

Haciendo **[[copias de seguridad]]**
de todo el **[[conocimiento humano]]**
con **[[Python]]** y **[[software libre]]**

por Jaime Crespo (@jynus)



WIKIMEDIA
FOUNDATION

2021-10-02

Algunos datos

- [22 mil millones de páginas vistas](#) en agosto de 2021
- 47 millones de [ediciones](#) en agosto de 2021
- [300+ idiomas](#) / [900+ wikis](#)
- [6,3 millones de páginas wiki](#) en inglés
- [77 millones de archivos libres](#) en Commons
- [95 millones](#) de “items” en Wikidata

¿Cuánto pesa la Wikipedia?



WIKIMEDIA
FOUNDATION

CC by SA 3.0, victorgrigas

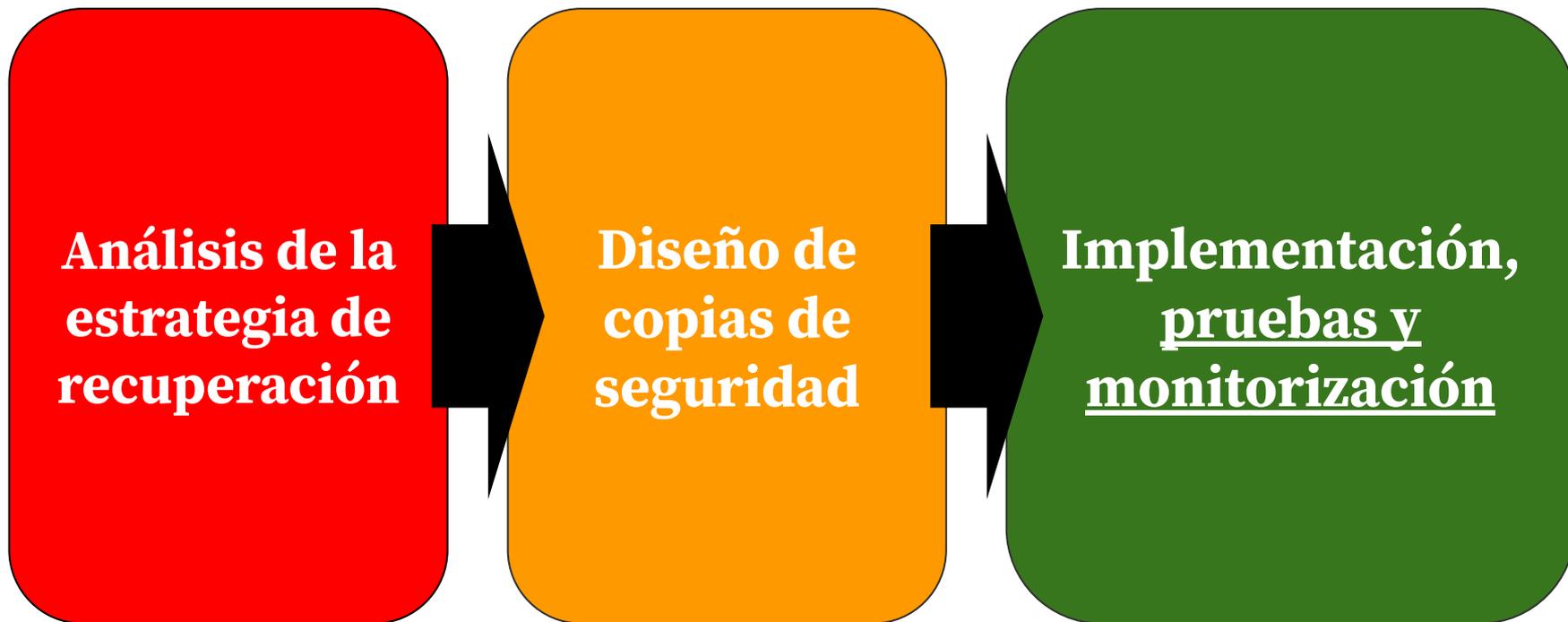
Tamaño aproximado*

CATEGORÍA	TAMAÑO (2021-09-25)	CRECIMIENTO (últimos 30 días)
Código (git)	56 GiB	1,2 GiB
Metadatos (sql)	12 TiB	<i>Negativo[†]</i>
Texto (sql)	78 TiB	120 GiB
Multimedia (Swift)	380 TiB	1,3 TiB
Varios	1.4 TiB	320 GiB

