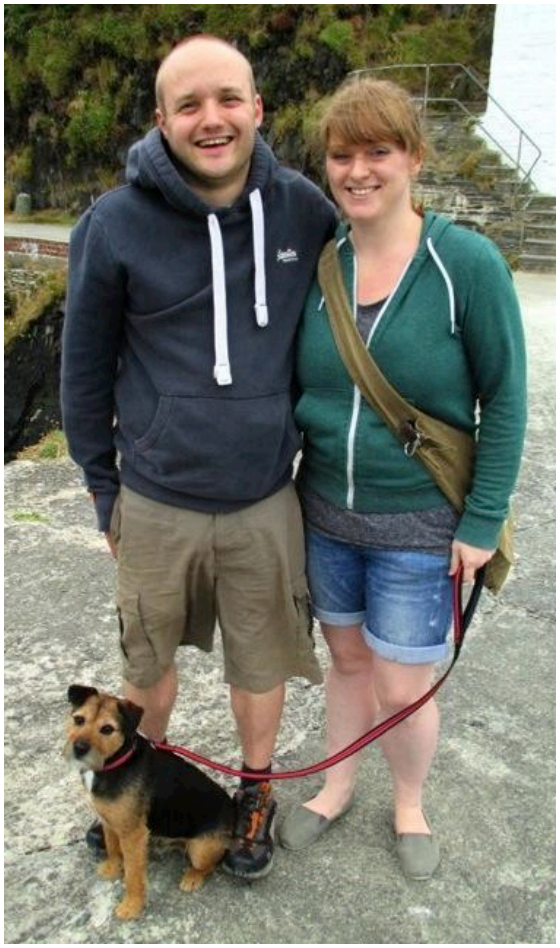
Anycast on a shoestring

Nat Morris - @natmorris

Tuesday 4th November 2014 - RIPE69, London

About me

- Consultant, **Esgob Ltd**
- Team leader, **Cumulus**
- Board member, **UKNOF**
- Based in West Wales
- Aspiring lighthouse keeper



Esgob ltd

cumulus®

UKNOF

Anycast 101

- Simple concept:
  - Announce the same address space from multiple locations
  - Multiple paths in BGP, best one selected based on policy
- Benefits:
  - Increased reliability
  - Load balancing
  - Improved performance
  - Localized impact of DoS attacks

Summary

- Fancied deploying a DNS Anycast service
- Motivated by:
  - Bill Woodcock, PCH
    - SANOG8: “Best Practices in DNS Anycast Service-Provision”
  - Dave Knight, ICANN (now Dyn)
    - RIPE64: “Dense Anycast Deployment of DNS Authority Servers”
- Gain more experience automating distributed environments.

Thoughts

- Can I do it without spending too much?

- What to offer?
  - Initially only secondary DNS
  - IPv4 + IPv6
  - Free service - no SLA, no Revenue

- Does Claire need to find out?

