Wikimedia Infrastructure

Roan Kattouw

WikiWhat?

WIKIPEDIA
The Free Encyclopedia

- Encyclopedia
- Volunteer-run
- No staff editors

MediaWiki
Because ideas want to be free.

- Wiki software
- GPL
- Volunteer and staff devs
- 3rd party users

WIKIMEDIA

- Movement
- WP + sister projects
- Non-profit foundation

Fundraising

**Please read:
A personal appeal from
Wikipedia programmer Brandon Harris**

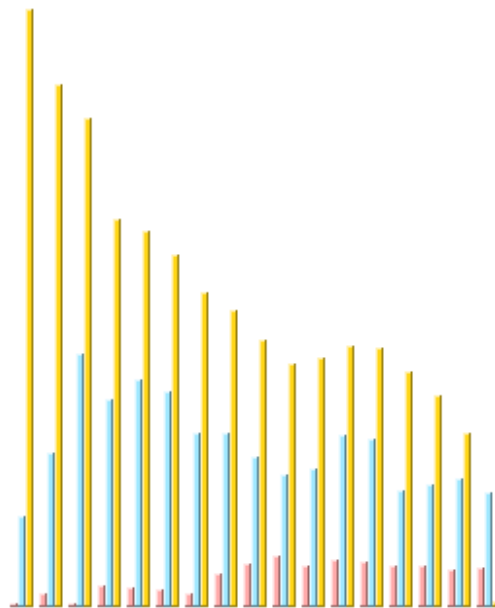## Linux User Group

From Wikipedia, the free encyclopedia

Small budget

- Budget: $28.3 million (2011-2012)
- Tech budget: $12.4 million (44%)
- ~480 servers (+210 not yet in production)
- ~95 staff, ~40 in technology
- This is *nothing* compared to Google, Microsoft, Yahoo, Facebook
- 2008-2009: $6M budget, 26 staff
- Infrastructure focus: **cheap**
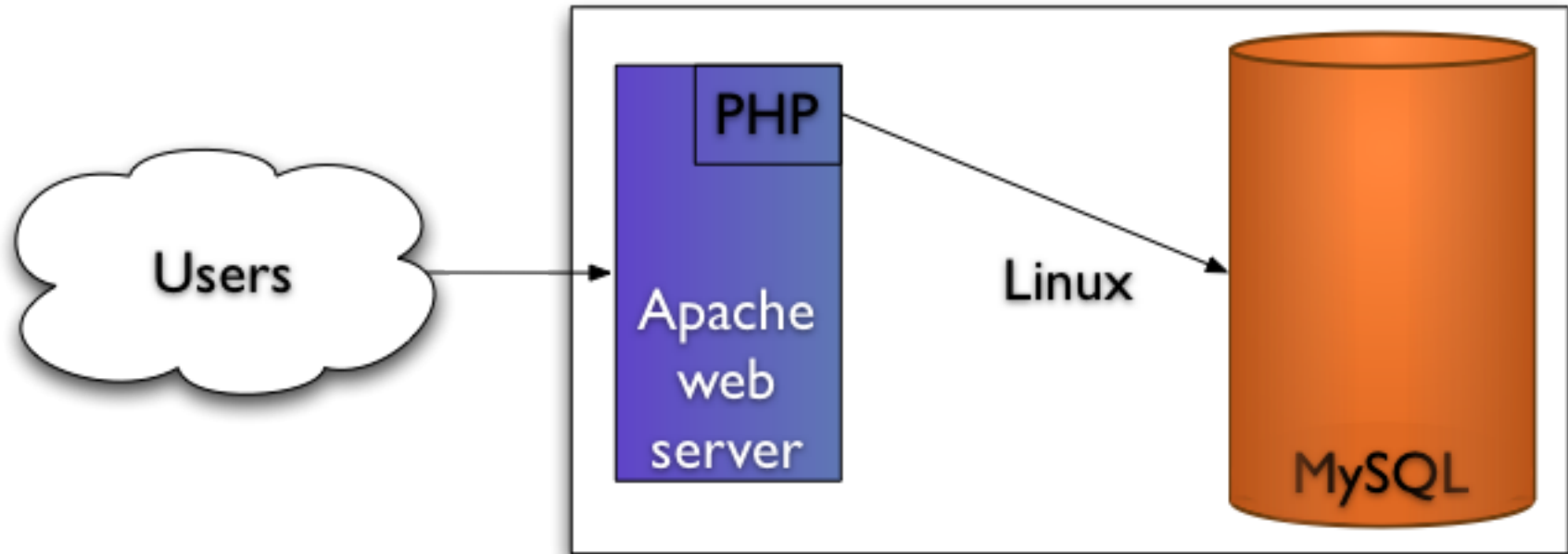
Open source philosophy

- Wiki content is CC-BY-SA / GFDL
    - Dual-licensed for historical reasons
    - Both are copyleft licenses (like GPL) for text that isn't code
- MediaWiki is open source (GPLv2+)
- This presentation is CC-BY-SA
    - It's down in my contract
- Servers run Ubuntu
- Everything we run is open source (almost)

