

Introducción



WIKIMEDIA
FOUNDATION

¿Quién es Jaime Crespo?

- Empleado de la Wikimedia Foundation desde hace 6 años
- Persistence Team (1 de 4), Site Reliability Engineering
- Equipo encargado de bases de datos, backups y media storage

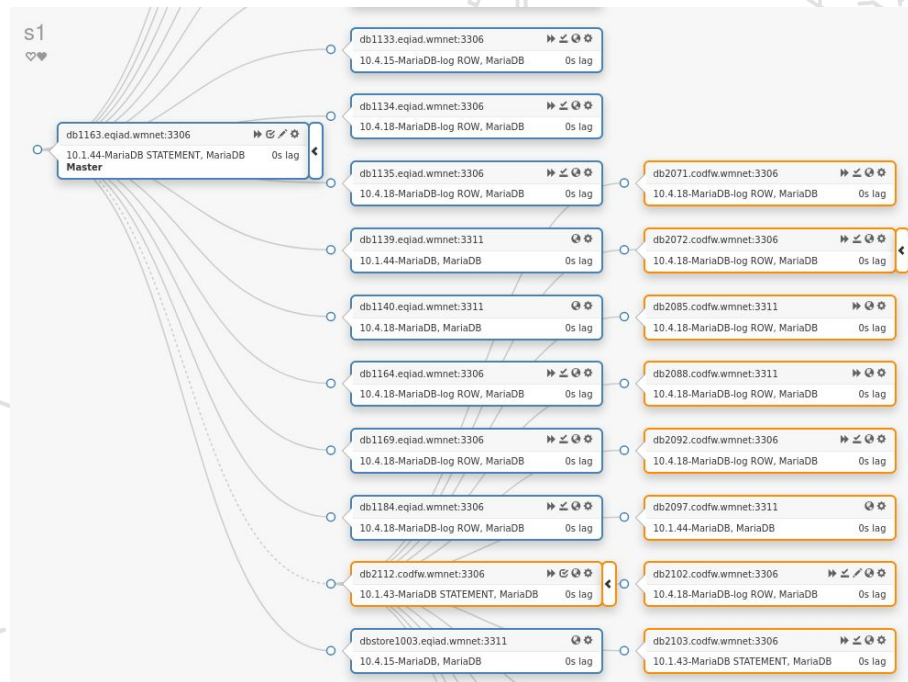


Algunos datos sobre nuestros proyectos

- Wikipedia: top 10 de páginas más populares
 - [23 mil millones de páginas vistas](#) en abril de 2021
 - Entre 100-200 mil [“GET”s por segundo](#)
- 51 millones de [ediciones](#) en Marzo de 2020
- [>300 idiomas](#) / [>900 wikis](#)
- [>6 millones de páginas wiki](#) en inglés
- [>70 millones de archivos libres](#) en Commons (300 TB)
- [Casi 100 millones](#) de “items” en Wikidata

Algunos datos de las bases de datos

- ~500 TB de datos en bases de datos relacionales
 - +300 TB en sus backups
- [316 instancias](#) de MariaDB 10.4 (y de 10.1)
- 24+ grupos de replicación
- [450 - 600 mil consultas SQL por segundo](#)



¿Qué haces?

- Mantenimiento de la infraestructura de bases de datos y backups
 - Monitorización de salud de servicios
 - Respuesta a incidentes
- Compra y setup de equipos
- Automatización de tareas
- Consultoría en temas de rendimiento y seguridad
 - Optimización de código y consultas SQL
 - Despliegue de cambios de esquema

¿Cómo has terminado aquí?

- Miembro fundador de Púlsar, Linux geek
- Prácticas de empresa en BIFI
- Oficina de Software Libre de UZ
- Instructor/Desarrollador @Warp Networks
- Instructor/Consultor MySQL @Percona
- Consultor independiente

Sobre esta charla

¡Disclaimer! Esto refleja solo nuestra experiencia, que puede variar mucho en otros entornos

- Enfocada a estudiantes de grado de la asignatura de BBDD
- Nuestro entorno:
 - Software libre
 - Web
 - OLTP
 - “Grandes” clusters
 - Metal
 - Maximizar rendimiento/\$

Diseñando el esquema de las bases de datos de Wikipedia



WIKIMEDIA
FOUNDATION

Diseñando un esquema para Wikipedia

- ¿Qué entidades necesitaríamos?
- ¿Qué relaciones?
- ¿Cuáles serían las consultas más frecuentes?
- ¿Cuáles crees que son los principales problemas de escalado?

Funcionalidad Básica



BIQUIPEDIA
A enciclopedia libre

- Portalada
- A tabierna
- Actualidat
- Zaguers cambeos
- Una pachina a l'azar
- Aduya
- Donativos

Ferramientas

- Pachinas que enlazan con ista
- Cambios relacionatos
- Cargar fichero
- Pachinas especiales
- Información d'a pachina
- Elemento de Wikidata

En otros proyectos

Wikimedia Commons

En atras luengas

JCrespo (WMF) [Discusión](#) [Preferencias](#) [Beta](#) [Lista de seguimiento](#) [Contribuciones](#) [Salir d'a sesión](#)

Pachina **Discusión** [Leyer](#) [Editar](#) [Modificar codigo](#) [Amostrar l'historial](#) Más ▾

Párrafo ▾ Citar ▾ Insertar ▾ ? **Publicar cambios...**

Universidat de Zaragoza

(Reendrezau dende [Universidad de Zaragoza](#))

A **Universidat de Zaragoza** ye a única universidat d'Aragón, con seu en Zaragoza y campus en Zaragoza, Uesca, Chaca, Teruel y L'Almudena de la Reina Godina.

Establita en 1542, aculle actualment (2004) amás de 32.000 estudiants y 3.500 miembros docents, en 22 centros y 60 titulacions.^[1] L'actual rector ye José Antonio Mayoral Murillo.

Historia

L'orichen d'a Universidat en o territorio que hue clamamos Aragón gosa situar-se en a escuela de letras que o cheneral romano y pretor uescano Quinto Sertorio creyó en o sieglo I a. C. en Osc ta os fillos d'os mainates locals. Ista escuela se considera a simient d'a futura universidat medieval. En 1354, Pero IV d'Aragón decreta a creyación d'a primer universidat d'o reino: a Universidat Literaria de Uesca, conoixita posteriorment como Universidat Sertoriana de Uesca. «En Zaragoza, as radices más viellas d'a Universidat se troban en l'antiga escuela catedralicia de Zaragoza, creyata en o sieglo VII por o bispe Braulio, que dimpués fue nombrato patrón d'a Universidat de Zaragoza. Entre 1474 y 1476, o prencipe don Ferrando de Aragón, posteriorment Ferrando II d'Aragón torna o Estudio d'Artes de Zaragoza en un Estudio Cheneral d'Artes. Ya en o sieglo XVI Carlos I sinya en as Cortz de Monzón (

Universidat de Zaragoza
Universidad de Zaragoza



Universidad
Zaragoza



Entidad básica: Página

- page_title
- page_content
- Otros atributos? (length, timestamp, ...)

```
INSERT INTO page (page_title, page_content)
VALUES (
    'Universidat_de_Zaragoza',
    '{{Institución academica\n|nombre = Universidat de Zaragoza ...'
);
```

```
SELECT page_content
FROM page
WHERE page_title = 'Universidat_de_Zaragoza';
```

Añadiendo revisiones



BIQUIPEDIA
A enciclopedia libre

Portalada
A tabierna
Actualidat
Zaguers cambeos
Una pachina a l'azar
Aduya
Donativos

Ferramientas

Pachinas que enlazan con ista
Cambios relacionatos
Atom
Cargar fichero
Pachinas especiales
Información d'a pachina
Elemento de Wikidata

En atras luengas

JCrespo (WMF) [Discusión](#) [Preferencias](#) [Beta](#) [Lista de seguimiento](#) [Contribucions](#) [Salir d'a sesión](#)

Pachina

[Discusión](#)

[Leyer](#)

[Editar](#)

[Modificar codigo](#)

[Amostrar l'histrial](#)



Más ▾

Mirar en Wikipedia



Historial de versiones de "Universidat de Zara"

[Aduya](#)

[Veyer os rechistros d'ista pachina \(ver registro de abusos\)](#)