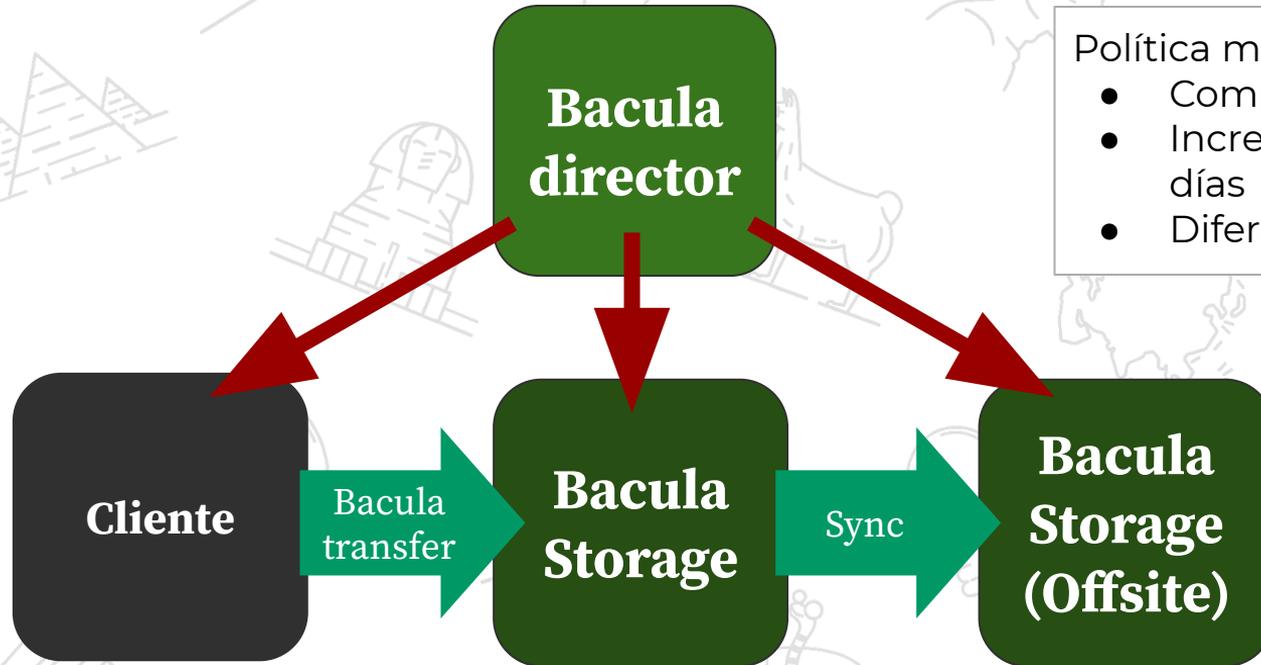
* Sin tener en cuenta compresión, redundancia, deduplicación o nivel de RAID. No incluidos: monitorización, estadísticas, índices de búsqueda, alojamiento externo, hosting para la comunidad, ficheros temporales, caché, renders, miniaturas, etc.

† El mantenimiento permitió ahorrar más de 400GB de almacenamiento

Las copias de seguridad no es lo importante, sino tu plan de recuperación



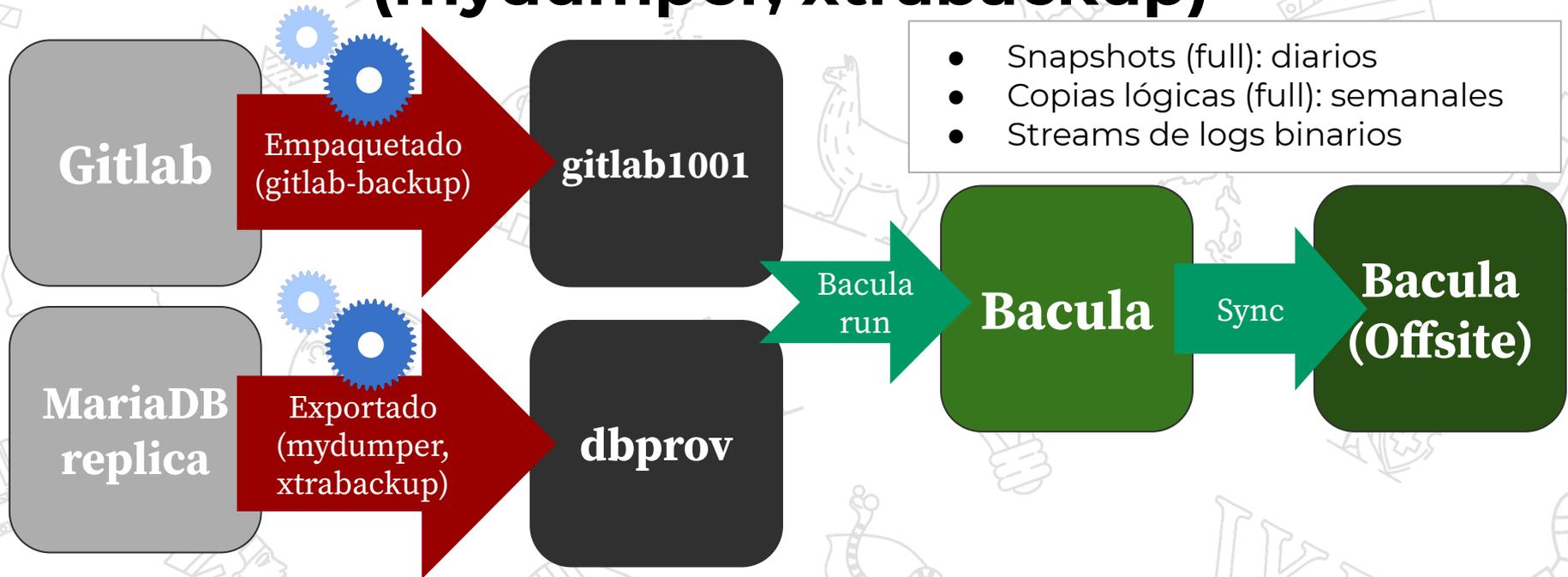
Software libre en backups varios (Bacula)