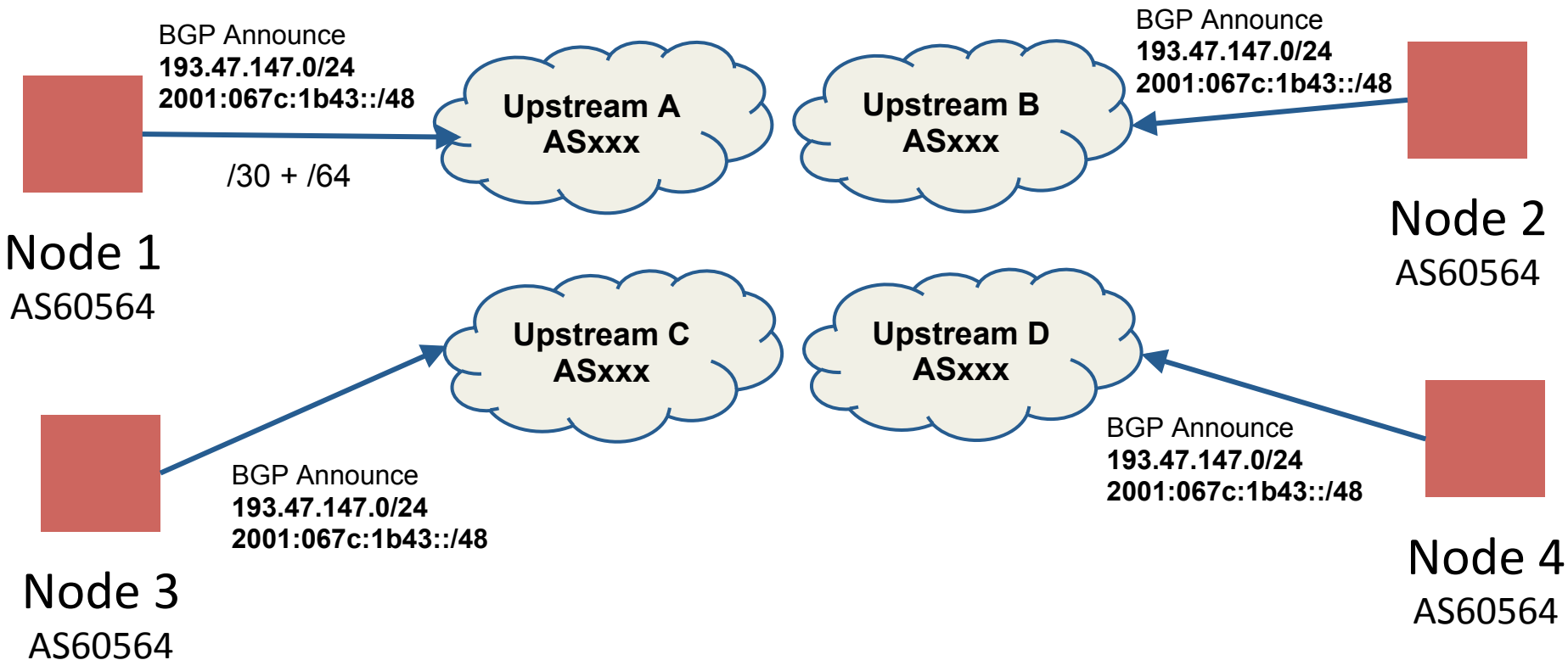
  Must be sub $1000/yr running cost

Requirements

- Separation from existing management network - AS30746
- Had spare PI /24 + /48
- Applied for new ASN - got AS60564
- Need highly automated framework
- Wanted to play with new tools
- Open source everything
  - https://github.com/esgob

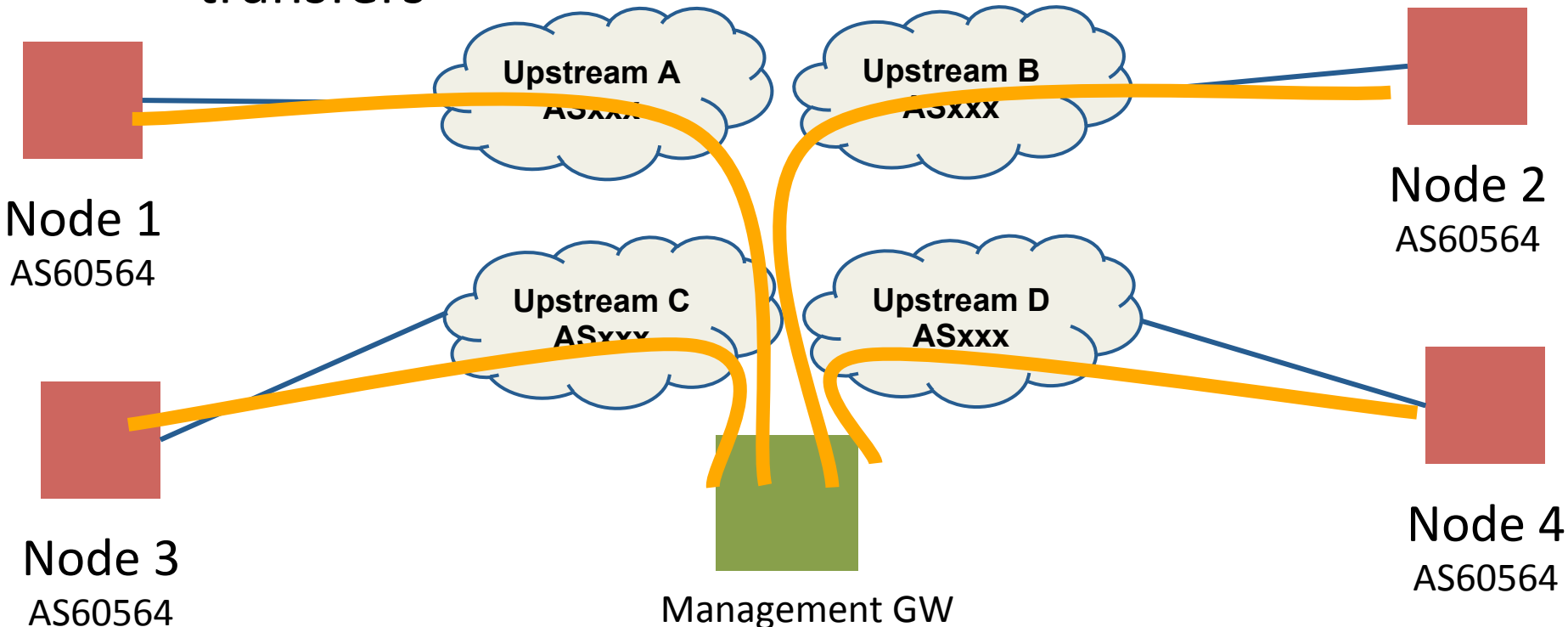
Network

- Each anycast node:
  - Announce /24 + /48 via BGP
  - Static default route to the provider (accept no routes)