Infrastructure

Datacenter locations
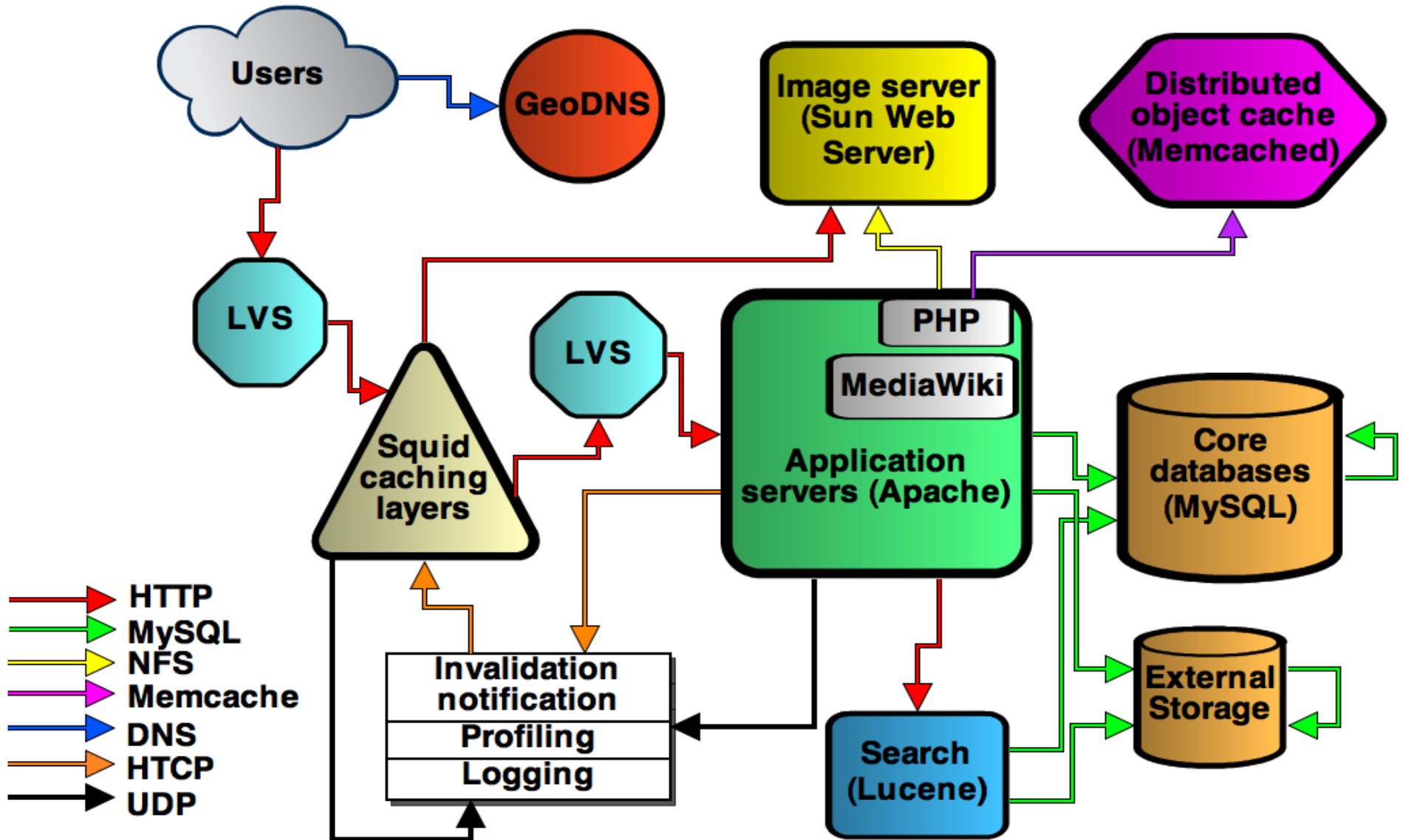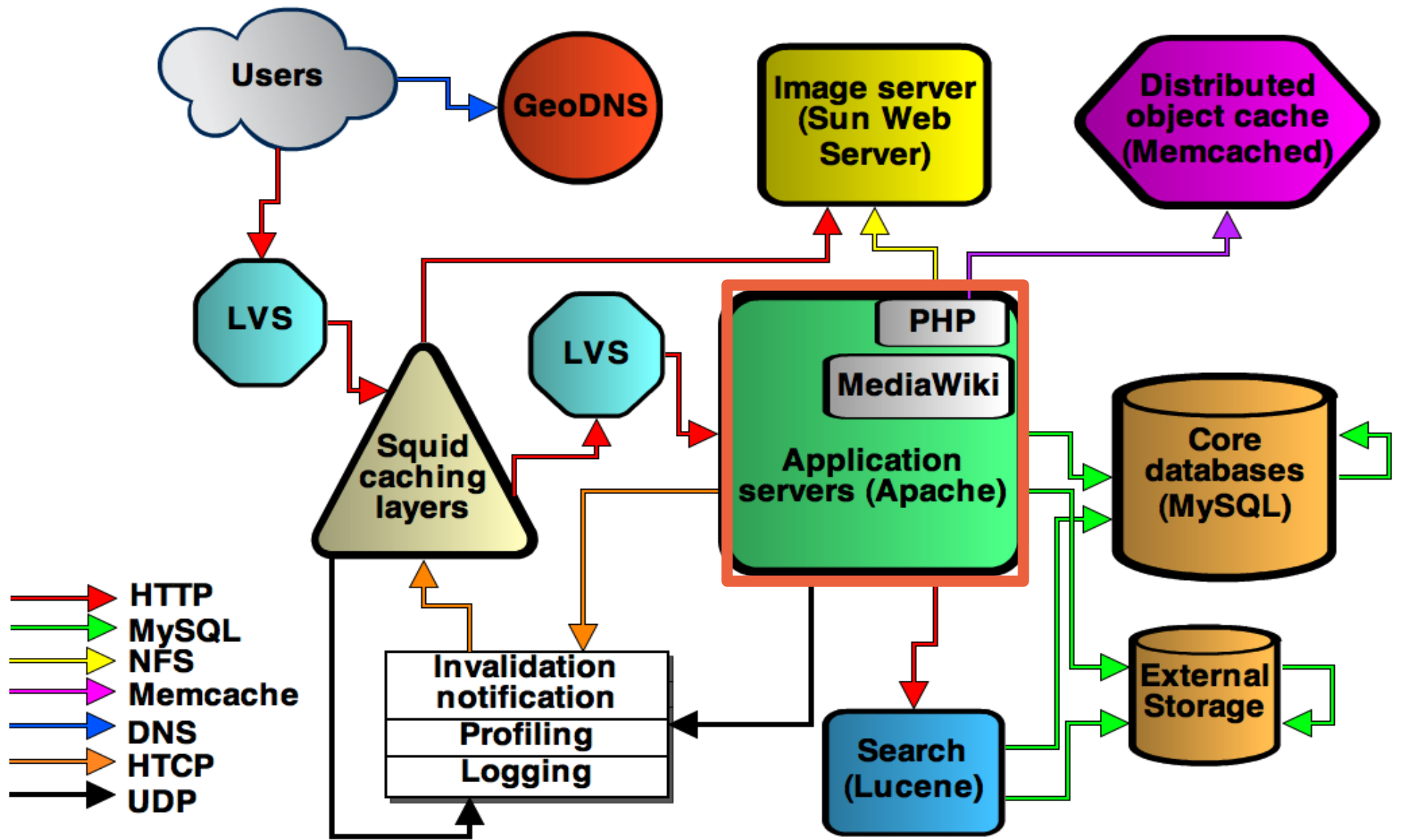
- Tampa, Florida, USA (pmpta & sdtpa)

  - Primary datacenter

- Amsterdam, Netherlands (knams & esams)

  - Caching datacenter

  - Servers now in esams (Haarlem), peering in knams

- Ashburn, Virginia, USA (eqiad)

  - Deployment ongoing, started Feb 2011

  - Proper redundancy, better peering, cheaper transit

  - Will have all Tampa services for redundancy

Architecture: LAMP...

...on steroids

Users → GeoDNS

LVS → Squid caching layers → LVS → Application servers (Apache) [PHP, MediaWiki]

Image server (Sun Web Server)

Distributed object cache (Memcached)

Core databases (MySQL)

External Storage

Search (Lucene)

Invalidation notification
Profiling
Logging

Legend:
- → HTTP
- → MySQL
- → NFS
- → Memcache
- → DNS
- → HTCP
- → UDP

Users

GeoDNS

Image server (Sun Web Server)

Distributed object cache (Memcached)

LVS

LVS

PHP

MediaWiki

Application servers (Apache)

Squid caching layers

Core databases (MySQL)

