Running bitcoin (in production)

Baltic Honeybadger 2019 - September 15, 2019

Who I am

twitter/telegram/keybase: @ketominer

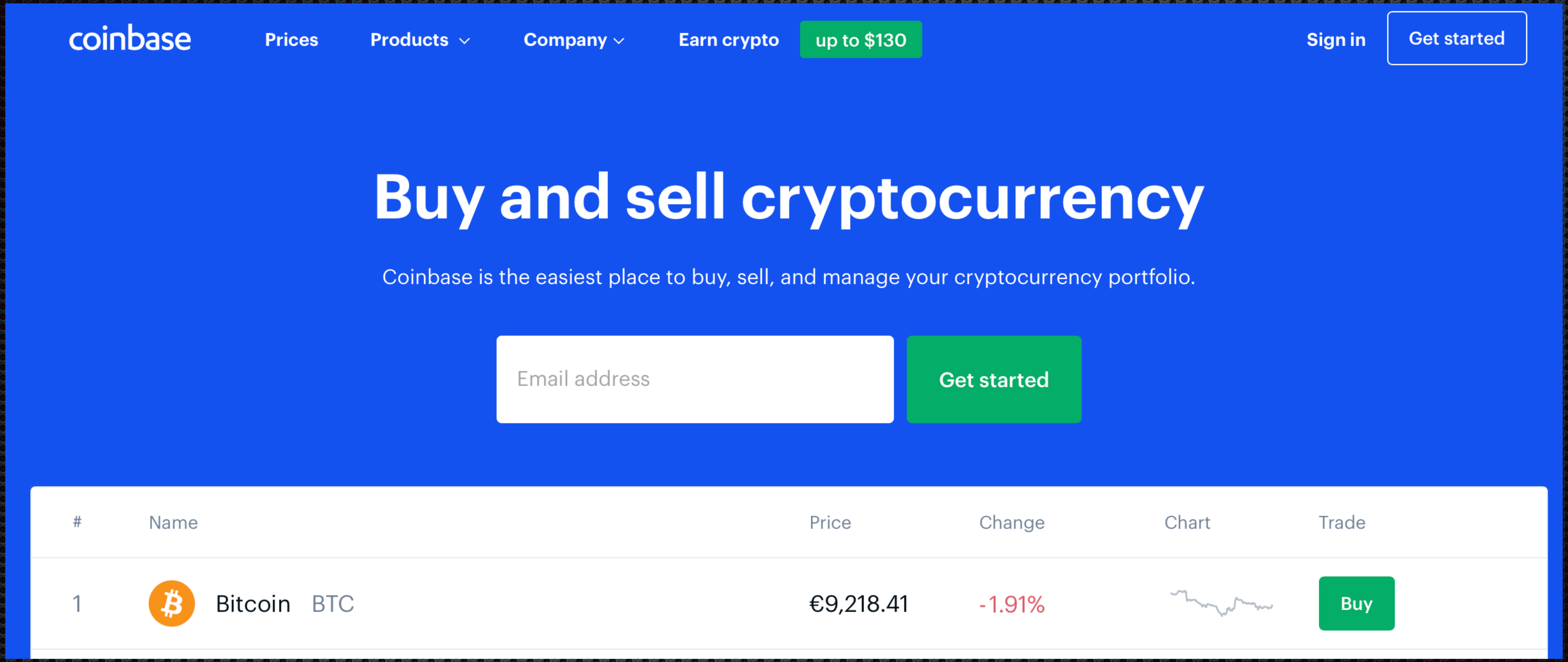
```
GPG: B5F6 FEBB 4D88 0398 7C60
     3E3F EAD1 75FA A6C4 B7DE
```

ketominer@nodl.it

- “ketominer”
  - electronics
  - (low level) code
  - networks
  - systems
  - putting weird things together and making them work
- Doing the **nodl box**, **nodl hosted services** and **host4coins**

How people run bitcoin

Custodial



Buy and sell cryptocurrency

Coinbase is the easiest place to buy, sell, and manage your cryptocurrency portfolio.

