

Joost Network Architecture

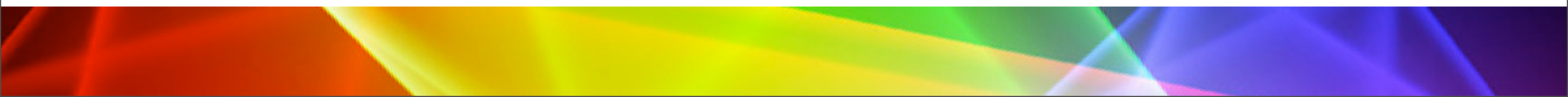
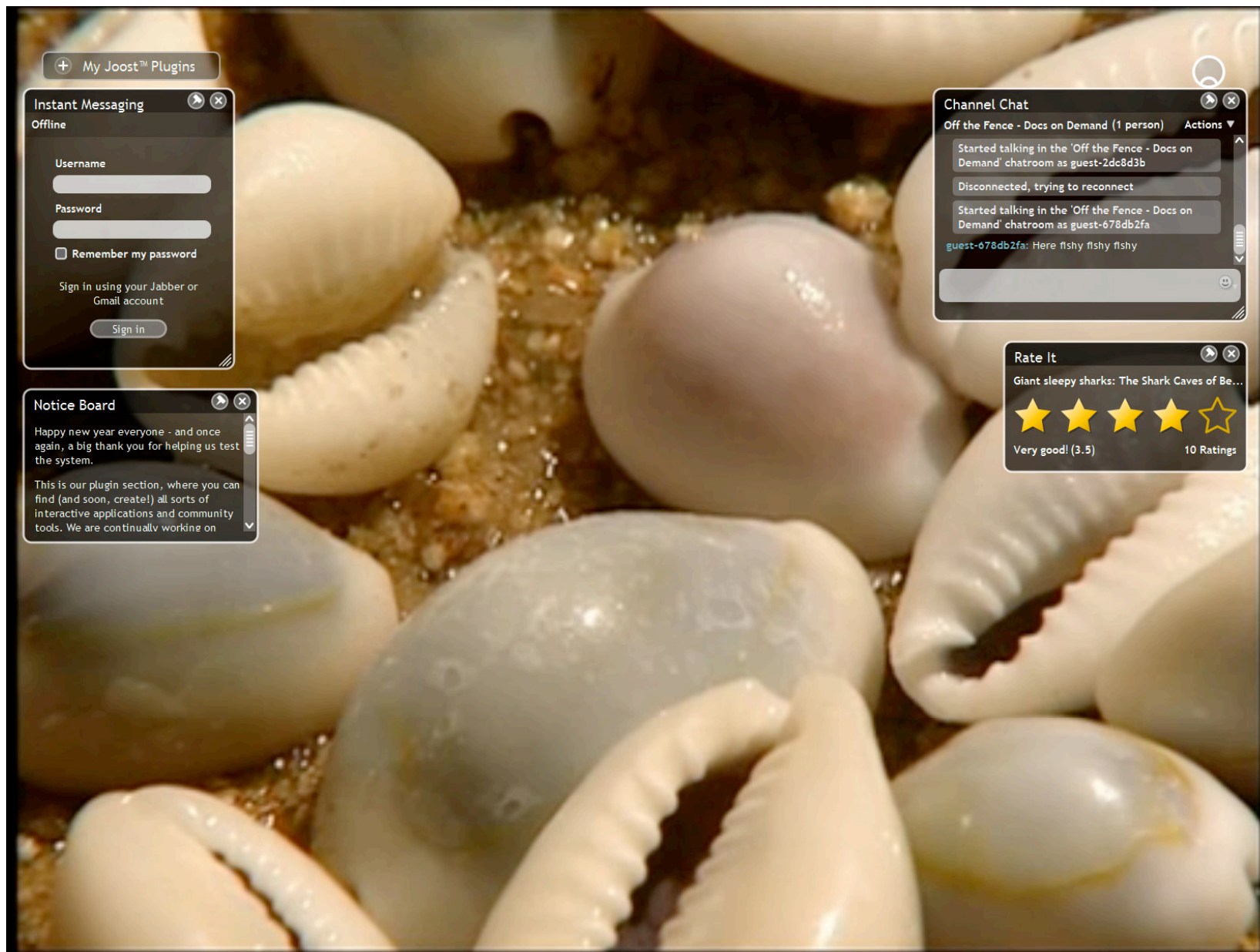
April 4th, 2007

colm.maccarthaigh@joost.com

CM2064-RIPE

AS42072



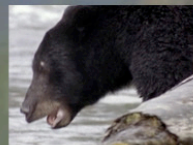




My Channels



My Joost™



0:15:22 / 0:56:10

Last Stand of the Great Bear

National Geographic embarks on a 250-mile adventure through unspoiled territory along the coast of British Columbia called the Great Bear Rain Forest. It is here that bear-hunting wolves take to the sea, grizzlies clash in titanic battles, and wild salmon are the pulsing lifeblood of an entire ecosystem.

