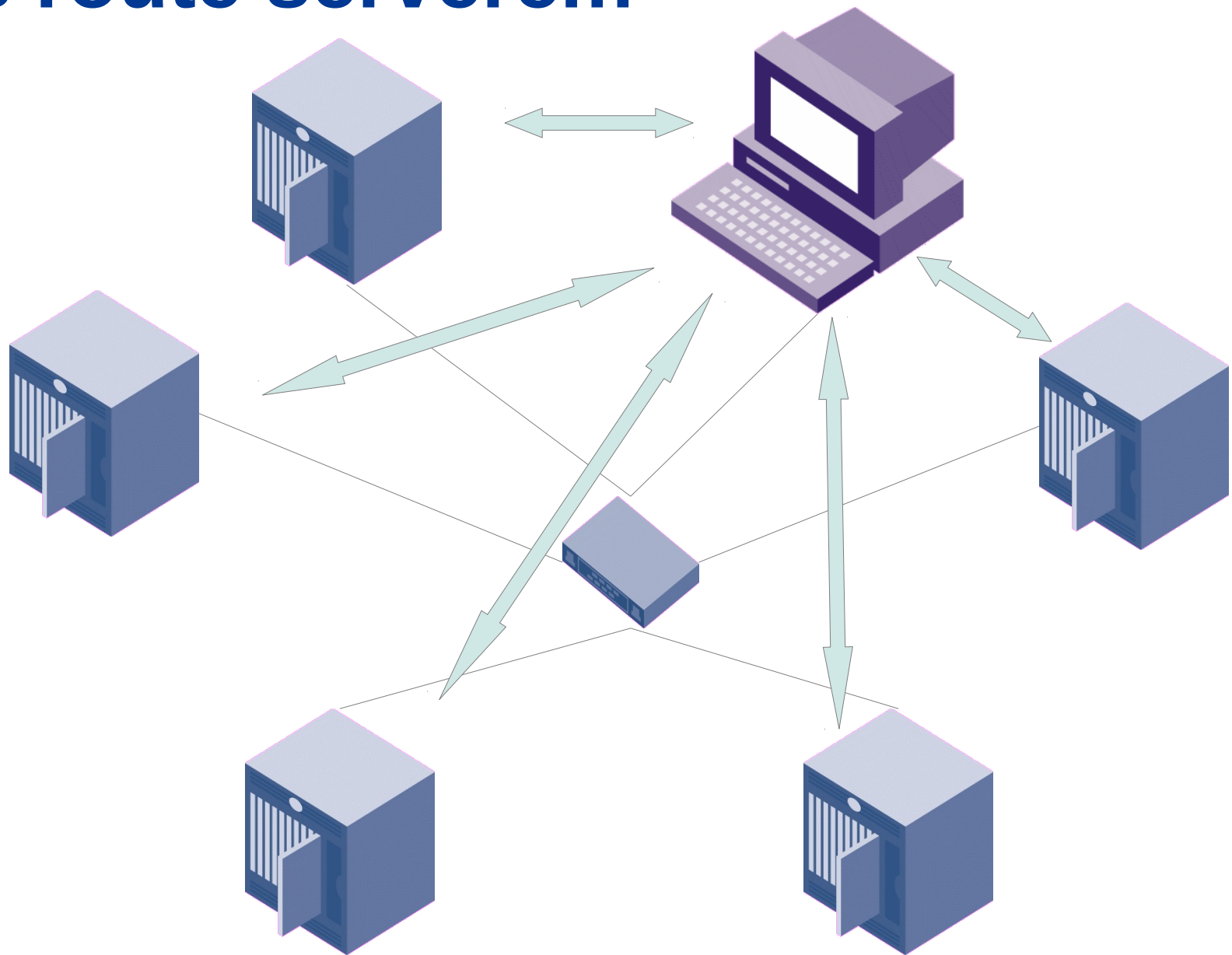


- Filtry kompilování do byte kódu
- Proměnné (!), Sety
- Datové typy: bool, int, ip, prefix, enum, quad, string, bgppath, bgpmask, clist
- Operátory: +, -, *, /, comparisons, logical, element_of_set (~)
- Kontrolní struktury – if/else, case
- Sady implementovány pomocí vyvážených binárních stromů (nebo podobných) –
logaritmická časová složitost

IXP bez route serveru



IXP s route serverem



Route servery

- Bez RS
 - Každá síť – $n-1$ BGP relací - celkově $n*(n-1)/2$
 - Sestavování peeringu z každým nově příchozím
 - Zátěž CPU routerů (především s TCP-MD5)
- S RS
 - Pouze jedna BGP relace na směrovač (případně 2)
 - Snižuje počet BGP relací
 - IXP je atraktivní pro nově příchozí

Příklad filtru - RS

- Politika route serveru

| Pořadí | Komunita | Akce |
|--------|-----------------|----------------------------|
| 1 | 0:<peer-as> | Nepropaguj routu <peer-as> |
| 2 | 47200:<peer-as> | Propaguj routu <peer-as> |
| 3 | 0:47200 | Nepropaguj nikomu |
| 4 | 47200:47200 | Propaguj všem |