#	Name	Price	Change	Chart	Trade
1	Bitcoin BTC	€9,218.41	-1.91%		Buy

Old computers

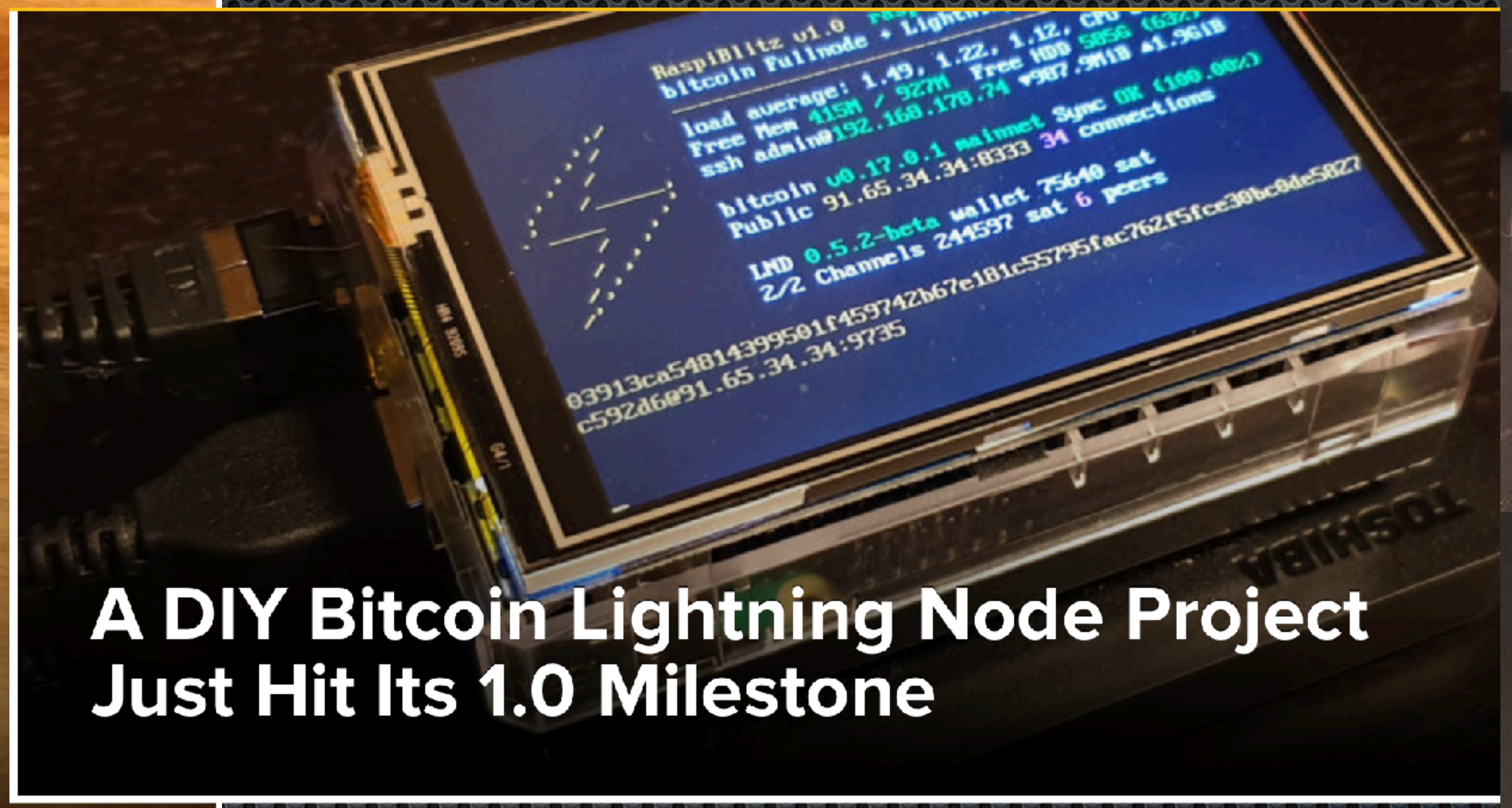
* not actual full nodes

Little metal and plastic boxes

Medium

My first impressions of the Casa Bitcoin Node

By Caprehab - November 28, 2018 · 57 · 0

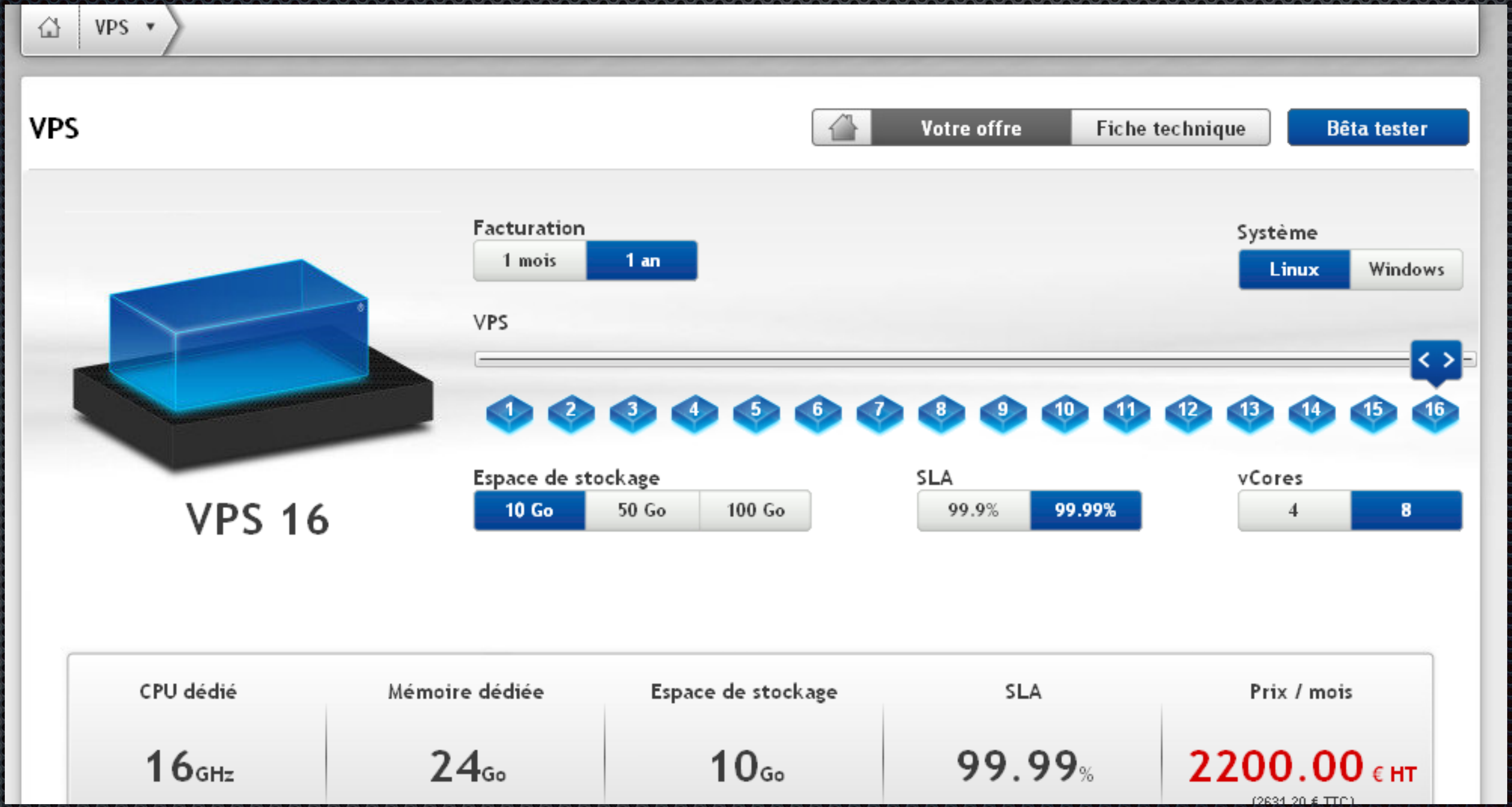


```
RaspiBlitz v1.0
bitcoin Fullnode + Lightn
load average: 1.49, 1.22, 1.12, CPU
Free Mem 415M / 927M  Free HDD 585G (63%)
ssh admin@192.168.178.74 ▼987.9MiB ▲1.96iB
bitcoin v0.17.0.1 mainnet Sync OK (100.00%)
Public 91.65.34.34:8333 34 connections
LND 0.5.2-beta wallet 75640 sat
2/2 Channels 244597 sat 6 peers
03913ca5481439950 1f459742b67e181c55795fac762f5fce30bc0de5827
c592d6@91.65.34.34:9735
```

