Infrastructure monthly meeting minutes

Important documents

Current PRs - What can I work on ? - Project Board - Infrastructure on the wiki

Main repositories:
- https://github.com/openfoodfacts/openfoodfacts-infrastructure/
- https://github.com/openfoodfacts/openfoodfacts-monitoring/

Documentation: https://openfoodfacts.github.io/openfoodfacts-infrastructure/

Infrastructure catalog
- CT and VM list of OFF infrastructure

Roadmap
(outdated) OFF days infrastructure workshop
(outdated) Infrastructure roadmap

Post Mortems
see
https://github.com/openfoodfacts/openfoodfacts-infrastructure/tree/develop/docs/reports

Manual Deployments
DevOps - Deployment center - TODO: move to git
DevOps - Calendrier de maintenance - TODO: move to git

Other stuff
- ~~OpenFoodFacts CICD - copie: OpenFoodFacts CICD~~
- ~~Live reload of ProductOpener - copie: Live reload of ProductOpener~~
- ~~DevOps - Static pages - Command Generator~~

Copy paste me to this month's meeting

News of the month

Issue review/triage/points of concern

Points to discuss

9 oct. 2025 | Infrastructure monthly meeting

Participants: Alex,

News of the month

- work in progress on Ansible / servers install (but slow, very few availability of Alex past month)
- bad blocks on a disk on off1
    - CQuest saved the day by marking bad blocks and asking for a ZFS scrub
- created the superset container
    - this is a good example of container creation with
- added ansible lint in CI
- folksonomy engine staging installed with ansible
- sept 16: Christian installed SSDs in the 3 new servers: sc01, sc02, sc03

Issue review/triage/points of concern

- Open Food Facts Explorer Install - hetzner-02
- Update Wiki (mediawiki), Odoo (done by a freelancer), Matomo
    - first move to hetzner Matomo to a new container with a newer debian
        - migrate db
        - then update matomo
        - it's a bit complex - so not the first thing to do
- Investigations Keycloak (mediawiki)
    - wait for full migration of product opener

Points to discuss

- we will use VirtioFS on VM for docker deployment

- plan
    - move everything from OVH to Hetzner
        - Thomas: migrate the blog
            - use ansible
    - then upgrade OVH
    - then remove Hetzner
    - in // move off1 + off2 to sc1/sc2/sc3
        - off1 and off2 won't be in the same proxmox cluster
    - then move off1 and off2 servers to the dedicated bay, and possibly use them for secondary things
    - goal:
        - always have 1 spare server, with 3 servers, but able to run on 2
- sc1/2/3: get one of them ready if we lose off1?
    - Thomas
        - install proxmox on sc1/2/3 (need ip + credentials from Christian)
            - we use a vlan for internal addresses
        - prepare backups from current prod
        - install the reverse proxy nginx + stunnel (server part)
        - install mongodb in a container / or migrate the container
        - attention: we have stunnel setups
- upgrade
- (Raphaël) launch a GPU instance on Google Cloud to plan the change of osm45 GPUs?
    - use ansible, maybe
    - use google cloud + firewall + ip whitelist to secure the access to the services

11 sept. 2025 | Infrastructure monthly meeting

Participants: Alex,

News of the month

- we had a strange sudden augmentation of requests (crawling) for one night (see this slack thread and this one)
- two new servers where transported to our new bay at scaleway
    - we are waiting for SSDs to put in them
- Work on installing new servers with ansible is progressing (not as fast as wanted)
    - PR for container configuration
    - folksonomy staging in progress (a good test of a container install)
- We resolved a double disk failure on ovh3

Issue review/triage/points of concern

Points to discuss

Infrastructure monthly meeting

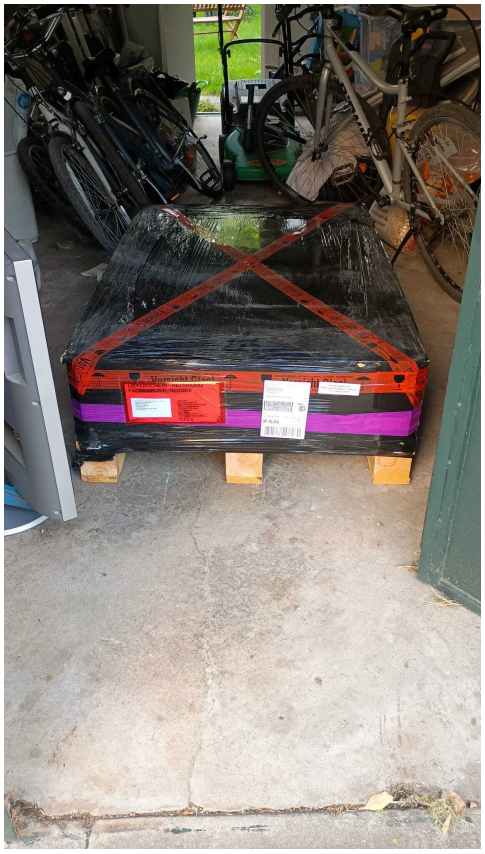
14 août 2025 |

Participants: Milo, Raphaël, Alex, Vincent

Introductions

- Milo is working on devops

News of the month



- we received new servers - still not in the bay
    - we still needs the SSD
- ovh3 has a disk problem again
- Raphaël has installed GPU on a VM on osm45 aka moji
- Hetzner server install has only progressed a bit, but we should be ready to deploy real stuff there
- made virtiofs tests (report still to commit)

Issue review/triage/points of concern

- Alex will try to concentrate on hetzner install
    - install reverse proxy container (nginx / stunnel)
        - https://openfoodfacts.github.io/openfoodfacts-infrastructure/explain-server-config-in-git/
    - install stunnel client container
        - https://openfoodfacts.github.io/openfoodfacts-infrastructure/explain-server-config-in-git/
    - install a docker ready VM (ansible)
        - use virtiofs on the VM
    - move search-a-licious staging there
- Possible work by milo
    - connexion osm45 prometheus
    - vpn -
        - because at each datacenter we have private network
        - replace stunnel with a more flexible tool
        - issue: https://github.com/openfoodfacts/openfoodfacts-infrastructure/issues/481
        - we have to explore candidates
            - Netbird maybe a good candidate but maybe too complex ?
            - interesting discussion here: https://github.com/openfoodfacts/openfoodfacts-infrastructure/pull/485#issuecomment-2902059505
- keycloak status:
    - repo: https://github.com/openfoodfacts/openfoodfacts-auth/
    - Product Opener integration with service level
        - normally on monday we switch to level 1 in prod
        - … and up to level 5
    - in the future:
        - propose 0Auth for third party app using API
        - branch community tools on keycloak: forum, wiki,…
        - propose 2FA to members
- STO to JSON migration on off
    - starting monday
    - will ask to closely monitor disk space

Points to discuss

- Updates:
    - mediawiki, odoo, matomo, more ?
    - some of those would be better done with a re-install
- Installs:
    - obf, opf, opff staging - that's docker compose work essentially

10 juil. 2025 | Infrastructure monthly meeting

Participants: Alex, Stéphan, Thomas, and more…

News of the (two) month

- migration of monitoring to google cloud machine
    - installation done with ansible
    - we had proxy the access to exporters on ovh
    - TODO: osm45 / Moji not monitored by prometheus (configuration of reverse proxy in ipv6 required)
- Ovh3: we had to replace two disks
- recipe estimator: deployed to staging and production
- sanoid_check was sometime triggered every 15 min → switch back to 2h
- problems with DNS on osm45 → use TCP instead of UDP (sometimes failing)
- (Raphael) GPU accessible from a vm, that can hold docker containers on osm-45, first model (clip) seems to be running smoothly

Issue review/triage/points of concern

- next step / pending issues on monitoring
    - add a proxy to let prometheus access exporters on osm45
- we haven't yet installed nothing productive (but backups) on all the new servers (hetzner included) ! While we have big problems with existing ones (lack of availability)
    - search-a-licious could be a good candidate to move to hetzner-02/03
    - staging could be a good candidate to move to hetzner-02/03
- new servers delivery:
    - open ticket on Scaleway to get the procedure to let Serverschmiede deliver new machines directly on DC1 datacenter (discussed with cquest)
    - configuration change to buy machines SPF+ network cards instead of RJ45
    - decision to buy a SPF+ network card to avoid adaptations RJ45 <-> SPF+
- GPU on osm45:
    - next step: move triton to this new VM
    - we may not have enough memory to run every models on GPU, but at least the most resource consuming (eg. categories) or the most useful (because it will run faster)
    - will lower memory and CPU usage

Points to discuss

- next steps on server installs
    - have docker vm on hetzner
        - push search-a-licious on that
        - use direct access to datasets
- VPN: we add discussion on this PR
    - we use stunnel:

        - we have unsecured elements to access mogodb / redis
        - it's bit cumbersome to configure: you have to configure it for each service / port with client and server
    - a VPN, enables transparency on that, with a big plus if it has name resolution
        - transform exchanges on the PR in a discussion
- firewalls:
    - proposal of OpenSense https://github.com/openfoodfacts/openfoodfacts-infrastructure/pull/485
        - pros:
            - fully featured
            - good UI
        - cons:
            - use OpenBSD (new technology for the team)
            - need to be deployed on a VM
            - can't be deployed on normal host (eg google cloud machine)
    - maybe proxmox firewall is the way to go
        - Charles use to
    - regular scans could be a thing ?
- Simplifying infrastructure

7 mai 2025 | Infrastructure monthly meeting

Participants: Alex, Charles, Thomas, Shashi, Areeb, Pierre

News of the month

- ovh1 crashes two times because of disk full (during week-ends)
    - disk goes back to normal because snapshots gets removed by sanoid
    - a restart of VM and container must be done in this case
    - we really need to install new servers to free space
    - monitoring is on ovh1 so there way no real alerts
    - maybe also **move the batch schedule to week days**
- slow (but steady ?) progression on ansible recipe:
    - openfoodfacts-ops was merged into openfoodfacts-infrastructure in ansible repository
        - enable more "atomic" changes to infrastructure
    - Thomas is progressing on monitoring deployment on a Google server instance (will alleviate ovh1 and also make monitoring more robust as located outside of main infrastructure)
        - https://github.com/openfoodfacts/openfoodfacts-infrastructure/pull/474
        - almost ready
        - make prometheus exporters of OVH reachable through nginx
        - we need a migration plan
    - sanoid / syncoid deployment
        - we now have ovh3 backup on hetzner-01 (as well as osm45)