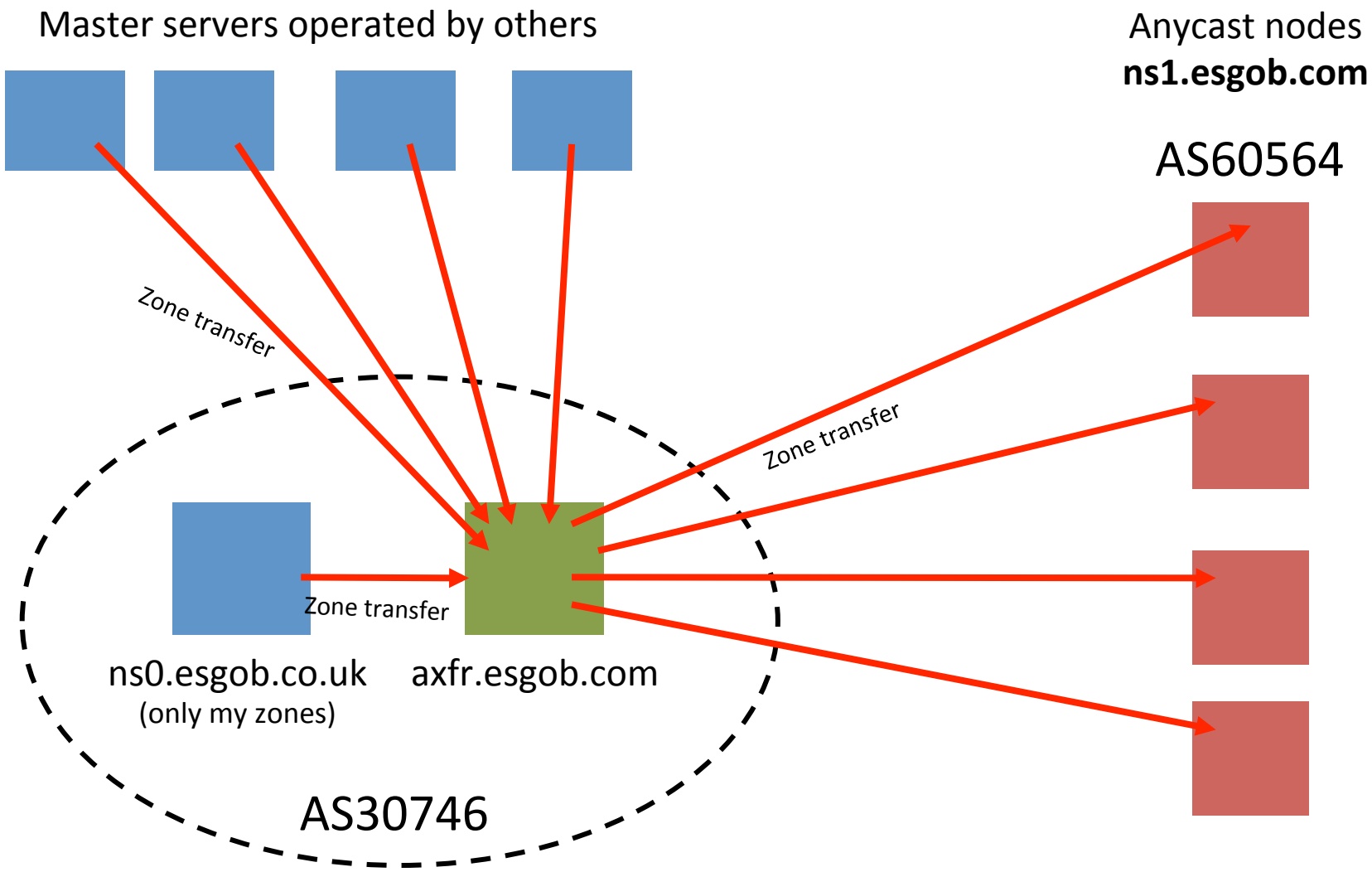


Network

- All management traffic inside OpenVPN
  - Additional RFC1918 loopback per node
  - No need for TSIG from AXFR to anycast nodes
  - Protects beanstalk traffic, API, collectd, zone transfers



DNS zone transfers



Supports notifies from masters

Store everything in JSON

- Data stored in **RethinkDB**
  - Distributed JSON document database
  - Free form, easy to add attributes