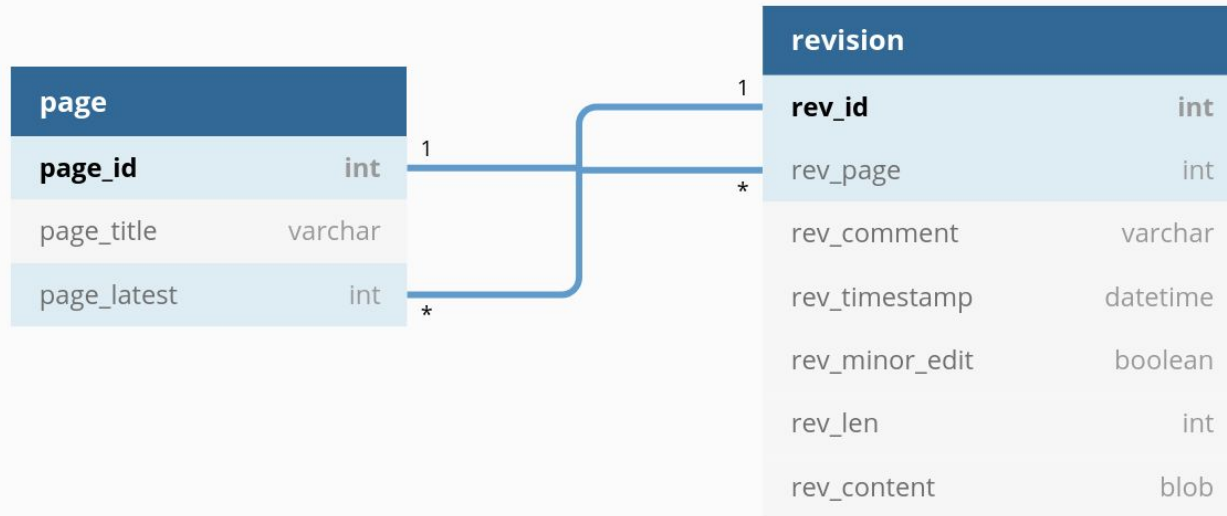
▼ Filtrar as revisiones

Selección de diferencias: sinyale as versiones a comparar y prete "enter" u o botón .
Leyenda: **(act)** = esferencias con a versión actual, **(ant)** = esferencias con a versión anterior, **m** = edición menor
(zagueras | primeras) Veyer (50 más recientes | 50 más antigas) (20 | 50 | 100 | 250 | 500)

Confrontar as versions trigadas

- [\(act | ant\)](#) 13:06 20 ago 2016 [Abián](#) ([Descutir](#) | [contrebuacions](#)) .. (9396 bytes) (+6) .. *(Logo vectorial e imagen más representativa)* ([desfer](#) | [agradeixer](#))
- [\(act | ant\)](#) 23:04 17 mar 2016 [85.251.33.253](#) ([Descutir](#)) .. (9390 bytes) (+28) .. ([desfer](#))
- [\(act | ant\)](#) 17:16 16 nov 2015 [Chusec](#) ([Descutir](#) | [contrebuacions](#)) **m** .. (9362 bytes) (0) .. ([desfer](#) | [agradeixer](#)) (Etiqueta: **Edición visual**)
- [\(act | ant\)](#) 17:12 16 nov 2015 [Chusec](#) ([Descutir](#) | [contrebuacions](#)) **m** .. (9362 bytes) (+70) .. *(correccions, vinclos)* ([desfer](#) | [agradeixer](#)) (Etiqueta: **Edición visual**)
- [\(act | ant\)](#) 17:04 16 nov 2015 [Chusec](#) ([Descutir](#) | [contrebuacions](#)) .. (9292 bytes) (+2) .. *(corrección ta vinclo)* ([desfer](#) | [agradeixer](#))
- [\(act | ant\)](#) 16:41 10 feb 2014 [EBRO](#) ([Descutir](#) | [contrebuacions](#)) .. (9290 bytes) (+4) .. ([desfer](#) | [agradeixer](#))

revision y su relación con page



Añadiendo usuarios



WIKIPEDIA
The Free Encyclopedia

- [Main page](#)
- [Contents](#)
- [Current events](#)
- [Random article](#)
- [About Wikipedia](#)
- [Contact us](#)
- [Donate](#)
- [Contribute](#)
- [Help](#)
- [Learn to edit](#)
- [Community portal](#)
- [Recent changes](#)
- [Upload file](#)
- [Tools](#)
- [Upload file](#)
- [Special pages](#)
- [Printable version](#)

JCrespo (WMF) [Talk](#) [Sandbox](#) [Preferences](#) [Beta](#) [Watchlist](#) [Contributions](#) [Log out](#)

Special page

Search Wikipedia



Preferences

[Help](#)

For information about the settings on this page, see [User preferences](#).

User profile

Appearance

Editing

Recent changes

Watchlist

Gadgets

Search

Banners

Beta features

Notifications

Basic information

Username: JCrespo (WMF)

Member of groups: [Autoconfirmed users](#), [Users](#)

Connected apps: [Manage 5 connected applications](#)

Number of edits: 91

Registration time: 21:29, 8 May 2015

Password: [Change password](#)

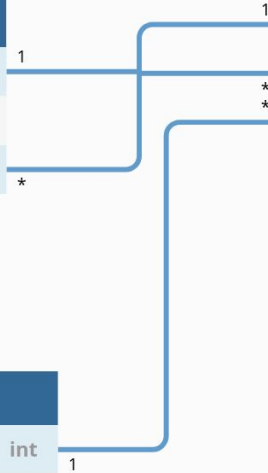
Account recovery is enabled. See [Email options](#) for more settings.

user y su relación con revision

page	
page_id	int
page_title	varchar
page_latest	int

user	
user_id	int
user_name	varchar
user_password	varchar
user_email	varchar
user_editcount	int

revision	
rev_id	int
rev_page	int
rev_user	int
rev_comment	varchar
rev_timestamp	datetime
rev_minor_edit	boolean
rev_len	int
rev_content	blob



Optimización física I: Cambios recientes



WIKIPEDIA
The Free Encyclopedia

Main page
Contents
Current events
Random article
About Wikipedia
Contact us
Donate

Contribute

Help
Learn to edit
Community portal
Recent changes
Upload file

Tools

Atom
Special pages
Printable version

JCrespo (WMF) Talk Sandbox Preferences Beta Watchlist Contributions Log out

Special page

Search Wikipedia

Recent changes

Help

This is a list of recent changes to Wikipedia.

Other review tools

Saved filters

Active filters

Hide

Human (not bot) Page edits Page creations Logged actions



Filter changes (use menu or search for filter name)

Namespaces Tags

Live updates

View new changes since 10:33, 16 May 2021

500 changes, 30 days

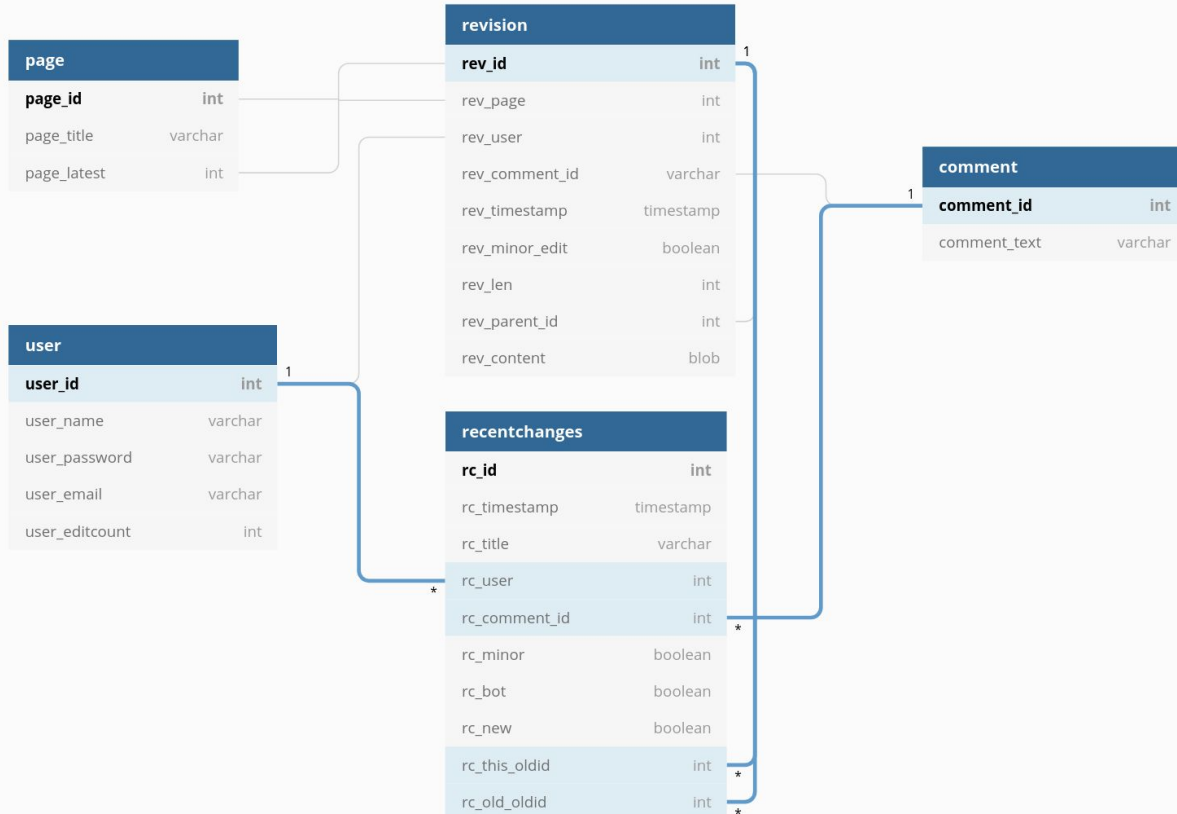
- ... BBC Three (online); 10:33 .. (+58) .. Dannycreaire (talk | contribs)
(→Approval and launch) (Tags: Mobile edit, Mobile web edit, Advanced mobile edit)
- ... Manchu Restoration; 10:33 .. (+28) .. 49.195.227.54 (talk) (Tags: Mobile edit, Mobile web edit)
- ... m Cherier; 10:33 .. (+261) .. Markusep (talk | contribs)
- ... 1995 WTA Tour Championships; 10:33 .. (-1) .. Wolbo (talk | contribs)