[more >](#)



National Geographic



Last Stand of the Great Bear



...

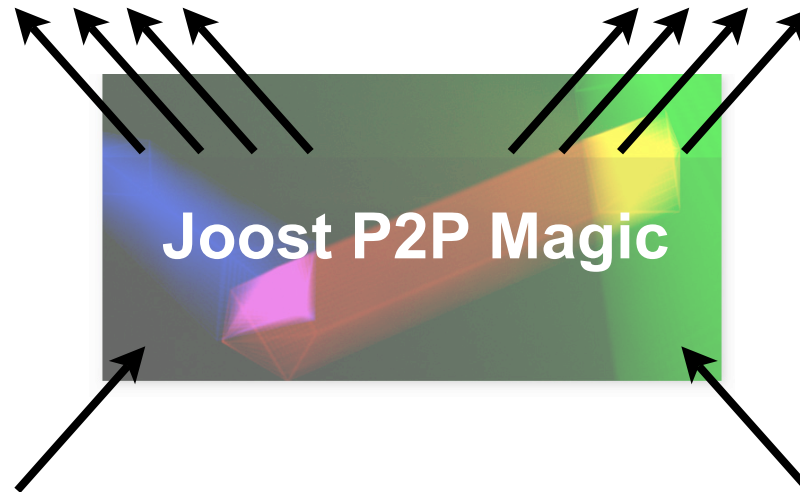
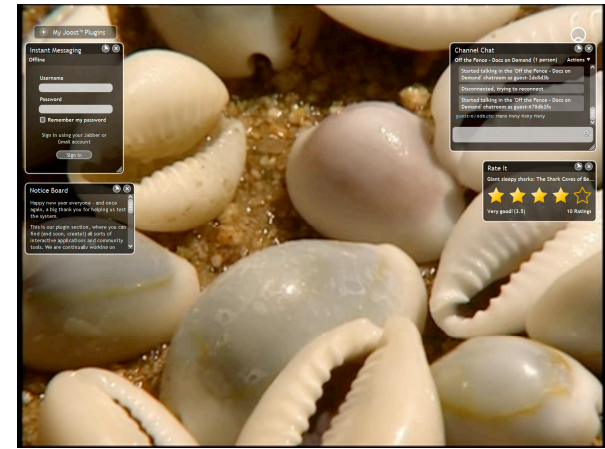
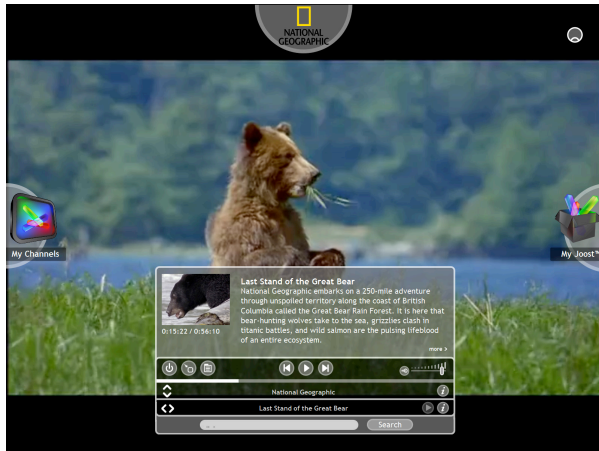
Search

Joost Network Architecture

- **Ground rules**
 - **No firewalls**
 - **No hardware load-balancers**
 - **High availability (this is TV)**
 - **Lots of bandwidth (this is TV)**
 - **Ethernet only**
 - **Rapidly provisionable**
 - **Business requirements**
 - **Cost-effective**
- **4 main service topologies ...**