A DIY Bitcoin Lightning Node Project Just Hit Its 1.0 Milestone

In the cloud(s)

On VPS (thanks BTCPay!)



VPS

Facturation: 1 mois [✓ 1 an]

Système: [✓ Linux] Windows

VPS: 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 [✓ 16]

VPS 16

Espace de stockage: [✓ 10 Go] 50 Go 100 Go

SLA: 99.9% [✓ 99.99%]

vCores: 4 [✓ 8]

CPU dédié	Mémoire dédiée	Espace de stockage	SLA	Prix / mois
16GHz	24Go	10Go	99.99%	2200.00 € HT (2631.20 € TTC)

There’s nothing wrong about it…

…but how could we make it different?

What do we want to run?

- The “full stack” (as of now)
  - a bitcoin full archiving node
  - a lightning node
  - a mixer
  - a payment server
  - a wallet backend

How do we want to run it?

- Factor what can be factored
- Provide reasonable redundancy
- Keep it simple

Factor what can be factored

- a bitcoin full archiving node -> YES
- a lightning node -> NO
- a mixer -> NO
- a payment server -> YES
- a wallet backend -> it depends (let’s say NO)

Provide reasonable redundancy

- a bitcoin full archiving node -> easy (active/active)
- a lightning node -> tricky (active/passive)
- a mixer -> useless
- a payment server -> easy
- a wallet backend -> it depends (we’ll deal with that later)

Factoring

bitcoind

- pointless to run many full archiving nodes in a single network range
- one (or two, max) node for every range and/or ASN

BTCPay Server

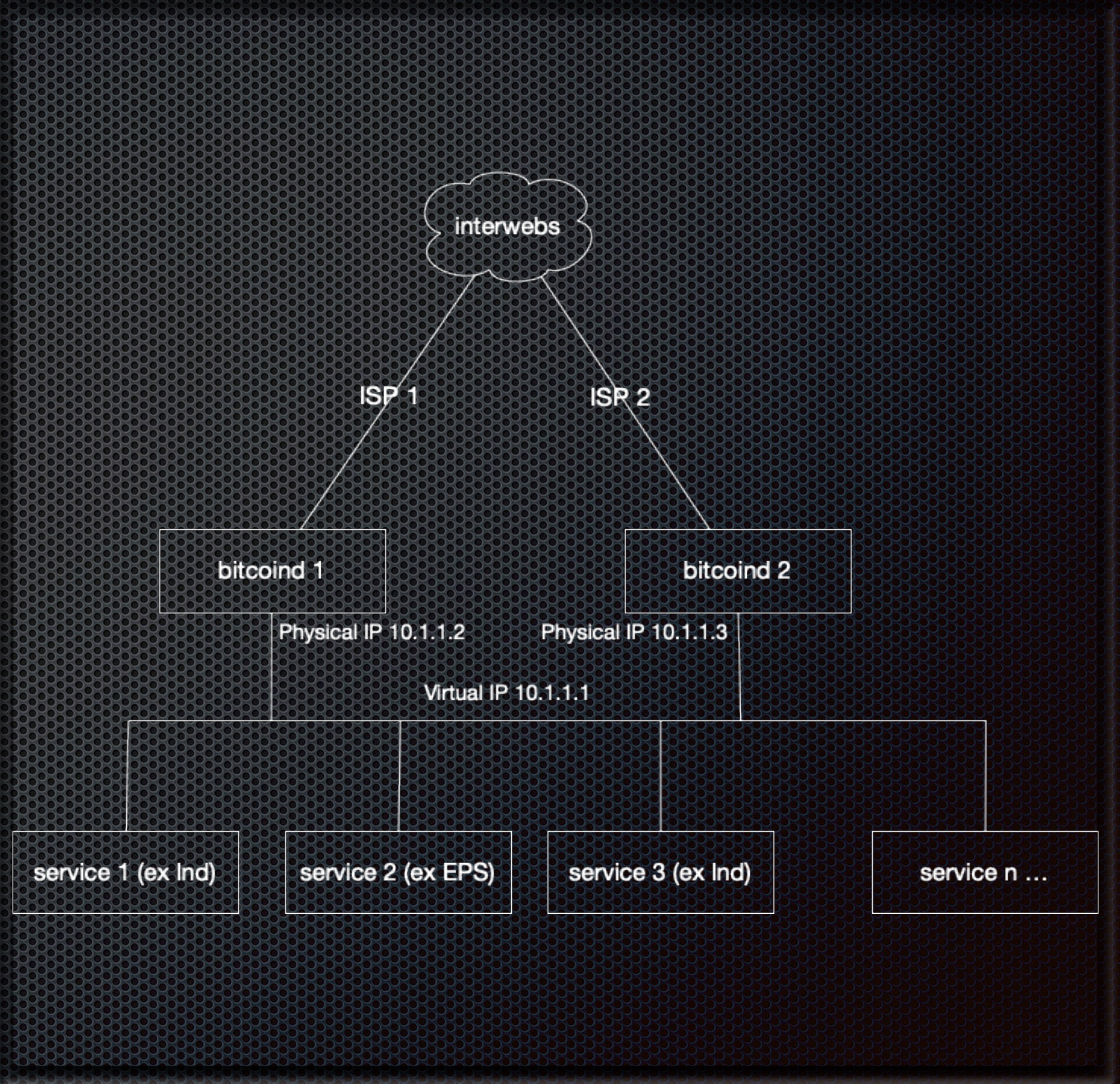
- BTCPay Server can’t be multitenant for lightning, but BTCPay Server can*
- Running one btcpayserver.dll, multiple LNDs (one per merchant/store)

*see what I did here? Please disambiguate the name!