Invalidation notification

Profiling

Logging

Search (Lucene)
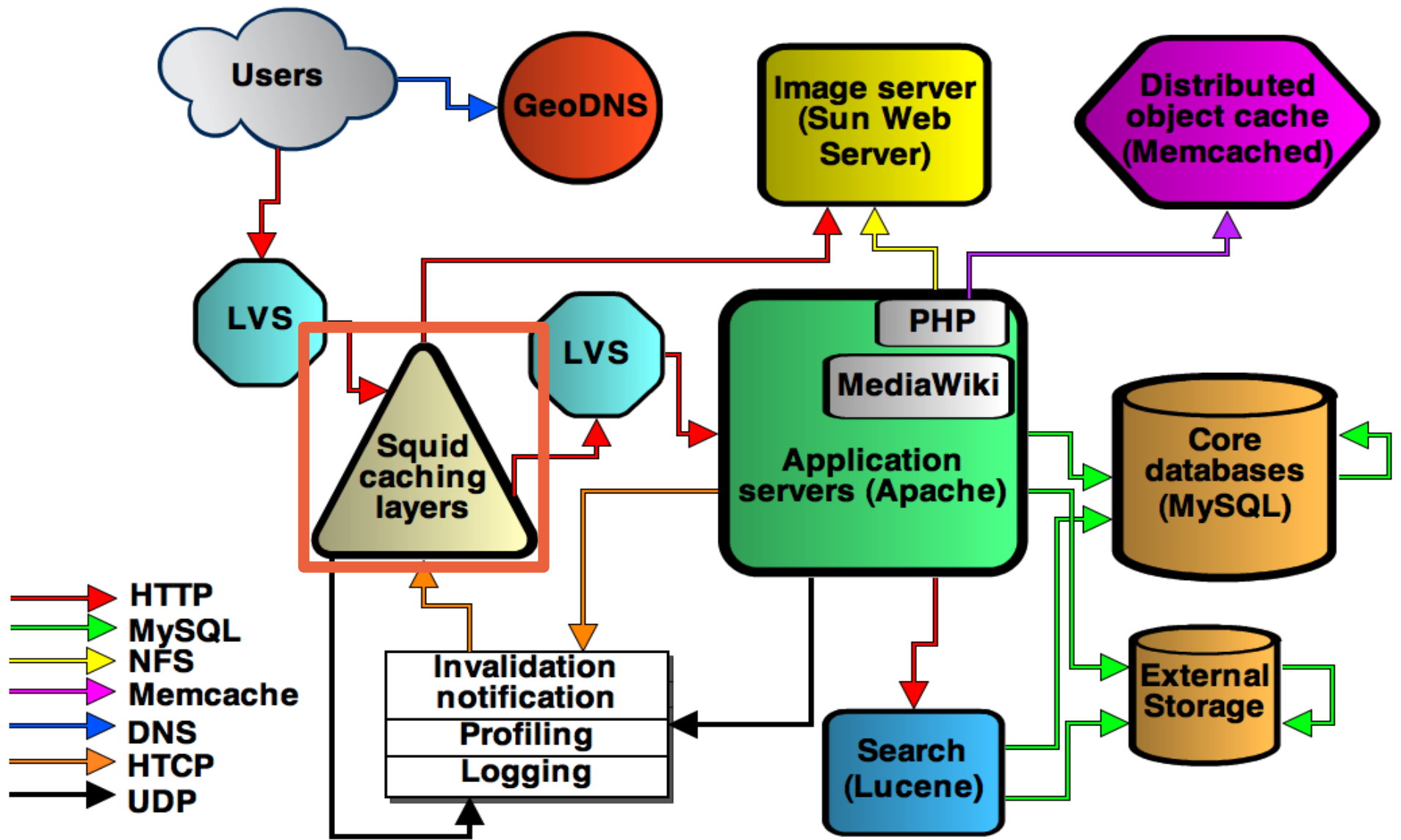
External Storage

HTTP
MySQL
NFS
Memcache
DNS
HTCP
UDP

Apache servers

- Apache servers run PHP & MediaWiki

- Page views and almost everything else is rendered here

- Logged-in users hit Apaches directly, anonymous users go through Squid

- Dedicated pools for image scaling, api.php and JS/CSS bundling

- 182 Apaches, all in Tampa

Users → **GeoDNS**

Image server (Sun Web Server)

Distributed object cache (Memcached)

LVS

Squid caching layers

LVS

PHP

MediaWiki

Application servers (Apache)

Core databases (MySQL)

Invalidation notification

Profiling

Logging

Search (Lucene)

External Storage

HTTP
MySQL
NFS
Memcache
DNS
HTCP
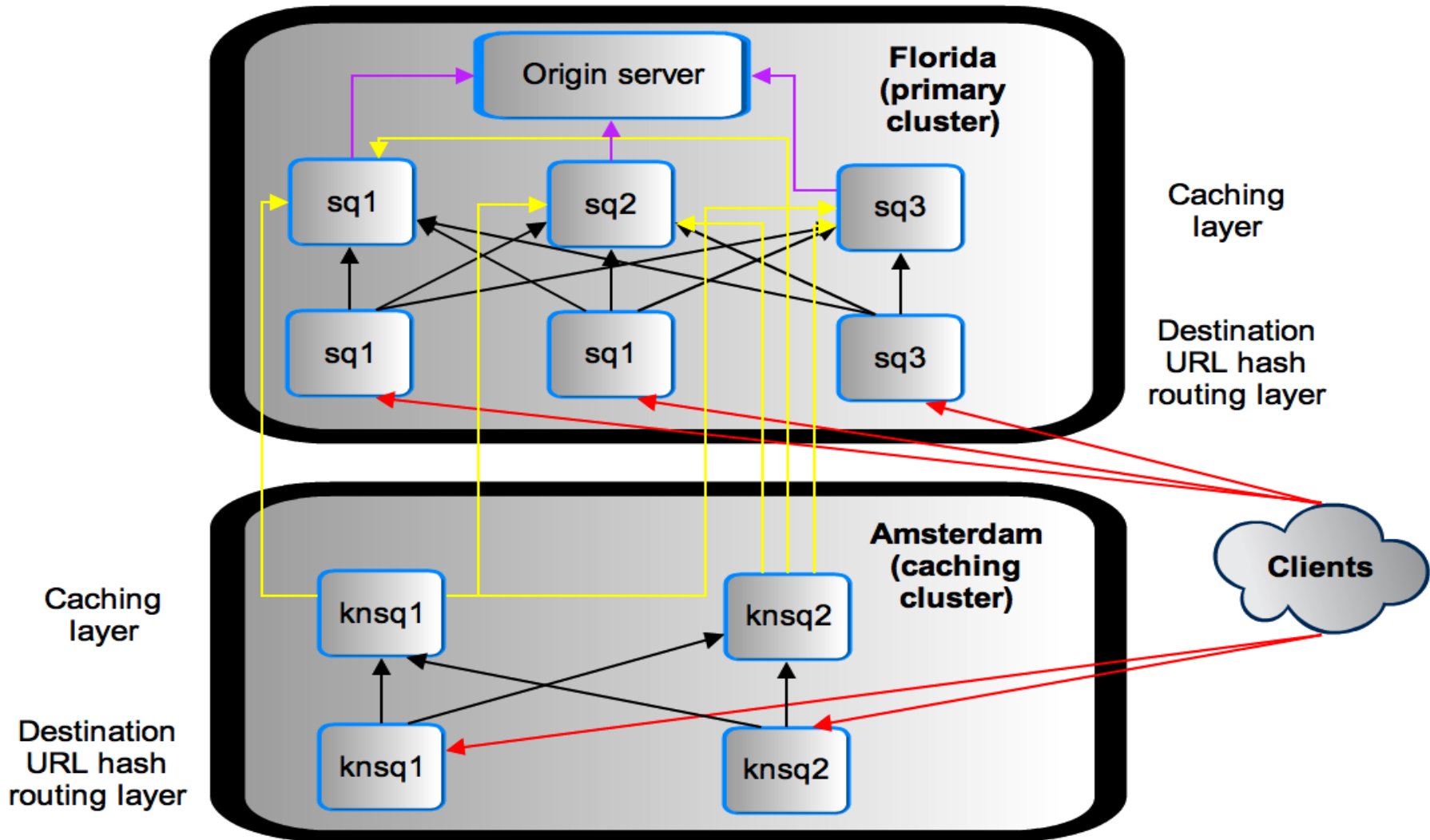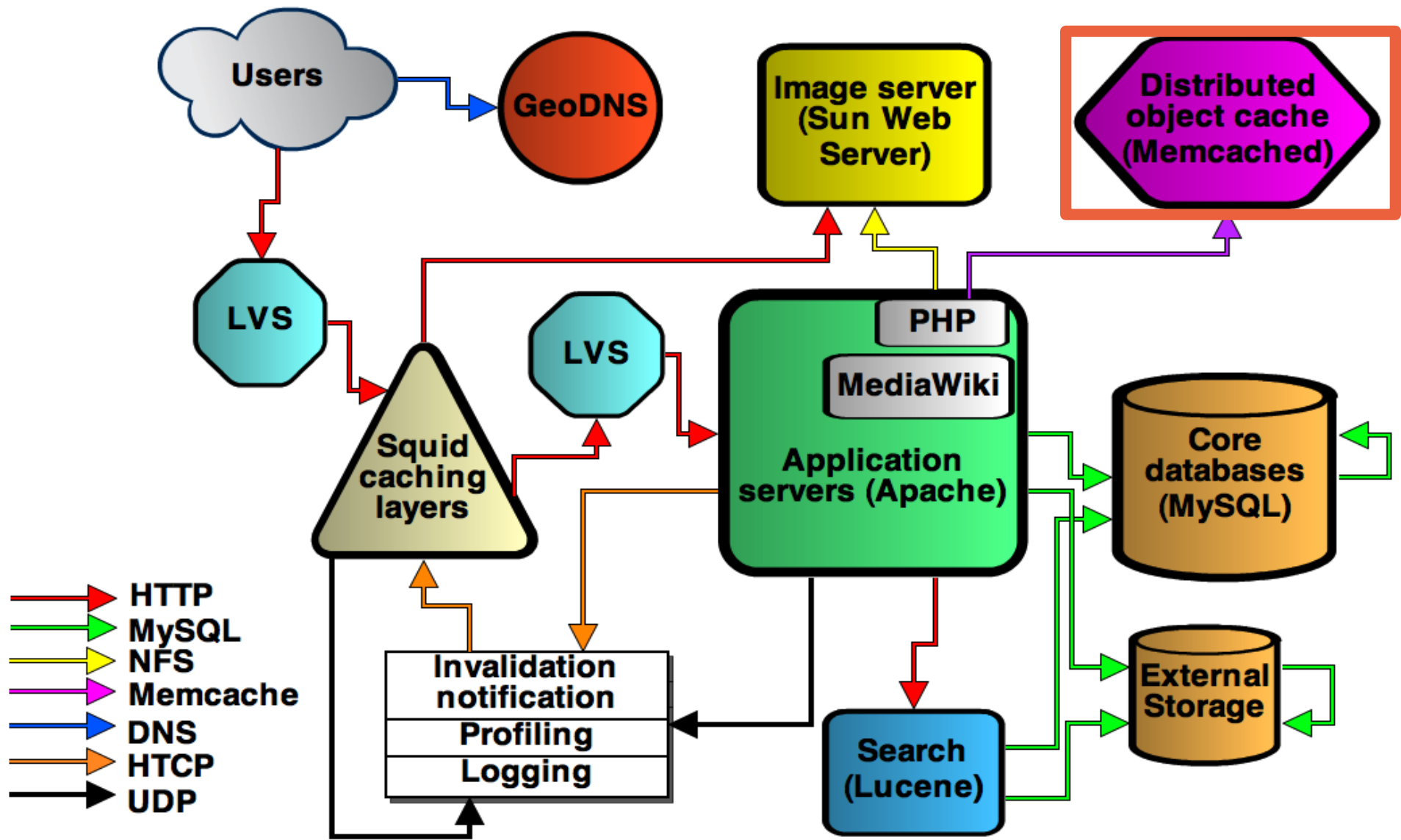UDP