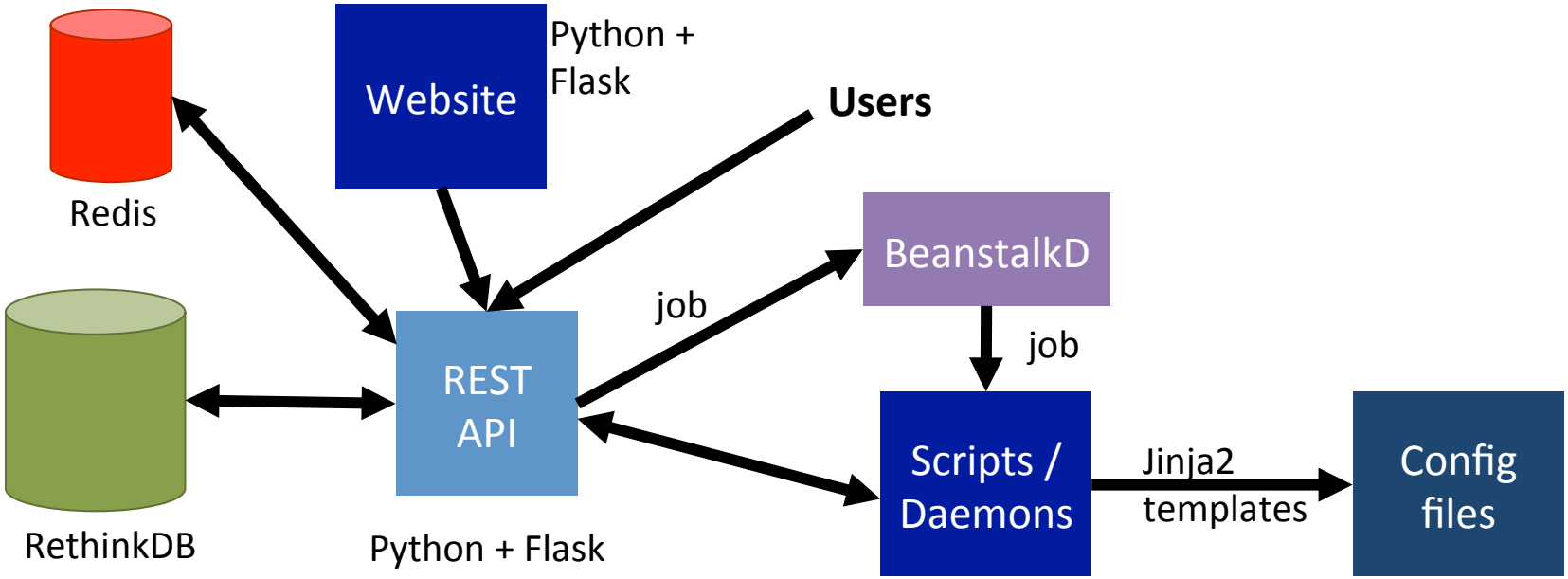

```
{
    "account": "esgob" ,
    "added": 1391807801 ,
    "id": "esgob.co.uk" ,
    "masterip": "195.177.253.166" ,
    "type": "slave"
} ,
{
    "account": "mike" ,
    "added": 1392340310 ,
    "id": "grepular.com" ,
    "masterip": "2001:470:1f15:d7::2" ,
    "type": "slave"
} ,
```



```
{
    "city": "Edinburgh" ,
    "country": "Scotland" ,
    "countryiso": "gb" ,
    "flag": "scotland" ,
    "fqdn": "alban.esgob.com" ,
    "geolat": 55.929959 ,
    "geolng": -3.294011 ,
    "hostdesc": "VPS at Fluency" ,
    "id": "alban" ,
    "livedisplay": "25th Oct 2013" ,
    "locationdisplay": "Scotland, Edinburgh" ,
    "mgmtip4": "46.226.2.182" ,
    "mgmtip6": "2a00:a600:0:6::2" ,
    "peers4": [
        {
            "asn": 56595 ,
            "communities": [
                "56595:9500" ,
                "56595:9510" ,
                "56595:9551" ,
                "60564:10003" ,
                "60564:20044" ,
                "60564:30001"
            ] ,
            "name": "fluency" ,
            "neighbor": [
                {
                    "ip": "46.226.2.181"
                }
            ]
        }
```

Application design

- REST API - Python + Flask framework
- cronjob + rsync free zone
- Beanstalk - Message queue
  - Add/change/removal of zones via API triggers job to be added to queue
  - Daemons on AXFR + Anycast nodes wait for jobs
  - New zones live on all nodes with 5 seconds



Where to host the anycast nodes?

- Can't cost too much each month
  - Ruled out renting co-lo / dedicated servers
- Need a wide reach
- Virtual machines
  - Swap with friends?
- Low cost hardware

Mythic Beasts

- UK clueful hosting company
- £7/month = £84/yr = $141/yr
  - 512mb RAM, 1 CPU, 10gb HDD, 75gb BW
- Bought a VM, opened a support ticket
  - BGP not listed on the website.
  - Me: "I have a /24 + /48 of PI, can you set me up a BGP session? see AS60564 + AS-ESGOB-ANYCAST"
  - Them: "Peer with these IPs, sessions are ready!"
- Too easy!

mythic beasts

Quest to find more friendly VM hosts

- LowEndBox.com
  - Blog / adverts for VM providers, all < $10 month
  - Usually OpenVZ based
  - Need KVM/XEN/VMWare to support Quagga/BIRD
- LowEndTalk.com
  - Message board, various small scale VM hosts
- Google
  - “vps bgp session”

Anynode in Detroit

anyNode <support@anynode.net> 06/10/2013

to Nat

anynode.net

Dear Nat Morris (ESGOB LTD),

We have received your order and will be processing it shortly. The details of the order are below:

Order Number: **6758645372**

Product/Service: Detroit KVM - DK-512
Domain: esgob
First Payment Amount: $80.00 USD
Recurring Amount: $0.00 USD
Billing Cycle: Annually

Total Due Today: $80.00 USD

- Leaps
  - Nev
- $120/

Leapswitch | Billing <billing@leapswitch.com> 30/09/2013

to Nat

Dear Nat Morris (Esgob Ltd),

This is a notice that an invoice has been generated on 29/09/2013.