Let’s add some redundancy

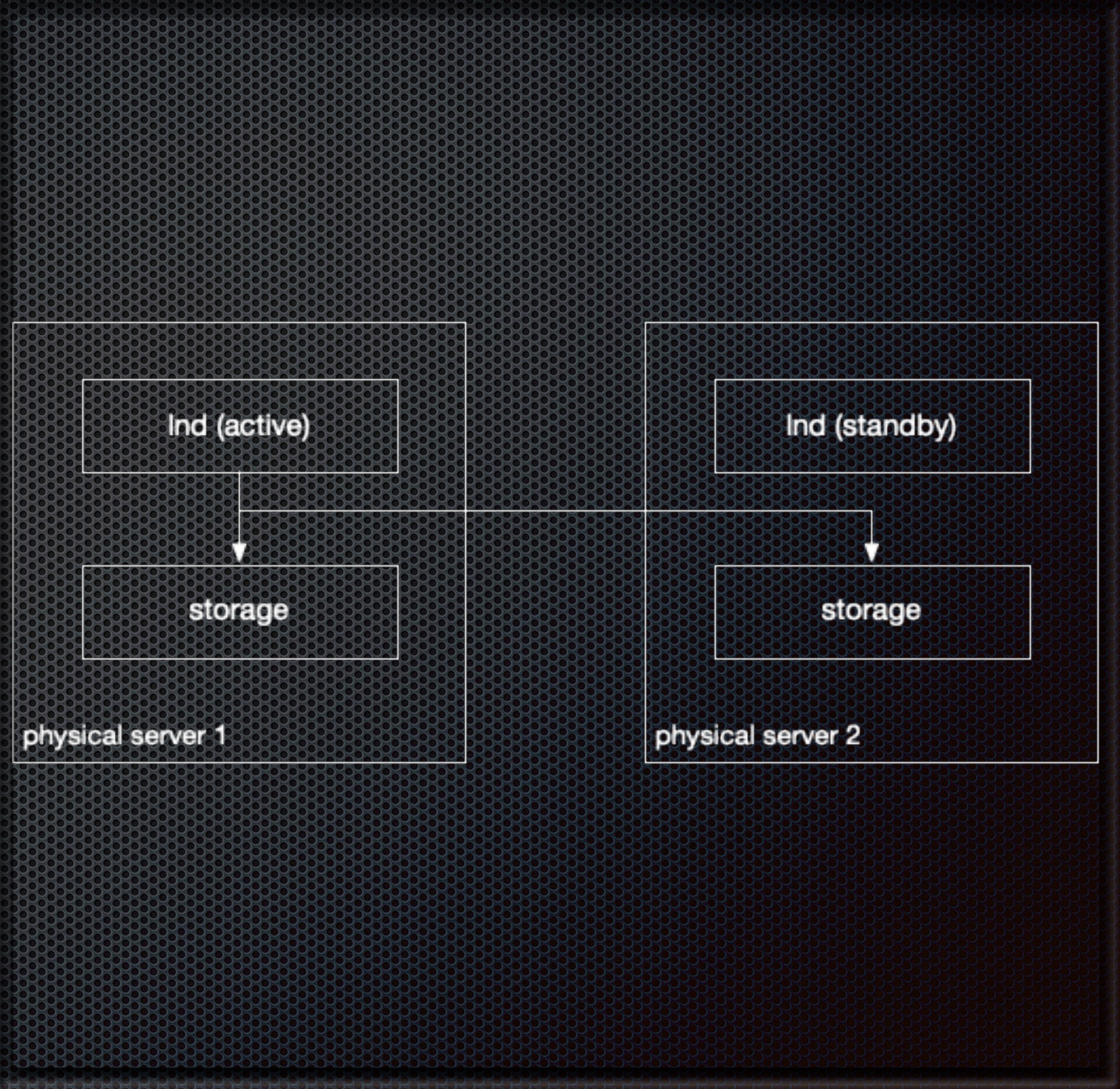
bitcoind

- run 2 (or more)
- expose RPC and ZMQ over shared VIP (Virtual IP)
- run them on separate public networks (AS) to make attacks (DDoS) harder