Squid/Varnish caching

- All requests go through a caching layer
- Page views are cached in Squid for anonymous users. Logged-in users bypass the cache
- Images are cached in Squid for everyone
- Resources (JS, CSS, icons) are cached in Varnish
- High hit rates: ~75-85% for text, ~98% for media and resources
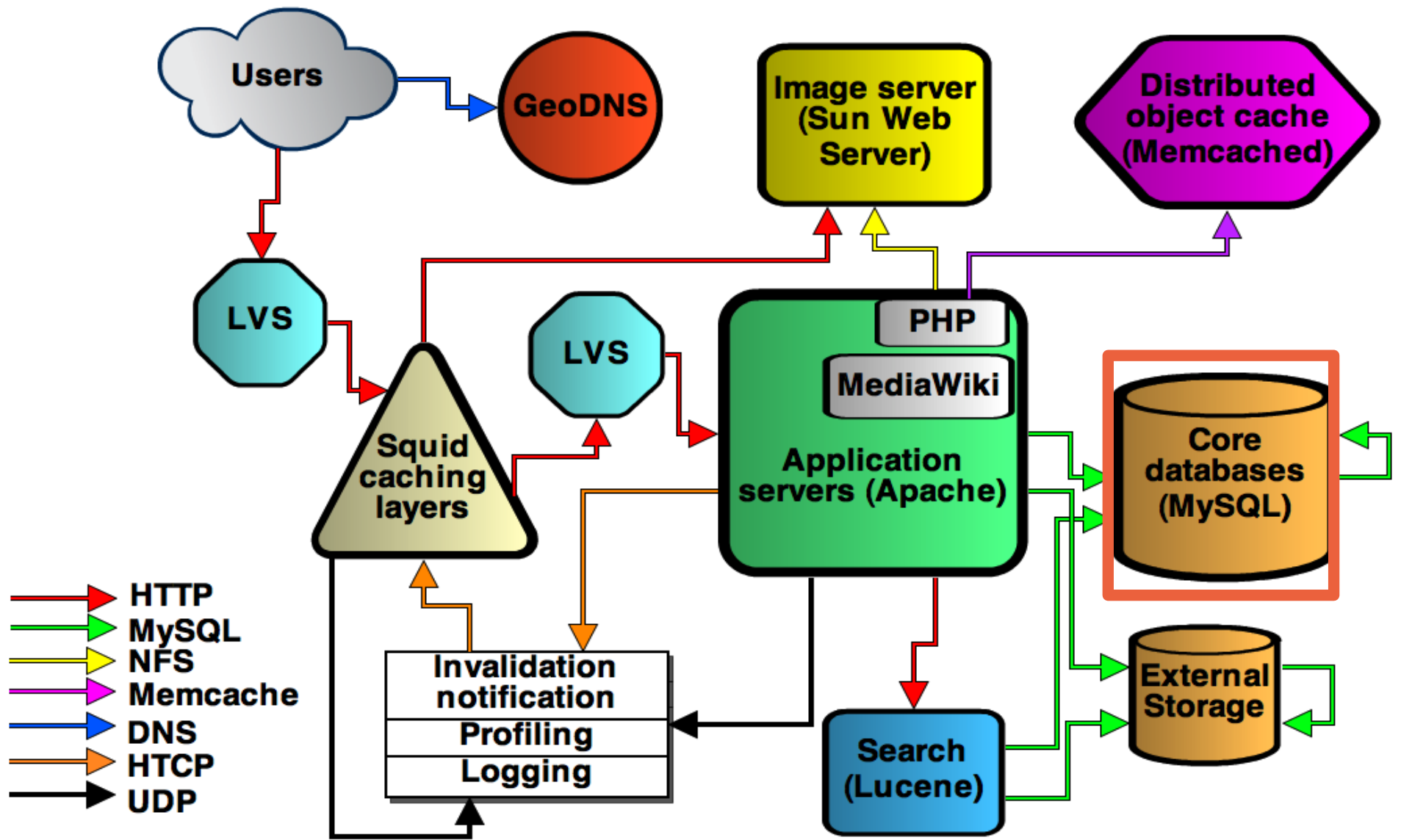- Servers: 51 in Tampa, 53 in Amsterdam

CARP

Squid invalidation

- Pages and media can be edited at any time

- Showing stale versions is not acceptable

- Use on-demand purging

  - Apache server in Tampa sends HTCP purge messages over UDP

  - Multicast to all Tampa Squids

  - Relay copies all purges to Amsterdam

Users

GeoDNS

Image server (Sun Web Server)

Distributed object cache (Memcached)

LVS

LVS

PHP

MediaWiki

Application servers (Apache)

Core databases (MySQL)

Squid caching layers

Invalidation notification

Profiling

Logging

Search (Lucene)