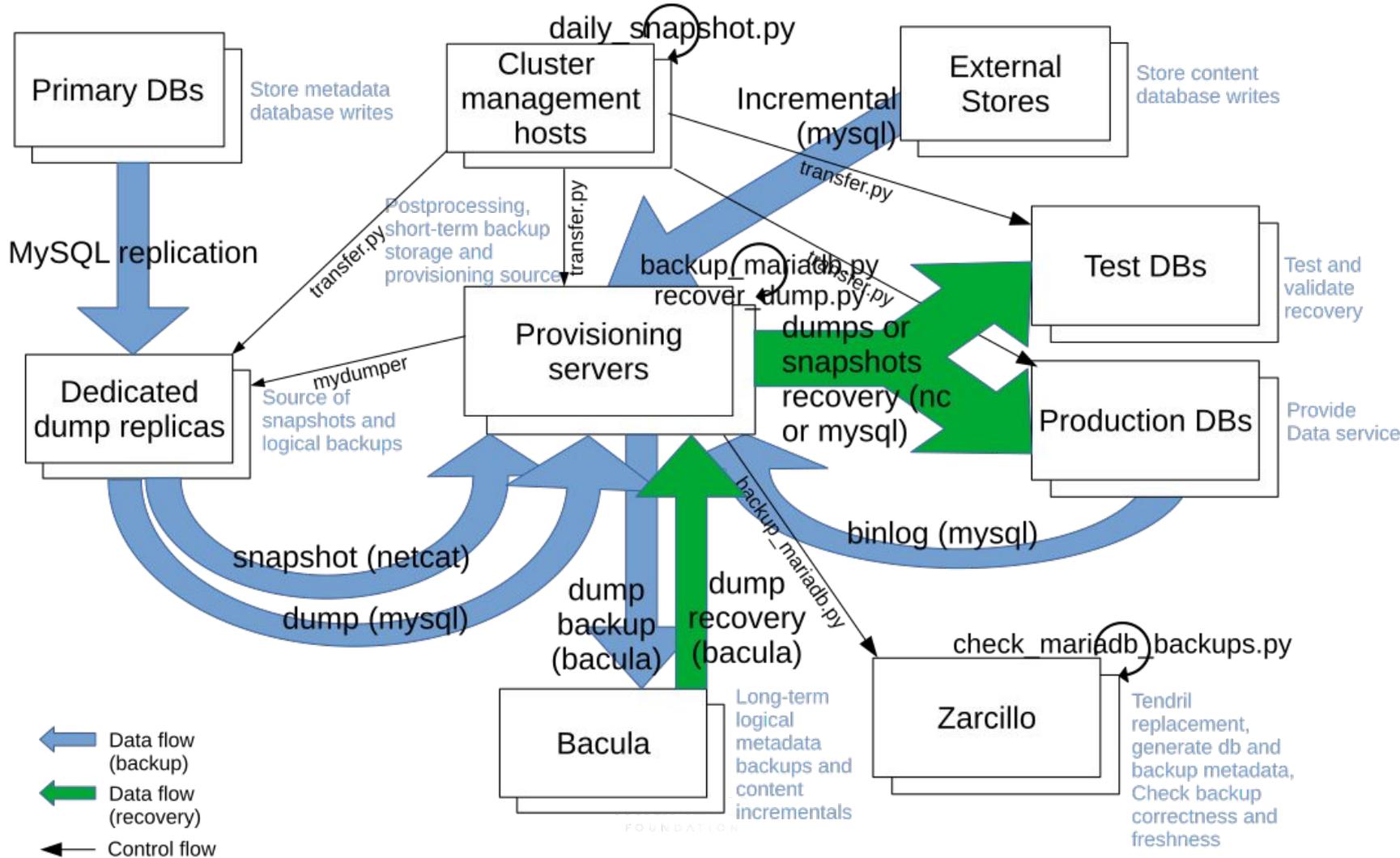


Política más común:

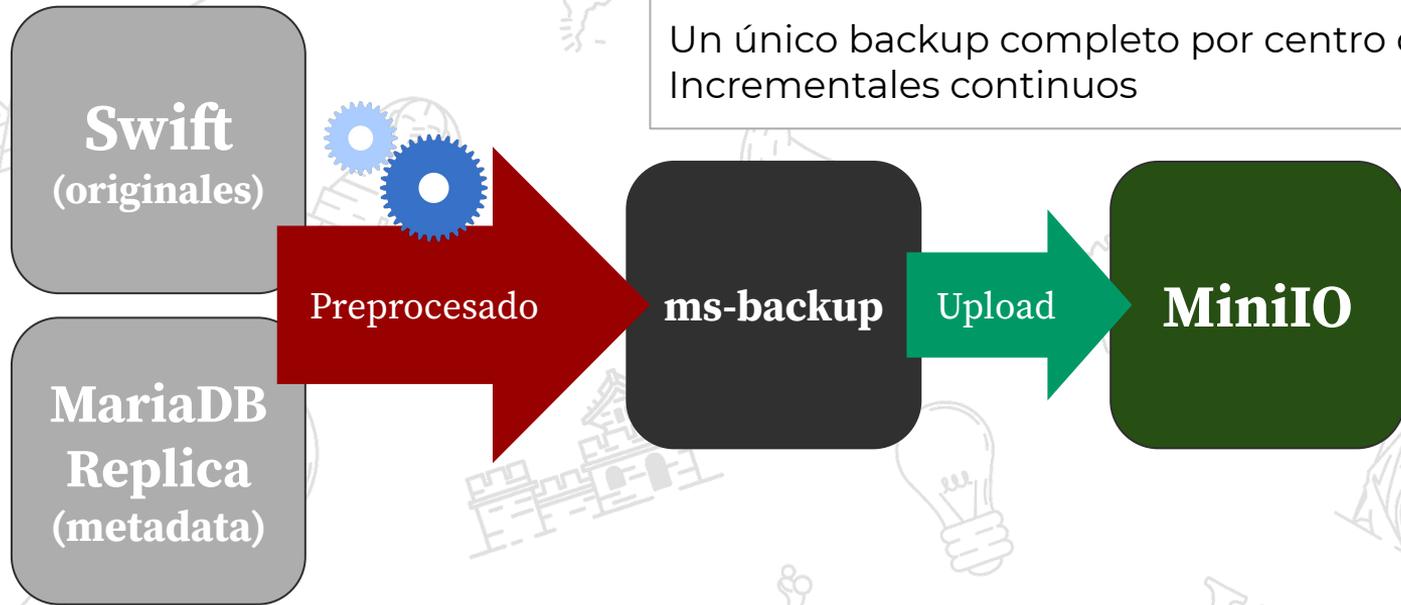
- Completos cada mes
- Incrementales cada 15 días
- Diferenciales cada día

Software libre en backups de código y bases de datos (mydumper, xtrabackup)





Software libre en backups multimedia (MiniIO)



Un único backup completo por centro de datos +
Incrementales continuos

Swift
(originales)

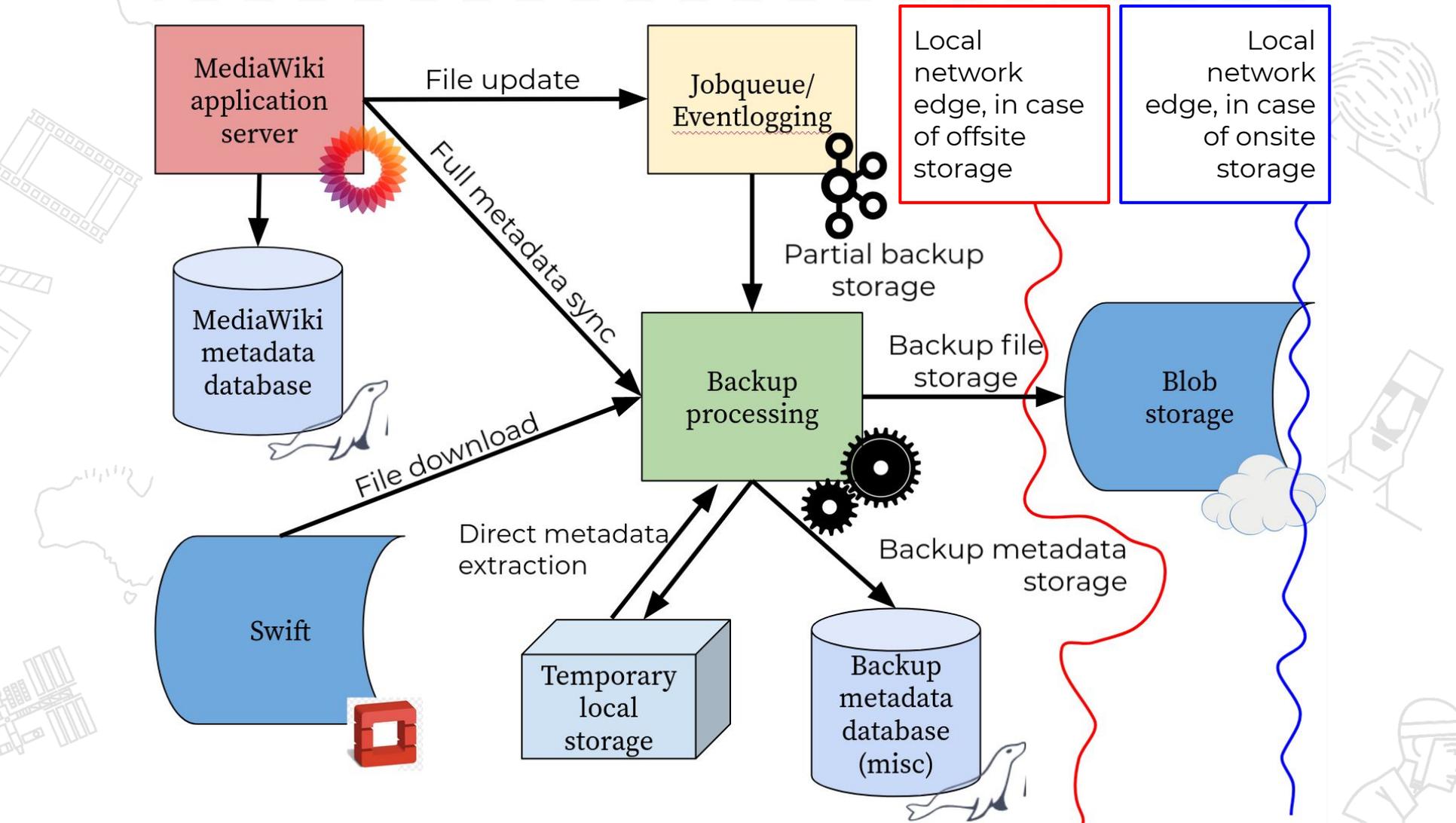
MariaDB
Replica
(metadata)