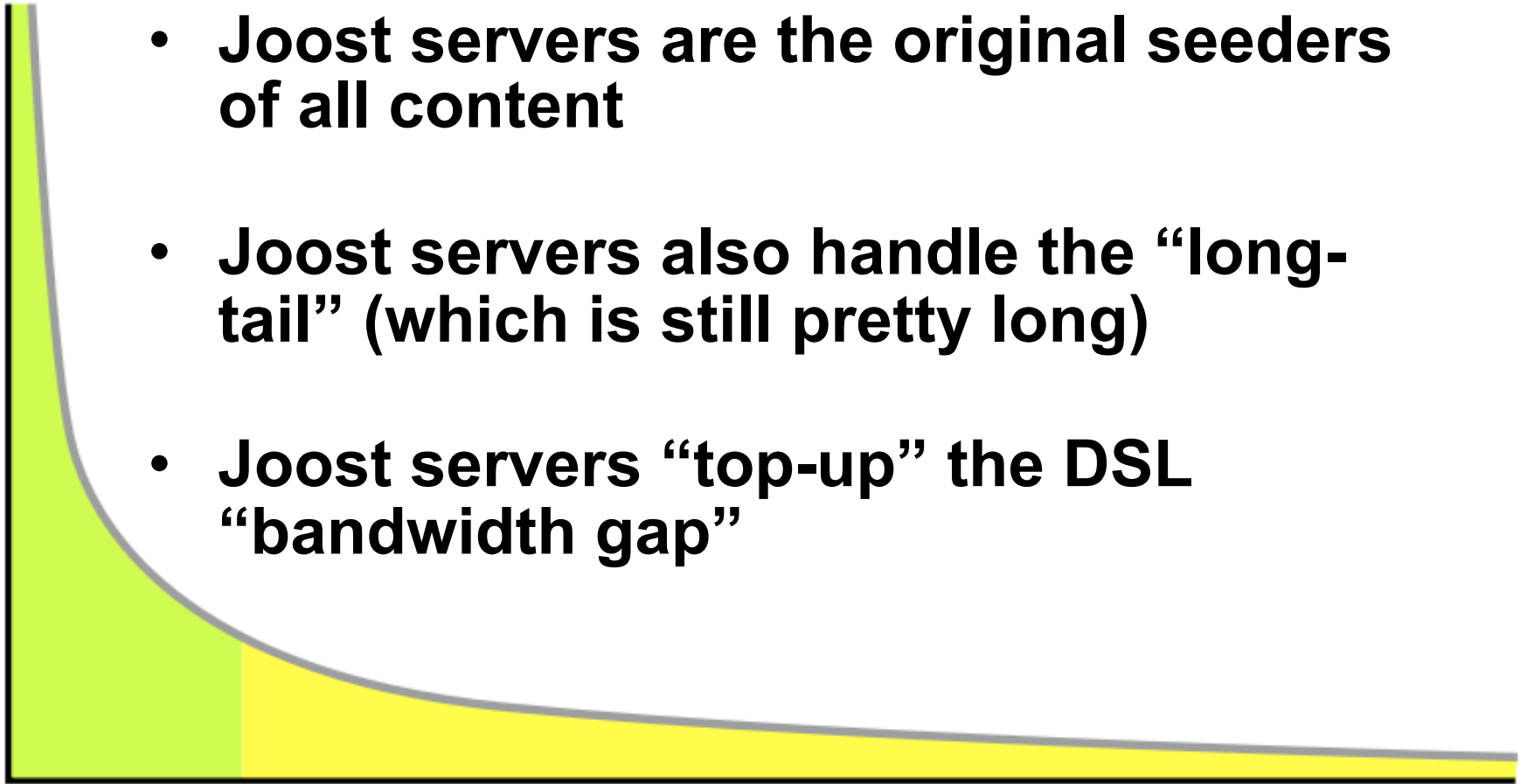


Video Streaming



Joost P2P Magic

- **Joost servers are the original seeders of all content**
- **Joost servers also handle the “long-tail” (which is still pretty long)**
- **Joost servers “top-up” the DSL “bandwidth gap”**



Joost P2P Magic

- **A long-tail server cluster is 11 servers**
 - **1 control server**
 - **10 long-tail servers**
- **2x Cisco 3560's (crypto image)**
 - **HSRP**
 - **EBGP**
 - **802.1q**
 - **OSPF (for anycast services)**