External Storage

HTTP

MySQL

NFS
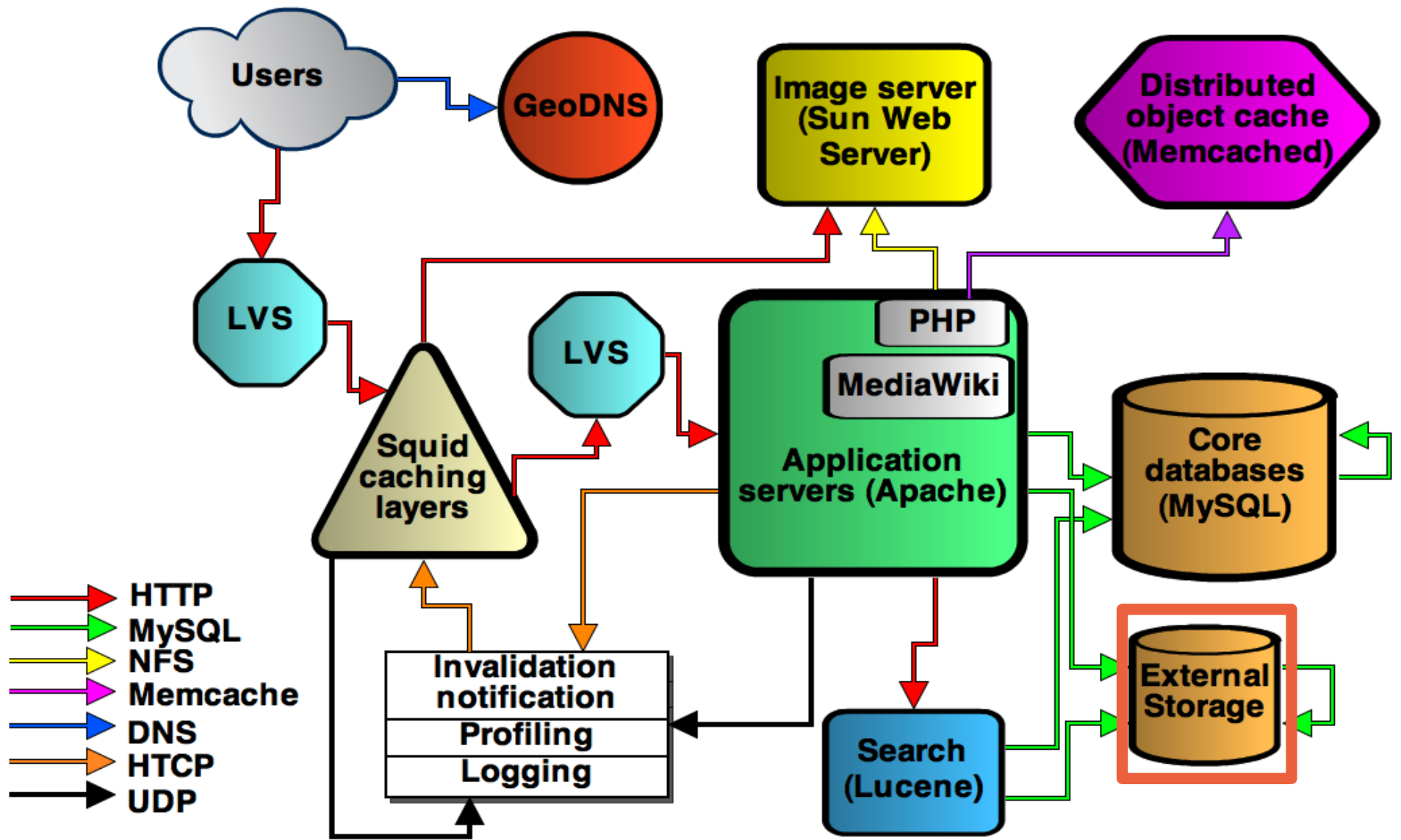
Memcache

DNS

HTCP

UDP

Memcached

- Object cache written in C

- Very simple, very fast

- Sharded based on key hash

- Runs on ~1/3 of Apache servers

- MediaWiki uses memcached to store:

  - Parse results (parser cache)

  - Translations (localization cache)

  - ...and much more

Users

GeoDNS

Image server (Sun Web Server)

Distributed object cache (Memcached)

LVS

LVS

PHP

MediaWiki

Application servers (Apache)

Core databases (MySQL)

Squid caching layers

Invalidation notification

Profiling

Logging

Search (Lucene)

External Storage

HTTP
MySQL
NFS
Memcache
DNS
HTCP
UDP

MySQL

- Separate database for each wiki

- Databases split over 7 clusters

- Each cluster has one read/write master server and 2-4 read-only slave servers

- MediaWiki does load balancing, accounts for replication lag

External storage

- Stores page contents (wikitext)
- Content of all historical revisions is stored
- Optimized for space, not speed
- Delta-based compression, saves 93% compared to gzipping individual revisions
- Fetches are expensive, cached in memcached
- Used to live on Apache servers, now on dedicated DB servers

Users

GeoDNS

Image server (Sun Web Server)

Distributed object cache (Memcached)

LVS

LVS

PHP

MediaWiki

Application servers (Apache)

Core databases (MySQL)

Squid caching layers

Invalidation notification

Profiling

Logging

Search (Lucene)

External Storage

HTTP
MySQL
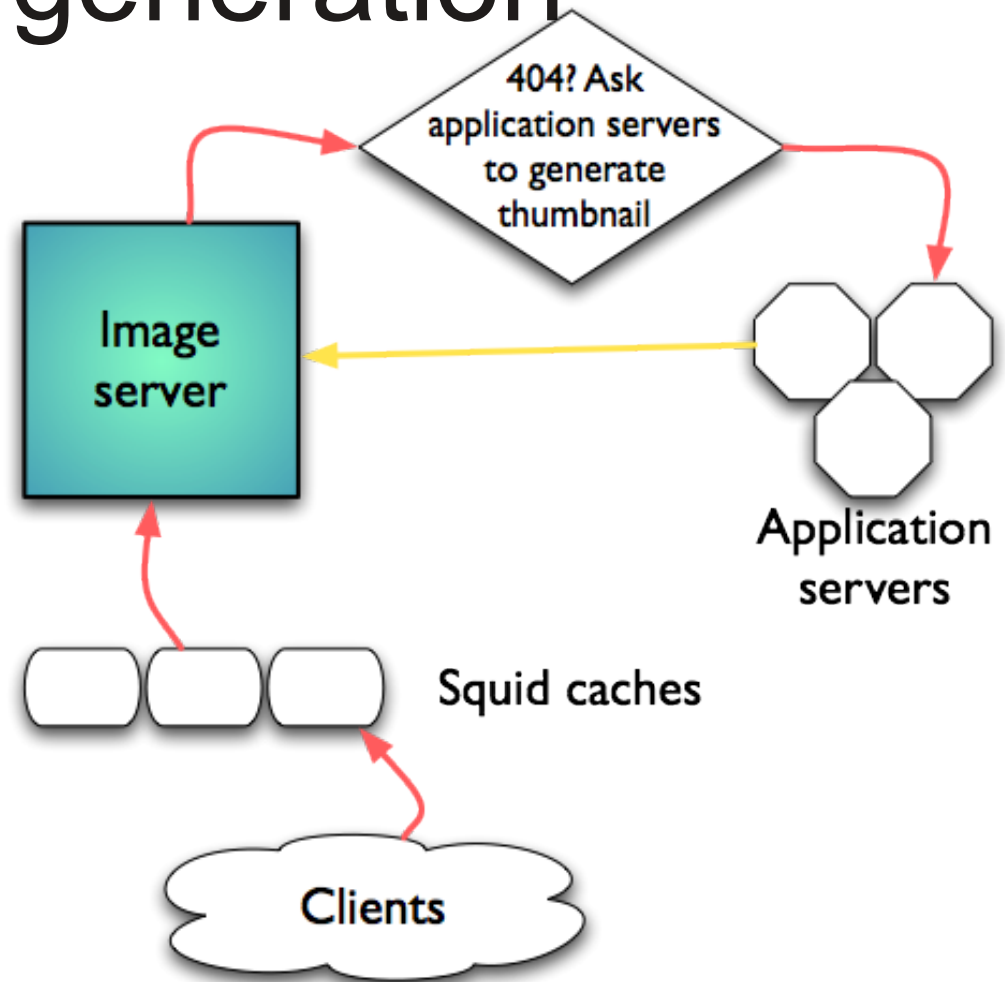NFS
Memcache
DNS
HTCP
UDP
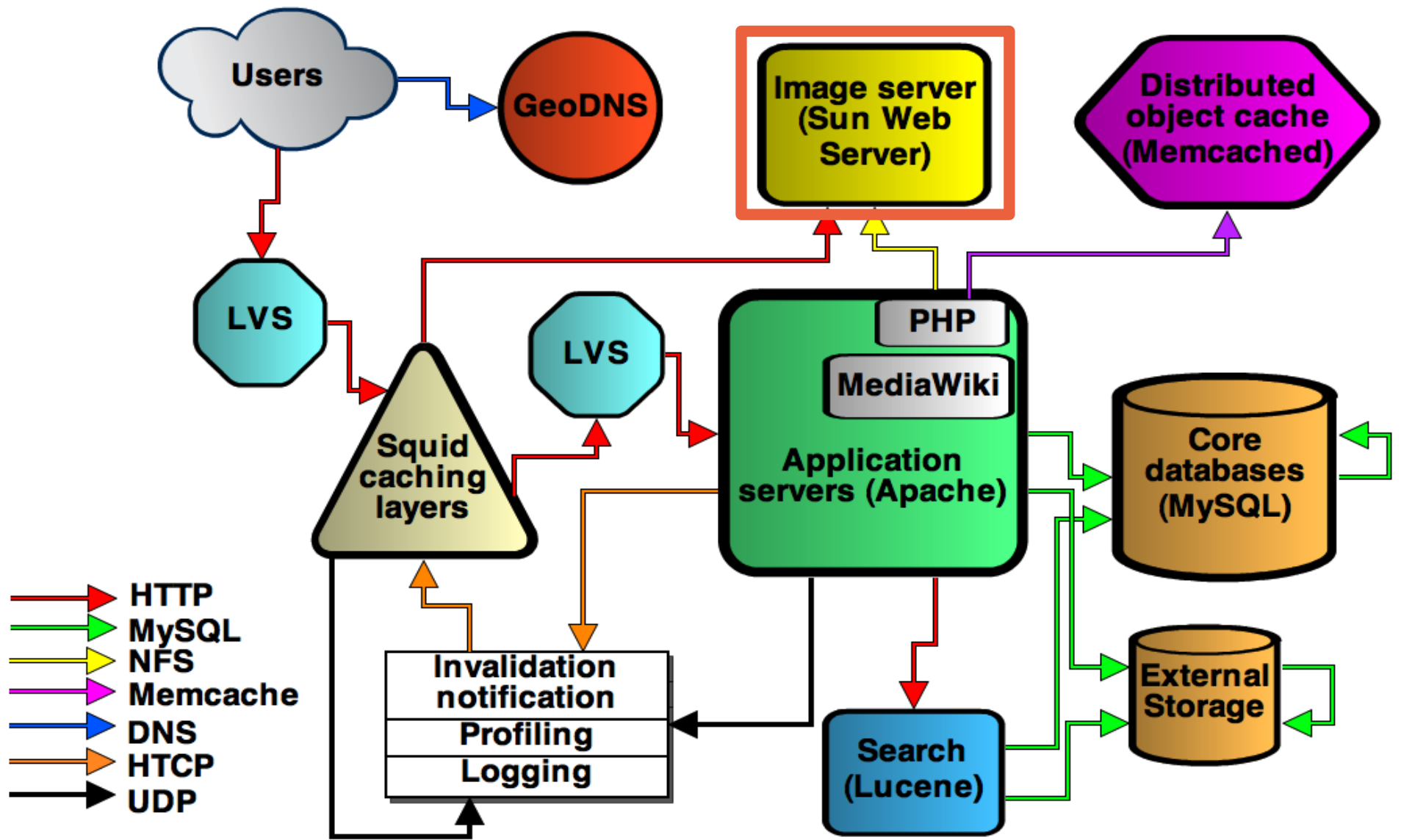
Media storage