Preprocesado

ms-backup

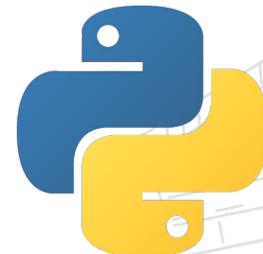
Upload

MiniIO



Lenguajes de programación en Wikimedia

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



BASH
THE BOURNE-AGAIN SHELL

...

SRE elegimos Python para:

- Orquestación
- Automatización
- Monitorización
- "Pegamento" entre aplicaciones
- Estadísticas
- Dashboards



Cumin

- Framework de automatización usando clustershell
- Incluye un lenguaje de consultas, diferentes capas de transporte y gestión de errores
- Simplifica acciones como "reinicia los servidores de aplicaciones del grupo "API" en tandas de 5, espera 1 minuto entre tanda y tanda y párate si más del 25% fallan"

```
$ sudo cumin --batch-size 2 --batch-sleep 5
--success-percentage 95 --timeout 30 --mode async \
'(P{R:class = role::puppetmaster::backend} or
P{R:class = role::puppetmaster::frontend}) and not
D{rhodium.eqiad.wmnet}' \
'date' 'ls -la /tmp/foo'
4 hosts will be targeted:
puppetmaster[2001-2002].codfw.wmnet,puppetmaster[1001-10
02].eqiad.wmnet
Confirm to continue [y/n]? y
===== NODE GROUP =====
(2) puppetmaster[2001-2002].codfw.wmnet
----- OUTPUT -----
Thu Nov  2 18:45:18 UTC 2017
===== NODE GROUP =====
(1) puppetmaster2002.codfw.wmnet
----- OUTPUT -----
ls: cannot access /tmp/foo: No such file or directory
===== NODE GROUP =====
(1) puppetmaster2001.codfw.wmnet
----- OUTPUT -----
-rw-r--r-- 1 root root 0 Nov  2 18:44 /tmp/foo
=====
PASS: | ██████████
| 25% (1/4) [00:05<00:01, 2.10hosts/s]
FAIL: | ██████████
| 25% (1/4) [00:05<00:01, 2.45hosts/s]
25.0% (1/4) of nodes failed to execute command 'ls -la
/tmp/foo': puppetmaster2002.codfw.wmnet
25.0% (1/4) success ratio (< 95.0% threshold) of nodes
successfully executed all commands. Aborting.:
puppetmaster2001.codfw.wmnet
```

Cookbooks

- Centralización de "recetas" de automatización comunes
- Incluye una librería de servicios comunes (spicerack)
- Permite realizar tareas automáticas o interactivas, de varios pasos, así como de detectar fallos y revertir
- Ejemplos de uso: reiniciar máquinas físicas, instalación automática, desmantelar hosts, failover del centro de datos, ...

```
$ cookbook -lv
cookbooks
|-- sre: SRE Cookbooks
|   |-- sre.dns: DNS Operations
|   |   |-- sre.dns.netbox: Update and deploy the DNS re
|   |-- sre.ganeti: Ganeti Cookbooks
|   |   |-- sre.ganeti.addnode: Add a new node to a Gane
|   |   |-- sre.ganeti.makevm: Create a new Virtual Mach
|   |-- sre.hosts: Generic hosts Cookbooks
|   |   |-- sre.hosts.decommission: Decommission a host
|   |   |-- sre.hosts.downtime: Downtime hosts and all t
|   |   |-- sre.hosts.ipmi-password-reset: Reset the hos
|   |   |-- sre.hosts.reboot-cluster: Perform a rolling
|   |   |-- sre.hosts.reboot-single: Downtime a single h
|   |   |-- sre.hosts.remove-downtime: Remove the Icinga
|   |   |-- sre.hosts.upgrade-and-reboot: Upgrade all pa
|   |   |-- sre.hosts.upgrade-varnish: Upgrade/downgrade
|   |-- sre.mediawiki: Mediawiki Cookbooks
|   |   |-- sre.mediawiki.restart-appservers: Restart se
|   |-- sre.switchdc: Switch Datacenter
|   |   |-- sre.switchdc.mediawiki: Switch Datacenter fo
|   |   |-- sre.switchdc.mediawiki.00-disable-puppet
|   |   |-- sre.switchdc.mediawiki.00-reduce-ttl: Re
|   |   |-- sre.switchdc.mediawiki.00-warmup-caches
|   |   |-- sre.switchdc.mediawiki.01-stop-maintenan
|   |   |-- sre.switchdc.mediawiki.02-set-readonly:
|   |   |-- sre.switchdc.mediawiki.03-set-db-readonl
|   |   |-- sre.switchdc.mediawiki.04-switch-mediawi
|   |   |-- sre.switchdc.mediawiki.05-invert-redis-s
|   |   |-- sre.switchdc.mediawiki.06-set-db-readwri
|   |   |-- sre.switchdc.mediawiki.07-set-readwrite
|   |   |-- sre.switchdc.mediawiki.08-restart-envoy-
|   |   |-- sre.switchdc.mediawiki.08-start-maintena
```