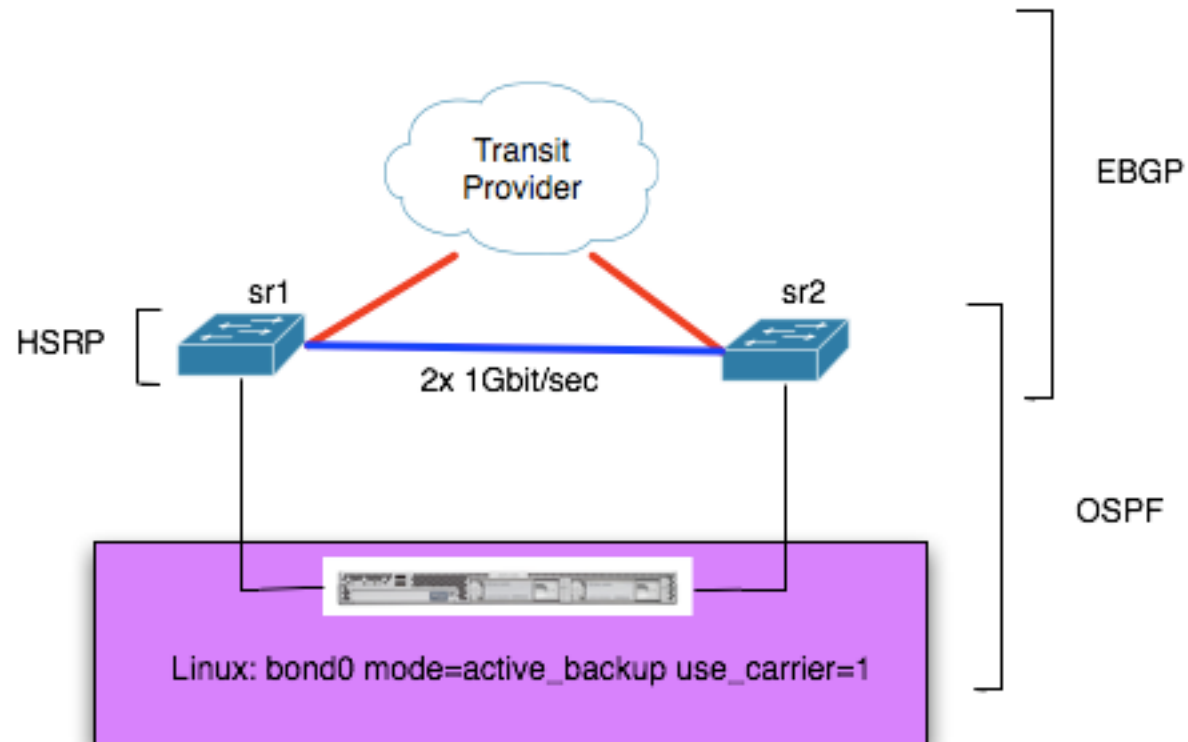


Joost P2P Magic

- **Each cluster gets 1Gbit/sec of IP transit**
- **/26 from transit provider PA space**
 - **we're not sensitive to re-numbering**
 - **avoids prefix-filtering and dampening**
- **Each cluster is an “island” with no connectivity back to Joost HQ apart from an IP tunnel for server lights-out management**
- **Preference is to add sites when scaling**
 - **Decreases average latencies**



Joost P2P Magic



Joost P2P Magic

- **Pros**
 - **Highly cost-effective for 1Gbit/sec of usable, resilient, bandwidth**
 - **No multi-gig network scalability issues, just add another cluster**
 - **Rapidly provisionable, almost anywhere**
- **Cons**
 - **May make future peerings pointless**
 - **3560's can't hold a full table, or do Netflow**



Joost P2P Magic

- **UDP-based (Port 33333)**
 - **Client will perform STUN, ICE in progress.**
- **TCP congestion control would kill any video. Buffering usually significantly increases bandwidth usage.**
- **Packets are generally 1k in size.**



Joost P2P Magic

- **Client first contacts super-node, which handles control traffic only, and direct clients to peers. Peers are re-negotiated frequently.**
- **Each video stream comes from multiple peers, with FEC to handle live peer loss.**
- **1 hour of video \leq 320 MB down, 105MB up.**



Joost P2P Magic

- **No hardware or DNS load-balancing.**
- **All done natively in the p2p code, load-balancing and fault-tolerance is shifted directly into the client.**
- **This is a huge operational saving.**
- **Code is highly-efficient at load-distribution.**