- Sun storage servers, also run web server
- Apaches and scalers access storage via NFS
- This is not scalable and not open source
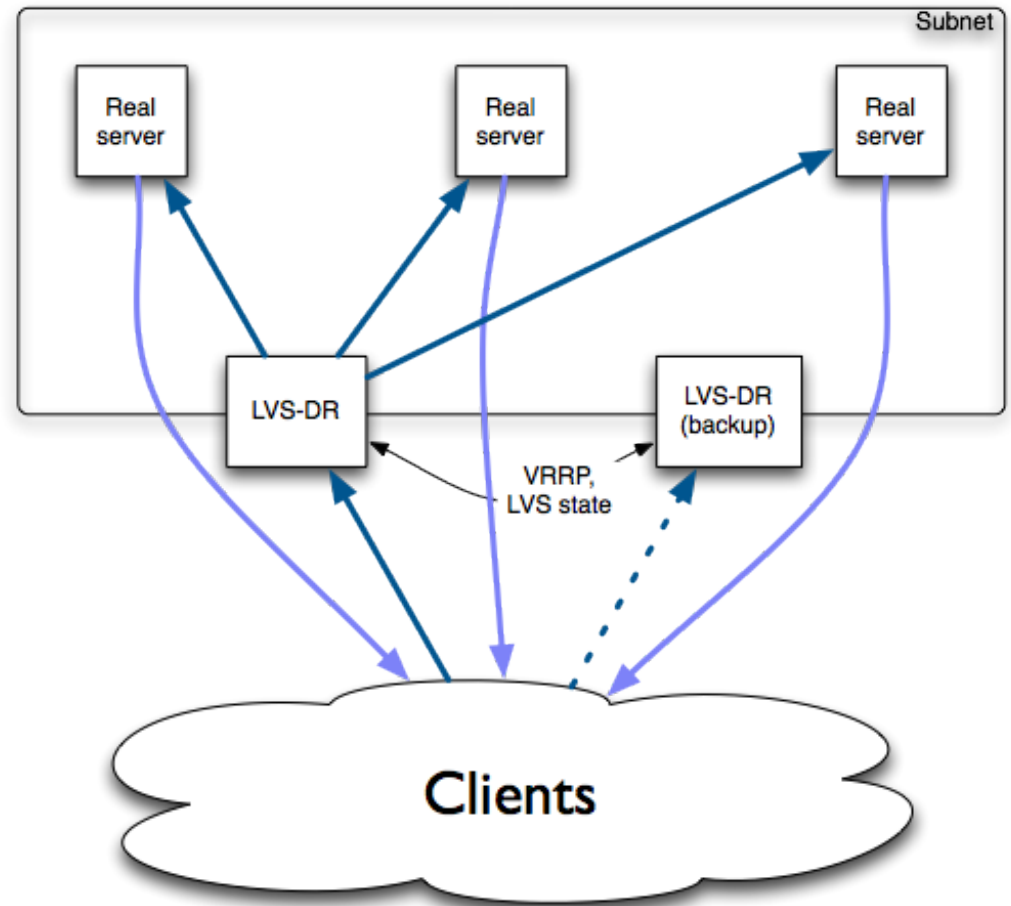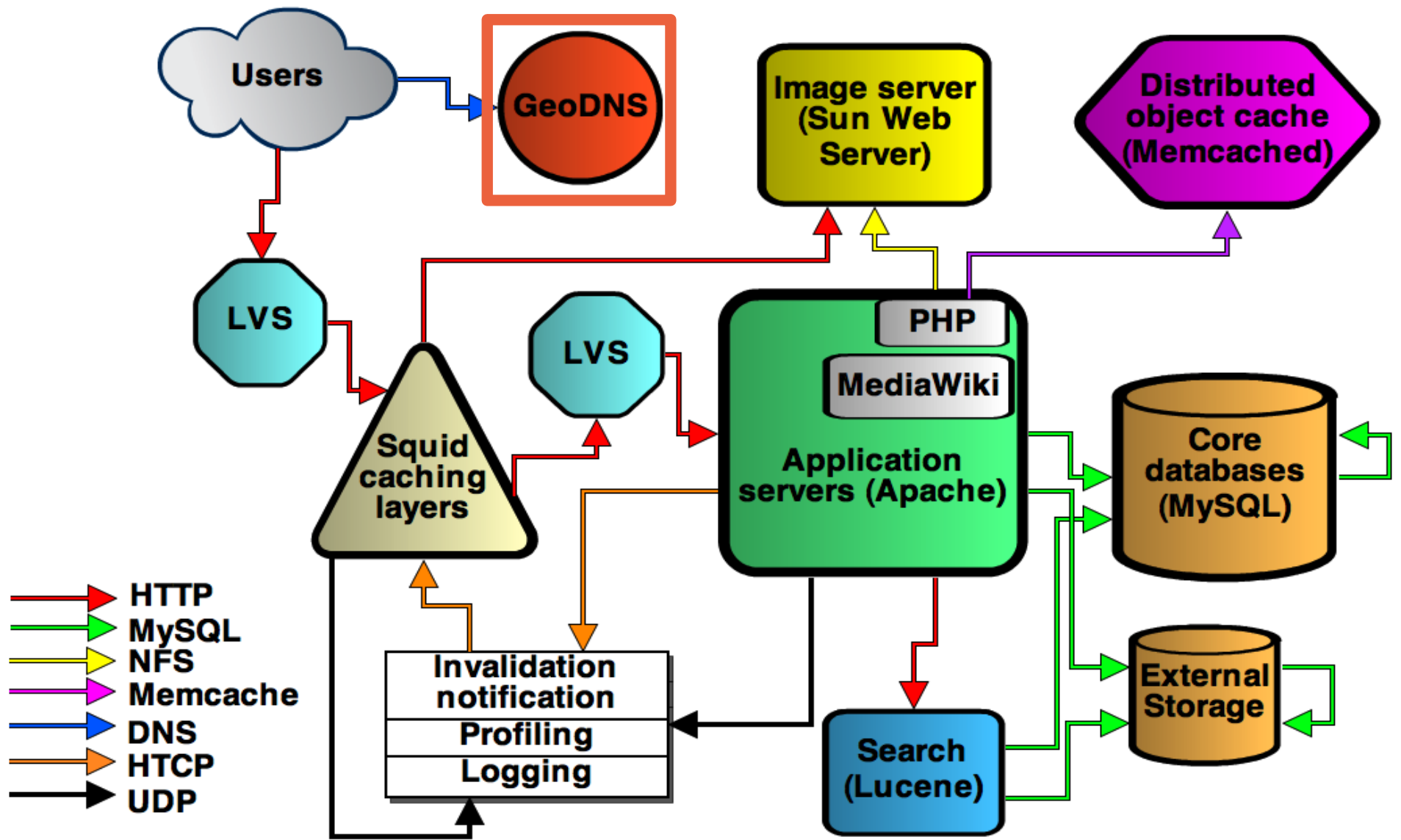- Will be replaced with OpenStack Swift

Thumbnail generation

- HTML generation assumes all thumbnails exist

- 404 handler tells scalers to generate missing thumbnails

Users → **GeoDNS**

Image server (Sun Web Server)

Distributed object cache (Memcached)

LVS

LVS

PHP

MediaWiki

Application servers (Apache)

Squid caching layers

Core databases (MySQL)

Invalidation notification

Profiling

Logging

Search (Lucene)

External Storage

Legend:
- → **HTTP**
- → **MySQL**
- → **NFS**
- → **Memcache**
- → **DNS**
- → **HTCP**
- → **UDP**

LVS-DR

- Linux Virtual Server in Direct Routing mode

- All servers listen on the same public IP

- Return traffic does not go through LVS!