WMFbackups y transfer.py

- Framework que resuelve la programación, orquestación (usando cumin), transferencia de archivos, ejecución de varios tipos de backups de bases de datos, gestión de errores & logging, restauración y generación de estadísticas para las necesidades de la WMF

```
$ cat /etc/wmfbackups/remote_backups.cnf
type: 'snapshot'
rotate: True
retention: 5
compress: True
archive: False
stats_file: '/etc/wmfbackups/statistics.cnf'
sections:
  s1:
    host: 'db1139.eqiad.wmnet'
    port: 3311
    destination: 'dbprov1001.eqiad.wmnet'
    stop_slave: True
    order: 2
  s2:
    host: 'db1102.eqiad.wmnet'
    port: 3312
    destination: 'dbprov1002.eqiad.wmnet'
    order: 4
  s3:
    host: 'db1102.eqiad.wmnet'
    port: 3313
    destination: 'dbprov1002.eqiad.wmnet'
    stop_slave: True
    order: 8
  s4:
    host: 'db1150.eqiad.wmnet'
    port: 3314
    destination: 'dbprov1003.eqiad.wmnet'
    stop_slave: True
    order: 3
```

check_bacula.py

- Pequeño script que permite:
 - Consultar el estado de los backups en línea de comandos
 - Crear alertas de icinga
 - Es un exporter de métricas de prometheus

```
$ check_bacula.py
```

```
== jobs_with_all_failures (0) ==
```

```
== jobs_with_fresh_backups (103) ==
```

```
an-master1002.eqiad.wmnet-Monthly-1st-Mon-production-hadoop-namenode-backup
```

```
[...]
```

```
seaborgium.wikimedia.org-Monthly-1st-Fri-production-openldap
```

```
serpens.wikimedia.org-Monthly-1st-Sun-production-openldap
```

```
webperf1002.eqiad.wmnet-Monthly-1st-Wed-production-arclamp-application-data
```

```
webperf2002.codfw.wmnet-Monthly-1st-Thu-production-arclamp-application-data
```

```
== jobs_with_no_backups (0) ==
```

```
== jobs_with_stale_backups (0) ==
```

```
== jobs_with_stale_full_backups (0) ==
```

```
$ check_bacula.py
```

```
dbprov1001.eqiad.wmnet-Weekly-Wed-DatabasesCodfw-mysql-srv-backups-dumps-latest
```

```
[...]
```

```
id: 371115, ts: 2021-09-15 02:07:32, type: F, status: T, bytes: 561946592960
```

```
id: 373235, ts: 2021-09-22 02:05:07, type: F, status: T, bytes: 563897821872
```

```
$ check_bacula.py --icinga
```

```
Fresh: 103 jobs
```

Monitorización y pruebas

- Monitorización del proceso:
 - Fallo de transferencia (checksum y ficheros)
 - Fallo de la preparación
 - Fallo de ejecución
 - Comprobación de no errores en logs
- Monitorización post-backup
 - Backup existe, tiene menos de X horas y su tamaño no varió más de un 5%
- Pruebas de recuperación
 - Recuperación a host de prueba y resincronización

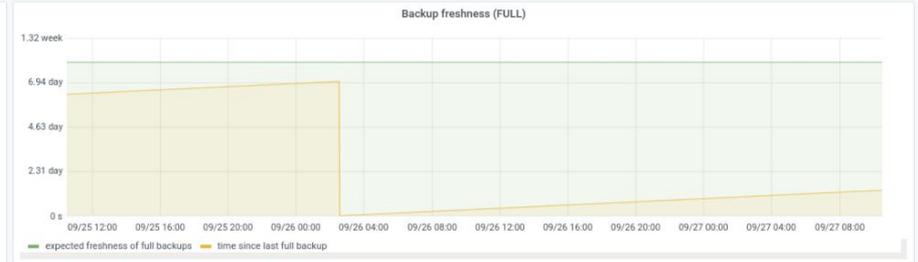
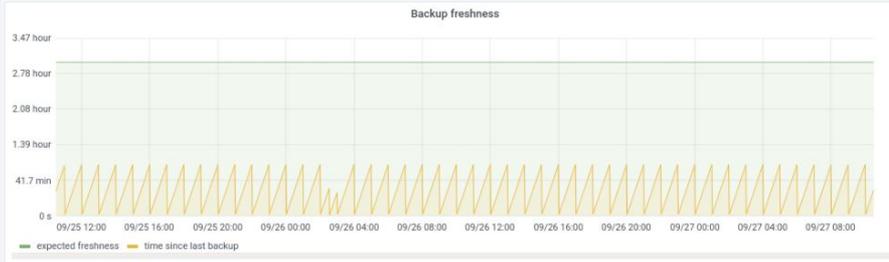
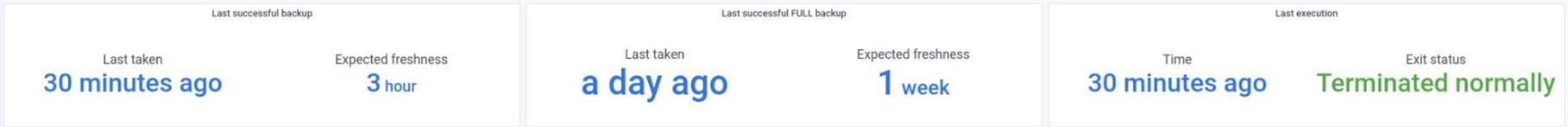