Příklad filtru - RS

- Každý ISP zapojený do route serveru neztrácí svobodu definice své vlastní směrovací politiky
- Příklad 1 – chci posílat svou síť pouze CZ.NICu
 - 10.0.0.0/8 community 47000:25192, 0:47000
- Příklad 2 – chci posílat všem krom CZ.NICu
 - 10.0.0.0/8 community 0:25192, 47200:47200



Příklad filtru – RS@BIRD

```
define myas = 47200;
```

```
function bgp_out(int peeras)
```

```
{  
    if ! (source = RTS_BGP ) then return false;  
    if (0,peeras) ~ bgp_community then return false;  
    if (myas,peeras) ~ bgp_community then return true;  
    if (0, myas) ~ bgp_community then return false;  
    return true;  
}
```

```
protocol bgp R25192x1 {  
    local as myas;  
    neighbor 194.50.100.13 as 25192;  
    import where bgp_in(25192);  
    export where bgp_out(25192);  
    rs client;  
}
```



Příklad filtru – RS@Quagga

```
ip community-list standard C-0-10001 permit 0:10001
ip community-list standard C-0-47200 permit 0:47200
ip community-list standard C-47200-10001 permit 47200:10001

route-map Policy10001 deny 10
  match community C-0-10001
!
route-map Policy10001 permit 20
  match community C-47200-10001
!
route-map Policy10001 deny 30
  match community C-0-47200
!
route-map Policy10001 permit 40
!
```



BIRD v NIX.CZ

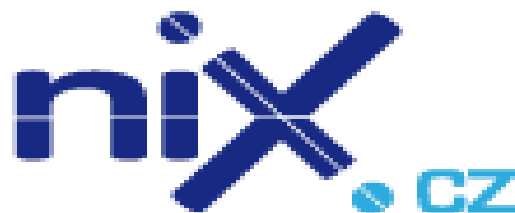


NIX.CZ

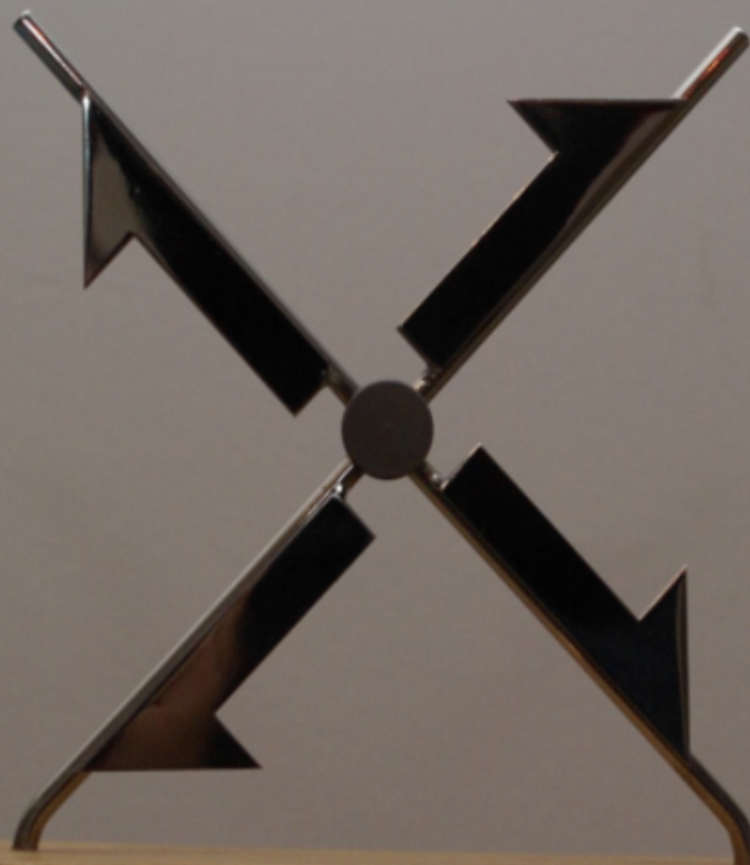
- AS112
- Duální route server
- Linux & FreeBSD, IPv4 & IPv6
- Okolo 130 BGP relací
- Konfigurace z intranetu a RPSL (RIPE DB)
- Filtrování dle prefixu a AS path
- Rekonfigurace každé 2 hodiny
- 4M znaků



Nasazení BIRDa - >50% IXP



Ocenění



Conspicuous Contribution to LINX Award 2010
To Ondrej Filipp and his talented Czech colleagues,
supported by NIX.CZ and CZ.NIC,
for outstanding effort in creating BIRD, an open source
route server platform of great benefit to IXPs



Jiná nasazení

- Netflix Open Connect – CDN/Cache
- BGP/OSPF router u menších ISP
- Router pro některé anycastové uzly CZ.NICu
- Používán v malých embedded systémech – součást firmware některých WiFi AP (OpenWRT)
- Součást analyzátoru BGP tabulek



Program podpory

- Open source
- Smlouvy na podporu
 - Velké či významné světové společnosti
- Projekt BIRD v zisku a pomáhá financovat ostatní projekty Laboratoří CZ.NIC
- Hledáme vývojáře!!!



Budoucí vývoj

- Integrace IPv4 a IPv6 kódu
- IS-IS
- Plná podpora RPKI

- Záleží na největších partnerech





Děkuji!



Ondřej Filip • ondrej.filip@nic.cz • <http://bird.network.cz>