- Used for load balancing Apaches

Users

GeoDNS

Image server (Sun Web Server)

Distributed object cache (Memcached)

LVS

LVS

PHP

MediaWiki

Application servers (Apache)

Core databases (MySQL)

Squid caching layers

Invalidation notification

Profiling

Logging

Search (Lucene)

External Storage

HTTP
MySQL
NFS
Memcache
DNS
HTCP
UDP

GeoDNS

- Geographic load balancing using CNAMEs
- *.wikipedia.org → wikipedia-lb.wikimedia.org
  - → wikipedia-lb.esams.wikimedia.org or
  - → wikipedia-lb.pmtpa.wikimedia.org
- -lb CNAME points to esams if the DNS resolver's IP is European, pmtpa otherwise
- Using PowerDNS with a Geobackend
- Extensible to more than 2 locations

SSL termination

- Old secure gateway: http://en.wikipedia.org → https://secure.wikimedia.org/wikipedia/en

- Single server, bypassed caching

- Since October 2011: https://en.wikipedia.org supported

- 4 SSL termination gateways in each datacenter
  - SSL termination with nginx
  - Happens before Squid, so benefits from caching
  - Support added to MediaWiki 1.18

Wikimedia Labs

Puppet

- Configuration management tool
- Written in Ruby
- Manifests written in declarative language
  - Describe classes of hosts
  - List classes each host is in
- puppetd runs periodically on each host
  - Queries puppet master for config changes
  - Applies config changes (packages, files, etc.)

Puppet example

```
55 class exim::simple-mail-sender {
56         $exim_queuerunner = 'queueonly'
57
58         require exim::packages
59
60         file {
61                 "/etc/exim4/exim4.conf":
62                         require => Package[exim4-config],
63                         owner => root,
64                         group => root,
65                         mode => 0444,
66                         source => "puppet:///files/exim/exim4.minimal.conf";
67         }
68
69         include exim::service
70 }

2731 node /virt[2-4].pmtpa.wmnet/ {
2732         include standard,
2733                 exim::simple-mail-sender,
2734                 openstack::compute
2735 }
```

Puppet in git

- With puppet, architecture = code
- So manage it like code
- Puppet manifests are in a git repo
- Repo is now public

operations/puppet.git		
27 min ago	asher	adding required time unit `26/1226/1` `production`
75 min ago	asher	varnish: don't cache if backend resp code > 400 `25/1225/4`
2 hours ago	asher	graphite: adding code deploy metric type `24/1224/1`

Gerrit

- Code review tool for git
- Web UI for merging revisions

I21b78f85	Create system group (now we have puppet 2.7), and fix modes (MERGED)	Mark Bergsma	operations/puppet	production	Nov 30	✓	✓
Ice158ecd	redo. hoepfully passes lint check.	Pyoungmeister	operations/puppet	production	Nov 30	✓	
I488d0548	working on mark's comments for exim4.conf template (ABANDONED)	Pyoungmeister	operations/puppet	production	Nov 30	✗	

Reviewer		Verified	Code Review
gerrit2	☒	✓	
Demon	☒		+1
Hashar	☒	✓	✓

Ops as a software project

- Our site operations have now become an open source software project!

- Anyone can create a gerrit account and submit changes

- Submitted changes go to the review queue

- When accepted, they're merged into the test branch

- From the test branch, changes are cherry-picked to the production branch

Labs VMs

- VM cluster managed by OpenStack Nova

  - Nova is a cloud computing controller

  - API-compatible with Amazon EC2

- Managed from MediaWiki using the OpenStackManager extension

  - Web interface for creating and managing VMs

  - Also manages users, projects, groups, etc.

Test cluster

- Clone of the Wikimedia cluster in VMs
- Built using 'test' branch of puppet repo
- Workflow
    - Submit change for review in Gerrit
    - When approved, merge into test branch
    - Test change on test cluster
    - Cherry-pick to production branch
    - Deploy to live site
- Essentially this is continuous integration for ops

VMs for everyone!

- Anyone with an account can create VMs
- Easily clone a Wikimedia server, or the entire cluster
- Run your own branch of the puppet repo
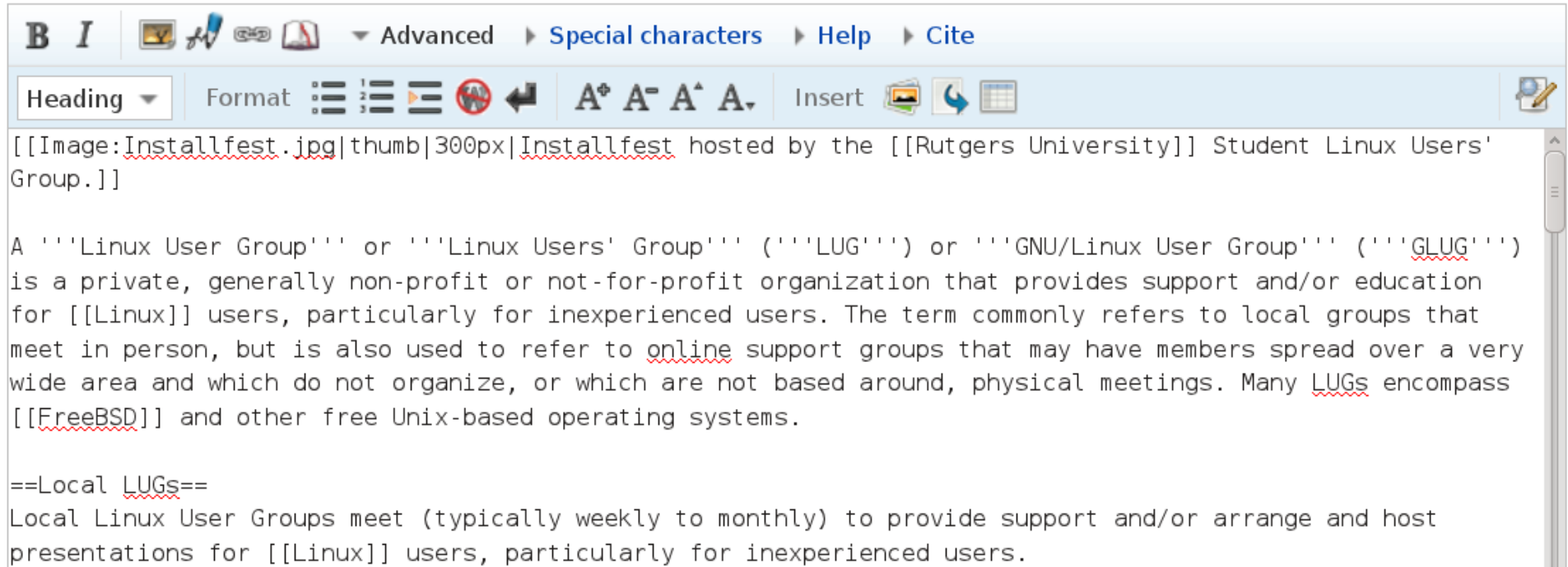- Test and experiment before submitting changes to gerrit

Progress

- The Wikimedia labs project is still in progress
- You can already:
  - read our puppet manifests
  - submit puppet changes to gerrit
  - create VMs based on selected server types
- TODO
  - puppetize all of our servers
  - create test cluster
  - support creating VMs for any server type