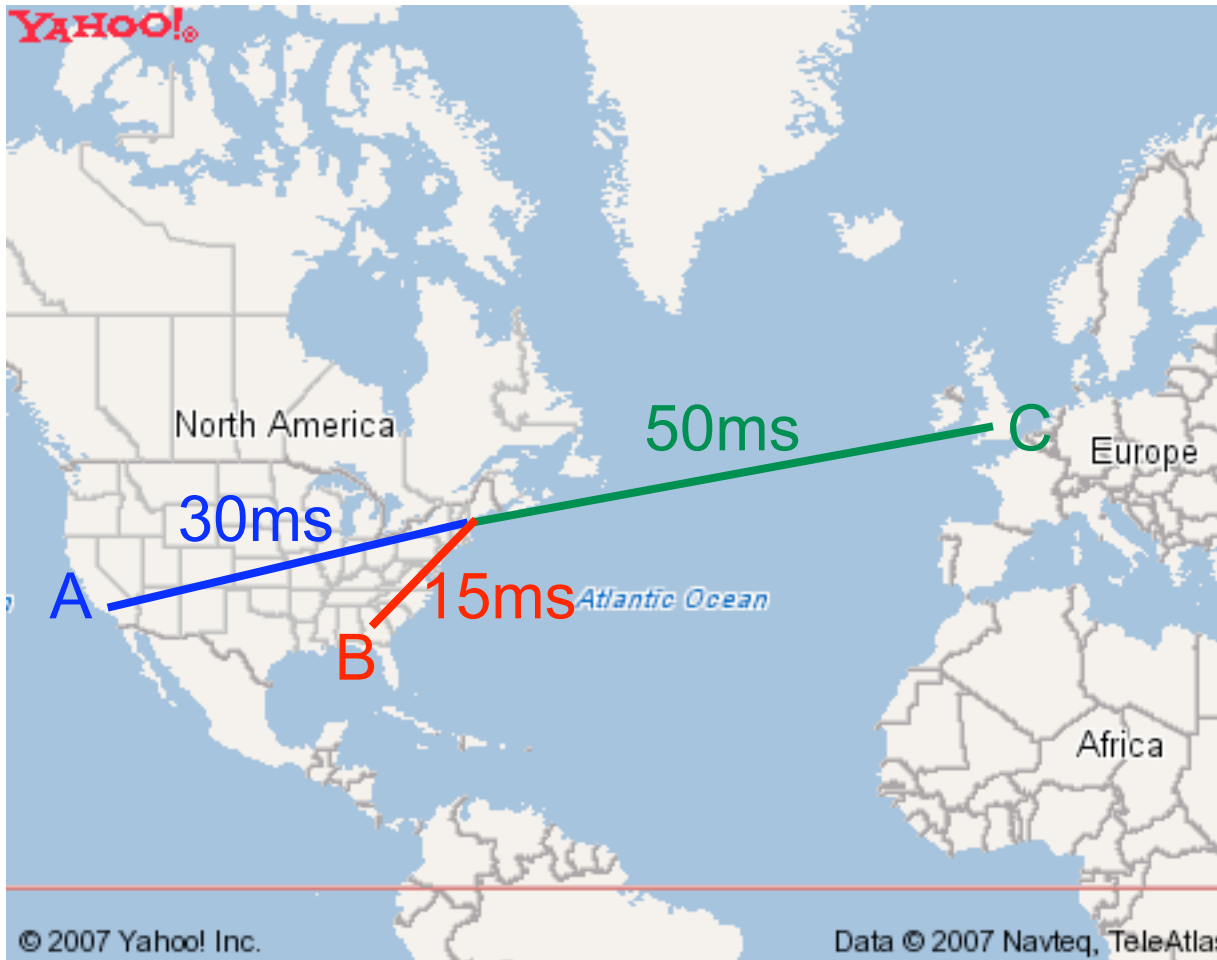


Joost P2P Magic

- **p2p code is prefix aware now, will prefer peers in same /24, /16 etc ...**
- **adding real AS-level awareness is in progress.**
- **latency-based decisions are something to watch out for.**



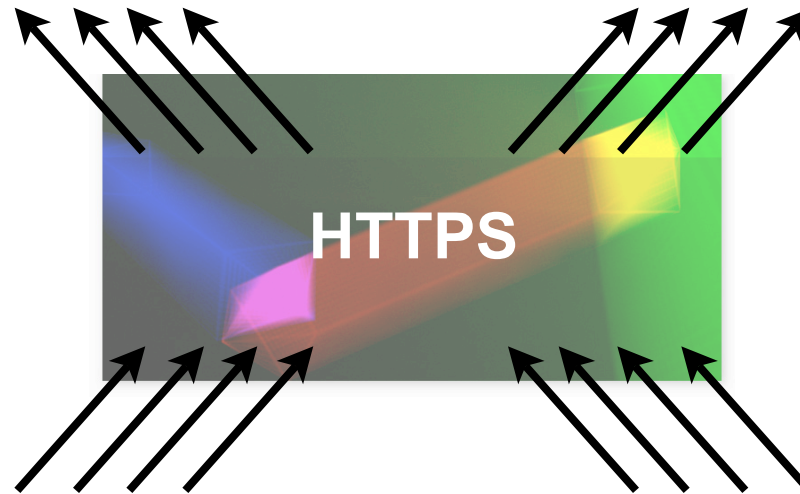
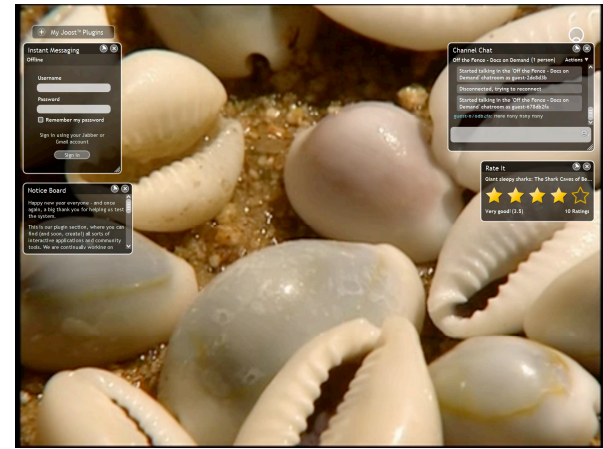
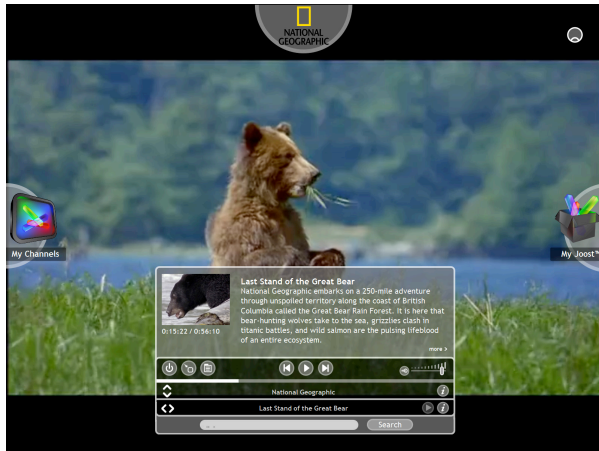
Joost P2P Magic



What if a supernode at C has to coordinate A and B?

