About me

• Consultant, Esgob Ltd
• Team Lead, Cumulus
• Board member, UKNOF
• Based in West Wales
• Aspiring lighthouse keeper
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)
• All management traffic inside OpenVPN
  – Additional RFC1918 loopback per node
  – No need for TSIG from AXFR to anycast nodes
  – Protects beanstalk traffic, collectd, zone transfers
DNS zone transfers

Master servers operated by others

AS30746
ns0.esgob.co.uk
(axfr.esgob.com
(only my zones)

AS60564

Anycast nodes
ns1.esgob.com

Supports notifies from masters
Store everything in JSON

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

```json
```
Application design

- REST API - Python + Flask framework
  - Add/change/removal of zones via API triggers job to be added to queues for each node
  - Daemons on AXFR + Anycast nodes monitor for jobs
  - New zones live on all nodes with 5 seconds
  
  ![Diagram of application design]

RethinkDB

Website

Python + Flask

REST API

Python + Flask

BeanstalkD

Scripts

Jinja2 templates

Config files
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
• 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?”
  – Them: “Peer with these IPs, sessions are ready!”
• Too easy!
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”
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
LeapSwitch, based in Pune, India.

- New VPS host, saw an offer advertised
  - $120/yr, 1gb RAM, 20gb HDD, 500gb BW

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

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.com/billing.
Growing

• 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
### Friends

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

<table>
<thead>
<tr>
<th>#</th>
<th>Location</th>
<th>State</th>
<th>Protocols</th>
<th>In service</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>United States, Detroit</td>
<td>Live</td>
<td>IPv4, IPv6</td>
<td>9th Oct 2013</td>
</tr>
<tr>
<td>4</td>
<td>United States, Boston</td>
<td>Live</td>
<td>IPv4, IPv6</td>
<td>13th Oct 2013</td>
</tr>
<tr>
<td>5</td>
<td>India, Pune</td>
<td>Live</td>
<td>IPv4, IPv6</td>
<td>14th Oct 2013</td>
</tr>
<tr>
<td>6</td>
<td>Germany, Bremen</td>
<td>Live</td>
<td>IPv4</td>
<td>23rd Oct 2013</td>
</tr>
</tbody>
</table>
• “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)
• 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

- 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
  - Layer42, Lookout
Pi problems

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

- 11 live nodes
  - London C, VPS from Thomas Greer @ TSONE
- Bahrain on its way
Fun along the way

• First BGP customer for some of the LowEndTalk advertisers
  – 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
Market for VMs with BGP sessions

Virtual Server
A dedicated single virtual machine for any purpose, on demand, ready in minutes. Read More.

Processor (CPU)
Memory (RAM)
Storage
Operating System
Bandwidth
Location

BGP Service
If you want to assign your own IP ranges to a ZettaGrid Cloud Server then you will need this option. Please note that you will require your own AS (Autonomous System) Number and IP Range.

Summary
AS Number
IP Range

MONTHLY COST
Setup Cost
Total Cost

$0.00
$0.00
$0.00

Add to Order
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

• Many tunnel all DNS traffic back to a single location
  – Looks like anycast
  – Poor DNS performance
• Simple HTTP/HTTPS service to assist debugging
  – local.esgob.com
  – local4.esgob.com / local6.esgob.com
• Returns JSON via lighttpd on each node

```bash
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 API + web interface
• Open source more code
• Support: Multi master, TSIG inbound
• Public SOA checker
• Mix of routing and DNS daemons
  – configurable per node via JSON in RethinkDB
  – BIND, NSD4, KNOT, Quagga + BIRD
• Looking for friendly hosts:
  – Africa, India, Asia, South America, anywhere!
As of today...

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