Visual editor

This is a bit scary

Editing Linux User Group

B I ☑ ♫ ∞ 📖 ▾ Advanced ▸ Special characters ▸ Help ▸ Cite

Heading ▾ Format ≔ ≔ ⇥ 🚫 ↵ A⁺ A⁻ A A▾ Insert 🖼 🔗 ▦

```
[[Image:Installfest.jpg|thumb|300px|Installfest hosted by the [[Rutgers University]] Student Linux Users'
Group.]]

A '''Linux User Group''' or '''Linux Users' Group''' ('''LUG''') or '''GNU/Linux User Group''' ('''GLUG''')
is a private, generally non-profit or not-for-profit organization that provides support and/or education
for [[Linux]] users, particularly for inexperienced users. The term commonly refers to local groups that
meet in person, but is also used to refer to online support groups that may have members spread over a very
wide area and which do not organize, or which are not based around, physical meetings. Many LUGs encompass
[[FreeBSD]] and other free Unix-based operating systems.

==Local LUGs==
Local Linux User Groups meet (typically weekly to monthly) to provide support and/or arrange and host
presentations for [[Linux]] users, particularly for inexperienced users.
```

This is very scary

```
{{About|the city and county in California}}
{{Use mdy dates|date=June 2011}}
{{pp-move-indef}}
{{Infobox settlement
|name = San Francisco
|official_name = City and County of San Francisco
|nickname      = The City by the Bay<br />Fog City<br/>Frisco (''[[wiktionary:deprecated|deprecated]]'')<ref name="Don't Call It Frisco">{{cite news
|title = Don't Call It Frisco
|url = http://www.sfgenealogy.com/sf/history/hgoe82.htm
|accessdate = July 11, 2011
|newspaper = San Francisco Examiner, San Francisco Chronicle
|date = April 3, 1918
|page = 6
}}</ref><ref>Although many residents still maintain that the nickname "Frisco" is taboo, many residents, especially younger and working-class natives, have kept
the term alive and well. In any case, this is a matter of ongoing debate that reflects certain cultural divisions within The City. {{cite news
| last = Sullivan | first = James
| title = Frisco, that once-verboten term for the city by the bay, is making a comeback among the young and hip. Herb Caen is spinning at warp speed.
| url = http://www.sfgate.com/cgi-bin/article.cgi?f=/c/a/2003/10/14/DD67721.DTL
| work = San Francisco Chronicle
| page = D-1
| date = October 14, 2003
| accessdate = June 12, 2008
}}</ref><ref>Some tourists refer to San Francisco as "Frisco." However, locals discourage this. Samuel D. Cohen writes that many credit "Friscophobia" to
newspaper columnist [[Herb Caen]], whose first book, published in 1953, was "Don't Call it Frisco" after a 1918 newspaper article of the same name. Caen was
considered by many to be the recognized authority on what was, and what was not, beneath the city's dignity, and to him, Frisco was intolerable. {{cite web
| last = Cohen | first = Sam
```

A few screenfuls later...

'''San Francisco''' ({{IPAc-en|ˌ|s|æ|n|_|f|r|ə|n|'|s|ɪ|s|k|oʊ}}), officially the '''City and County of San Francisco''', is the financial, cultural, and transportation center of the [[San Francisco Bay Area]], a region of 7.15 million people which includes [[San Jose, California|San Jose]] and [[Oakland]].<ref>{{cite web
| url = http://www.census.gov/popest/metro/tables/2008/CBSA-EST2008-02.xls
| publisher=U.S. Census Bureau
| accessdate =June 30, 2008
| title = Annual Estimates of the Population of Combined Statistical Areas: April 1, 2000 to July 1, 2008
|| archiveurl = http://www.webcitation.org/619qhElJd |archivedate = 2011-08-23| deadurl=no}}</ref> The only [[Consolidated city–county|consolidated city-county]] in California,<ref name="CityCounty"/> it encompasses a land area of about {{convert|46.9|sqmi|km2}}<ref>{{cite web
| url = [http://www.census.gov/geo/www/gazetteer/files/Gaz_places_national.txt U.S. Census]
| publisher=US Census Bureau
| accessdate = 2011
| title = U.S. Census
}}</ref> on the northern end of the [[San Francisco Peninsula]], giving it a density of about 17,179 people per square mile (6,632 people per km²). It is the most densely settled large city (population greater than 200,000) in the state of California and the [[List of United States cities by population density|second-most densely populated]] large city in the United States after [[New York City]].<ref name="SF_Population_Density"/> San Francisco is the [[List of cities in California (by population)|fourth most populous city]] in California and the [[List of United States cities by population|13th most populous city]] in the United States, with a population of 805,235 as of the [[2010 United States Census|2010 Census]].<ref name="SFCensus2010">{{cite web
| title = U.S. Census Bureau Delivers California's 2010 Census Population Totals

What we want

- A visual editor that's easy to use
  - Kind of WYSIWYG-like but not really
- Written in JavaScript
- Able to parse and render wikitext in JS

New parser

- Wikitext is not a well-defined language
  - No real grammar or specification
  - Test library with ~600 cases
  - "What the parser does"
- Parser implementation has issues
  - Lacks separation between parsing and rendering
  - Doesn't output a parse tree
  - Can only output HTML
- We're now writing a new parser in JavaScript
- Lots of legacy content, so minimize incompatibility

WikiDOM

- Essentially a parse tree for wikitext

- Structured representation in JSON

- WikiDOM can be rendered to multiple formats

  - back to wikitext

  - HTML for page view

  - Others? HTML for mobile? PDF?

- Problem: wikitext → WikiDOM → wikitext round-trips not always clean

  - Proposed solution: normalize pages before enabling

Edit surface

- We need a flexible in-browser editor

  - \<textarea\> does not allow display of rich content

  - \<iframe\> with contentEditable gives us too little control over selection, cursor movement, etc.

- Solution: write an edit surface *from scratch* in JavaScript

  - Con: you have to implement all the basics yourself

  - Pro: after that, the sky is the limit

Tying them together

- New parser parses wikitext, generates a WikiDOM

- Edit surface renders WikiDOM, allows user interaction

- Transaction layer manipulates WikiDOM

- After editing, WikiDOM is rendered back to wikitext

Progress

- Pretty far from completion
- First demo of edit surface expected in Dec 2011
- Eventual deployment will be long and painful

Demo

B *I* ⊘ ↶ ↷

Direct manipulation interface

In computer science, direct manipulation is a human-computer interaction style which involves continuous representation of objects of interest, and rapid, reversible, incremental actions and feedback. The intention is to allow a user to directly manipulate objects presented to them, using actions that correspond at least loosely to the physical world. An example of direct-manipulation is resizing a graphical shape, such as a rectangle, by dragging its corners or edges with a mouse.

Having real-world metaphors for objects and actions can make it easier for a user to learn and use an interface (some might say that the interface is more natural or intuitive), and rapid, incremental feedback allows a user to make fewer errors and complete tasks in less time, because they can see the results of an action before completing the action, thus evaluating the output and compensating for mistakes.

The term was introduced by Ben Shneiderman in 1983 within the context of office applications and the desktop metaphor.[1] Individuals in academia and computer scientists doing research on future user interfaces often put as much or even more stress on tactile control and feedback, or sonic control and feedback than on the visual feedback given by most GUIs. As a result the term direct manipulation interface has been more widespread in these environments.[citation needed]

In contrast to WIMP/GUI interfaces

Direct manipulation is closely associated with interfaces that use windows, icons, menus, and a pointing device (WIMP GUI) as these almost always incorporate direct manipulation to

Links

- https://wikitech.wikimedia.org/view/Presentations
- https://labsconsole.wikimedia.org/
- http://tinyurl.com/wmlabsblog
- http://ganglia.wikimedia.org/
- https://www.mediawiki.org/wiki/Visual_editor

Credits

- Slide 1: Wikimedia logo circle
  https://commons.wikimedia.org/wiki/File:Wikimedia_logo_family.png

- Slide 1: CC-BY-SA logo https://creativecommons.org/licenses/

- Slide 7: datacenter information derived from Rob Halsell's "Wikimedia Operations Overview" presentation (2011), see
  https://wikitech.wikimedia.org/view/Presentations

- Slide 8, 9, 10, 12, 14, 16, 18, 20, 22, 24, 25, 26, 27: diagrams copied from Ryan Lane's "Wikimedia Architecture/Community/The Site Architecture you can edit"