This makes us highly sensitive to latency.

Backend services (search ...)

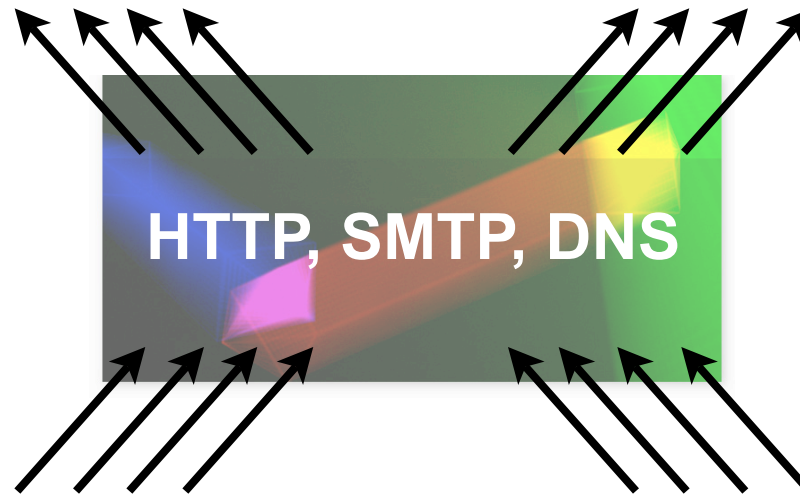


Joost Backend Services

- **Accessed via HTTPS**
- **Provided using Apache Lucene, Hadoop, and many internally-developed services**
- **Each IP is a wack-a-mole virtual-ip**
- **Geographic fail-over provided by DNS**



General Internet Services



Joost General Services

- **Joost.com website and e-mail**
- **Content Owner Website (COW)**
- **Provided by resilient servers with wackamole**
- **Geographic fail-over provided by DNS**

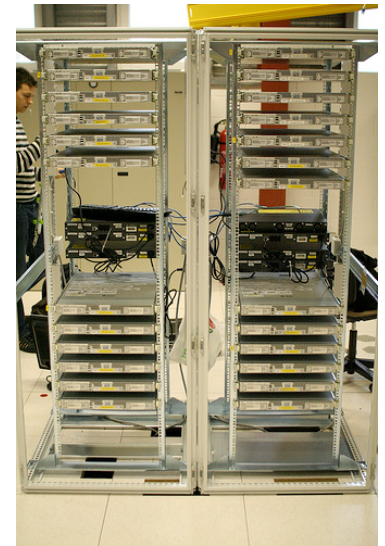
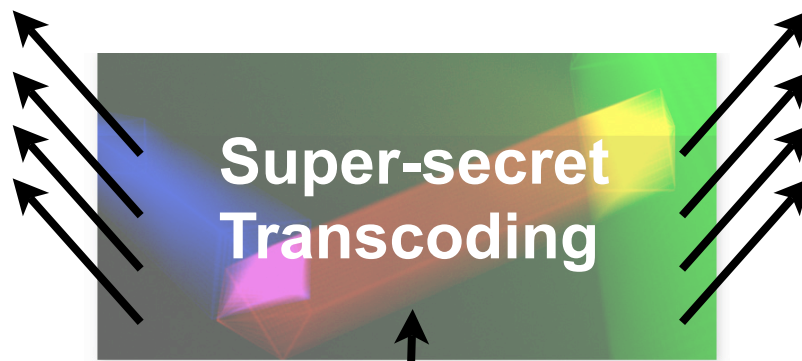


Joost Internal Services

- **Recursive DNS**
- **Syslog**
- **Some authentication services**
- **All provisioned via IP anycast, at each site include long-tail cluster sites (the addresses are internally reachable only).**