Your payment method is: PayPal

Invoice #44717
Amount Due: $48.00
Due Date: 29/09/2013

Invoice Items

KVM-UVPS-1 - cyri.esgob.com (29/09/2013 - 28/09/2014)
Control Panel: NONE(Default)
FTP Backup: No Backup
IP Address: 0 x IP Addresses (In addition to default 1 IP) $20.00
cPanel Addon: None
Monitoring: No Monitoring
~~Operating System: Debian 7.0.0 64bit $120.00~~
Promotional Code: INKVMVPS512MB48 - $72.00 Recurring Discount $-72.00

Sub Total: $48.00
Credit: $0.00
Total: $48.00

You can login to your client area to view and pay the invoice at https://service.leapswitch

Growing - Oct 2013

- Up to 4 VMs
  - London A (on existing KVM server)
  - London B
  - Detroit
  - India
- Spinning up VM taking about 10 minutes
  - Add JSON entry into RethinkDB
  - Install Debian
  - Install Puppet (Collectd, BIND, Quagga, custom daemons etc)
  - Done
  - Pub!

Friends - Oct 2013

- Offers to host VMs for free from:
  - Edinburgh @ Fluency
  - Boston @ TorwardEx
- One swap:
  - Bremen, Germany @ Fremaks
- Quickly up to 7 nodes after 1 month

FLUENCY

TOWARDEX® Carrier Services

Fremaks
Providing, Hosting & Software

#		Location	State	Protocols	In service
1		England, London, A	Live	IPv4 IPv6	24th Sept 2013
2		England, London, B	Live	IPv4 IPv6	2nd Oct 2013
3		United States, Detroit	Live	IPv4 IPv6	9th Oct 2013
4		United States, Boston	Live	IPv4 IPv6	13th Oct 2013
5		India, Pune	Live	IPv4 IPv6	14th Oct 2013
6		Germany, Bremen	Live	IPv4	23rd Oct 2013
7		Scotland, Edinburgh	Live	IPv4 IPv6	25th Oct 2013

Friends

- "We can host something, but not a VM"
- Various issues:
  - Network engineers didn't have access to VM hosts
  - No VM infrastructure at all
  - All VMs routed by hypervisor, can't bridge to BGP routers

- Solution…
  - Raspberry PI, $35
  - 512mb / 16gb SD Card
  - 800mhz (overclocked)

Raspberry Pi

- Sent RPI to Belfast – David Farrell @ Tibus
- Ran NSD3 great, ~200qps
- Adding / removing zones with NSD3 required service restart ☹
- Swapped to PowerDNS, testing went ok
- pdns_control segfault'ing on ARM
- Gave in and moved to BIND

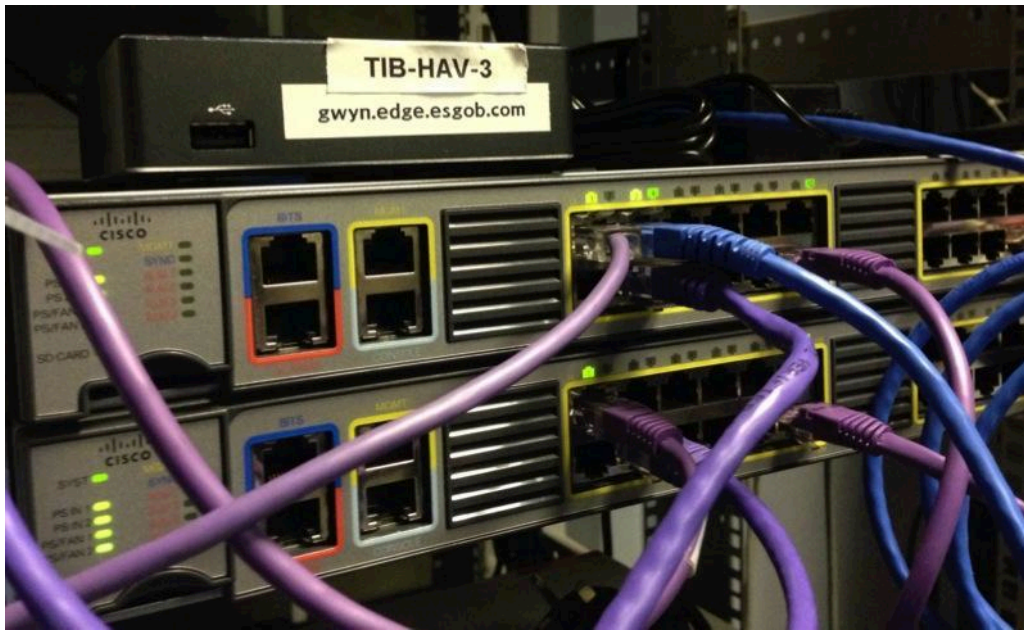
Time to peer - Jan 2014

- Offer of a node at SFMIX from Matt Peterson
- San Francisco based IXP
- Couldn't turn this down.
- Zero U install
- Needed 2 NICs, IX + OOB
  - FitPC2i - perfect ☺
  - Atom 1.6ghz, 1gb RAM, 16gb SSD
  - $150 on eBay
- Peering with:
  - HE.net, ISC, Unwired, PCH
  - Layer42, Lookout, DigitalOcean