List of abbreviations (help):[hide]

- D** Edit made at Wikidata
- r** Edit flagged by ORES
- N** New page
- m** Minor edit
- b** Bot edit
- (±123)** Page byte size change
- Temporarily watched page



Entidad recentchanges



Optimización física II: Contenido

```
SELECT * FROM text LIMIT 1\G
***** 1. Row *****
      text_id: 1
      text_content: DB://cluster3/1
      text_flags: utf-8,gzip,external
1 row in set (0.032 sec)
```

revision	
rev_id	int
rev_page	int
rev_user	int
rev_comment_id	varchar
rev_timestamp	timestamp
rev_minor_edit	boolean
rev_len	int
rev_parent_id	int
rev_text_id	bigint

text	
text_id	bigint
text_content	blob
text_flags	tinyblob

1

1

Resultado final (I)

<p>User</p> <ul style="list-style-type: none">user<ul style="list-style-type: none">user_id INTuser_name VARCHAR(255)user_real_name VARCHAR(255)user_password TINYBLOBuser_newpassword TINYBLOBuser_newpass_time BINARY(14)user_email TINYTEXTuser_touched BINARY(14)user_token BINARY(32)user_email_authenticated BINARY(14)user_email_token BINARY(32)user_email_token_expires BINARY(14)user_registration BINARY(14)user_editcount INTuser_password_expires VARBINARY(14)user_properties<ul style="list-style-type: none">up_user INTup_property VARBINARY(255)up_value BLOBuser_newtalk<ul style="list-style-type: none">user_id INTuser_ip VARBINARY(40)user_last_timestamp VARBINARY(14)bot_passwords<ul style="list-style-type: none">bp_user INTbp_app_id VARBINARY(32)bp_password TINYBLOBbp_token BINARY(32)bp_restrictions BLOBbp_grants BLOB	<p>Permissions</p> <ul style="list-style-type: none">user_groups<ul style="list-style-type: none">ug_user INTug_group VARBINARY(255)ug_expiry VARBINARY(14)page_restrictions<ul style="list-style-type: none">pr_id INTpr_page INTpr_type VARBINARY(60)pr_level VARBINARY(60)pr_cascade TINYINTpr_user INTpr_expiry VARBINARY(14)protected_titles<ul style="list-style-type: none">pt_namespace INTpt_title VARCHAR(255)pt_user INTpt_reason_id BIGINTpt_timestamp BINARY(14)pt_expiry VARBINARY(14)pt_create_perm VARBINARY(60)iplocks<ul style="list-style-type: none">ipb_id INTipb_address TINYBLOBipb_user INTipb_by_actor BIGINTipb_reason_id BIGINTipb_timestamp BINARY(14)ipb_auto BOOLipb_anon_only BOOLipb_create_account BOOLipb_enable_autoblock BOOLipb_expiry VARBINARY(14)ipb_range_start TINYBLOBipb_range_end TINYBLOBipb_deleted BOOLipb_block_email BOOLipb_allow_undertalk BOOLipb_parent_block_id INTipb_sitewide BOOLiplocks_restrictions<ul style="list-style-type: none">ir_ipb_id INTir_type TINYINT(1)ir_value INT	<p>Logging</p> <ul style="list-style-type: none">logging<ul style="list-style-type: none">log_id INTlog_type VARBINARY(32)log_action VARBINARY(32)log_timestamp BINARY(14)log_actor BIGINTlog_namespace INTlog_title VARCHAR(255)log_page INTlog_comment_id BIGINTlog_params BLOBlog_deleted TINYINTlog_search<ul style="list-style-type: none">ls_field VARBINARY(32)ls_value VARCHAR(255)ls_log_id INTcomment<ul style="list-style-type: none">comment_id BIGINTcomment_hash INTcomment_text BLOBcomment_data BLOB	<p>Tags</p> <ul style="list-style-type: none">change_tag<ul style="list-style-type: none">ct_id INTct_rc_id INTct_log_id INTct_rev_id INTct_params BLOBct_tag_id INTchange_tag_def<ul style="list-style-type: none">ctd_id INTctd_name VARBINARY(255)ctd_user_defined TINYINT(1)ctd_count BIGINT
<p>Recent changes</p> <ul style="list-style-type: none">recentchanges<ul style="list-style-type: none">rc_id INTrc_timestamp VARBINARY(14)rc_actor BIGINTrc_namespace INTrc_title VARCHAR(255)rc_comment_id BIGINTrc_minor TINYINTrc_bot TINYINTrc_new TINYINTrc_cur_id INTrc_this_oldid INTrc_last_oldid INTrc_type TINYINTrc_source VARCHAR(16)rc_patrolled TINYINTrc_ip VARBINARY(40)rc_old_len INTrc_new_len INTrc_deleted TINYINTrc_logid INTrc_log_type VARBINARY(255)rc_log_action VARBINARY(255)rc_params BLOBwatchlist<ul style="list-style-type: none">wl_id INTwl_user INTwl_namespace INTwl_title VARCHAR(255)wl_notificationtimestamp VARBINARY(14)	<p>Pages</p> <ul style="list-style-type: none">page<ul style="list-style-type: none">page_id INTpage_namespace INTpage_title VARCHAR(255)page_restrictions TINYBLOBpage_is_redirect TINYINTpage_is_new TINYINTpage_random REALpage_touched BINARY(14)page_links_updated VARBINARY(14)page_latest INTpage_len INTpage_content_model VARBINARY(32)page_lang VARBINARY(35)page_props<ul style="list-style-type: none">pp_page INTpp_propname VARBINARY(60)pp_value BLOBpp_sortkey FLOATarchive<ul style="list-style-type: none">ar_id INTar_namespace INTar_title VARCHAR(255)ar_restrictions TINYBLOBar_actor BIGINTar_timestamp BINARY(14)ar_minor_edit TINYINTar_rev_id INTar_deleted TINYINTar_len INTar_page_id INTar_parent_id INTar_sha1 VARBINARY(32)redirect<ul style="list-style-type: none">rd_from INTrd_namespace INTrd_title VARCHAR(255)rd_interwiki VARCHAR(32)rd_fragment VARCHAR(255)category<ul style="list-style-type: none">cat_id INTcat_title VARCHAR(255)cat_pages INTcat_subcats INTcat_files INT		

https://www.mediawiki.org/wiki/Manual:Database_layout

Resultado final (II)

Revisions

revision

- rev_id INT
- rev_page INT
- rev_comment_id BIGINT
- rev_actor BIGINT
- rev_timestamp BINARY(14)
- rev_minor_edit TINYINT
- rev_deleted TINYINT
- rev_len INT
- rev_parent_id INT
- rev_sha1 VARBINARY(32)

revision_comment_temp

- revcomment_rev INT
- revcomment_comment_id BIGINT

revision_actor_temp

- revactor_rev INT
- revactor_actor BIGINT
- revactor_timestamp BINARY(14)
- revactor_page INT

slots

- slot_revision_id BIGINT
- slot_role_id SMALLINT
- slot_content_id BIGINT
- slot_origin BIGINT

slot_roles

- role_id SMALLINT
- role_name VARBINARY(64)

ip_changes

- ipc_rev_id INT
- ipc_rev_timestamp BINARY(14)
- ipc_hex VARBINARY(35)

content

- content_id BIGINT
- content_size INT
- content_sha1 VARBINARY(32)
- content_model SMALLINT
- content_address VARBINARY(255)

content_models

- model_id SMALLINT
- model_name VARBINARY(64)

text

- old_id INT
- old_text MEDIUMBLOB
- old_flags TINYBLOB

Link tables

pagelinks

- pl_from INT
- pl_from_namespace INT
- pl_namespace INT
- pl_title VARCHAR(255)

iwlinks

- iwl_from INT
- iwl_prefix VARBINARY(20)
- iwl_title VARCHAR(255)

externallinks

- el_id INT
- el_from INT
- el_to BLOB
- el_index BLOB
- el_index_60 VARBINARY(60)

langlinks

- ll_from INT
- ll_lang VARBINARY(35)
- ll_title VARCHAR(255)

imagelinks

- il_from INT
- il_from_namespace INT
- il_to VARCHAR(255)

templatelinks

- tl_from INT
- tl_from_namespace INT
- tl_namespace INT
- tl_title VARCHAR(255)

categorylinks

- cl_from INT
- cl_to VARCHAR(255)
- cl_sortkey VARBINARY(230)
- cl_sortkey_prefix VARCHAR(255)
- cl_timestamp TIMESTAMP
- cl_collation VARBINARY(32)
- cl_type ENUM(...)