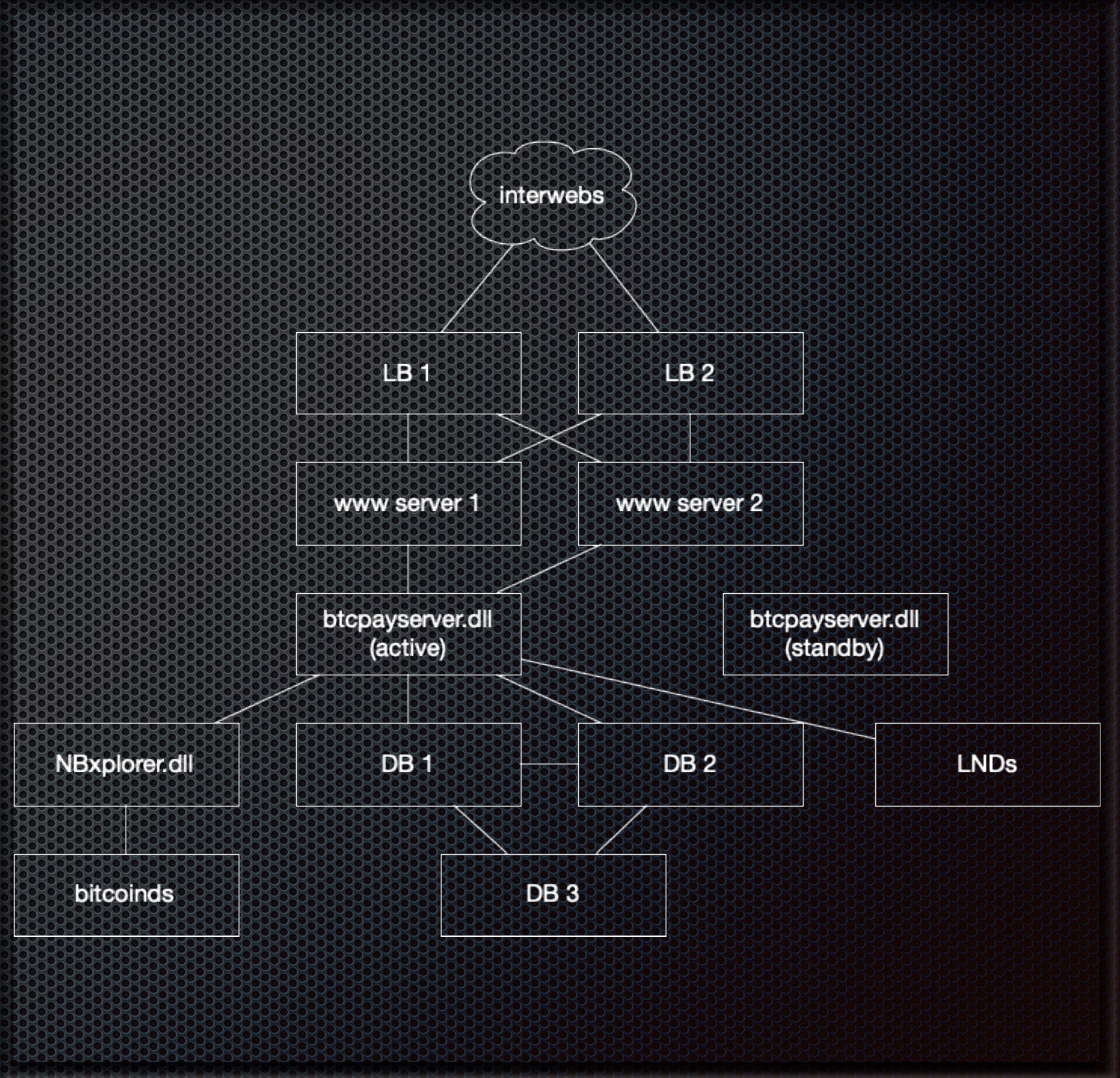
LND

- backup / restore -> not ideal (closing channels)
- store .lnd on a distributed storage (ceph, glusterFS, DRBD, …)
- hope that the data will not be corrupted by crash of active instance
- restart another LND from same storage



BTCPay Server*

- Classic multi-tier web application
  - load balancer(s) (active/backup)
  - web server(s) (active/active)
  - app server(s) (active/backup)
  - middlware server(s) (TBD)
  - database server (cluster)



*please disambiguate the name!

Minimal setup (also for SMB)

- BGP Router
- Switch
- nodl “rack dual”
- two x86 based servers in one rack

- Runs 2 bitcoind’s, up to 10 LNDs, BTCPay Server, a Galera database cluster, and all the little stuff around

(actual picture of the nodl cloud infrastructure)