        - proof of concept of mixing ansible and having confs files in git (more easy to tweak live)
    - vswitch on hetzner servers → have a private network for proxmox
- delivery of the new scaleway bay sponsored by Free Foundation and shared with OSM France
    - half a bay (~25U, meaning possibility to install more than 10 2U servers)
    - first server put in the bay with the switch (for public network), thanks to Christian
    - now operational, needs deployment
- we still had some down-time
    - we had a bot crawling our sites with a very huge number of IP address, so rate limiting did not work
- docs
    - are now searchable
    - are also editable directly from the docs thanks to logo  at the upper right

Issue review/triage/points of concern

- could a solution like anubis help with bots
    - we need someone to investigate it
    - how does it works for an API ?
        - for API we should enforce being logged in maybe ?
- we would like to have dashboards (maybe apache superset) but we would need postgres and mongodb replication for that
    - we need someone to investigate / do it (docker containers)
    - MongoDB own replication is quite simple but I'm a bit afraid about latency and IO costs
        - Charles : Maybe it could it be done once a daywith a bash script
        - Alex : it should be less costly with MongoDB own replication mechanism
        - Vincent : replication on Postgre is very efficient

Points to discuss

- Migration plan of off1 / off2 to the new bay at scaleway bay
    - [x] off0-dc1 is physically installed
    - finalize ansible experiments on Hetzner
    - install private network on off-dc1 (need to buy and install a small switch)
    - then deploy on off0-dc1
    - then buy off1-dc1 and off2-dc1
    - then migrate off1 and off2 on off1-dc1 and off2-dc1
    - then use off1 and off2 to relieve ovh1 and ovh2 servers (or to substitute hetzner servers)
- Migration plan for OVH servers

    - Why:
        - our convention with OVH is not formally reviewed, we still got the server but this could end in an uncontrolled matters (happens to OSM France), so it's important to be ready to migrate services to other servers
        - also OVH servers currently have their disks full, we must alleviate this
    - What: where do we move services ?
    - We will use hetzner server
        - better setup (ansible)
        - fresh proxmox
    - Then wipe and reinstall ovh servers
- [Jagjeevan] : Any update on installation of new machines as memory is saturated fast and VM's crashing always after few interval of days mostly usually on holidays and no one is on 'On call during this time'.
    - the answer is above

5 mai 2025 | mini meeting

participants: Alex, Charles, Stéphane

Discussion:

- on hetzner hosts, we need to do:
    - private network (use vrack, not that easy because compute server have only one interface)
    - proxmox cluster
- in between we have a new server at free (and we plan to have two more)
    - they are more planned to replace off1 & off2
- we also have ovh servers that are really over capacity
- Proposal:
    - use hetzner (3 servers) to migrate current OVH
        - try to setup the private lan between machines and see how difficult it is
    - install OFF server to get actual production
- we can wait a bit to get new servers

10 avr. 2025 | Infrastructure monthly meeting

Participants

- Thomas - contributor from France, interested in Ansible
- Mitali - contributor, getting news
- Hao - researcher on metabolism - started a company for personal healthcare - computational bio engineer, familiar with infrastructure - GitHub - GitHub actions
- Jagjeevan - contributed to OFF on CI / ops, testing OpenAPI doc generation

- Areeb - working on existing issues assigned. would like to explore issues mentioned in notes later
- Abraham
- Alex, Pierre, Stéphane

News of the month

- not too many crashes 😅
- we decided we will move ops to infrastructure repository
    - but Alex needs to complete his PR first
    - this will enable linking files on servers as we do for now from ansible, also simplify discoverability
- 3 VMs (Staging, forum etc.) were down last week-end
    - due once-again to a shortage in storage (ovh1 is low in storage)
    - we really have to free space there
- not much done otherwise
- we are about to receive a server rack at Scaleway thanks for Free
    - we currently have 2 servers provided by the Free Foundation
    - we will get a complete rack to host those 2 servers + other servers that we are ordering
    - the rack will also be shared with openstreetmap france
    - there will be work to get everything set up
        - Christian from OSM will do the initial setup for networking etc.

Issue review/triage/points of concern

- it's very urgent to work on deploying new servers:
    - moving monitoring to google cloud could save space on ovh1 (will be done by Thomas)
        - docker / docker compose install
        - ensure a off user with a specific private ssh key
        - thomas
            - provide me a github handle
            - gpg public key see docs
        - change address of server where we want to deploy in the CI of monitoring project https://github.com/openfoodfacts/openfoodfacts-monitoring (on a deploy-xxx branch, it will run CD on every commit)
    - hetzner installs still need ansible work, help really welcome
        - proxmox install
        - other things that you think is important (talk about it first)
- we would like to have replications on database (useful to run dashboards on a database that is not the production database)
- lots of reusers / heavy contributors had their IPs address blocked, maybe raise a bit the limits, or reduce the ban to 1 hour?
    - French TV, French Museum of Science banned
    - it's an argument for Product Opener maybe ?

        - moved to Product Opener - Dev Notes
- Jagjeevan is testing idx usage for product opener
- openfoodfacts-metrics project: for business data (dashboard etc.)
    - that's the project where it would be good to have database replication
    - some metrics are currently generated by robotoff, but they should be moved to openfoodfacts-exports
    - influx db is for metrics, not monitoring
- monitoring: for monitoring servers and services
    - uses prometheus
- to take a stale issue (assigned to someone, but no progress for several weeks): comment on the issue and communicate on Slack
- docker swarm: currently not used as each service runs on only 1 server
    - in the future, it may change
- credentials to login to .net domain: off / off
- product opener log rotation issue: https://github.com/openfoodfacts/openfoodfacts-server/issues/5970#issuecomment-2727013044 (Jagjeevan)

Points to discuss

- github project kanban for infrastructure needs to be created / reviewed
- don't hesitate to read the documentation (or improve it!) https://openfoodfacts.github.io/openfoodfacts-infrastructure/
- CICD topics related to specific projects (e.g. Product Opener, Robotoff etc.) can be discussed in the specific project meeting

13 mars 2025 | Infrastructure monthly meeting

Participants:

- Alex, Stéphane, Pierre, Charles (permanent type)
- Areeb - devops intern in an open source startup
- Bhavishya - in a noisy environment; can't open his mic
- Jagjeevan - third meeting; devops in offline (kube, grpc, docker, python etc.)
- Alper - full stack dev, created https://github.com/AlperMulayim/openfoodfacts-spring-boot-starter/
- Bid out (left)

News of the month

- work in progress on Hetzner server install using Ansible:
    - firewall, fail2ban, users, base proxmox install => https://github.com/openfoodfacts/openfoodfacts-ops/pull/9 (wairting for review from Vince)
    - https://github.com/openfoodfacts/openfoodfacts-ops

- we had a meeting on 19/02 to prepare free / scaleway migration (see below)
    - new dedicated rack available soon (April?) at Free foundation
        - still no news from Scaleway (cquest wrote to them on 2025-03-07)
        - see notes about dedicated meeting:
          Notes - Infrastructure monthly meeting - PUBLIC
        - we are ordering 2 new servers

Issue review/triage/points of concern

- we had two incidents with ovh1 blocked during the week-end
    - linked to off-query imports on staging which was creating big snapshots and saturating the disk
- https://openfoodfacts.github.io/openfoodfacts-infrastructure/explain-server-config-in-git/
    - we still need a script to check every conf is symlinked at the right place

Points to discuss

- Is it ok to mix the old approach (on infrastructure) with the new one
    - for example: symlink for systemd services definition, or for sanoid configuration, instead of doing it all with ansible
        - this enable quick interventions on servers
    - I'm not sure if file would stay in -infrastructure
        - big drawback: it means we need to commit in two repository for a change
        - advantage: we commit from servers only in -infrastructure repository and not in -ops
            - **suggestion**: put the content of openfoodfacts-ops in a directory of openfoodfacts-infrastructure
                - **action:** check with vince
    - maybe we can achieve same results by limiting some roles to copy servers in a specific location, and good use of tags (to ease retro-porting modifications)
- discussion today on metrics / dashboard
    - the first step to do seems to have data sources available on a server
    - we can then eventually make different choices
    - can we go on osm45, or should we wait hetzner server (already quite high load on osm45)
- what should go asap on hetzner install ?
    - still todo: networking, proxmox clustering, sanoid/syncoid
    - replicating what we have on off1 and off2 in Free datacenter: off, opf, opff etc + mongo ?
        - needed to make the switch from old servers to new servers / new rack
    - migration won't be that easy because of centralized product database (.sto files and images)
        - we will need to migrate OFF, OPF, OBF, OPFF at the same time
            - but we can test it beforehand (with zfs clones)
        - potential performance issue?

            - Hetzner servers are in theory far more powerful (cpu, memory, disk..) than off1 and off2
                - we will need to use NFS for images (not enough space on the Compute Hetzner servers)
        - it should a good motivation to move to another storage solution

13 mars 2025 | Metrics and dashboard tool choice

Participants: Charles, Alex, Manon

See: Metrics and dashboard tool choice

19 févr. 2025 - Meeting about server migration at Free

- Participants: Alex, Charles, Christian, Stéphane
- Dedirack at Bezons (Free datacenter) should be available in 2 or 3 weeks
-
- Migration strategy
    - worse case: shutdown the service for some time to migrate the off1 and off2 servers
    - best case: first install new Hetzner servers and migrate off1 and off2 to them first
- Rack will be delivered “raw”
    - we need a switch (in the old rack, it is the switch of the Free foundation)
        - in: 10Gb : out, min 1Gb, 10gb better
            - ~ 1.5k
        - network: fiber?
            - Christian: will ask what the incoming network is
        - 2 switches or 1?
            - OSM in Telehouse
                - 1 switch 10G with 2 fibers
                - 1 switch 1G
                - VLANs to separate traffic
    - need a PDU or a piloted plug
    - steps
        - ask scaleway (**Christian**)
            - what is the type of the network connector (for the switch)
            - what is the depth of the bay? e.g. 80cm servers?
            - badges how-to?
            - where is the bay located?
            - goal: 28th February
            - when will the bay delivered
        - buy a switch: ~1500€ => **Christian** (choisi) + Charles (payment)
            - + 1 small spare switch 1Gb from Christian (provided by OSM)

            - goal: 8th March
        - buy a server: 2 servers ~ 5-6 K€
            - goal: 8th March
        - preinstall / tests the two servers (Christian)
            - goal: 25th
            - possibly: precopy images from ovh3
        - install the switch and the first two servers
            - goal: 30th Mars
        - move all software from off1 off2 to new server (not the same day)
            - write migration plan
                - need to plan what will be reinstalled from scratch
            - use ansible config that will be done for hetzner servers
        - physically move off1 & off2 to new rack
- What will be in the rack ?
    - existing machines
        - off1, off2
        - osm machines
    - new machines
        - R740xd : Notes - Infrastructure monthly meeting - PUBLIC
            - depth compatible with the bay? server: 71.5cm
        - **or 2 medium servers instead of 1 large?**
            - **5k to 6k**
        - decision: 2 servers
- Alex:
    - focus on migration
    - current
        - 1/3d of Alex time on investigating alerts (zfs synchs, docker issues on staging, etc.)
        - Antoine to manage: intégration en web component des modèles de ML développés par Raphaël
        - NutriSight report: missing points
    - delayed:
        - search-a-licious
- Charles:
    - helps Alex: support level one for alerts
        - [x] subscribe root@openfoodfacts.org ?
    - helps Alex with Ansible
    - new helpers/volunteers
        - [x] contact Jeremy L