Statistics

site_stats

- ss_row_id INT
- ss_total_edits BIGINT
- ss_good_articles BIGINT
- ss_total_pages BIGINT
- ss_users BIGINT
- ss_active_users BIGINT
- ss_images BIGINT

Search

searchindex

- si_page INT
- si_title VARCHAR(255)
- si_text MEDIUMTEXT

Maintenance

job

- job_id INT
- job_cmd VARBINARY(60)
- job_namespace INT
- job_title VARCHAR(255)
- job_timestamp VARBINARY(14)
- job_params MEDIUMBLOB
- job_random INTEGER
- job_attempts INTEGER
- job_token VARBINARY(32)
- job_token_timestamp VARBINARY(14)
- job_sha1 VARBINARY(32)

Multimedia

image

- img_name VARCHAR(255)
- img_size INT
- img_width INT
- img_height INT
- img_metadata MEDIUMBLOB
- img_bits INT
- img_media_type ENUM(...)
- img_major_mime ENUM(...)
- img_minor_mime VARBINARY(100)
- img_description_id BIGINT
- img_actor BIGINT
- img_timestamp VARBINARY(14)
- img_sha1 VARBINARY(32)

oldimage

- oi_name VARCHAR(255)
- oi_archive_name VARCHAR(255)
- oi_size INT
- oi_width INT
- oi_height INT
- oi_bits INT
- oi_description_id BIGINT
- oi_actor BIGINT
- oi_timestamp BINARY(14)
- oi_metadata MEDIUMBLOB
- oi_media_type ENUM(...)
- oi_major_mime ENUM(...)
- oi_minor_mime VARBINARY(100)
- oi_deleted TINYINT
- oi_sha1 VARBINARY(32)

filearchive

- fa_id INT
- fa_name VARCHAR(255)
- fa_archive_name VARCHAR(255)
- fa_storage_group VARBINARY(16)
- fa_storage_key VARBINARY(64)
- fa_deleted_user INT
- fa_deleted_timestamp BINARY(14)
- fa_deleted_reason_id BIGINT
- fa_size INT
- fa_width INT
- fa_height INT
- fa_metadata MEDIUMBLOB
- fa_bits INT
- fa_media_type ENUM(...)
- fa_major_mime ENUM(...)
- fa_minor_mime VARBINARY(100)
- fa_description_id BIGINT
- fa_actor BIGINT
- fa_timestamp BINARY(14)
- fa_deleted TINYINT
- fa_sha1 VARBINARY(32)

uploadstash

- us_id INT
- us_user INT
- us_key VARCHAR(255)
- us_orig_path VARCHAR(255)
- us_path VARCHAR(255)
- us_source_type VARCHAR(50)
- us_timestamp VARBINARY(14)
- us_status VARCHAR(50)
- us_chunk_inx INT
- us_props BLOB
- us_size INT
- us_sha1 VARCHAR(31)
- us_mime VARCHAR(255)
- us_media_type ENUM(...)
- us_image_width INT
- us_image_height INT
- us_image_bits SMALLINT

Interwiki

sites

- site_id INT
- site_global_key VARBINARY(64)
- site_type VARBINARY(32)
- site_group VARBINARY(32)
- site_source VARBINARY(32)
- site_language VARBINARY(35)
- site_protocol VARBINARY(32)
- site_domain VARCHAR(255)
- site_data BLOB
- site_forward BOOL
- site_config BLOB

interwiki

- iw_prefix VARCHAR(32)
- iw_url BLOB
- iw_api BLOB
- iw_wikiid VARCHAR(64)
- iw_local BOOL
- iw_trans TINYINT

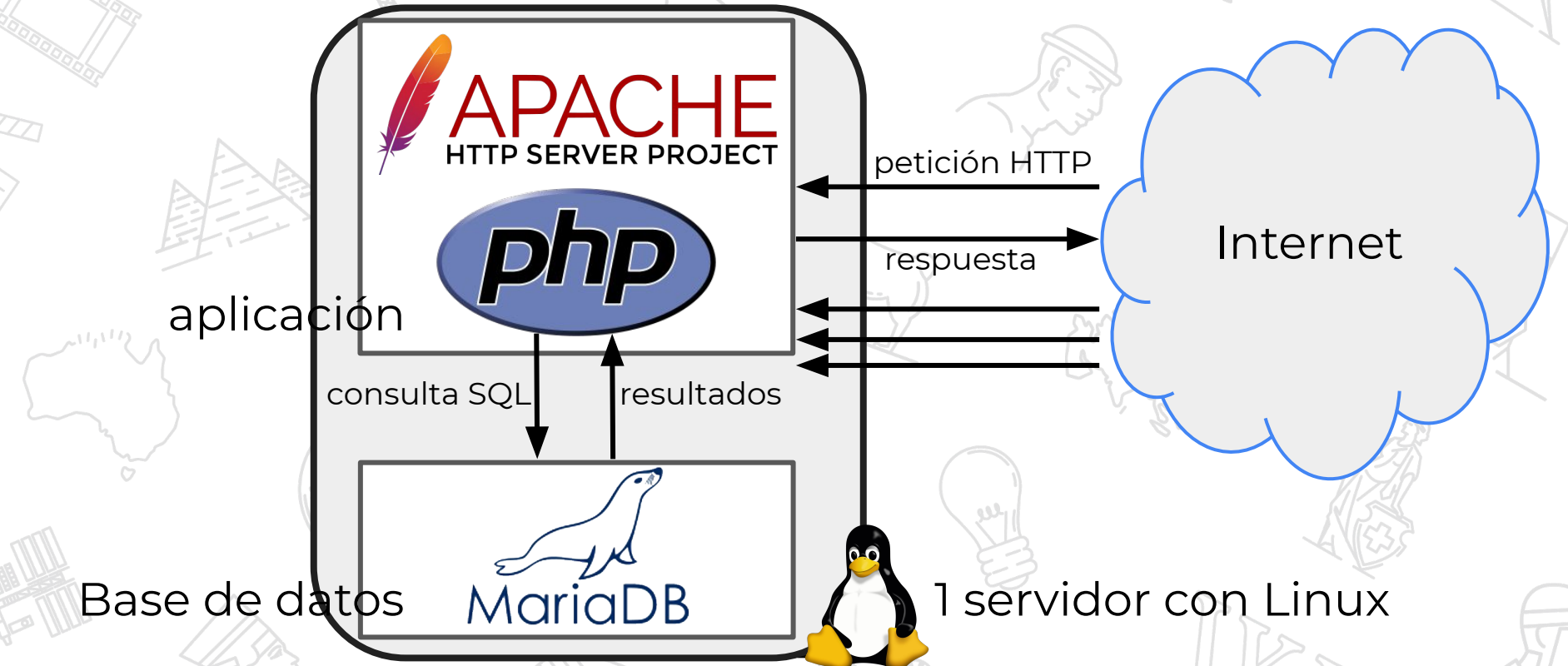
https://www.mediawiki.org/wiki/Manual:Database_layout

