what happens when you type
en.wikipedia.org

effie mouzeli • alexandros kosiaris
Did you know...

- ... the Wikipedia infrastructure is run by the Wikimedia Foundation, an American nonprofit charitable organisation?
- ... and we are ~370 people?
- ... and we have no affiliation with Wikileaks?
- ... all content is managed by volunteers?
- ... we support 304 languages?
- ... Wikipedia is 18 years old?
- ... Wikipedia hosts some really really weird articles?
- ... which can’t be read in Turkey (2017) nor China (2019)?
Wikimedia Projects

- Wikiquote
- Wikidata
- Wikisource
- Wikibooks
- Wikinews
- Wikivoyage
- Wikiversity
- Wiktionary
- Wikispecies
Wikimedia Infrastructure

- Open source software
- 2 Primary Data Centres
- 3 Caching Points of Presence
- ~17 billion pageviews per month*
- ~300k new editors per month
- ~1300 bare metal servers

* it’s complicated
The SRE team is a **globally distributed team** of 26 people responsible for developing and **maintaining** Wikimedia's production systems.

The Foundation has more SREs in other teams as well!
Application Layer
MediaWiki

- Our core application
- PHP, Apache, MySQL. Yes.*
  - PHP7.2 since Sept 2019
- Wiki web pages - app servers cluster
- API cluster
- Jobrunners/Videoscalers cluster

MediaWiki is a free server-based software, licensed under the GNU GPL. It is an extremely powerful, scalable software, and a feature-rich wiki implementation that uses PHP to process and display data stored in a database, such as MySQL.

* it’s complicated
Application Layer Caches
From a Monolith to Microservices
From a Monolith to Microservices

- Elasticity
- Hardware fault mitigation
- Deployments
- Migration is not easy, and still ongoing
Microservices!

- Thumbor
- Mathoid
- ORES
- Mobile Content Service (MCS)
- And many more

**Thumbor** is used for imagescaling

**Mathoid** renders LaTeX, and returns JSON with PNG, SVG or MathML renderings of the formula

**ORES** scores edits using Machine Learning (anti-vandalism effort)

**MCS** modifies page content on the fly, tailoring it for mobile
Kubernetes
Kubernetes

- Bare metal
- Calico as a CNI plugin
- Helm for deployments
- 2 clusters + 1 staging one
- Docker as a CRE

We have been running it successfully for the last 2 years! Currently, 11 services on it. Got a pipeline in the works.

Powers all mathematical formulas on Wikipedia!!!
Message Queueing
Yes, we use Apache Kafka

We are sending events like:

- wikitext templates refresh
- edge caches purging
- cross wiki links
- create new thumbnails
- re-encoding videos to open source formats

**Apache Kafka**: stream processing platform for real-time data feeds

One message queue to rule them all; started as a service for Analytics only. Now, it is our de facto solution.
Database clusters are divided into sections
Sections have masters and replicas*
MediaWiki reads from replicas and writes to master
Clusters:
- Wikitext (compressed)
- Metadata
- Parsercache

* MariaDB: fork of MySQL, migrated from MySQL in 2013*

Have a go at
https://quarry.wmflabs.org

* it’s complicated
MariaDB

- Online schema migrations*
- Cross DC replication
- TLS across all DBs
- Snapshots and local dumps for Backups

- ~570 TB total data
- ~150 DB servers
- ~350k queries per second (qps)
- ~70 TB of RAM

* it’s complicated
Elasticsearch

You guessed it right, we use it for search. That box on your top right. Run by a team surprisingly called Search Platform!
Storage
All our media are stored on Swift
- It has frontends
  ... and backends
- 1 billion objects
- ~390 TB of media!

**OpenStack Object Storage**: a scalable storage system that stores and retrieves data via HTTP
Network

The Internet

It's complicated

LVS

CDN
We have our own content delivery network

We direct traffic to a location on demand (via GeoDNS)
- Pooling/Depooling DCs
- 10 min TTL

LVS as a Layer 3/4 Linux loadbalancer*

**gdnsd**: GeoDNS is written and maintained by one of us

**peering**: interconnection with other internet networks

**Linux Virtual Server**: an advanced L3/L4 load balancing solution for Linux, supports consistent hashing

**pybal**: LVS manager, developed in-house

* it’s complicated
LVS-DR

The Internet \[\rightarrow\] It's complicated \[\rightarrow\] Gateway

server 1 \[\text{VIP}\] \[\text{in}\] \[\text{LVS}\] \[\text{VIP}\]

server 2 \[\text{VIP}\] \[\text{out}\]

server 3 \[\text{VIP}\]
Nginx-: Highly performant HTTP webserver/proxy with excellent TLS support

Varnish: Reverse HTTP caching proxy
**CDN** (coming soon)

- **ATS TLS**
  - in memory
- **ATS backend**
  - local store (SSDs)
- **ATS text**
  - HTML, CSS, JS etc
- **ATS upload**
  - media, media, media
- **ACME-chief**

**Apache Traffic Server**: Reverse and forward proxy with excellent caching support

**ACME-chief**: handles all the process of issuing and renewing Let’s Encrypt certificates (dns-01)
what happens when you type
en.wikipedia.org
Read (cached)
Read (uncached)

Internet → EDGE → NginX- → Varnish

Mх [88] → Swift

gdnsd

Application Cache Layer
Edit - Media Upload

The Internet → [EDGE] → [TLS] → [NginX] → [Varnish] → [Jobrunners] → [KAFKA] → [MW] → [Microservices] → [Swift] → [S1] → [Microservices]
Managing to Manage
Managing to Manage

- Infrastructure as code
- Configuration management
- Kubernetes
- Testing/CI/CD
- Orchestration tooling

**Puppet**: configuration management system for servers/services
- ...~50k lines of puppet code
- ...~100k lines of Ruby/ERB

**Cumin**: in-house automation and orchestration tool
In a Nutshell
Want to sell encyclopedias?

https://jobs.wikimedia.org

https://grafana.wikimedia.org/
https://github.com/wikimedia/operations-puppet
https://phabricator.wikimedia.org/
https://wikitech.wikimedia.org/

SREcon19 Dublin
@kosiaris • @manjiki