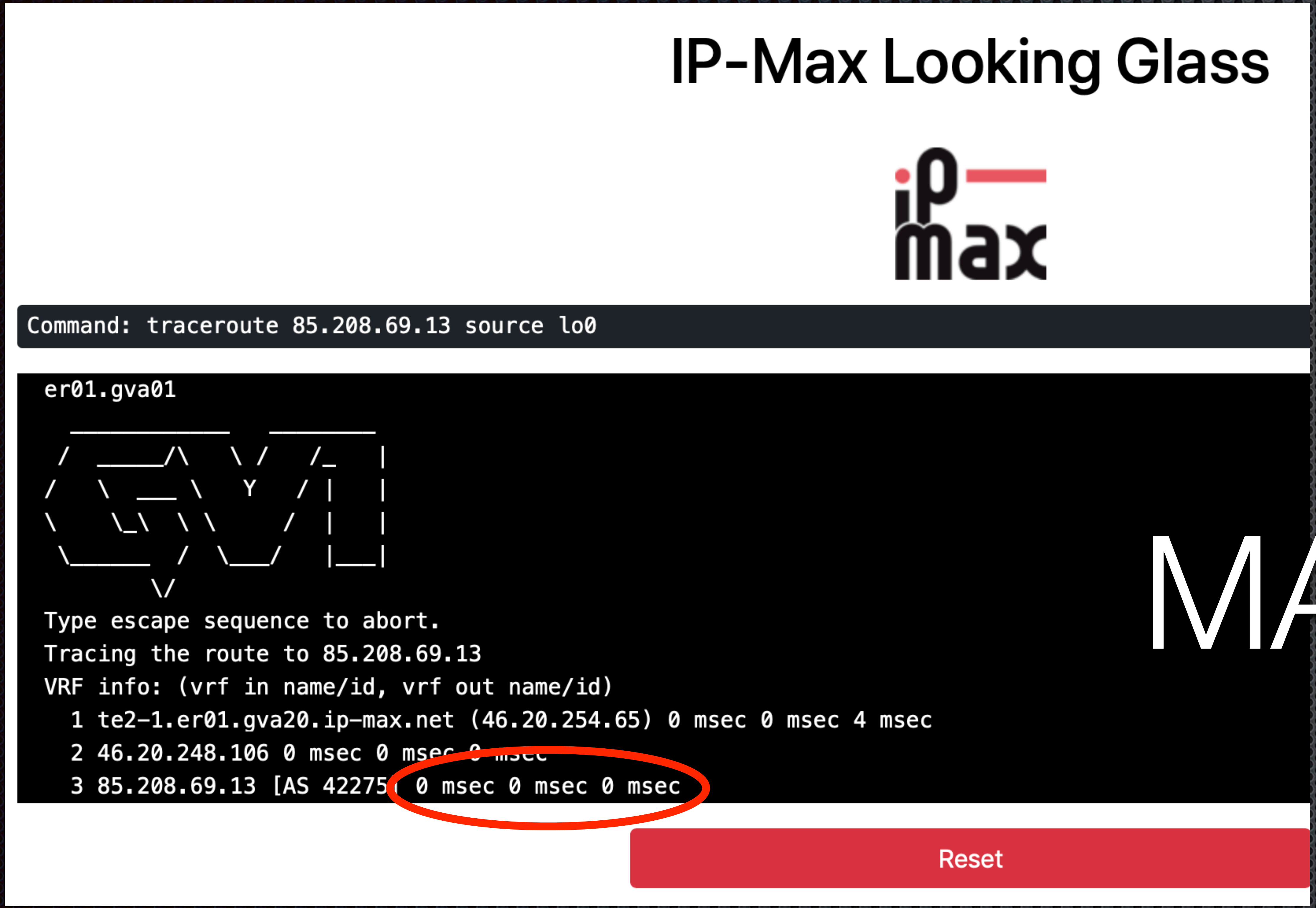
Go big or go home

- multi 10Gbps fabric
- lots of cores
- lots or RAM
- lots of SSD
- scales up to thousands of lightning nodes

(actual picture of the nodl cloud infrastructure)

So you’re running your own servers.
Cool story bro

The curious case of 85.208.69.13

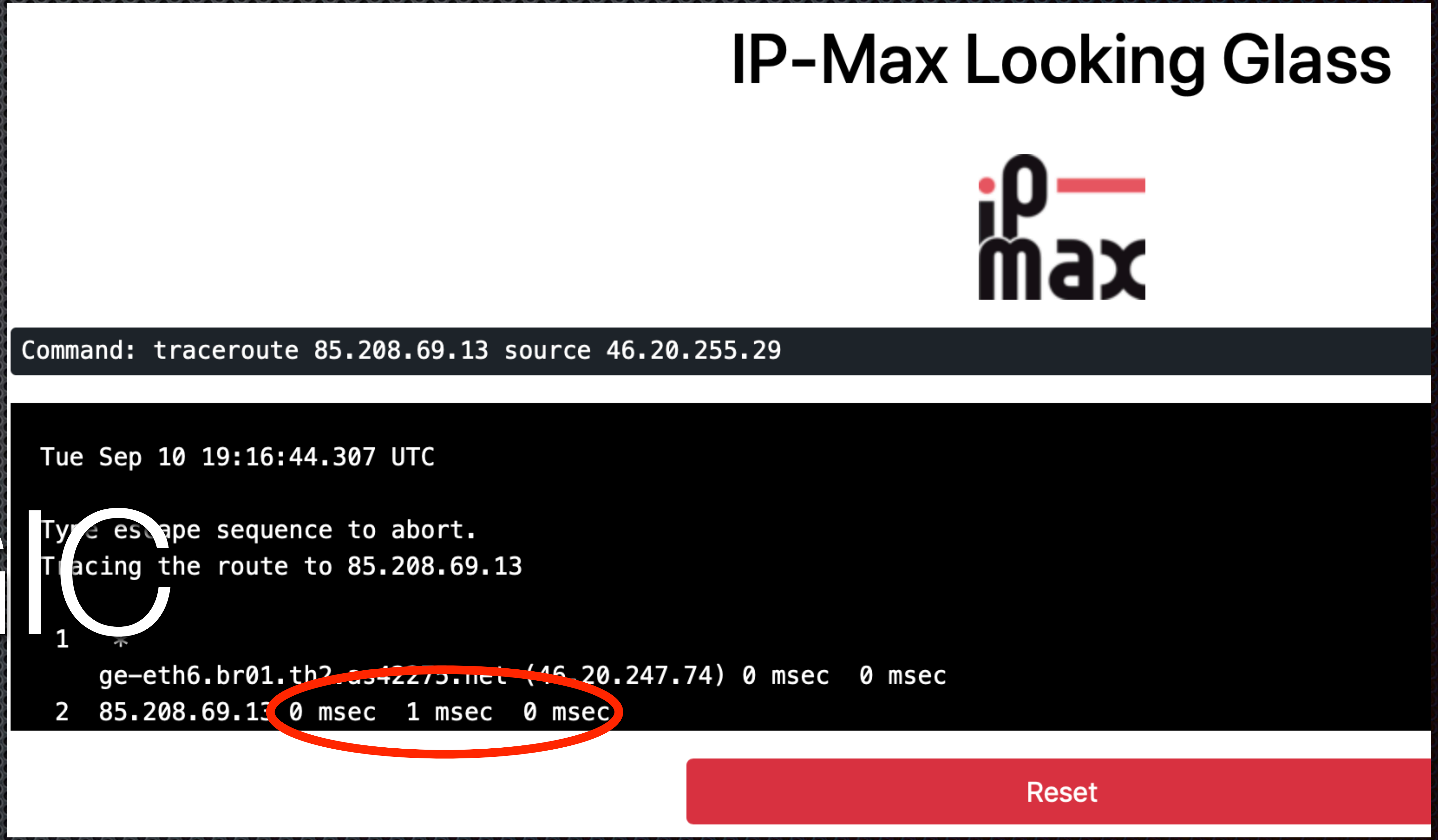


IP-Max Looking Glass

```
Command: traceroute 85.208.69.13 source lo0
```

```
er01.gva01
  _____________   _________
 /   _____/\   \ /   /_   |
/   \  ___ \   Y   / |   |
\    \_\  \ \     /  |   |
 \______  /  \___/   |___|
        \/
Type escape sequence to abort.
Tracing the route to 85.208.69.13
VRF info: (vrf in name/id, vrf out name/id)
  1 te2-1.er01.gva20.ip-max.net (46.20.254.65) 0 msec 0 msec 4 msec
  2 46.20.248.106 0 msec 0 msec 0 msec
  3 85.208.69.13 [AS 42275] 0 msec 0 msec 0 msec
```