Escalando la arquitectura de las bases de datos de Wikipedia



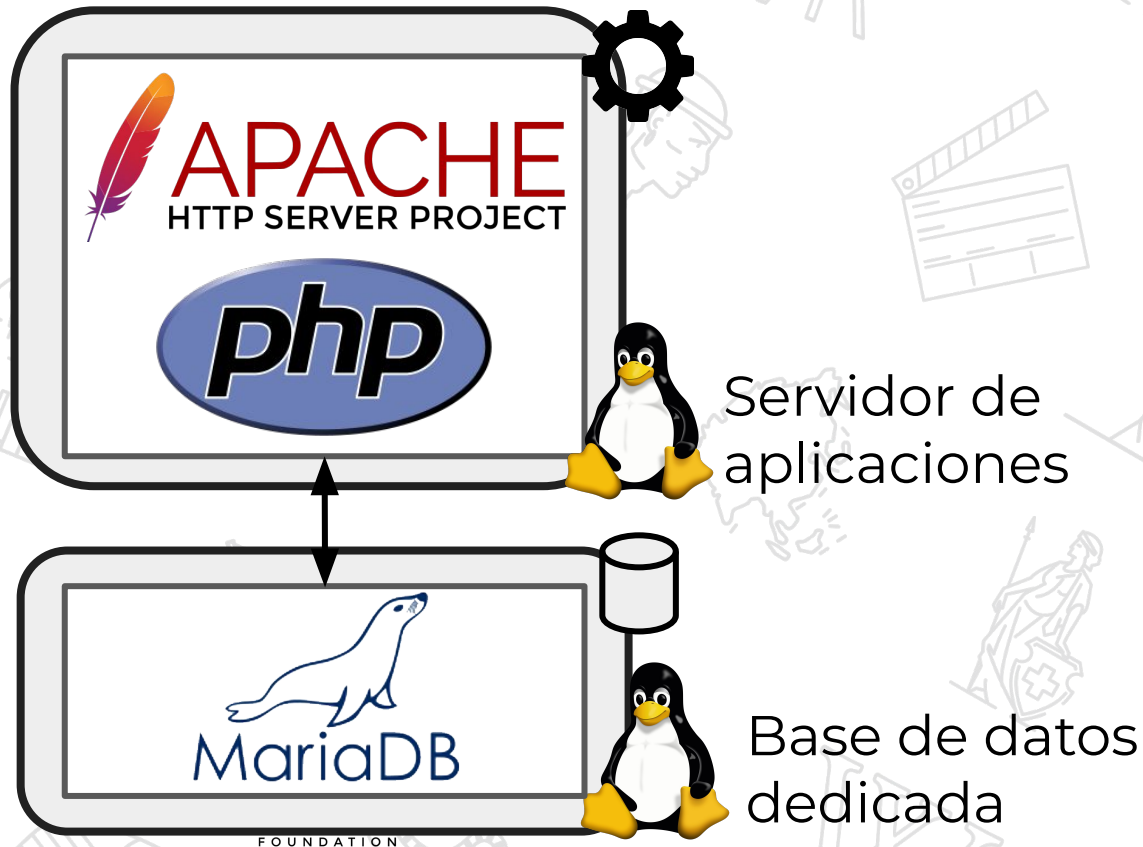
WIKIMEDIA
FOUNDATION

Primera posible implementación*

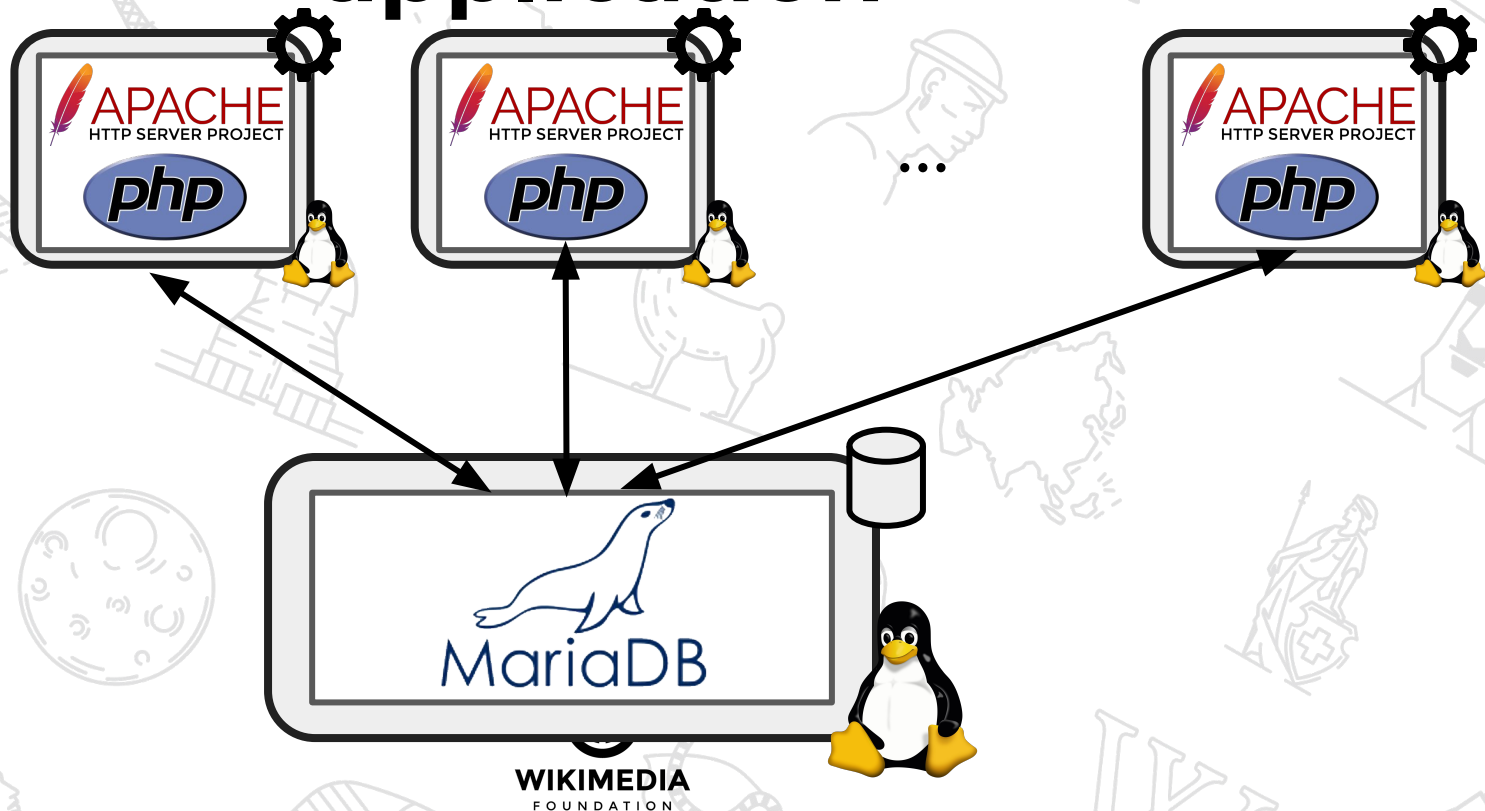


* En su primera versión, Wikipedia usaba Perl + archivos locales, posteriormente, PHP + MySQL: https://www.mediawiki.org/wiki/MediaWiki_history