- Dell 740dx: https://www.serverschmiede.com/konfigurator_bulk/en/dell-poweredge-r740xd-19-2u-12x-35-lff-2x-intel-xeon-scalable-lga3647-ddr4-ecc-raid-2x-psu-server?p=e8c3c45d17a3628825d9020e6c770080
    - Dell PowerEdge R740xd 19" 2U 12x 3,5" LFF 2x Intel XEON Scalable LGA3647 DDR4 ECC Raid 2x PSU Server. Rationale:

        - 740 machines have plenty of space and three risers for PCI cards.
        - Allows to install full height GPU card, see this video
        - Allows to install 12 HDD disks
    - CPU: 2 Intel Xeon Platinum **28 cores** / 56 threads (**165w** tdp each). Rationale:
        - CPU Benchmark : 56381 (x2)
        - Facing current Xeon Silver 4110 (off1/off2): 10028
    - **512GB** Registered ECC DDR4 SDRAM (8x 64GB DIMM). Rationale:
        - We never have too much RAM (ZFS cache is using all the RAM that remains) and each 64 GB is cheap (~110€).
        - 12/24 banks would still be empty
    - 1 x NVMe Bifurcation Switch 4x M.2 PCIe x8 / x16 Controller incl. **4x 1,92TB** Enterprise Performance NVMe. Rationale:
        - 4 NVMe for both disk space and redundancy.
        - All is shipped in one card.
        - NVMe is quite faster than SSD.
    - **6 x 3,5" 20 TB** 7,2k 12G Raid Enterprise Storage 24/7 SAS HDD. Rationale:
        - plenty of HD space for both disk space and redundancy
        - allows to backup OFF photos for years
    - 1 x Dell Broadcom 57416 Quad Port **2x 10Gb** + 2x 1Gb copper RJ45. Ethernet Network Daughter Card 01224N. Rationale:
        - 10 Gb seems to be the norm now
        - a daughter card does not waste place.
    - **2 x DELL 1100 Watt** Netzteil PSU - PowerEdge R630 R640 R730 R740 R930 Tx30 Tx40. Rationale:
        - 1100w PSU are needed to support GPU according to this thread
        - 2 PSU for redundancy

13 févr. 2025 | Infrastructure monthly meeting

Participants: Alex, Stéphane, Thomas, Vincent

News of the month

- fail2ban better configuration (e.g. proxmox interfaces) + ban on 429 + ban of url scans (wp-admin etc.)
- firewall - work to improve them on
- we log answer time on off - needs analyzing
- we banned Amazon CloudFront on ks1 (based on user agent)
    - https://prometheus.openfoodfacts.org/graph?g0.expr=nginx_connections_active%7Bapp%3D%22off%22%2C%20env%3D%22prod%22%2C%20instance%3D%22images-ks1-nginx%22%7D&g0.tab=0&g0.stacked=0&g0.show_exemplars=0&g0.range_input=1d
- Alex must start installing new servers
- restored some backups from ovh3 to osm45

- New repo for ops: https://github.com/openfoodfacts/openfoodfacts-ops

Issue review/triage/points of concern

- Ongoing performance issues and downtime (web, API, images)
    - Some software improvements done / coming
        - 2 Apache servers: 1 for important fast queries (API, edits etc.), another for long queries
        - prefix facet queries with /facets/ so that we can better differentiate
        - work on fail2ban: ban when we have too many 403/409 etc.
        - have off-queries serve list of products queries (e.g. home page, /labels/organic etc.)
        - static robots.txt (currently 2% of queries, and generated dynamically)
- Install of new servers
    - not really started yet
- Vincent : working on using nullmailer to send email via mailjet
    - mailjet account created
        - needs to create a group
    - git-crypt deployed to encrypt secrets
        - advantage: secrets are encrypted automatically
        - everyone needs a gpg key
- Vincent also small problem on hostname
- **new repo for ansible scripts etc.: https://github.com/openfoodfacts/openfoodfacts-ops**

Points to discuss

- How do we use ansible on old server / distinguish
    - to be prudent, not to run playbooks:
        - the best is to have different inventories to
    - we only use ansible for new servers (fully managed servers)
    - we really separate base install from “sites” (specific install)
    - the playbook can list groups to which it applies
- Firewall: are we ok to switch from iptables to nftable?
    - V: ansible has a rock solid iptables module, not sure about nftable (maybe ipr-cnrs package)
    - A: nftable more easy to read
    - T: nftable not supported by docker: docker iptables rules need to be added manually
        - this allows us to know which ports are opened, and not let docker-compose create its own input rules
    - decision:
        - try nftables and go back to iptables if we have problems
- Name of the ansible account:
    - debian - because it was the name of the user on OVH servers
    - proposed: **conf-ops**
    - we use debian only

- giving access:
    - portainer is a good way to gives access to containers
- mail root
    - lots of alerts that Alex knows are not important
        - https://openfoodfacts.github.io/openfoodfacts-infrastructure/how-to-handle-alerts/
        - osm45 Too many snapshots for hdd-zfs/off-backups/ovh3-rpool/subvol-101-disk-0 (238)
            - there should be a script that removes older snapshots
        - ovh3 Last snapshot for rpool/staging-clones/images is too old:
            - those are staging clones, we don't want to back them up
            - we need to change the configuration so that they are ignored
    - but it may hide important alerts
    -

9 janv. 2025 monthly meeting

Participants

- Alex
- Pierre
- Vince

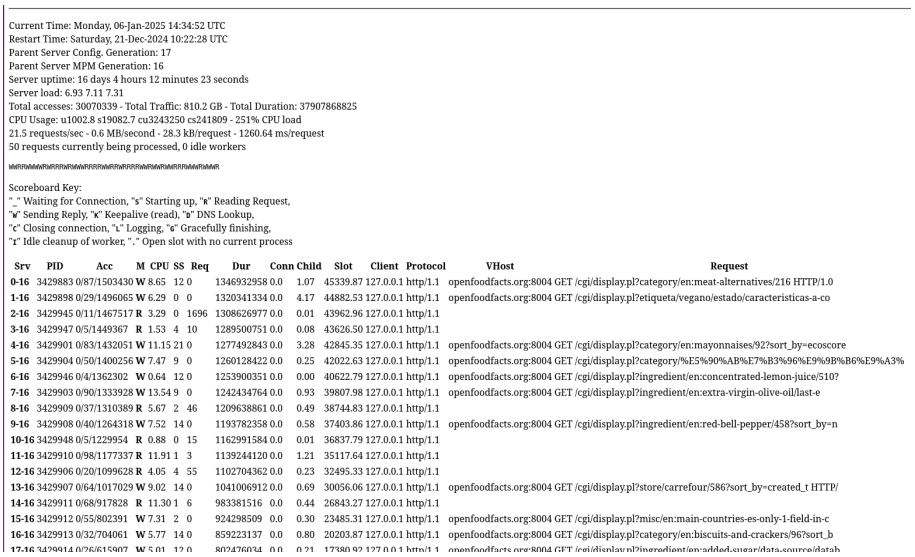
News of the month

- Alex made his first ansible request, and feels ready to work with it
- documentation of product opener release process
- small doc on new service deployments
- (minor) removed -new names on o*f instances
- keycloak -

Issue review/triage/points of concern

- we are back with availability issues on off
    - saturation for search requests
        - we did see it by looking at the server-status page showing which page each process is processing

(https://app.slack.com/client/T02KVRT1Q/C1FPYCWM7)



Current Time: Monday, 06-Jan-2025 14:34:52 UTC
Restart Time: Saturday, 21-Dec-2024 10:22:28 UTC
Parent Server Config. Generation: 17
Parent Server MPM Generation: 16
Server uptime: 16 days 4 hours 12 minutes 23 seconds
Server load: 6.93 7.11 7.31
Total accesses: 30070339 - Total Traffic: 810.2 GB - Total Duration: 37907868825
CPU Usage: u1002.8 s19082.7 cu3243250 cs241809 - 251% CPU load
21.5 requests/sec - 0.6 MB/second - 28.3 kB/request - 1260.64 ms/request
50 requests currently being processed, 0 idle workers

WWRRWWWRWRRRRWWWRRRRWWRRWRRRRWWRWRWWRRRWWWRWWR

Scoreboard Key:
"_" Waiting for Connection, "s" Starting up, "R" Reading Request,
"W" Sending Reply, "K" Keepalive (read), "D" DNS Lookup,
"C" Closing connection, "L" Logging, "G" Gracefully finishing,
"I" Idle cleanup of worker, "." Open slot with no current process