SFMIX

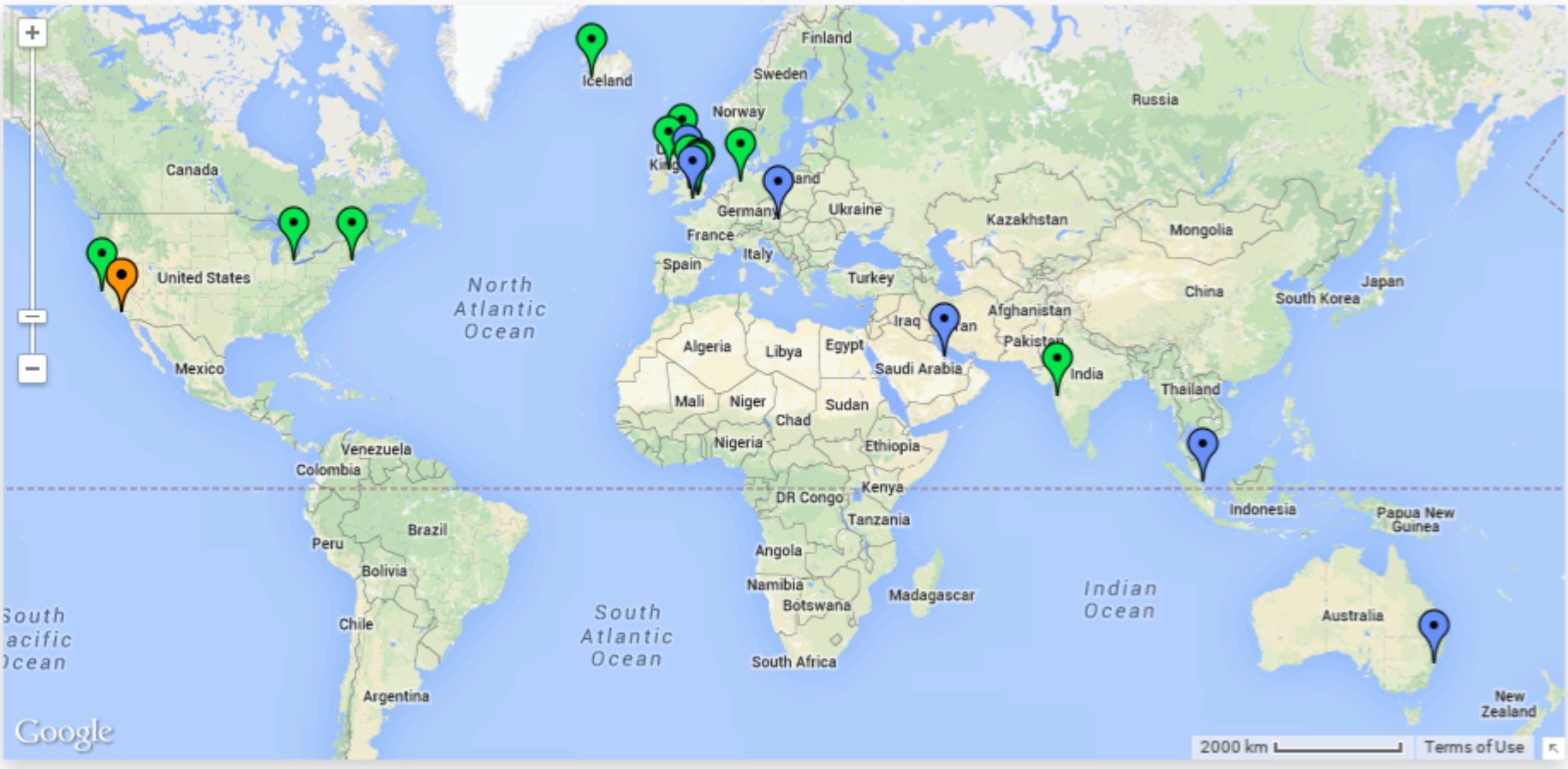
Pi problems

- Puppet slow + loading zones taking too long
- More offers to host h/w nodes
  - Manchester, Andy Davidson @ Allegro
- Gigabyte BRIX
  - Quad Core, 1.8ghz, 4gb RAM, 30gb MSATA
- Swap out Belfast Pi