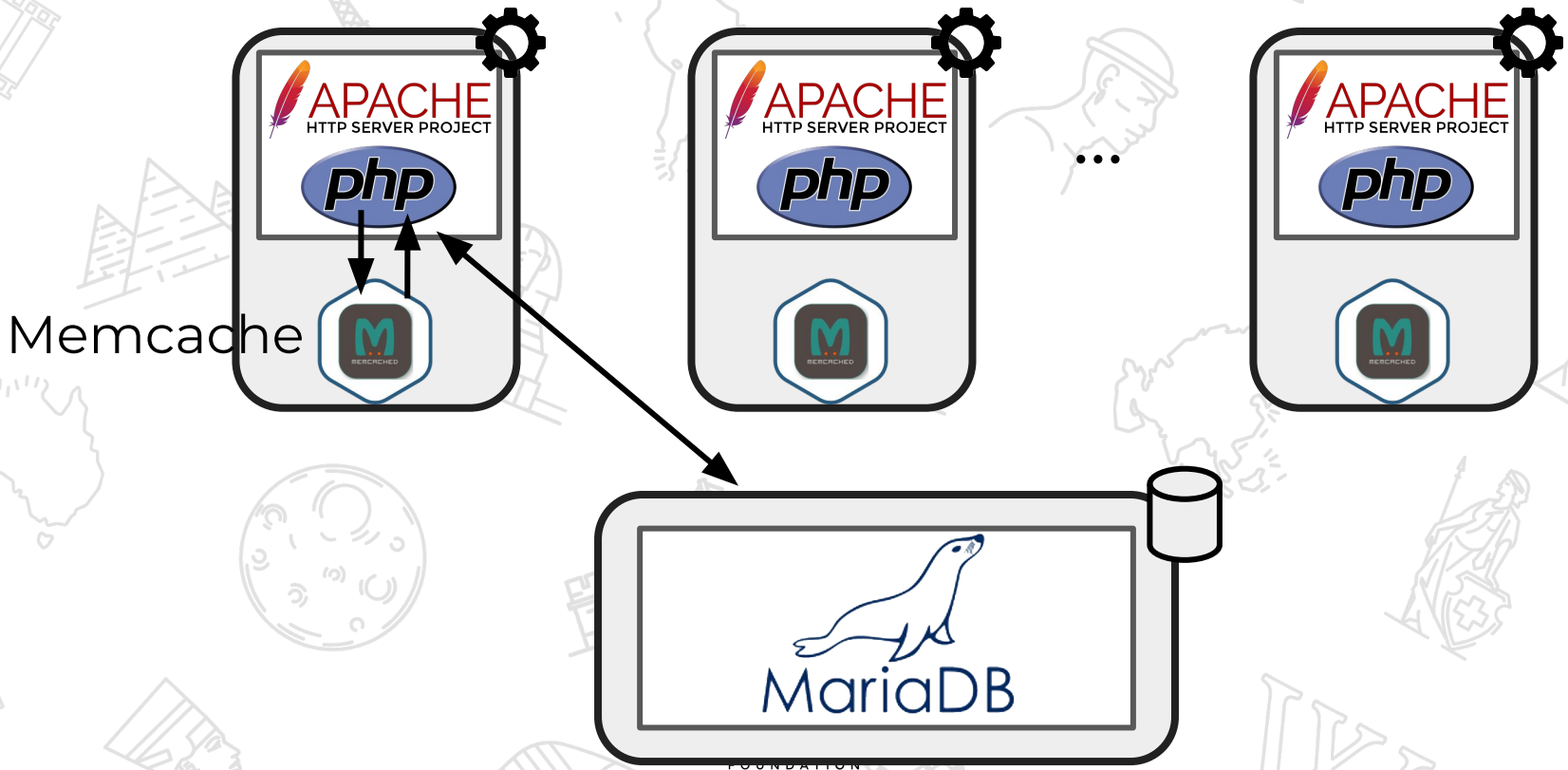
¿Cómo escalar (I)?: BBDD dedicada



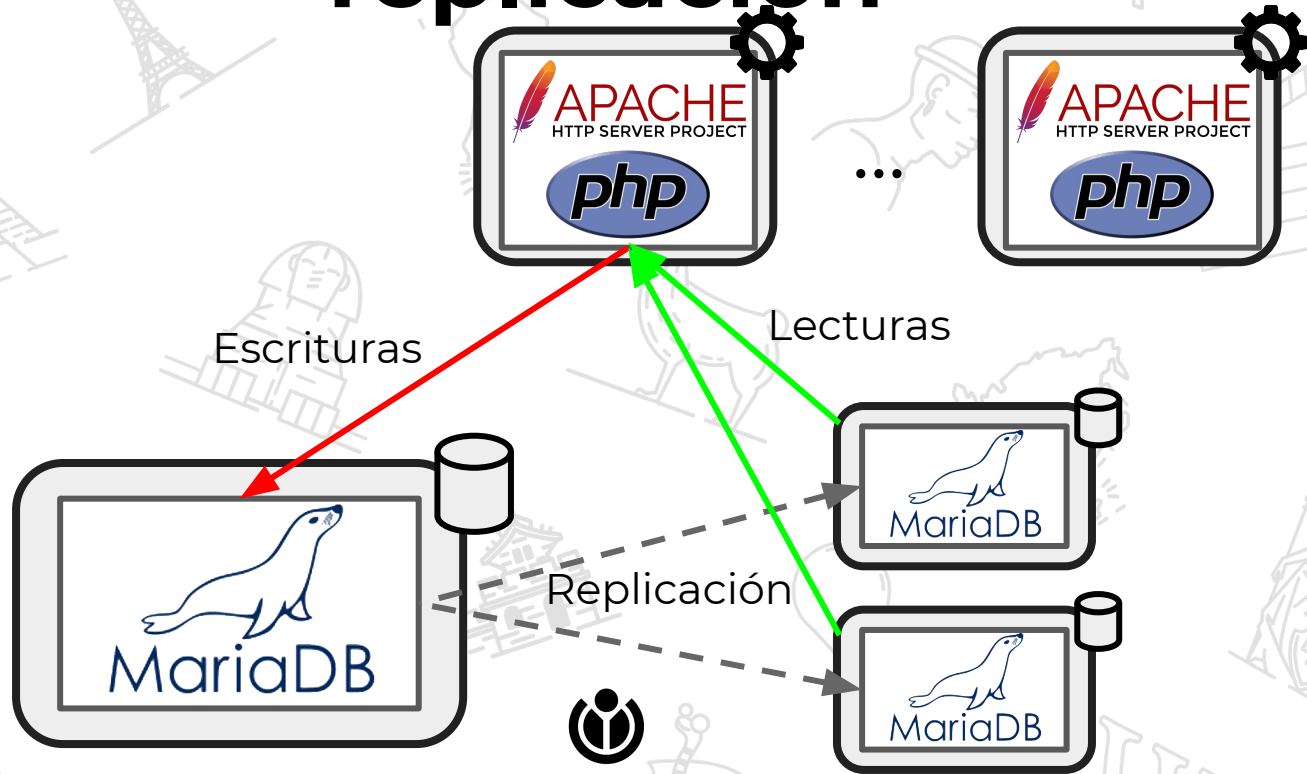
¿Cómo escalar (II)?: Stateless application