Srv	PID	Acc	M	CPU	SS	Req	Dur	Conn	Child	Slot	Client	Protocol	VHost	Request
0-16	3429883	0/87/1503430	W	8.65	12	0	1346932958	0.0	1.07	45339.87	127.0.0.1	http/1.1	openfoodfacts.org:8004	GET /cgi/display.pl?category/en:meat-alternatives/216 HTTP/1.0
1-16	3429898	0/29/1496065	W	6.29	0	0	1320341334	0.0	4.17	44882.53	127.0.0.1	http/1.1	openfoodfacts.org:8004	GET /cgi/display.pl?etiqueta/vegano/estado/caracteristicas-a-co
2-16	3429945	0/11/1467517	R	3.29	0	1696	1308626977	0.0	0.01	43962.96	127.0.0.1	http/1.1		
3-16	3429947	0/5/1449367	R	1.53	4	10	1289500751	0.0	0.08	43626.50	127.0.0.1	http/1.1		
4-16	3429901	0/83/1432051	W	11.15	21	0	1277492843	0.0	3.28	42845.35	127.0.0.1	http/1.1	openfoodfacts.org:8004	GET /cgi/display.pl?category/en:mayonnaises/92?sort_by=ecoscore
5-16	3429904	0/50/1400256	W	7.47	9	0	1260128422	0.0	0.25	42022.63	127.0.0.1	http/1.1	openfoodfacts.org:8004	GET /cgi/display.pl?category/%E5%90%AB%E7%B3%96%E9%9B%B6%E9%A3%
6-16	3429946	0/4/1362302	W	0.64	12	0	1253900351	0.0	0.00	40622.79	127.0.0.1	http/1.1	openfoodfacts.org:8004	GET /cgi/display.pl?ingredient/en:concentrated-lemon-juice/510?
7-16	3429903	0/90/1333928	W	13.54	9	0	1242434764	0.0	0.93	39807.98	127.0.0.1	http/1.1	openfoodfacts.org:8004	GET /cgi/display.pl?ingredient/en:extra-virgin-olive-oil/last-e
8-16	3429909	0/37/1310389	R	5.67	2	46	1209638861	0.0	0.49	38744.83	127.0.0.1	http/1.1		
9-16	3429908	0/40/1264318	W	7.52	14	0	1193782358	0.0	0.58	37403.86	127.0.0.1	http/1.1	openfoodfacts.org:8004	GET /cgi/display.pl?ingredient/en:red-bell-pepper/458?sort_by=n
10-16	3429948	0/5/1229954	R	0.88	0	15	1162991584	0.0	0.01	36837.79	127.0.0.1	http/1.1		
11-16	3429910	0/98/1177337	R	11.91	1	3	1139244120	0.0	1.21	35117.64	127.0.0.1	http/1.1		
12-16	3429906	0/20/1099628	R	4.05	4	55	1102704362	0.0	0.23	32495.33	127.0.0.1	http/1.1		
13-16	3429907	0/64/1017029	W	9.02	14	0	1041006912	0.0	0.69	30056.06	127.0.0.1	http/1.1	openfoodfacts.org:8004	GET /cgi/display.pl?store/carrefour/586?sort_by=created_t HTTP/
14-16	3429911	0/68/917828	R	11.30	1	6	983381516	0.0	0.44	26843.27	127.0.0.1	http/1.1		
15-16	3429912	0/55/802391	W	7.31	2	0	924298509	0.0	0.30	23485.31	127.0.0.1	http/1.1	openfoodfacts.org:8004	GET /cgi/display.pl?misc/en:main-countries-es-only-1-field-in-c
16-16	3429913	0/32/704061	W	5.77	14	0	859223137	0.0	0.80	20203.87	127.0.0.1	http/1.1	openfoodfacts.org:8004	GET /cgi/display.pl?category/en:biscuits-and-crackers/96?sort_b
17-16	3429914	0/26/615907	W	5.01	12	0	802476934	0.0	0.21	17380.92	127.0.0.1	http/1.1	openfoodfacts.org:8004	GET /cgi/display.pl?ingredient/en:added-sugar/data-source/datab

    - work in progress to have two apache instances on off
        - one for priority request (home page + read product)
        - one for the rest
        - code PR - docs PR
- No progress yet on hertzner re-installs while we are saturating on servers
    - Vince will be able to spend some time working on that this weekend
- rate limiting on nginx instead of apache?
    - we have tons of 429 responses
    - delay nginx response? tarpit module?
    - too many 429 responses for x minutes → ban ip for x days ?
        - **do-able with fail2ban**
    - → use kong when we have keycloak (on the api)
- robots
    - ai blacklist: https://github.com/anthmn/ai-bot-blocker
    - top user agents
        - off@off:/srv/off/logs$ (off) tail -n 100000 access_log | ./concatenate_by_user_agent.pl | tail -n 20
        - Lenus/166 CFNetwork/1568.200.51 Darwin/24.1.0 679
        - Mozilla/5.0 (Linux; Android 6.0.1; Nexus 5X Build/MMB29P) AppleWebKit/537.36 (KHTML, like Gecko) Chrome/131.0.6778.85 Mobile Safari/537.36 (compatible; GoogleOther) 691
        - **Mozilla/5.0 (Macintosh; Intel Mac OS X 10_15_7) AppleWebKit/537.36 (KHTML, like Gecko) Chrome/131.0.0.0 Safari/537.36; compatible; OAI-SearchBot/1.0; +https://openai.com/searchbot 753**
        - Mozilla/5.0 (Linux; Android 6.0.1; Nexus 5X Build/MMB29P) AppleWebKit/537.36 (KHTML, like Gecko) Chrome/131.0.6778.204 Mobile Safari/537.36 (compatible; AdsBot-Google-Mobile; +http://www.google.com/mobile/adsbot.html)791
        - **Mozilla/5.0 (Linux; Android 7.0;) AppleWebKit/537.36 (KHTML, like Gecko) Mobile Safari/537.36 (compatible; PetalBot;+https://webmaster.petalsearch.com/site/petalbot) 824**

        - Mozilla/5.0 (compatible; YandexBot/3.0; +http://yandex.com/bots) AppleWebKit/537.36 (KHTML, like Gecko) Chrome/108.0.0.0 838
        - Mozilla/5.0 (Windows NT 10.0; Win64; x64) AppleWebKit/537.36 (KHTML, like Gecko) Chrome/131.0.0.0 Safari/537.36 Edg/131.0.0.0 913
        - open-prices/0.1.0 921
        - Python/3.7 aiohttp/3.8.0 1106
        - Mozilla/5.0 (iPhone; CPU iPhone OS 18_1_1 like Mac OS X) AppleWebKit/605.1.15 (KHTML, like Gecko) Version/18.1.1 Mobile/15E148 Safari/604.1 1258
        - Mozilla/5.0 (Macintosh; Intel Mac OS X 10_15_7) AppleWebKit/537.36 (KHTML, like Gecko) Chrome/131.0.0.0 Safari/537.36 1279
        - Mozilla/5.0 (Windows NT 6.1; Win64; x64) AppleWebKit/537.36 (KHTML, like Gecko) Chrome/60.0.3112.113 Safari/537.36 1349
        - Mozilla/5.0 (Linux; Android 6.0.1; Nexus 5X Build/MMB29P) AppleWebKit/537.36 (KHTML, like Gecko) Chrome/131.0.6778.85 Mobile Safari/537.36 (compatible; Googlebot/2.1; +http://www.google.com/bot.html) 1641
        - okhttp/4.12.0 1738
        - Dart/3.5 (dart:io) 2077
        - MarkenDetektive App - Version 3.0 - https://markendetektive.de 2206
        - **com.teacapps.QRbot - iOS - Version 2.5.4 https://qrbot.net 2472**
        - - 2601
            - “-” user agent: reject them?
        - Mozilla/5.0 (Linux; Android 10; K) AppleWebKit/537.36 (KHTML, like Gecko) Chrome/131.0.0.0 Mobile Safari/537.36 6593
        - Mozilla/5.0 (Windows NT 10.0; Win64; x64) AppleWebKit/537.36 (KHTML, like Gecko) Chrome/131.0.0.0 Safari/537.36 7045

Points to discuss

12 déc. 2024 monthly meeting

Participants

- Alex
- Charles
- Vince

News of the month

- we installed a temporary instance of opff to test keycloak (work done by john)
    - https://world.new.openpetfoodfacts.org/ (off off pour le basic auth)
    - we need tests
        - log in with existing account
        - create test account etc.
        - delete account: needs confirmation
        - need to test admin account
        - https://auth.openfoodfacts.org
        - https://auth.openfoodfacts.org/realms/master/protocol/openid-connect/auth?client_id=security-admin-console&redirect_uri=https%3A%2F%2Fauth.openfoodfacts.org%2Fadmin%2Fmaster%2Fconsole%2F&state=8445fc14-2c74-418e-b989-5eb89257df4a&response_mode=query&response_type=code&scope=openid&nonce=dc271088-0ea8-4fff-835f-eafcb301f535&code_challenge=HLIOUEKrT4mO5uBcd0VbO93Vg9DFaotJ-a81FSFozaI&code_challenge_method=S256
- we installed keycloak docker compose in a container
    - https://github.com/openfoodfacts/openfoodfacts-infrastructure/pull/434
- we have some new servers to install (Hetzner)
    - Christian did minimum install on two of them (store and compute)
    - 3 others needs to be installed: 1 store and 2 compute
        - goal: do it with ansible
- Vince started an repository for ansible
    - https://github.com/openfoodfacts/openfoodfacts-ops/
    - next step: email configuration with mailjet
- started a document on handling alerts
    - it would be great to enrich it
- we updated the firmware 2 off2 disks
    - some user reported huge benefits in term of latency on the mobile app
- New bugs:
    - ovh3 NMI error received for unknown reason on CPU
- We merged products folders between flavors
- Google renewed their credits for this year 🎉

Issue review/triage/points of concern

- do we shutdown / cleanup the old opf/opff/obf containers (before migration to new Product Opener version)
    - TODO
- after products merge:
    - we must cleanup mounts on off/opf/obf/opff now that products are in the same dataset
    - remove datasets that are not used anymore at some point

Points to discuss

- Decision: send emails from an external provider for new emails

- Alex: would like to work on making a script to send back data about servers (/etc/pve folder, existing zfs datasets, etc.)
    - enables checking datasets are replicated and where
- will need secrets handling
    - ansible-vault or git crypt
    - start with ansible-vault eventually go for git crypt next
- how to test ansible deployments ?
    - there are not so many evolutions
- a target for deployment is blue / green deployment
    - this means really separating data on our deployments

12 nov. 2024 - Planification meeting

Participants

- Pierre
- Vincent
- Stéphane
- Alex (cannot make it, unplanned issue)

News of the month

-

Issue review/triage/points of concern

- Servers
    - Free foundation ok to dedicate a Dedirack XL for OSM + OFF in DC2 (Bezons, near Paris, where current off1 and off2 server are hosted)
    - 1 Hetzner server ordered and received
        - need to do some testing
            - disks
                - Christian: disk models ok, up to date firmware, smart tests ok
            - zfs performance
- OVH3
    - sda disk done
    - sdb: seems ok even though ovh wants to change it
        - decision: do not change sdb
- moving monitoring VM to GCP
    - missing OVH account for DNS zone /vince (.net + .org)

Points to discuss

- Scaleway: describe the servers we have, don't ask for temporary servers

9 oct. 2024 - Planification meeting

Agenda

Topics

- security, firewall
- backup and resiliency
- server updates
- redundancy / spare machines
- systematization, documentation, complexity
- skills and human resources
- software architecture
- monitoring, observability
- environmental impact

Starting point

- constrained resources that are used almost at capacity
    - resources are currently under-estimated
- hardware issues
- needs that are going up
    - more and more data
    - more and more uses
    - more and more code
- backup to be improved
- monitoring to be improved
- security to be improved (software updates + more restrictive firewall)
    - security is an issue (firewall wide open)
    - some backups configuration and proxmox configuration are not in git
    - update servers difficult without redundancy or spares
- maintenance is too complex and time consuming
- ease of administration to be improved
    - reproduce setups
- see:
- At OSM:
    - spare machine: 1 for every production server

Where we want to land

- Enough capacity OR enough flexibility to handle the foreseen growth in data, uses and running code