Site: eqiad Job: gerrit1001.wikimedia.org-Hourly-Sun-production-gerrit-repo-data

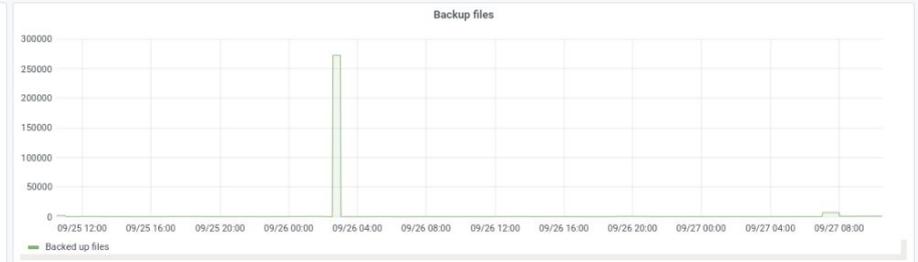
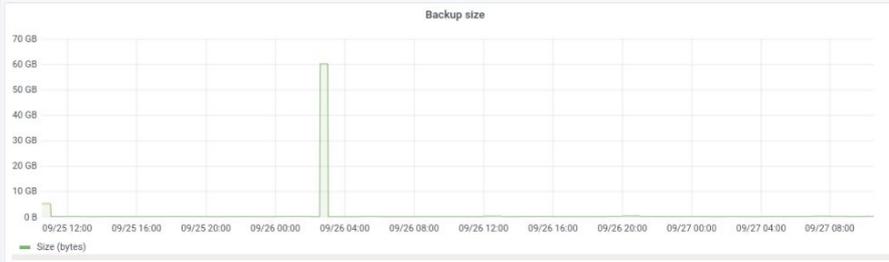
Global success ratio