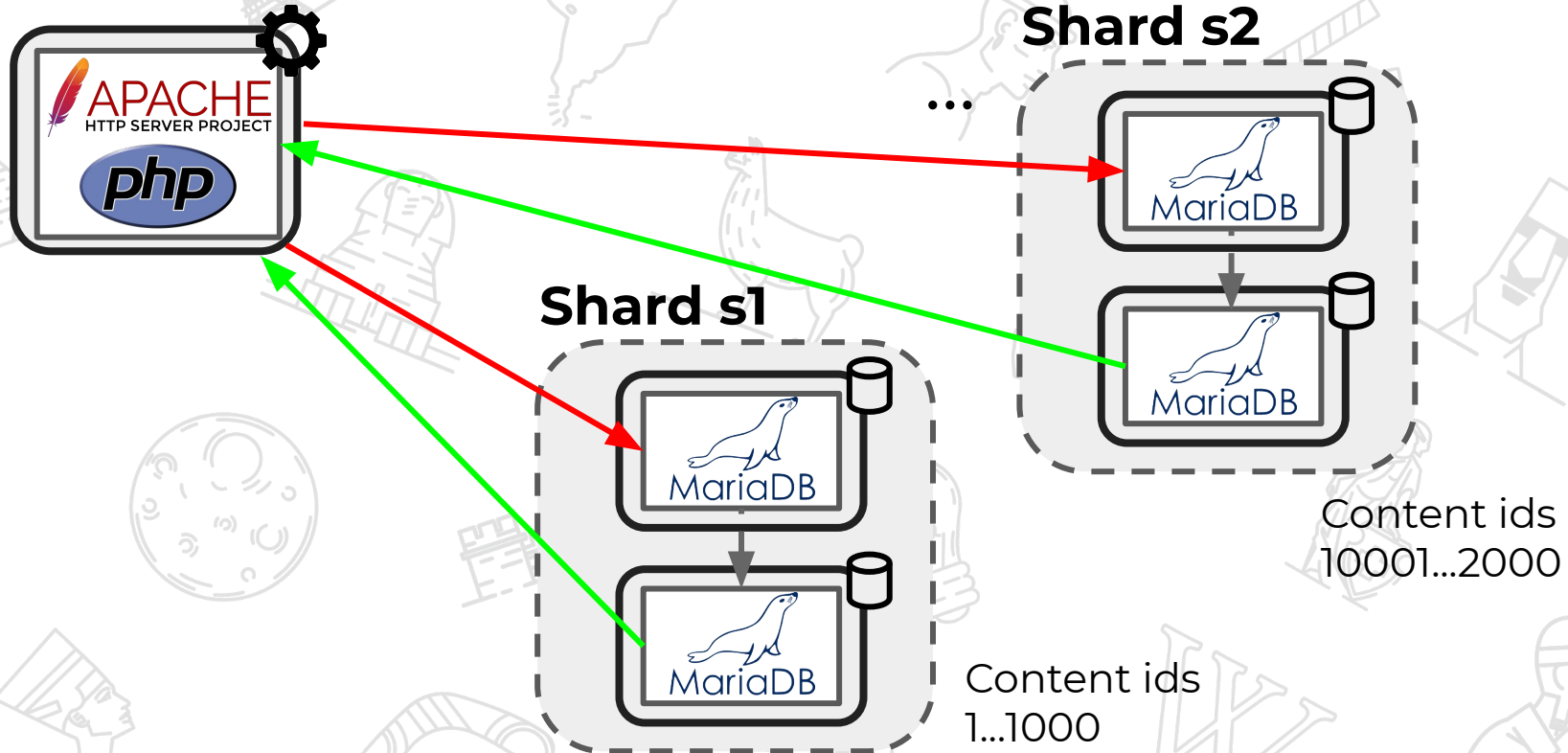
¿Cómo escalar (III)?: Lecturas - Cache



¿Cómo escalar (IV)?: Lecturas - replicación

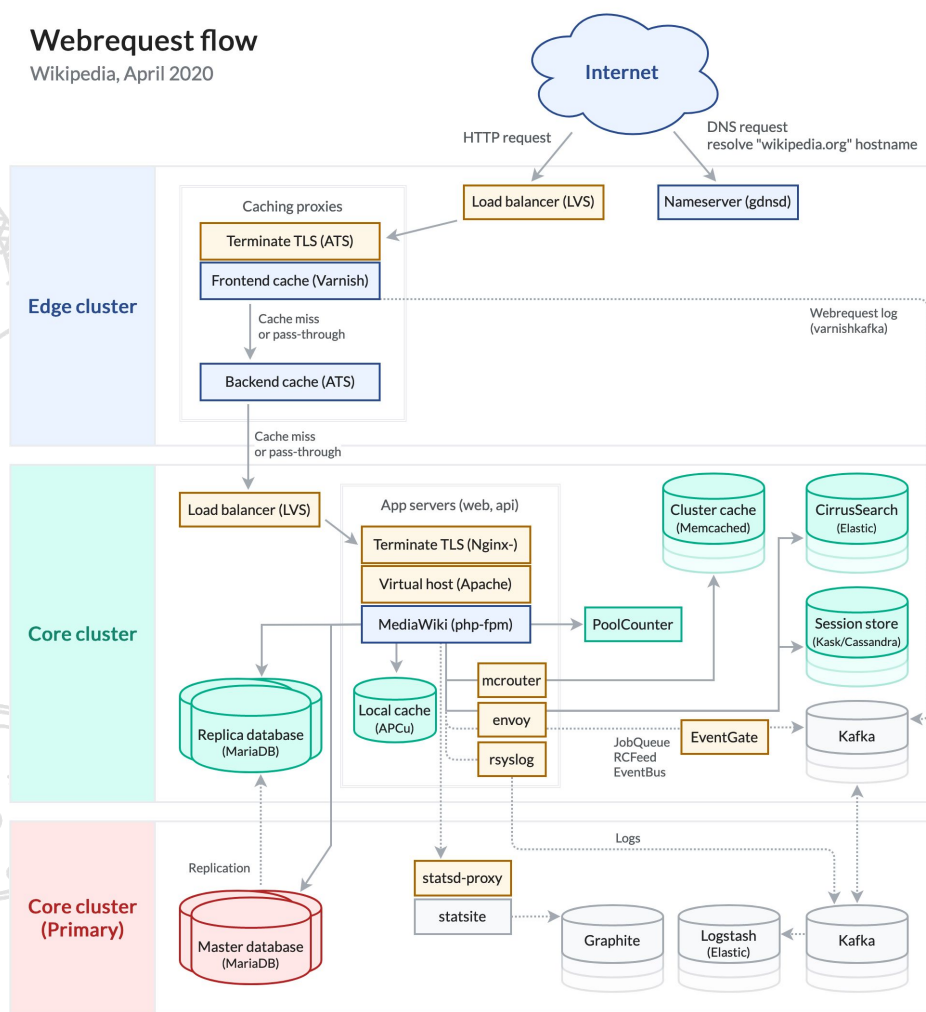


¿Cómo escalar (V)? Escrituras - Sharding



Webrequest flow

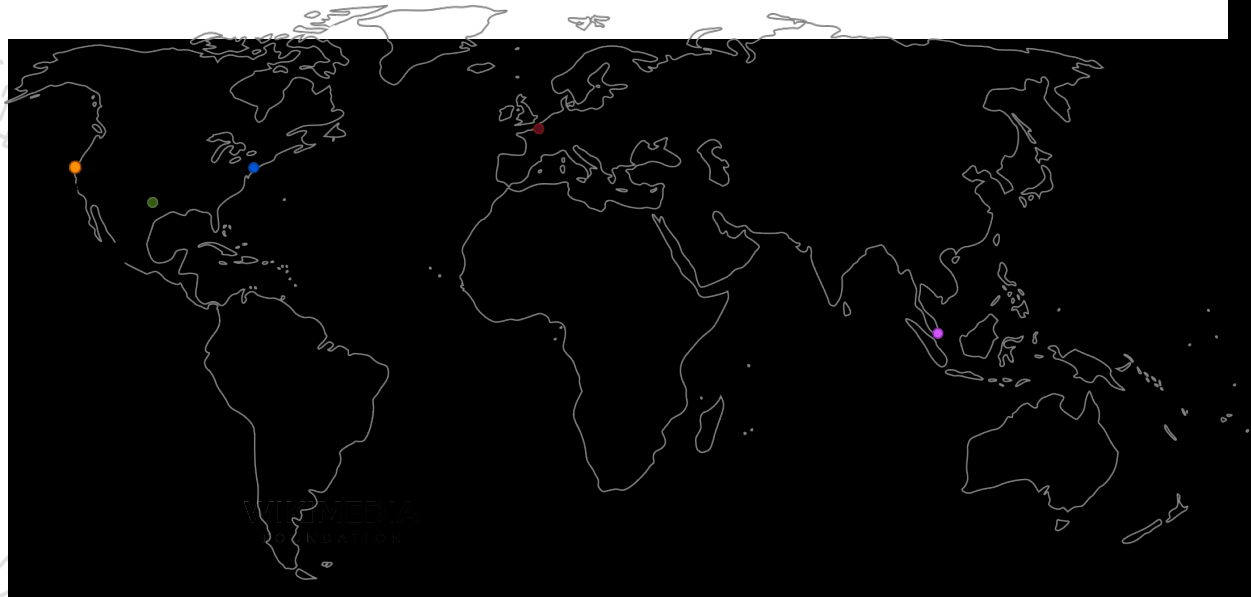
Wikipedia, April 2020



Legend: Blocking connect Async or post-send Storage Proxy Responds to user Responds to MW

Arquitectura física

- +1400 servidores físicos
- 5 centros de datos, 2 primarios



Hardware & configuración BD

- Servidores x86_64 (2P * 8C)
- 512 GB of RAM
- 10 * 2TB SSDs en RAID10
- BBDD entre 90GB y 12TB)
- 1 servidor puede realizar cómodamente 20K QPS
- Configuración MariaDB:
https://phabricator.wikimedia.org/source/operations-puppet/browse/production/modules/role/templates/mariadb/mysql_config/production.my.cnf.erb?blame=off



Pila de software

- **Debian Linux** Stretch/Buster (os)
- Linux Virtual Server (load balancing)
- ATS (cache)
- **Mediawiki** on top of PHP7 (application servers)
- Memcache
- Kafka (jobqueue, messaging)
- Elasticsearch (indexing & search)
- **MariaDB** 10.1/10.4 (data)
- OpenStack Swift (media)

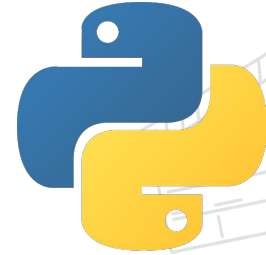
Otro software

- Puppet (configuration management)
- Git (code versioning)
- Hadoop (analytics)
- Cassandra (sessions/APIs)
- Jenkins (CI)
- K8s (orchestrator)
- Icinga + Prometheus/Grafana (monitoring)
- HAProxy (proxy)
- Gdnsd (DNS)

Lenguajes de programación

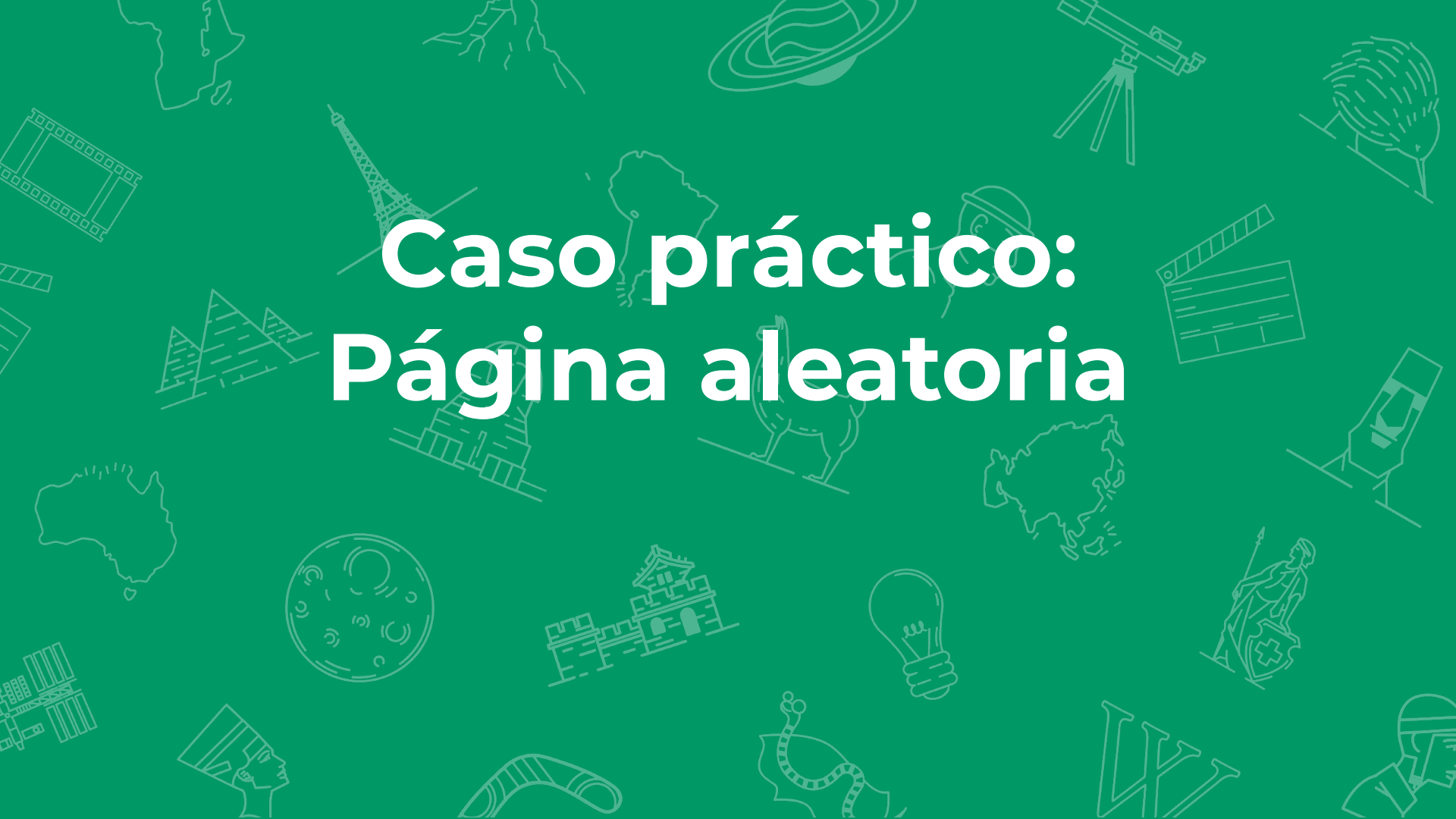
- ¡TODOS!
 - PHP
 - SQL
 - Java
 - Js/Node
 - Python
 - Ruby
 - Bash
 - Lua
 - SPARQL
 - C/C++

...