    - history
        - e.g. web users: +20% / year
        - mobile users: +80% / year
        - images: doubled in 3 years (~ similar than products)
        - https://analytics.openfoodfacts.org/index.php?module=CoreHome&action=index&idSite=2&period=day&date=yesterday#?period=month&date=2024-10-09&category=Dashboard_Dashboard&subcategory=1
    - new projects / features
        - open prices is just starting and stores additional images
        - when launching new products/services : need to evaluate about the impact on the infrastructure (new resources needed, additional load on existing services)
- cost per user needs to be under control

Path to go there

- Quick n Dirty estimation of needed resources
- Cultural changes
    - split things to scale horizontally,
    - maybe separate database completely from docker-composes (to scale horizontally)
    - see how critical services are and separate them
    - do not overload servers, keep a lot of spare (1-1 for critical services)
- Important point to tackle **redundancy and spare:** which strategy
    - ask more providers for **free bare metal machines**
        - limit: it's better for proxmox to have one cluster per data center
            - low latency required between proxmox hosts (it MAY be ok between same provider datacenters)
            - ask at least for 3 servers each time (for quorum requirement, and have 1 spare for 2 production servers)
        - "it costs nothing to ask" for resources, help, etc
    - better have our **own hardware** in big housing?
        - but it does not help us going to standard commercial offers
        - limit: it asks for more competence in hardware, it asks for physical
        - dedirack xl: 1579,00 € HT /month
    - have an **annual budget** that is offered by providers to take what we want
        - this offers us a lot of flexibility with the comfort of commercial service
        - we must secure this budget for multiple year
    - in between (waiting for budgets):
        - do we use pay for hosting ?
            - 3 servers proxmox
        - do we invest in existing rack
- Architecture
    - split the monolith
        - help have more part scale horizontally
    - avoiding sto?
        - object storage (S3 like API) or PostgreSQL
            - next step: do some benchmarking

Short term decision

- get 3 or 4 servers at a low cost provider to have one more cluster; to be confirmed
    - high number of cores (min. 16/32?)
    - 256 GB RAM (ECC, important or not)
    - SSD NVMe: 2 x 2 TB mini
    - 2 x 22 TB HD?
    - bandwidth: 1Gb/s is ok
    - **“Compute”**: https://www.hetzner.com/dedicated-rootserver/ex130-r/configurator/#/
        - **Xeon Gold 5412U — 24c/48t** — 3120/52000 cpubenchmark (currently 8c/16t on off1 & off2)
        - 256 GB RAM ECC
        - 2 x 2 TB NVMe SSD
        - 1 Gb/s
        - 166.80€ incl. 20 % VAT x 3 = 500.4€ / month => 6005€ / year
        - once-off setup per machine: + € 94.80
    - **“Store”**: https://www.hetzner.com/dedicated-rootserver/sx65/configurator/#/
        - AMD Ryzen™ 7 3700X — 8c/16t — 2660/22546 cpubenchmark
        - 64 GB RAM ECC => 128 GB RAM ECC = +30€ / month / machine
        - 2 x 1 TB SSD
        - **4 x 22 TB** SATA HDD
        - 1 Gb/s
        - 160.80€ TTC / month x 2 = 321.6€ / month => 3859 € / year
        - once-off setup per machine: + € 46.80
    - TOTAL = 9864€

    - To compare with current off1/off2
        - Intel Xeon Silver 4110 — 8c/16t — 1595/10216 cpubenchmark
        - 128 GB RAM ECC
        - 4 x 2TB SSD NVMe?
        - 2x4 TB SATA HDD
        - 2x14 TB SATA HDD
    - To compare with current ks1:
        - KS-STOR - Intel Xeon-D 1521 — 4 c / 8 t — 1685/5696 cpubenchmark
        - 16 Gb RAM
        - 4x 6 Tb HDD + 500 Gb SSD

- Scaleway?
    - budget or dedirack or new bare metal?
    - on a identifié 2 solutions possibles :
        - soit le budget annuel : 20K€ (par orga)
        - soit le dedirack XL + 6kw + 10Gb + 4 pass biométriques (Christian, Stéphane, Alex, ?)
        - sur 5 ans

- High priority on:
    - Infrastructure as code
    - Ansible because it's more easy to find contribution on it
- Backups designs
- Move monthly meeting to Thursday (more availability of Christian and Vincent)
- Write blueprints before deploying new methods
- Workshop related to risks: what are the 3 most important risks (https://cyber.gouv.fr/publications/agilite-et-securite-numeriques-methode-et-outils-lusage-des-equipes-projet)
- Create #infrastructure-fr (while removing 3 other channels)

Note: we have already done part of this job: Infra 2022-2030

News of the month

- We have a better handling of requests (less 503)
- moved images to ks1 - a paid small bare metal server
- making ovh3 backups better - duplicating some data on moji server
    - next action: ask for ovh3 sda disk replacement
    - use ovh3 + ks1 to serve images
- deployed the new instances of OpenXFacts
- Off-query moved to Moji
    - because off1 was constantly rebooting (seems to have hardware problems linked to Nvme disks or controller, but only when there are a lot of I/Os)

Points to discuss

- Have more people have access to ovh consoles (2 accounts: one sponsored, one for domain names + paid servers)
    - have a tech account on ovh.com with password in shared KeepassXC ?
    - or give access to Raphaël only
    - → **access given to Raphaël**

8 Oct 2024 | Infrastructure monthly meeting

News of the month

- a lot of performance issues
- a lot of performance improvements

Issue review/triage/points of concern

- hardware
- migration: normalization of product barcodes and paths for .sto and images

Points to discuss

13 Aug 2024 | Infrastructure monthly meeting

Attendees: Stéphane Gigandet Charles Népote Vincent Bataille Pierre Slamich cquest@openstreetmap.fr rorypowis@gmail.com c_9d0e7fc68f19f64fd91f60ac1936a0364750ec868272cefeb57464dba0ef906c@group.ca… Christian Quest ~~Alex Garel~~

Notes

-

Action items

- [ ]

8 août 2024 exposing services on moji

Good news:

- we can use stunnel between the reverse proxy on OVH and robotoff on moji: stunnel supports ipv6.
    - Request to Robotoff API: client – http/ipv4 → reverse proxy - ipv4 -> stunnel client on ovh - ipv6 -> stunnel server on moji - ipv4 -> robotoff on Moji (private)
- Add a container for stunnel client
    - should work in IPv4: it will be accessible only from other VM on Moji server, using local ipv4
- Add a container for stunnel server
    - must have a public IPv6 address. Reverse proxy on OVH1 (VM 101) will reach the stunnel server using the ipv6 address.
- robotoff nginx proxy config should forward to stunnel client on the right pid

16 juil. 2024 summer meeting ☀️

News of the month

- After performance problem we seems more or less back on track but still latency on some request
- Moji server install begun
- Big disk space problem on off2 due to sanoid not cleaning some old snapshots
    - this is a configuration issue, only partly resolved yet
- OVH: still waiting for a new server: Advance 4 with 128 GB + 2*1TB + 2x8TB SSD
- OVH2: creating docker volumes to prevent disk space shortage

Summer plan

- Move robotoff to Moji
- Ensure backups
    - improve script to check backups: also check snapshots removed
    - scripts to check we backup everything (needs fist a script to put states in git)
    - only use sanoid / syncoid (no proxmox tar/gz)
    - finish on OFF1/2 and do it on OVH1/2/3 and Moji
- upgrade OVH servers (Proxmox)
- resolve off2 hdd latency problems
    - 2 HDD de off2 qui déconnent (je pense qu'un upgrade de leur firmware résoudra ça)
        - Latency issues on 2 older disks of the 4 hdd on off2 **TODO**
        - firmware version is GB01 vs GB03
        - we could try to upgrade: https://www.dell.com/support/home/en-us/drivers/driversdetails?driverid=0942y
    - upgrade hdd driver on off2 ?
        - sdc, sdd: Firmware Version: GB01
    - replace off2 hard drives ?
- (Charles) See if we can install a 3rd server in free datacenter
    - a 1U server without GPU
    - goal is to enhance reliability in case of accident on one of the servers
- **Easy to do**:
    - Fight against the bursts
- Perf measurement on productopener
    - (Stéphane) see if it's easy to quickly create some stats on time/requests per type of views
        - (first quick implementation) just output request time in log
    - extend rate-limiter to facets

Later:

- ensure firewalls (try to use proxmox, enforce more restrictive rules on hosts)
- Monitoring server:
    - get a server

    - put monitoring dockers on it
    - create a munin instance on it
- prefix facets by facets/ to help easier rate limiting, robot.txt, etc. + only use ids in url (not translated names)

18 juin 2024 Monthly meeting

News of the month

- very hard time for our infrastructure
    - users and re-users are experiencing a lot of 500 errors
    - see below (11 juin 2024 performance issues review) and 2024-05 Performance issues - off2 - Product Opener
- we activated rate limiting
    - only on search and products ATM (not facets)
- optimizations on off-query postgresql configuration to avoid a lot of writes thanks to Christian and John
- we have contacted some of our big reusers
    - MCI will fix an issue that provide a dozen of request instead of one
    - Foodvisor now use a CDN for the images
        - Foodvisor was ~50% of our image requests, now ~20%
        - root@off2:/home/CharlesNepote# tail -n 1000000 /zfs-hdd/logs-nginx/static-access.log | grep -c Foodvisor
        - 200044
- removed NFS shares
    - it was far too much opened, it was a legacy usage due to migration

Issue review/triage/points of concern

- data upload is 160 G / 24h for html/data files ~ 2M / s

Points to discuss

- how to avoid replaying the scenario of a bad database config
    - have a real “go to production” phase
        - it’s a dialog within the team
        - to verify allocated resources (RAM, disk, SSDs, etc.)
        - to verify performances
- create documentation on diagnosis
    - nginx diagnosis
    - a page to recap all diagnosis
- make the list of banned ips public?
- changing disks on off2 ?
- add iptables rules to avoid NFS ?
- how to handle: iptables rules ?
    - flatfile of rules in the git might be the best practice

- on images:
    - moving serving images to off1
        - (Christian) still need to investigate the bad response times when images are served from off1 (host or container)
    - TODO: encourage big re-users to use cache on images etc.
    - Shan't we ask people to have a token so that we can identify big re-users and contact them
- Shan't we put some limitations on added values services like facets
- mongodb optimization
    - no progress
    - idea: add a single image_front_url field to all products, so that we don't need to get the huge "images" structure from mongodb
        - possibly: remove the "images" structure from mongodb?
            - impacts reusers who use it…
        - remove some other fields? e.g. big fields like nutriscore_data / ecoscore_data (if not needed for personal search)
- very slow commands
    - root@off2:/mnt# pct enter 101
    - root@proxy:/etc/nginx/sites-enabled# vi off
    - → took more than 10 seconds to display an empty file (off doesn't exist in fact)
- API cache not working?
    - check if the cookie changes on subsequent requests
    - see the following example into /var/log/nginx/openfoodfacts.org.log on@proxy

170.133.4.212 - - [11/Jun/2024:19:10:03 +0000] "GET /api/v3/product/4016241030603/?lc=de&tags_lc=de&app_name=MCI+-+Personal+Training+AI&app_version=1.1.990+1&app_platform=iOS&comment=use+website+for+contact+and+more+information HTTP/1.1" 200 40017 "-" "- MCI - Personal Training AI - 1.1.990 1 - iOS - https://mcisolutions.de - use website for contact and more information" "-" MISS [1.316] (repeated)

- create a separate log for scans / product views, to separate from the rest
- host exports elsewhere
    - ovh3?
    - decentralized solution: https://www.datprotocol.com/ ?