Content Ingestion



Content Ingestion

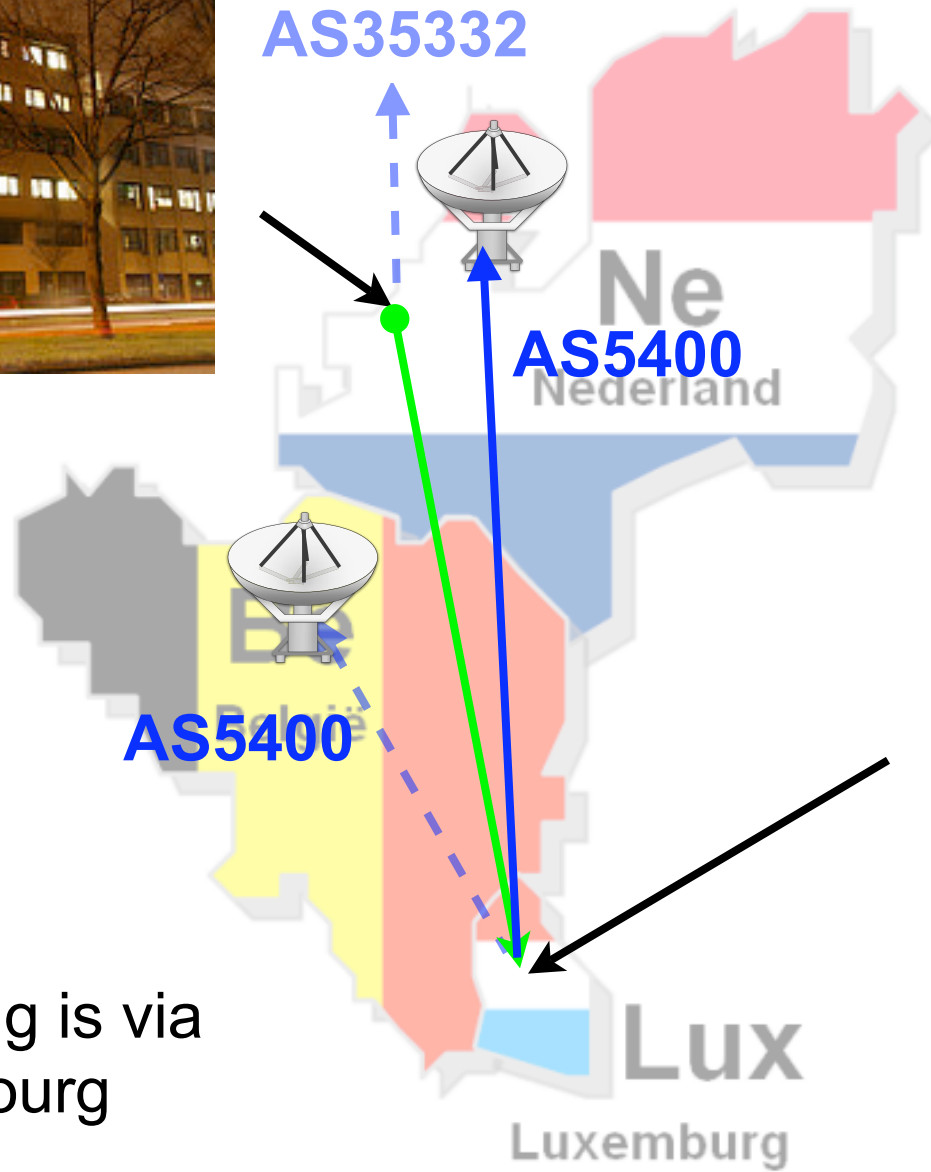
- **Content has to get to Leiden**
- **Currently investigating various network media-delivery options for this**
- **Transcoded in Leiden, and then sent to Luxembourg, and onward to all long-tail server sites**
- **Content-owner website for meta-data**