MAGIC



IP-Max Looking Glass

```
Command: traceroute 85.208.69.13 source 46.20.255.29
```

```
Tue Sep 10 19:16:44.307 UTC

Type escape sequence to abort.
Tracing the route to 85.208.69.13

 1  *
    ge-eth6.br01.th2.as42275.net (46.20.247.74) 0 msec  0 msec
 2  85.208.69.13 0 msec  1 msec  0 msec
```

traceroute from Geneva

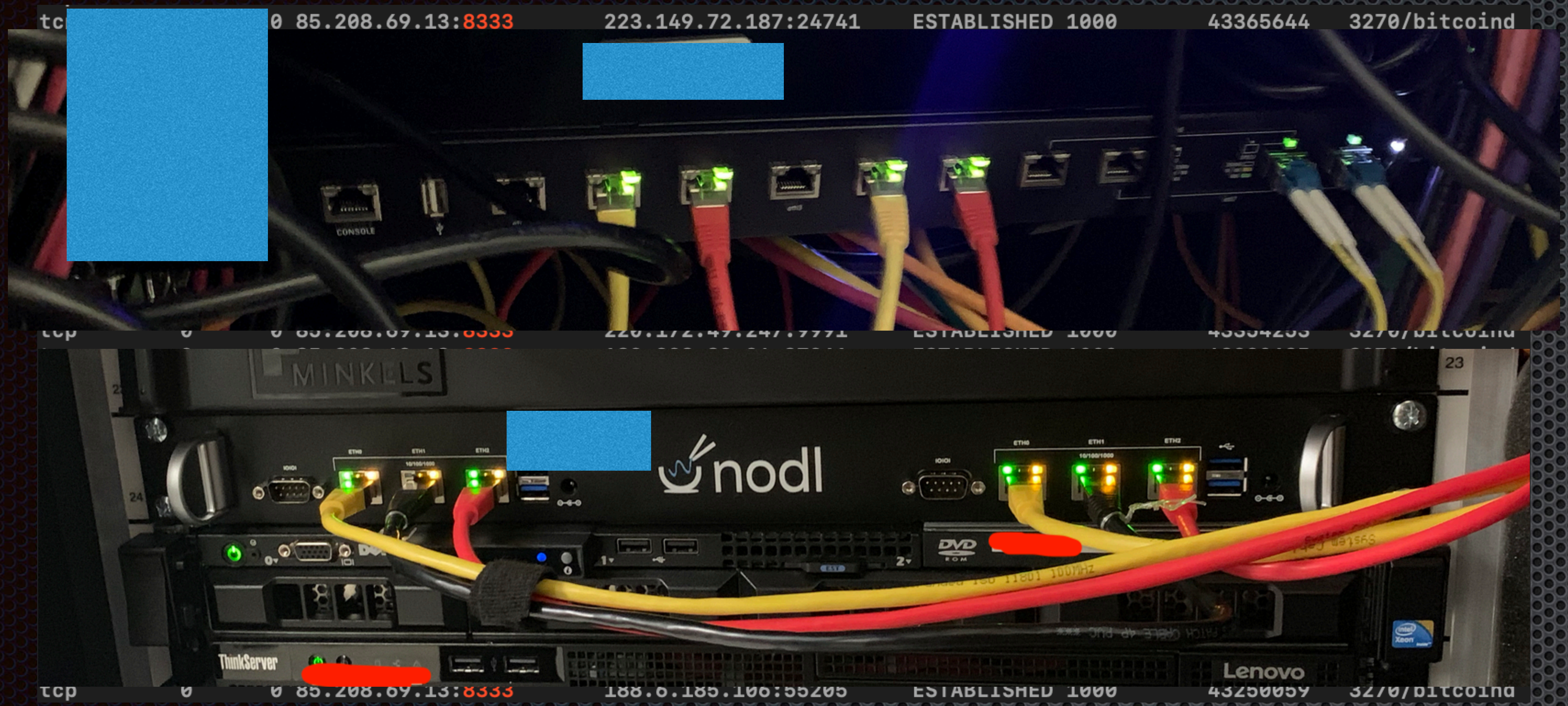
traceroute from Paris

Average latency from Geneva to Paris = ~9ms

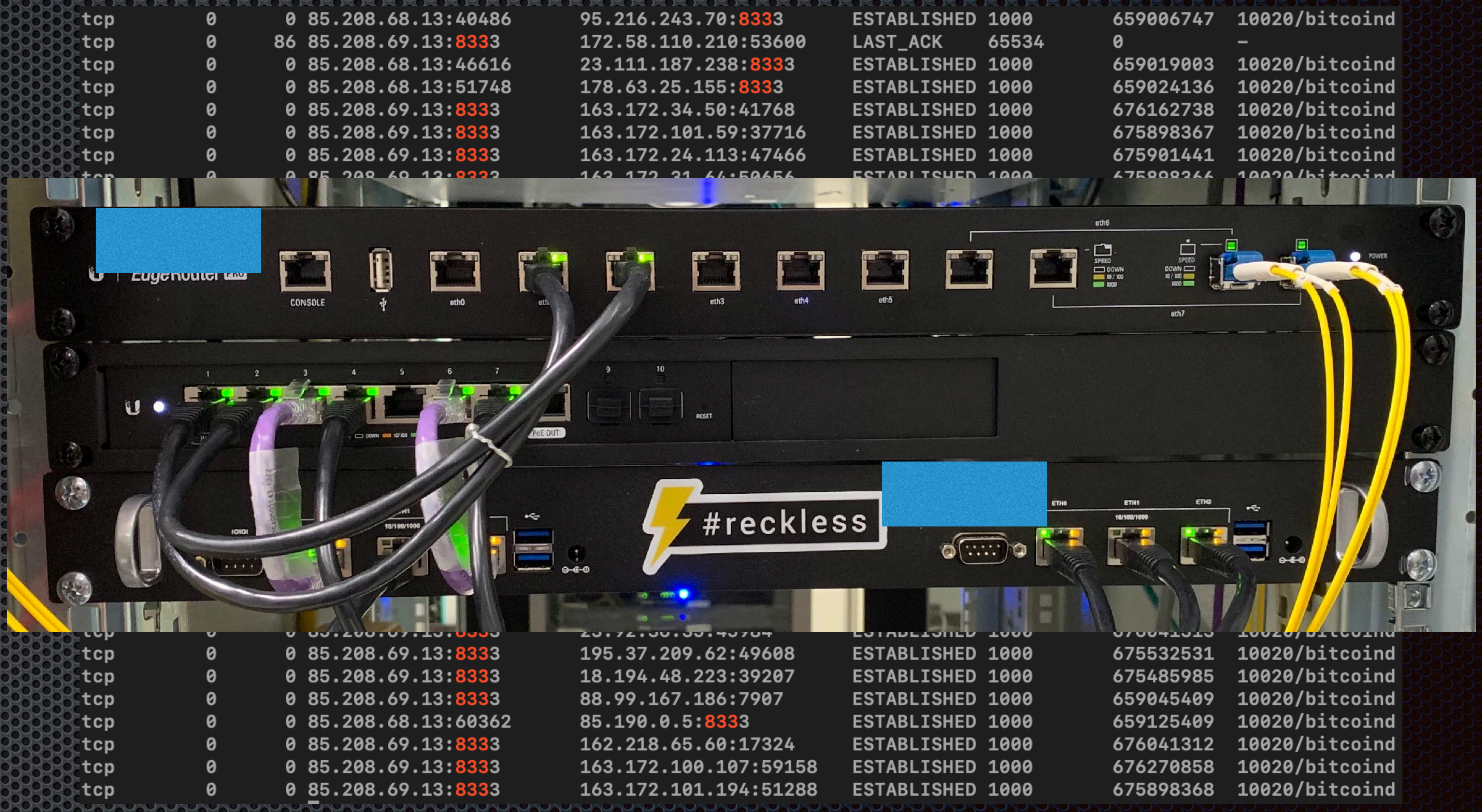
Enter anycast

- 85.208.69.0/24 -> AS42275's anycast range (for now)
- Same IP range announced from two locations: Geneva and Paris
- Other nodes connect indifferently to any of the nodes (usually the closest, from a network point of view)
- A node is announced (BGP) only if it's up and running (otherwise traffic goes to the next closest node)
- Actually, there may be multiple bitcoinds running behind this IP in each datacenter
- The nodes are inter-connected through private links (not Internet)
- The 8.8.8.8 (or 9.9.9.9, or 1.1.1.1) of bitcoin (except the cool IP = $$)

PoC||GTFO



```
tc[redacted]0 85.208.69.13:8333       223.149.72.187:24741     ESTABLISHED 1000       43365644   3270/bitcoind
tcp        0      0 85.208.69.13:8333       220.172.49.247:9991      ESTABLISHED 1000       43354253   3270/bitcoind
tcp        0      0 85.208.69.13:8333       188.6.185.106:55205      ESTABLISHED 1000       43250059   3270/bitcoind
```



```
tcp        0      0 85.208.68.13:40486      95.216.243.70:8333       ESTABLISHED 1000       659006747  10020/bitcoind
tcp        0     86 85.208.69.13:8333       172.58.110.210:53600     LAST_ACK    65534      0          -
tcp        0      0 85.208.68.13:46616      23.111.187.238:8333      ESTABLISHED 1000       659019003  10020/bitcoind
tcp        0      0 85.208.68.13:51748      178.63.25.155:8333       ESTABLISHED 1000       659024136  10020/bitcoind
tcp        0      0 85.208.69.13:8333       163.172.34.50:41768      ESTABLISHED 1000       676162738  10020/bitcoind