11 juin 2024 - Performance issues review

- Goals
    - Prioritize next actions
    - Decide who does what
- Symptoms as of 11/06/2024
    - Continuing periods of unavailability of OFF website and API
    - Search, facet queries etc. that don't return results
    - 2 reboots of off1 on 10/06/2024

        - Question: priority of resolving those vs everything else?
            - Searchalicious is a top priority in the short term (June)
- Review of issues and potential solutions
    - See 2024-05 Performance issues - off2 - Product Opener
- Prioritized list of actions
    - (to complete and discuss, just listing possible items for now)
    - (Stéphane) Investigating memory usage and OOMs
        - limit on container?
        - reason why apache not restarting
    - (Christian?) Investigating why serving images on off1 is so slow
        - does maybe include: Investigating slow disk (is it slow, where could it come from)
        - CDN? from Google?
        - move images to yet another server?
            - need zfs replication
    - (Stéphane) Move logs to nvme (reverse proxy, off)
    - Find a way to reduce the cost and time of facet and search queries
    - Monitoring
        - (Charles) install a new munin? (computel down)
            - munin.openfoodfacts.org
            - bare metal chez Google ? ou chez OVH ?
                - Charles donne la config à Raphaël
                - Vince installe ?
                - Condition : pas d'accès root depuis cette machine à des autres machines en root, idéalement pas d'accès d'aucun utilisateur (probe en https via le reverse proxy)
        - install prometheus node exporter
    - Reduce number of requests
        - (Raphaël) rate_limiting en Perl
            - it's for reading, not writting: products and searches
            - in production: not blocking, but logging
            - 10 req/min for search & 100/min for the products
            - facets are not yet included
            - need to eventually whitelists some users
        - proxy_cache_use_stale ?
    - Flush des logs d'Apache ?
    - Investigate why off1 reboots
        - cpu temperature?
        - munin graphs available locally on off1?

14 mai 2024 Monthly meeting

Participants: Pierre, Stéphane

News of the month

- Alex should be back in a week or two

Issue review/triage/points of concern

- Some unavailability of the OFF website and API in the past few weeks
    - Stéphane banned some bots (mostly ML engines like ClaudeBot and others), but load remains high
    - A dozen of ips that make ~100 requests for the same product, then ask for another product
        - [14/May/2024:16:15:02 +0000] "GET /cgi/display.pl?api/v0/product/4099200011370.json HTTP/1.0" 200 18824 "-" "Dart/3.2 (dart:io)" 0/12813

Points to discuss

- Cloud provider contacted us, Pierre to ask if they can offer us some credits

18 avr. 2024 internal team

Infos:

- Moji server is really close to be given
    - 8x8T HDD
    - 8T SSD
        - pour ref, 2*2T sur Free,
    - 2 GPU 1080
    - governance → fully OFF
    - “legal”? (contract ? convention ?) (=> Charles)
- OVH no news (4-5 follow-ups already) (=> Charles)
- Matomo still with problems but I’m on it (splitting archival process) (Alex)
- off-query deployed on off1 - docker within a container - test

Discussions:

- Disk problems on off2 very high latency on two disk (HD)
    - CQuest propose to change one disk with the spare we have to see if it resolves partly the problem (for this disk)
        - ok
    - TODO: Planifier une visite chez free
- Main arguments of concern for Alex
    - Backups
        - Good pattern on off1/off2/ovh3 (syncoid)
        - ovh1/2 less robust
        - tests of PRA
        - TODO: documentation of current / target
    - Firewalls - more restrictive policy
    - Updates
        - hosts: Proxmox 8 on OVH

        - containers
        - TODO: update checklist: 1. make snapshot, 2. update, 3. quick test (reboot?), 4. remove snapshot after one week?
- Release Product Opener to document
    - example on mobile: https://github.com/openfoodfacts/smooth-app/blob/develop/RELEASE.md
- Observability:
    - I have the pattern to proxy metrics exporters via nginx but lack time to deploy more → someone motivated ?
        - deployed redis exporter
            - using compiled binary of exporter from github
    - eg: latency on nginx servers
    - Add more exporters
        - Stéphane: add reverse proxy nginx exporter
- move nginx for static from off2 host → reverse proxy
    - mount images on reverse proxy
    - move config to reverse proxy
- activate rate limiting on images ?
- rate limiting by Raphaël for off (Product Opener) ?
    - dry run mode by default
    - use Redis in product opener
    - https://github.com/openfoodfacts/openfoodfacts-server/pull/10144
- should we really reserve moji server only for robotoff ? (it's a monster)
    - Moji server: proxmox?
- TODO: take an external VPS for monitoring ?
    - maybe at google cloud compute
- mail:
    - dmarc policy → reject ?
    - all outgoing mails go through the proxmox mail gateway cf https://openfoodfacts.github.io/openfoodfacts-infrastructure/mail/
- Bounty ? → add a page ? .md in infrastructure explaining that we thank them for their reports but that we can't offer monetary compensation
- what do we need to be **really** at ease ?
- ovh3 root mail not arriving currently → reboot ovh3 ?

9 avr. 2024 Monthly meeting

Participants: Stéphane, Pierre, Bohdan, Stéphane, Taxonomist

News of the month

- work done on backups to replicate between off1 / off2 / ovh3 using sanoid / syncoid
    - use a consistent pattern every-where
    - a service is checking that we have backups for everything needed
    - no more use of root access to sync backups (use ZFS delegation and guest accounts)
    - TODO need to remove root access to ovh3 from off1
- split VM disk for 200 and 201 separating docker volumes from the rest
    - mainly to avoid having too big disk + may be used to apply different backup strategies
- we had no space left of ovh1 during a week-end (linked to disk split)
- analytics seems under control after patching Matomo the code at two places and using 8 queues…
- new tool deployed: argilla : annotation software for ingredients
- we have a commissioned server sponsored by Moji

Issue review/triage/points of concern

- We need more disk space on OVH side… we are currently struggling with it
    - solutions:
        - move robotoff to the new moji server
        - move off-query to off1
- split VM disk, I wasn't able to do it cleanly, if someone can enlighten me on usage of grub ! (had to resort on debian disk in rescue mode, which means a downtime)
    - needs to do it for monitoring…
- backups not finished
    - I want to migrate to using sanoid / syncoid for all CT/VM at OVH
    - I'm not very secure about current mechanism (no snapshots)
- we need Moji server access to deploy robotoff
- getting zammad back to our infrastructure not done yet
- putting off-query on off1 not done yet
- we don't have enough available time
- two disks on off2 have very high latency
- (Pierre) we still have an issue with under-reporting in Matomo

Points to discuss

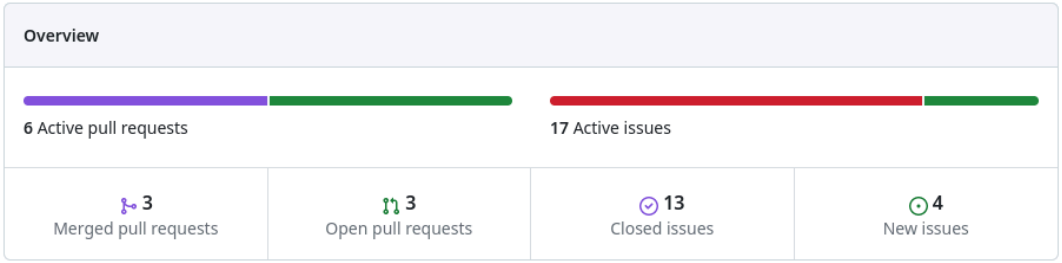
- have more prometheus metrics
    - we have a scheme to expose metrics through reverse proxy for servers off1 and off2
    - shall I store exporters binaries in github → no just store the url
- have more grafana dashboards
    - eg. mysql Matomo
- Helping on backups:
    - https://openfoodfacts.github.io/openfoodfacts-infrastructure/zfs-overview/
    - https://github.com/openfoodfacts/openfoodfacts-infrastructure/pull/323

12 mars 2024 Monthly meeting

Participants:

News of the month

- https://github.com/openfoodfacts/openfoodfacts-monitoring/pulse/monthly (1 PR)
- https://github.com/openfoodfacts/openfoodfacts-infrastructure/pulse/monthly



- Redis deployed to production with stunnel, used by off-search and off-query (soon)
- we met with OVH partnership, they told they would sponsor us a fourth server, and renew current convention
- work done on Matomo side but it's still failing
    - still unclear where the bottleneck is: mysql? php?
- added RESTART_POLICY to some Docker compose projects:
    - knowledgepanels-facets, openfoodfacts-query, taxonomy-editor
    - it's useful to have it to "no" on local dev, but "always" in staging and production
    - we should try not to forget this
- we have a first prometheus exporter for mongodb on free side
    - behind nginx proxy
    - for example you can monitor
- we now have an imperfect (some files are failing) google drive backup on a container on ovh3 using rclone
- started to continue work on backup
    - pool backups to OVH3, make it as simple as possible
    - use sanoid / syncoid everywhere !

Issue review/triage/points of concern

- help still needed on Matomo
- we would need some more dashboard on grafana (adapting existing templates) to render some metrics that we collect :
    - matomo mariadb (mysql) monitoring
    - production mongodb monitoring
- we must move zammad to our servers (from Christian server)
- I started experimenting to split VM disk between system and docker (doing it on docker-staging)
    - I will first start by putting all docker data on same volume
- ovh4 install: Stéphane in order to learn more about proxmox, ZFS, the off infra etc.

- todo: check that systemd services are configured to restart if killed (e.g. mongodb)
- just curious: why sanoid/syncoid instead of https://pypi.org/project/zfs-autobackup/3.1.2/ ?
    - no particular reason, but sanoid works well

Points to discuss

- Open X Facts preprod instances ?