Backup freshness for selected job



Backup history data for selected job



dump of m3 in codfw		OK	2021-09-28 08:03:26	306d 20h 31m 36s	1/3	Last dump for m3 at codfw (db2078.codfw.wmnet:3323) taken on 2021-09-28 06:26:03 (64 GB)	<input type="checkbox"/>
dump of m3 in eqiad		OK	2021-09-28 07:58:16	306d 20h 31m 36s	1/3	Last dump for m3 at eqiad (db1117.eqiad.wmnet:3323) taken on 2021-09-28 05:24:09 (64 GB)	<input type="checkbox"/>
dump of m5 in codfw		CRITICAL	2021-09-28 08:07:24	0d 0h 19m 45s	3/3	Last dump for m5 at codfw (db2078.codfw.wmnet:3325) taken on 2021-09-28 07:31:00 is 20 GB, but previous one was 37 GB, a change of 46.4%	<input type="checkbox"/>
dump of m5 in eqiad		CRITICAL	2021-09-28 07:50:15	0d 1h 38m 58s	3/3	Last dump for m5 at eqiad (db1117.eqiad.wmnet:3325) taken on 2021-09-28 06:24:33 is 20 GB, but previous one was 37 GB, a change of 46.4%	<input type="checkbox"/>
dump of matomo in eqiad		OK	2021-09-28 08:08:48	42d 3h 54m 13s	1/3	Last dump for matomo at eqiad (db1108.eqiad.wmnet:3351) taken on 2021-09-28 04:00:12 (0 GB)	<input type="checkbox"/>
dump of s1 in codfw		OK	2021-09-28 07:55:52	306d 20h 31m 36s	1/3	Last dump for s1 at codfw (db2097.codfw.wmnet:3311) taken on 2021-09-28 00:00:02 (160 GB)	<input type="checkbox"/>
dump of s1 in eqiad		OK	2021-09-28 07:55:52	306d 20h 31m 36s	1/3	Last dump for s1 at eqiad (db1139.eqiad.wmnet:3311) taken on 2021-09-28 00:00:02 (160 GB)	<input type="checkbox"/>
dump of s2 in codfw		OK	2021-09-28 08:19:15	306d 20h 31m 36s	1/3	Last dump for s2 at codfw (db2101.codfw.wmnet:3312) taken on 2021-09-28 00:00:01 (124 GB)	<input type="checkbox"/>
dump of s2 in eqiad		OK	2021-09-28 08:23:13	306d 20h 31m 36s	1/3	Last dump for s2 at eqiad (db1102.eqiad.wmnet:3312) taken on 2021-09-28 00:00:01 (124 GB)	<input type="checkbox"/>
dump of s3 in codfw		OK	2021-09-28 08:14:20	306d 20h 31m 36s	1/3	Last dump for s3 at codfw (db2139.codfw.wmnet:3313) taken on 2021-09-28 00:00:02 (118 GB)	<input type="checkbox"/>
dump of s3 in eqiad		OK	2021-09-28 08:02:10	306d 20h 31m 36s	1/3	Last dump for s3 at eqiad (db1102.eqiad.wmnet:3313) taken on 2021-09-28 00:00:01 (118 GB)	<input type="checkbox"/>
dump of s4 in codfw		OK	2021-09-28 08:11:36	42d 5h 29m 9s	1/3	Last dump for s4 at codfw (db2139.codfw.wmnet:3314) taken on 2021-09-28 00:00:02 (163 GB)	<input type="checkbox"/>
dump of s4 in eqiad		OK	2021-09-28 08:16:39	42d 5h 24m 8s	1/3	Last dump for s4 at eqiad (db1150.eqiad.wmnet:3314) taken on 2021-09-28 00:00:02 (163 GB)	<input type="checkbox"/>
dump of s5 in codfw		OK	2021-09-28 08:17:00	21d 5h 15m 17s	1/3	Last dump for s5 at codfw (db2101.codfw.wmnet:3315) taken on 2021-09-28 00:00:01 (94 GB)	<input type="checkbox"/>
dump of s5 in eqiad		OK	2021-09-28 08:06:31	21d 4h 48m 34s	1/3	Last dump for s5 at eqiad (db1150.eqiad.wmnet:3315) taken on 2021-09-28 00:00:02 (94 GB)	<input type="checkbox"/>
dump of s6 in codfw		OK	2021-09-28 08:21:27	0d 2h 35m 49s	1/3	Last dump for s6 at codfw (db2141.codfw.wmnet:3316) taken on 2021-09-28 04:23:35 (102 GB)	<input type="checkbox"/>
dump of s6 in eqiad		OK	2021-09-28 08:22:03	21d 4h 5m 49s	1/3	Last dump for s6 at eqiad (db1140.eqiad.wmnet:3316) taken on 2021-09-28 03:21:55 (102 GB)	<input type="checkbox"/>

Almacenaje y recuperación

- Retención de 2-3 meses
- TTR objetivo: 30 minutos por base de datos de 1-2 TB
- Hasta ahora, recuperación en **pequeños incidentes** (e.g. DROP TABLE) y contestar a **preguntas de desarrolladores** "podrías comprobar el estado de X el día Y" (debugging)

CATEGORÍA	TAMAÑO de backups por cada DC (2021-09-25)
Código (git)	1.7 TiB
Metadatos (sql)	66 TiB
Texto (sql)	85 TiB
Multimedia (Swift)	336 TiB
Varios	34 TiB

Utilidad más allá de la recuperación

- Provisionado automático de bases de datos
- Estadísticas de crecimiento detalladas y planificación de capacidad
- Detección de anomalías
- Fuente de datos de servicios no críticos

```
sql> SELECT file_name, size, file_date
FROM backup_files
JOIN backups
ON backups.id = backup_files.backup_id
WHERE section='s1' and
      type='snapshot' and
      host like '%eqiad%' and
      start_date > now() - INTERVAL 1 day and
      file_path='enwiki'
ORDER BY size DESC
LIMIT 10\G
```

```
***** 1. row *****
file_name: revision.ibd
      size: 193399357440
file_date: 2021-09-27 00:32:17
***** 2. row *****
file_name: pagelinks.ibd
      size: 169705734144
file_date: 2021-09-27 00:32:20
***** 3. row *****
file_name: templatelinks.ibd
      size: 105084092416
file_date: 2021-09-27 00:32:20
***** 4. row *****
file_name: revision_actor_temp.ibd
      size: 101149835264
file_date: 2021-09-27 00:32:17
***** 5. row *****
file_name: externallinks.ibd
      size: 73740058624
file_date: 2021-09-27 00:32:20
```

Para saber más...

- <https://www.usenix.org/conference/srecon19emea/presentation/mouzeli>
(video)
- <https://www.slideshare.net/jynus/backing-up-wikipedia-databases>
- <https://wikitech.wikimedia.org/wiki/Portal:Wikitech>
 - [Bacula](#)
 - [MariaDB/Backups](#)
 - [Media_storage/Backups](#)
- <https://grafana.wikimedia.org/d/413r2vbWk/bacula>
- Código:
 - [WMFBackups](#)
 - [Cumin](#)
 - [Transfer.py](#)
 - [Cookbooks](#) [[Spicerack](#)]
 - [Configuration management](#) ([backup module](#), [check_bacula.py](#))

GRACIAS

@jynus

Gracias especiales a: Manuel, Lukasz, Riccardo, Rubén, Ajumal, Mark, Faidon, Effie, Antoine (y muchos más)
y a la organización y sponsors de la PyConEs2021