Still growing today

- 12 nodes live, 6 in build

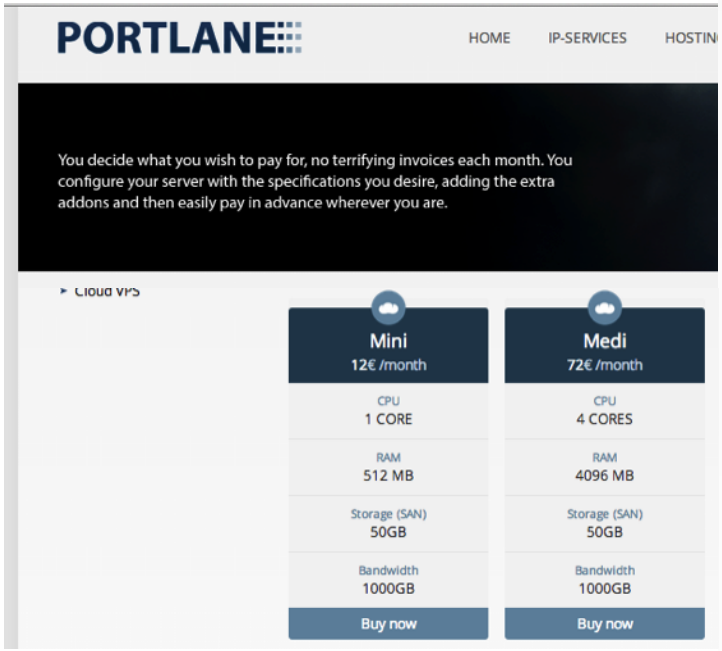


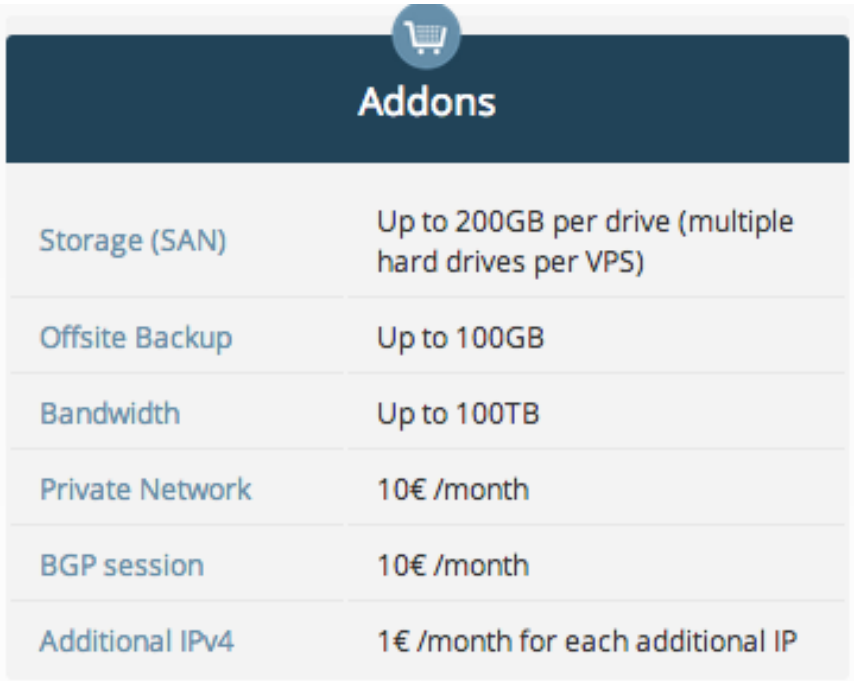
As of 3rd Nov 2014

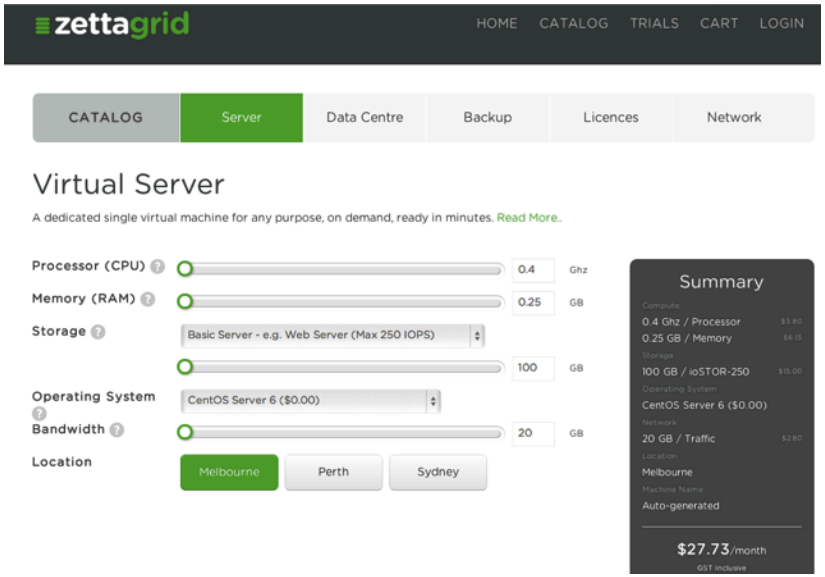
Fun along the way

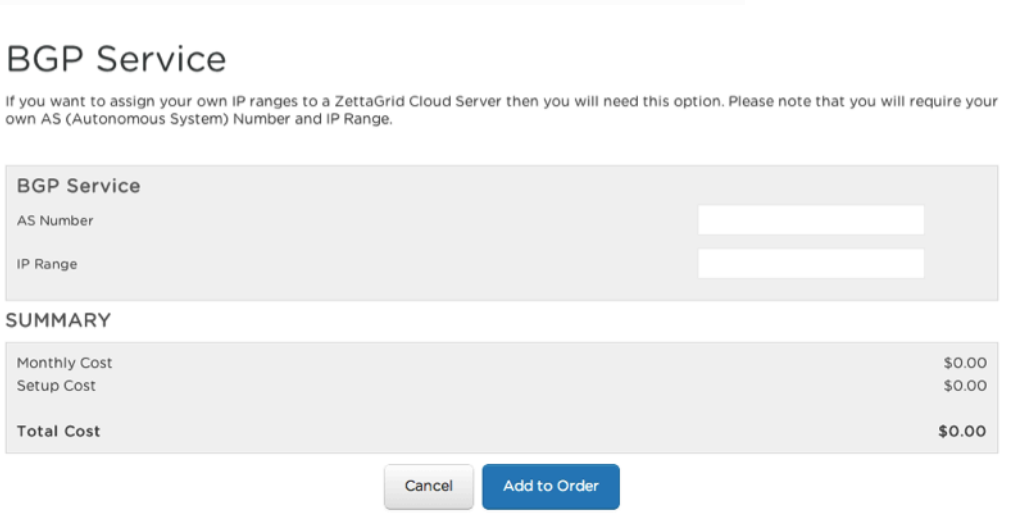
- First BGP customer for some of the VM hosts:
  - Assisted educating them…
  - Prefer transit routes over customers, eek
  - No BGP filters / route-maps or prefix lists - Ahhhhh!
- Some hosts don't have communities
  - Helping people get those implemented
- RIPE Atlas
  - Scheduled measurements to look at latency + CHAOS id.server
- RIPEstat
  - Using the REST API to query visibility of a /48 anchor from each instance, 15mins to get working.

Market for VMs with BGP sessions









Discoveries

- Easy to find budget anycasters
  - Webhost forums / LowEndTalk / bgp.he.net
- "Interesting" deployment methods
  - **Shared /24**
    - Dedicated IP, custom fwd/rev DNS, slaved zones
  - **Hosted /24**
    - Same as above but customer provides /24 to announce
  - **Shared /24, /32 tunneled**
    - Dedicated IP, tunneled by anycaster to customer via GRE
  - **Single /24**
    - Anycaster only has one /24, website + mail + mgmt in same space. DNS answered at edge, other IPs tunneled to another VM/dedicated box.

Discoveries

- Not all budget DNS hosting companies host every zone at the edge
  - Some host zones centrally and cache at the edge
- Some tunnel all DNS traffic back to a location:
  - Looks like anycast
  - Poor DNS performance
  - GRE from edge to a single auth server
  - Selling /32 GRE tunnels to customers!

Not just DNS?

- Simple HTTP/HTTPS service to assist debugging
  - local.esgob.com
  - local4.esgob.com / local6.esgob.com
- Returns JSON via lighttpd on each node

```
nat@enw:~ $ curl http://local.esgob.com
{
        "city": "London",
        "country": "England",
        "countryiso": "gb",
        "flag": "england",
        "locationdisplay": "England, London,
B",
    "ref": "ql7f823b"
}
```

What next?

- Keep honest – host every zone at the edge