Leiden

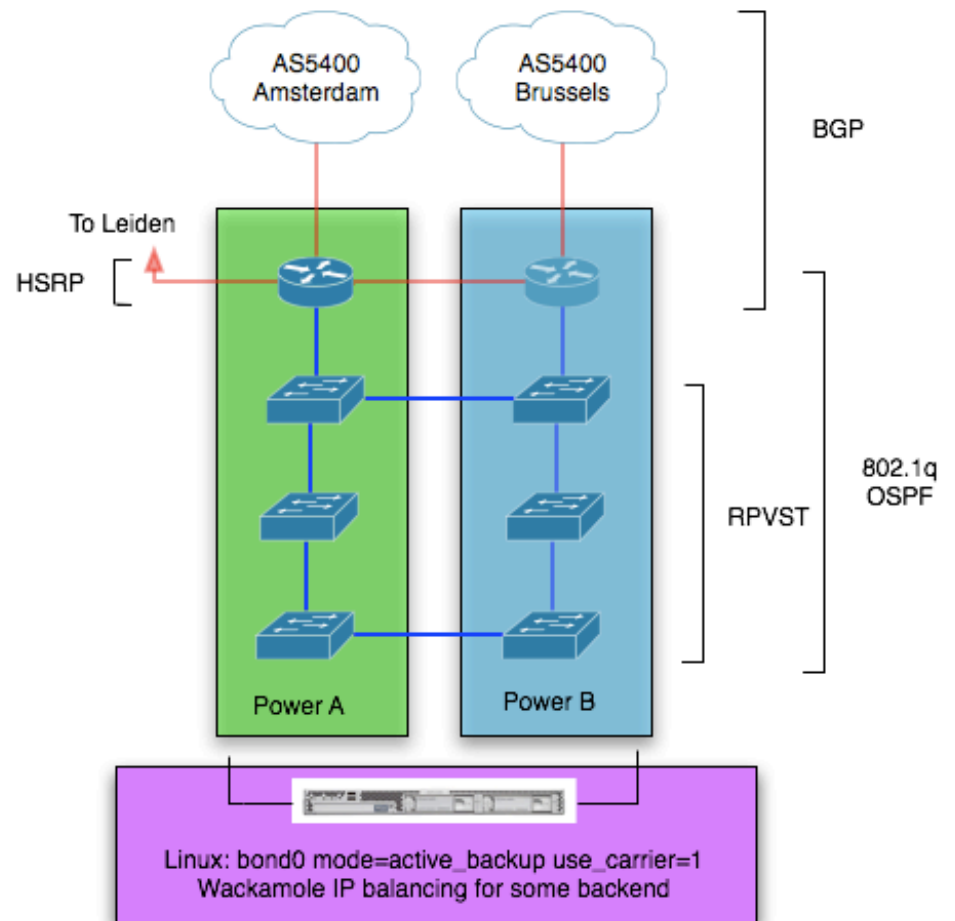
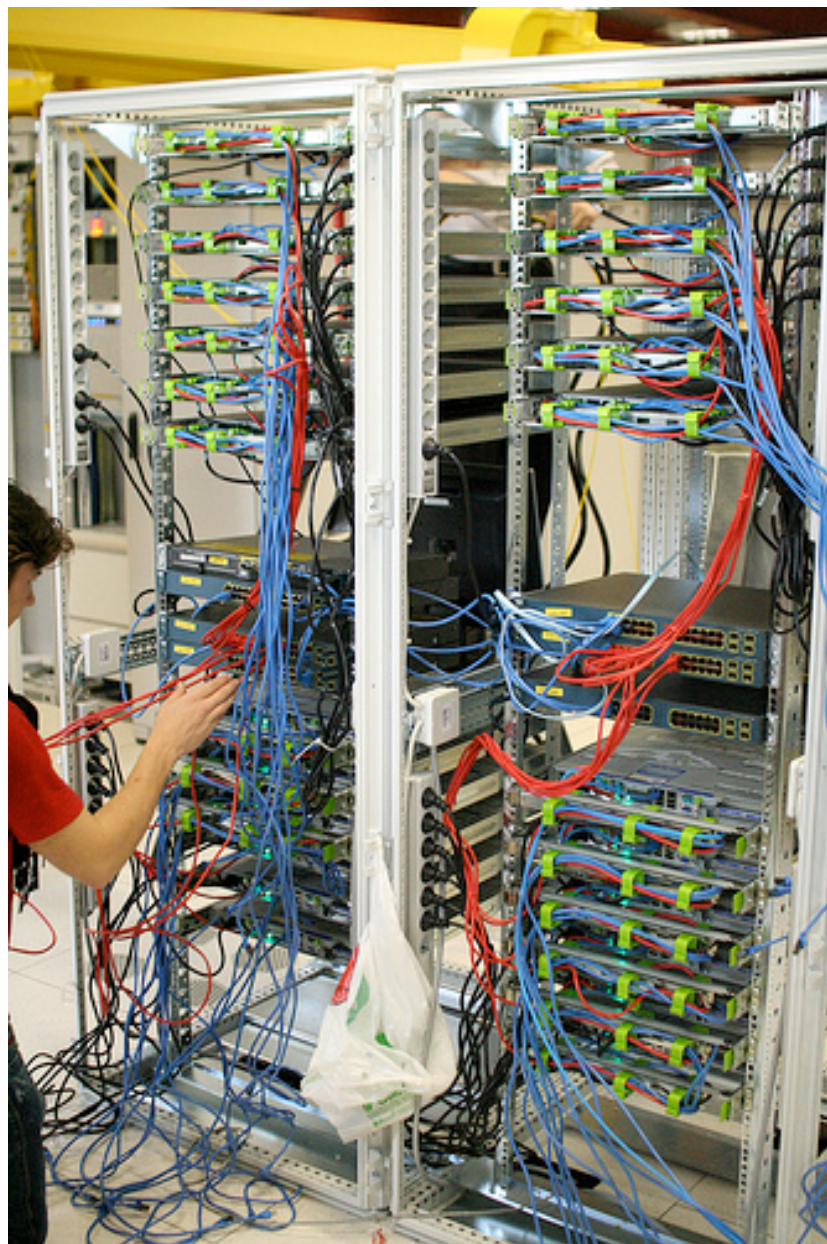
All routing is via
Luxembourg



Joost Benelux

- **Main Joost location is the Joost Benelux Network**
 - **Hosts Joost Leiden office, Luxembourg Datacentre, Primary Long-tail server site, Primary back-end site**
 - **89.251.0.0/20 (deaggregates during some outages)**
 - **AS42072**





- **Routers are Cisco 7301**
 - **Can do a Gig with 1k packets, just about**
 - **Full netflow support, anonymised via CryptoPan, for detailed analysis of p2p network performance**
- **Using OSPF as IGP**



Joost Network Management

- **RT and JIRA for tickets**
- **RANCID-SVN for config management**
- **NAGIOS and syslog for monitoring**
- **MRTG and Cricket for graphing**
- **SSH only**



So, what can we do for ISPs?

- **We're willing to peer, but is there much point? Only portions of the long-tail are peerable.**
- **In-ISP Long-tail servers?**
- **We're L3 and up, not much we can do about the last-mile**
 - **Any promising revenue share models?**



Any Questions?