13 févr. 2024 Monthly meeting

Participants: Stéphane, Vince, Pierre, Alex

News of the month

- worked on backups:
    - backups between off1 and off2
        - off1 -> off2 + ovh3
        - without root access ssh
    - backups of off1 to ovh3
    - backups from off2 to ovh3 needs to be reworked
    - sanoid
- Redis installed on off1
    - in a container
- matomo problem of performances, still coping
    - https://github.com/openfoodfacts/openfoodfacts-infrastructure/pull/300
    - moved one of the queue list in redis to cope on a specific system
        - PR still to open
    - possible alternative: https://plausible.io/ ?
        Raphaël: or countly (https://countly.com/)
    - there might be quick win
    - disable some plugins? e.g. shop
    - Look at GeoIp options ? “J'utilise le module serveur Geoip2 (Nginx, Apache...) et je veux configurer les variables du serveur”
- better backup checks see sanoid_check
- documentation as progressed (still needs more)

Issue review/triage/points of concern

- Matomo still not reducing gap ☹️
- Off staging disk is big:
    - ovh1 - dockers-staging
    - remediate
- mongodb down on Monday 13/2/2024

    - to investigate
    - export script (mongo_dump.sh) normally tries to only writes if process finished with 0 code
- monitoring CSV generation
    - it should fallback to yesterday's JSONLines, worst case scenario, symlinks, size checks ?
- munin notifications etc.
    - Vince will setup a VPS to have our own Munin
    - do it on OVH account where we have domain name
    - Add maybe an env variable to convey severity (1-5) urgency bool
        - modify script that send emails to take into account env variables
        - modify services to add env variable
- only use sanoid for backups (and as a replication)
    - missing backup VM / CT configuration
    - make a test for real
- problem on disk on off2: vince go have a look

Points to discuss

- backup strategy: shift all to sanoid, even replications ?
- hardening our configurations
    - should we disable password authentication in ssh
    - should we disable NOPASSWD for sudo (require password)
        - only on hosts ? It's better

9 janv. 2024 Monthly meeting

News of the month

- MongoDB moved to off1

Issue review/triage/points of concern

Points to discuss

- should we force 2FA on proxmox interface (I would advise so)
- ZFS tuning
    - Do we have enough memory for ZFS on our servers ? According to proxmox docs, *4GB plus 1GB RAM for each TB RAW disk space.* That would be 50T + 2T = 56G just for zfs
    - should we ask for a free subscription ?

12 déc. 2023 Monthly meeting

News of the month

- migration of off and off-pro from off1 to upgraded (double RAM, bigger HDDs, more SSDs) off2 (free data center) happened
    - proxmox
    - separated by container
    - it went quite well
    - producers imports still to finish automating
    - new sftp configuration
        - we did not handle communication well
- off1 hardware upgraded
    - (double RAM, bigger HDDs, more SSDs)
    - documentation still to come !
    - photo album to transcribe
- multiple disks problem on ovh3
    - seems solved after two disk replacement but might not !
    - documentation still in progress (various notes to aggregate)
    - we don't use it for images yet
- other ovh3 issues? this week-end unavailability
- added some more monitoring
    - more services
    - ping on servers

Issue review/triage/points of concern

- problems with Matomo analytics
    - upgraded memory and CPU
    - upgraded mysql configuration to use more memory
        - more hints needed
    - main problem was a hourly "core:archive" command launching multiple times
    - need to use plugin using Redis, hase we have a high write volume
    - must put prometheus exporter on nginx
    - (documentation in progress)
- we continue to have quite high CPU io wait on both ovh3 and off2
    - the one on ovh3 might be linked to scrub
    - we have quite a lot of ZFS sync too
    - move EAN8 images
        - https://github.com/openfoodfacts/openfoodfacts-server/pull/3915
    - use Amazon to serve images?
        - https://aws.amazon.com/fr/opendata/open-data-sponsorship-program/terms/
- we still needs more monitoring
    - prometheus exporters for production services
    - sensible grafana graph (thereafter)

- rclone of google drive - still working on it (problem with recursive shortcuts)
- still have to look at nginx rules for rate limiting (currently only logging)

Points to discuss

- we should move back images.openfoodfacts.org to proxy VM (mounting images zfs inside the container)
    - monitor first to see if it goes well
- can we remove /rpool/static on ovh3 ? (I think so)
    - obf-images off-pro-images opff-images opf-images
- do we need vz_dump on proxmox, can't we go for zfs snapshot + syncs on different servers (sanoid/syncoid to the rescue)
- upgrade of OVH servers
- backup of proxmox host server needed
    - zfs snapshot and sync there also ?(at least for off1 and off2)
- Images using a CDN
    - https://blog.openfoodfacts.org/en/news/open-food-facts-images-on-aws-open-dataset-the-ultimate-food-image-database
    - https://openfoodfacts-images.s3.eu-west-3.amazonaws.com/data/401/235/911/4303/1.jpg
    - https://aws.amazon.com/fr/opendata/open-data-sponsorship-program/terms/
- Make it harder for bot to access the full size image from the product page:
    - clicking on the CC icon: https://world.openfoodfacts.org/cgi/product_image.pl?code=5901939006048&id=front_pl
    - Clicking on the full size image
    - Could we return blank for bots ?
    - That would reduce IO for full size images

2023-12-05 Monthly meeting

participants: Alex, Stéphane

News of the month

Issue review/triage/points of concern

- the server with nginx images has very high io wait in CPU (be it off2 or ovh3) so we should find a better way to handle that.
    - the problem might be on small images there are a lot
    - can we find a proxy with good caching capabilities on NVME etc.
    - or use Ceph and so on
- our backup may not be safe enough (following ANSII recommendations)

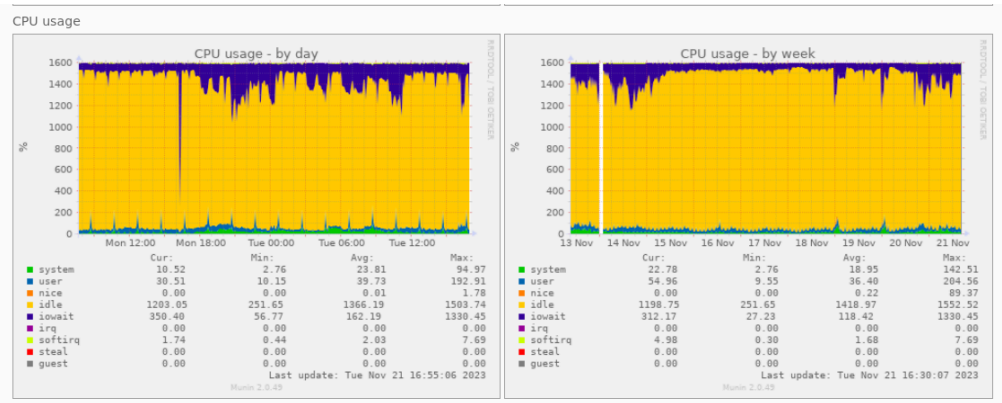
Points to discuss

-

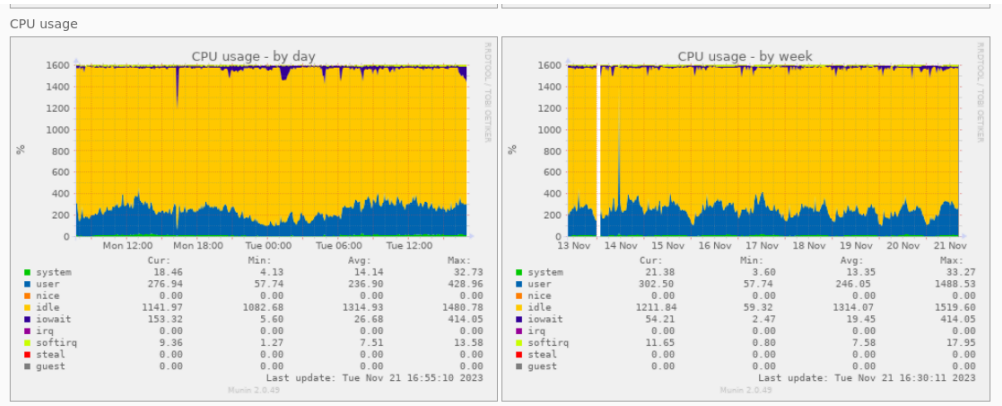
2023-11-21 Monthly meeting

News of the month

- seems we kind of resolved latency issues (but it's not that good when a scrub is taking place !). As of today
    - off2:



    - vs off1



- We had MongoDB problems - might be due to Pymongo

Issue review/triage/points of concern

- No prometheus metric exporter yet in production (but we have munin)
- ZFS going in Degraded mode yesterday (will it start again ?)
- we can't have proxmox replication on containers with bind mounts
- ELK stack for docker log still not resolved
- TODO: we should backup off2 host on ovh3
- we have to adjust: ➕ CT and VM list of OFF infrastructure

    - put 12 CPU on off and 80 Go RAM

Points to discuss

- how to structure the infra documentation, what would be ideal ?
    - a very general physical representation on readme / index
- if switch goes fine, what are the next steps ?
    - fix all producers imports
    - goes to data center to upgrade off1
- we have images.openfoodfacts.org on off2, should we move it back to ovh3 ?
    - move to ovh3 for at least a moment

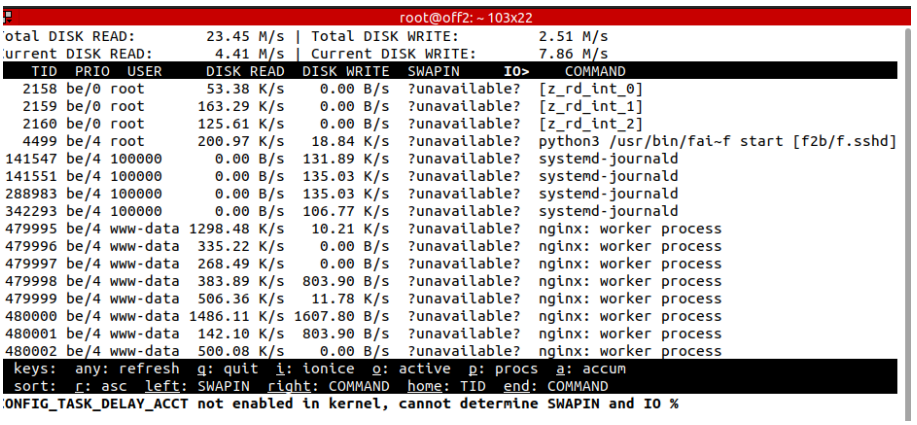
2023-10-10 Monthly meeting

News of the month

- images on: off2 - ovh3
    - certificates management on images (on off2) incident report
- little progress on off2 migration
    - import scripts
    - still some issues to fix (see fixme)
        - related PR: chore: Migration of Open Food Facts from off1 (bare metal) to off2 (proxmox)