tcp        0      0 85.208.69.13:8333       163.172.101.59:37716     ESTABLISHED 1000       675898367  10020/bitcoind
tcp        0      0 85.208.69.13:8333       163.172.24.113:47466     ESTABLISHED 1000       675901441  10020/bitcoind
tcp        0      0 85.208.69.13:8333       163.172.31.64:50656      ESTABLISHED 1000       675898366  10020/bitcoind
tcp        0      0 85.208.69.13:8333       23.92.36.33:45984        ESTABLISHED 1000       676041313  10020/bitcoind
tcp        0      0 85.208.69.13:8333       195.37.209.62:49608      ESTABLISHED 1000       675532531  10020/bitcoind
tcp        0      0 85.208.69.13:8333       18.194.48.223:39207      ESTABLISHED 1000       675485985  10020/bitcoind
tcp        0      0 85.208.69.13:8333       88.99.167.186:7907       ESTABLISHED 1000       659045409  10020/bitcoind
tcp        0      0 85.208.68.13:60362      85.190.0.5:8333          ESTABLISHED 1000       659125409  10020/bitcoind
tcp        0      0 85.208.69.13:8333       162.218.65.60:17324      ESTABLISHED 1000       676041312  10020/bitcoind
tcp        0      0 85.208.69.13:8333       163.172.100.107:59158    ESTABLISHED 1000       676270858  10020/bitcoind
tcp        0      0 85.208.69.13:8333       163.172.101.194:51288    ESTABLISHED 1000       675898368  10020/bitcoind
```

Connections to/from other nodes in GVA and PAR

Outgoing uses physical range (85.208.70/24 for GVA, 85.208.68/24 for PAR)

Future expansion

- Anycast requires as many direct peerings with other ISPs as possible
- We have SwissIX and FranceIX
- 2019Q4 - Adding Frankfurt (and DE-CIX - biggest exchange in the world)
  - thus covering 50+% of global ISPs and making a full triangle
- Later adding Moscow and NYC (or SF) for better latency and resiliency

Big thank you

Fred and IP-Max for sponsoring space and connectivity for this POC

Q ?

PS:

1/ Digging (way) deeper at BTCPay Server day

2/ nodl meetup on Tuesday
https://www.meetup.com/B-Markets/events/264792663/