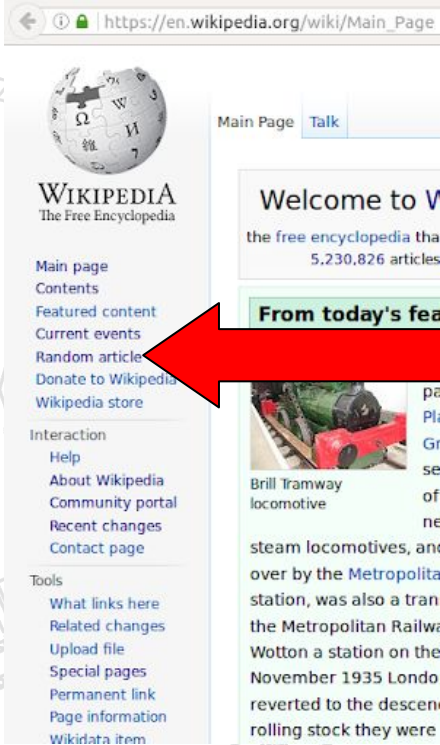
BASH
THE BOURNE-AGAIN SHELL

Caso práctico: Página aleatoria

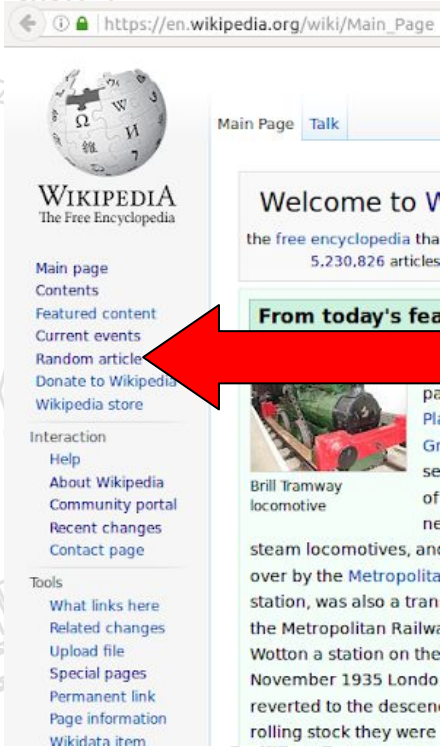


Descripción del caso

- A la izquierda de cada página hay un enlace a “Página aleatoria”
- Debe permitir filtrar por “espacio de nombres” (p.ej.: no todas las páginas son artículos)
- Es bastante importante, comparado con “voy a tener suerte” de Google, ya que tiene que dar una idea de los artículos a encontrar en una Wiki



Restricciones del caso



- Tiene que funcionar igual de rápido que una página normal, pero por razones obvias los resultados no se pueden cachear
- Tiene que ser creíblemente pseudoaleatoria
- Siempre tiene que devolver un resultado
- Tiene que funcionar siempre, independientemente del número de páginas, y escalar de 1 a millones de páginas

Potenciales soluciones

- ORDER BY rand() LIMIT 1
- Usar un entero bien distribuido y seleccionar uno al azar
- ¿Preguntas?
 - ¿Qué índices serían beneficiosos en cada caso?
 - ¿Cómo consultar el número total de ids?
 - ¿Cómo gestionar borrados?



Brainstorming time



Solución real: Diseño de la base de datos

(I)

```
CREATE TABLE /*_*/page (
```

```
[...]
```

```
-- A page name is broken into a namespace and a title.
```

```
-- The namespace keys are UI-language-independent constants,
```

```
-- defined in includes/Defines.php
```

```
page_namespace int NOT NULL,
```

```
-- The rest of the title, as text.
```

```
-- Spaces are transformed into underscores in title storage.
```

```
page_title varchar(255) binary NOT NULL,
```

```
-- 1 indicates the article is a redirect.
```

```
page_is_redirect tinyint unsigned NOT NULL default 0,
```

```
[...]
```

```
-- Random value between 0 and 1, used for Special:Randompage
```

```
page_random real unsigned NOT NULL,
```

Solución real: Diseño de la base de datos (II)

[...]

```
CREATE INDEX /*i*/page_random ON /*_*/page (page_random);
```

Solución real: Código

```
protected function getQueryInfo( $randstr ) {
    $redirect = $this->isRedirect() ? 1 : 0;
    $tables = [ 'page' ];
    $conds = array_merge( [
        'page_namespace' => $this->namespaces,
        'page_is_redirect' => $redirect,
        'page_random >= ' . $randstr
    ], $this->extra );
    $joinConds = [];

    // Allow extensions to modify the query
    Hooks::run( 'RandomPageQuery', [ &$tables, &$conds, &$joinConds ] );

    return [
        'tables' => $tables,
        'fields' => [ 'page_title', 'page_namespace' ],
        'conds' => $conds,
        'options' => [
            'ORDER BY' => 'page_random',
            'LIMIT' => 1,
        ],
        'join_conds' => $joinConds
    ];
}
```

From: [mediawiki/core/includes/specials/SpecialRandompage.php](https://mediawiki.org/core/includes/specials/SpecialRandompage.php)

Solución real: Consulta generada

```
SELECT    page_title, page_namespace
FROM      page
LEFT JOIN page_props
ON        page_id = pp_page AND
          pp_propname = ?
WHERE     page_namespace IN (...) AND
          page_is_redirect = 0 AND
          page_random >= $rand
ORDER BY  page_random
LIMIT    1;
```

Solución real: Rendimiento (I)

```
mysql> EXPLAIN SELECT ... \G
***** 1. row *****
      id: 1
  select_type: SIMPLE
        table: page
         type: range
possible_keys: name_title,page_random,page_redirect_namespace_len
           key: page_random
        key_len: 8
           ref: NULL
         rows: 20473233
       Extra: Using where
***** 2. row *****
      id: 1
  select_type: SIMPLE
        table: page_props
         type: eq_ref
possible_keys: PRIMARY,pp_propname_page,pp_propname_sortkey_page
           key: PRIMARY
        key_len: 66
           ref: enwiki.page.page_id,const
         rows: 1
       Extra: Using where; Using index; Not exists
```

Solución real: Rendimiento (II)

```
mysql> SELECT ... FROM sys.x$statement_analysis ...
```

```
***** 1. row *****
```

```
exec_count: 27126203
```

```
max_latency: 450802755000
```

```
avg_latency: 755698000
```

```
lock_latency: 2224515688000000
```

```
rows_sent: 27125869
```

```
rows_sent_avg: 1
```

```
rows_examined: 0
```

```
rows_examined_avg: 0
```

```
rows_affected: 0
```

```
rows_affected_avg: 0
```

```
tmp_tables: 0
```

```
tmp_disk_tables: 0
```

```
rows_sorted: 777598
```

```
sort_merge_passes: 0
```

```
+-----+
| sys.format_time(avg_latency) |
+-----+
| 755.70 us |
+-----+
```

Terminando...



Cambios de mentalidad más importantes trabajando “a escala”

- Programación/automatización de infraestructura
- Trabajo en equipo/comunicación
- La teoría mola y puede ser muy útil (estadística, estructura de datos, ...)
- A veces la mejor solución en papel no es la mejor solución en la práctica (rendimiento vs seguridad, estabilidad, simplicidad, ...)
- Aunque no os dediquéis a las BBDD, trabajaréis con ellas

Para saber más...

- <https://www.slideshare.net/jynus/mysql-schema-design-in-practice>
- <https://www.usenix.org/conference/srecon19emea/presentation/mouzeli> (video)
- <https://www.slideshare.net/jynus/wikipedia-burgos-devfest-2017>
- <https://www.slideshare.net/jynus/backing-up-wikipedia-databases>
- <https://wikitech.wikimedia.org/wiki/Portal:Wikitech>

¡Puedes ayudarnos!

- Ofertas de trabajo:
<https://wikimediafoundation.org/about/jobs/>
- GSoC
- “Becas” de desarrollo a contribuidores
- Voluntariado

¡GRACIAS!

Estoy en:

jcrespo@wikimedia.org



WIKIMEDIA
FOUNDATION