- Finish web interface
- Make the Github repos consumable:
  - –Produce install instructions
  - –Getting started guide
- Support: Multi master, TSIG inbound
- Mix of routing and DNS daemons
  - configurable per node via JSON in RethinkDB
  - BIND, NSD4, KNOT, Quagga, BIRD + ExaBGP
- Looking for friendly hosts:
  - Europe, Africa, India, Asia, South America, anywhere!

As of today...

#	Location	State	Protocols	In service	Notes
1	Los Angeles, US	Maintenance	IPv4 IPv6	18th Jun 2014	Hardware issue being investigated
2	England, London, A	Live	IPv4 IPv6	24th Sept 2013	
3	England, London, B	Live	IPv4 IPv6	2nd Oct 2013	
4	United States, Detroit	Live	IPv4 IPv6	9th Oct 2013	
5	United States, Boston	Live	IPv4 IPv6	13th Oct 2013	
6	India, Pune	Live	IPv4 IPv6	14th Oct 2013	
7	Germany, Bremen	Live	IPv4	23rd Oct 2013	
8	Scotland, Edinburgh	Live Serving you!	IPv4 IPv6	25th Oct 2013	
9	Northern Ireland, Belfast	Live	IPv4	28th Oct 2013	
10	England, London, C	Live	IPv4 IPv6	9th Jan 2014	
11	United States, San Francisco	Live	IPv4 IPv6	12th Jan 2014	
12	England, Studley	Live	IPv4 IPv6	25th May 2014	
13	England, Hemel Hempstead	Live	IPv4	26th Sep 2014	
14	Reykjavik, Iceland	Live	IPv4	16th Oct 2014	
15	England, Manchester	In Progress	IPv4		Partner to fix BGP filters
16	Manama, Kingdom of Bahrain	In Progress	IPv4 IPv6		Hardware to be delivered
17	Vienna, Austria	In Progress	IPv4 IPv6		Commissioning underway
18	Singapore	In Progress	IPv4 IPv6		Hardware to be delivered
19	Sydney, Australia	In Progress	IPv4 IPv6		Commissioning underway
20	England, Fareham	In Progress	IPv4 IPv6		Commissioning underway

Currently costing just under $480/yr - Claire compliant!

Take aways

- Low barrier to entry - VM with BGP full table $40/yr
- Don’t become part of the problem
  - –Ensure your customers aren’t hijacking prefixes
  - –Always use IRR prefix lists on cust BGP sessions
- Using anycast can improve service delivery to your customers
- Automate all things!
- Have fun and share your experiences

Questions?

https://noc.esgob.com

@esgobltd

https://nat.ms

@natmorris