Issue review/triage/points of concern

- erosion of disk space on off1 (srv2: 5%, srv: 7%).
    - BEWARE some folders are NFS mount of off2
    - S: will delete some old stuff (especially in imports, maybe /home/sftp)
    - can move stuff on backup folder on off2
- off2 performance - off2 is sometimes very slow
    - munin graphs: https://www.computel.fr/munin/openfoodfacts/off2.openfoodfacts/cpu.html
    - it would be better to have off2 capable of serving images + serving product opener
    - so we can images on off2 at the moment
    - Bohdan can volunteer to look into it
    - Stephane propose to cut services one by one and see impact on iowait / latency
        - use iotop / htop to see differences
    - lots of writes? output of iotop



- show load average by container on proxmox: would be useful
    - see discussion on https://forum.proxmox.com/threads/lxc-containers-shows-hosts-load-average.45724/

Points to discuss

- off1 -> off2 migration

2023-09-19 Monthly meeting

participants: Alex, Pierre

News of the month

- off migration to off2 is progressing
    - server for off and off-pro are working
    - there are now a list of fixme to finish before being able to migrate
        - most important part is import scripts
- obf/opf/opff migration on off2
    - seems to work fine
    - replication using sanoid / syncoid seems to be a win
- we had a problem on ovh3 because of disk I/O
    - there were too much bot trying to get images
    - we ban them (iptables)
    - we implemented rate limiting on off2 (because ovh3 do not support dry_run)
    - still have to analyze the logs

Issue review/triage/points of concern

- spam users creation (product opener)
    - we need to deploy on all instances
- still no prometheus exporter on new off2 installs

Points to discuss

- On off2 reverse proxy, I implemented ip ban using fail2ban for nginx + nftables, why not a deny list in nginx directly (through an include) ?

2023-08-11 - migration

Participants: Christian, Alex, Stéphane

- migrate off from off1 to off2: week of August 20 (or before August 15 if everything looks good)
- **Ask free to go to datacenter on August 21**
- upgrade off1
    - memory
        - maybe at Christian if we ordered it
    - disks
        - 14 TB disks: at Christian
            - we ordered 3 14Tb disks on
        - PCI card with 1 2TB SSD + 1 Octane: at Stéphane
        - missing: 1 2TB SSD
            - WD SN770 https://www.amazon.com/WD_BLACK-SN770-Internal-Gaming-Solid/dp/B09QV5KJHV/ref=sr_1_3?crid=2QG5EI7VR1NJ6&keywords=ssd+2tb+wd+black+sn770&qid=1691772374&sprefix=ssd+2tb+wd+black+sn770%2Caps%2C165&sr=8-3
        - missing: memory as for off1 (see what we got)
- idea on snapshot → could we add a hold

2023-08-08 - Monthly meeting

Participants: Bohdan B., Pierre S., Alex G.

News of the month

The main stake is that we are still migrating services from off1 to off2.

- migration of openfoodfacts main and producers instance still in progress
    - (ETA next week ?)
- replacement of a disk on ovh3
- rpool was full on off1 (see PR)
    - due to more snapshot on products
    - we shall investigate why snapshots take so much space
- a big work is done on bots (50% of traffic) by Raphaël
    - see fix: Improve web crawlers indexation
    - and return empty noindex webpage when crawlers hit specific pages
    - presentation of raphaël of monitoring data using http probes on main page

    - it will likely have a major impact on performances
- bug after Open Products Facts migration (client body size): The classic Open Products Facts gets a 413 HTTP Error

Issue review/triage/points of concern

- manually remove products snapshots on off1 (until off2 migration)
- https://github.com/openfoodfacts/openfoodfacts-infrastructure/issues/199
    - Bohdan will have a look
    - on production VM (201) we mainly have
        - robotoff
        - metrics (business metrics)
        - facets-knowledge panels
        - label studio
- Creating Docker images for Open Beauty Facts, Open Pet Food Facts and Open Products Facts using https://github.com/openfoodfacts/openfoodfacts-server/blob/main/Dockerfile.frontend, https://github.com/openfoodfacts/openfoodfacts-server/blob/main/docker-compose.yml (+ relevant.env files https://github.com/openfoodfacts/openfoodfacts-server/blob/main/env.pro) as an inspiration

Points to discuss

- rate limiting on our main website
    - we might a redis plugin for that
    - might be implemented in product opener
    - we shall limit based on request type: tags / search / product / static page
- should we immediately plan intervention to upgrade off1
- Logging for Open Food Facts - Prometheus / Grafana
- off1 - services still there:
    - to be moved:
        - off - off-pro
        - cestemballepresdechezvous - madenearme - madenearme-uk
        - combiendesucres - howmuchsugar (see how much it needs to be on same server as off)
    - can be removed:
        - foodbattle - old stuff
        - ~~hunger-game~~ - old stuff
        - supervision - we now use the munin instance of Christian
    - to be decided
        - labelme

2023-07-11 - Monthly meeting

News of the month

- Migration off openbeautyfacts.org (obf) and openproductsfacts.org (opf) to off2 server done (see PR and related PRs)
    - we now have a lot of thing in ZFS
        - off and off-pro products are still using specific script to sync (will change with off migration)
- Added some monitoring of ZFS snapshots and sync, see sanoid_check.sh

Issue review/triage/points of concern

- images.openfoodfacts.org is going down some times (added to https://status.openfoodfacts.org/)
- madenearme to be modified to use JSONL

Points to discuss

- can we upgrade off2 in August ?
- install prometheus exporters for nginx + apache on product opener instances + nginx frontend
    - do we expose them in https or through stunnel ?
- Fayez will help on ELK stack : https://github.com/openfoodfacts/openfoodfacts-infrastructure/issues/199

2023-06-13 − monthly meeting

Attendees: Alex, Stéphane.

News of the month

- Open Pet Food Facts moved to off2 (see infra PR and product opener PR)
- using sanoid / syncoid for all new datasets
- Images moved to ZFS on off2
- started the move of Open Product Facts and Open Beauty Facts
- some sync problems on products
- we added better backups for robotoff data
- recurring problems on ovh3 - ZFS datasets not accessible (conséquence: openfoodfacts.net and images.openfoodfacts.org down)
    - we still don't understand the causes
    - next time: to try
        - for device in sd{a..f};do smartctl -a /dev/$device; done
        - try fdisk /dev/sda (up to sdf)
        - zpool status ?
        - do we want to try failmode ?

            - "zpool set failmode=continue rpool"
        - could we do a zpool reopen
    - action: monitor images server with blackbox
- started an incident log by server
    - the idea is to help spot recurring issues

Issue review/triage/points of concern

- server incident log: is the format ok ?
- we are still stalled on the filebeat / logs in elasticsearch side
- still todo: updating products clone for staging (.net) daily
- we are still moving opf and obf to Matomo
- and we have to remove it from off
- Help needed: Matomo usage and making it performant
- ARP problem on proxmox ? See Slack message

Points to discuss

- using git and links on the servers to track configuration files
    - pulls / update might be tricky (but generally ok)
    - is there an easy way to manage authorship ? (maybe a file to source to set env variables)
- I would like to restructure the doc to match diataxis framework that we use for other projects
    - have a ref for each service (base info, link to how-to)
    - put how-to in their own files
    - keep a flat structure using prefixes in titles and file names

2023-05-15 — New server opportunity

We have the opportunity to add a new machine in the Free datacenter, thanks to their sponsoring. We are paying the machine but they're paying for the hosting.
It's interesting to group servers at the same place:

- easier maintenance
- very good networking performance: both bandwidth and latency
- open interesting usages such as production recovery.

What for?

1. AI computation (Robotoff).
2. production recovery — if either off1 or off2 (main production servers) fails.
3. backups (images)
4. Eventually, help in case of spikes?

Constrains:

- Not too costly in terms of power consumption.
- Under 10K€ budget.

Wishes:

- Installed in less than one month: the less things to setup, the better.
    - Ready-to-use server if possible (without “physical” things to manually add/setup).
- As much CPU as we can.
- Add a GPU.
- We should not have to upgrade the machine until 3-4 years (upgrades are costly in terms of time).
- We never have too much RAM (ZFS cache is using all the RAM that remains).

What?

- Dell 740dx: https://www.serverschmiede.com/konfigurator_bulk/en/dell-poweredge-r740xd-19-2u-12x-35-lff-2x-intel-xeon-scalable-lga3647-ddr4-ecc-raid-2x-psu-server?p=e8c3c45d17a3628825d9020e6c770080
    - Dell PowerEdge R740xd 19" 2U 12x 3,5" LFF 2x Intel XEON Scalable LGA3647 DDR4 ECC Raid 2x PSU Server. Rationale:
        - 740 machines have plenty of space and three risers for PCI cards.
        - Allows to install full height GPU card, see this video
        - Allows to install 12 HDD disks
    - CPU: 2 Intel Xeon Platinum **28 cores** / 56 threads (**165w** tdp each). Rationale:
        - CPU Benchmark : 56381 (x2)
        - Facing current Xeon Silver 4110 (off1/off2): 10028
    - **512GB** Registered ECC DDR4 SDRAM (8x 64GB DIMM). Rationale:
        - We never have too much RAM (ZFS cache is using all the RAM that remains) and each 64 GB is cheap (~110€).
        - 12/24 banks would still be empty
    - 1 x NVMe Bifurcation Switch 4x M.2 PCIe x8 / x16 Controller incl. **4x 1,92TB** Enterprise Performance NVMe. Rationale:
        - 4 NVMe for both disk space and redundancy.
        - All is shipped in one card.
        - NVMe is quite faster than SSD.
    - **6 x 3,5" 20 TB** 7,2k 12G Raid Enterprise Storage 24/7 SAS HDD. Rationale:
        - plenty of HD space for both disk space and redundancy
        - allows to backup OFF photos for years
    - 1 x Dell Broadcom 57416 Quad Port **2x 10Gb** + 2x 1Gb copper RJ45. Ethernet Network Daughter Card 01224N. Rationale:
        - 10 Gb seems to be the norm now
        - a daughter card does not waste place.
    - **2 x DELL 1100 Watt** Netzteil PSU - PowerEdge R630 R640 R730 R740 R930 Tx30 Tx40. Rationale:
        - 1100w PSU are needed to support GPU according to this thread
        - 2 PSU for redundancy
- GPU:
    - Dell 740dx, and it seems the vast majority of servers, does not support 3 slots cards.

    - Dell 740dx seems to support 2 slots “full height” and “full length”.
    - Full length means 111.15 x 312.00 x 20.32 according to: https://en.wikipedia.org/wiki/PCI_Express
    - ~~NVIDIA GeForce RTX 4090~~. Rationale:
        - KO because it takes 3 slots with full height and full length.
        - ~~Better performance than Tesla TA or even Tesla V100 (!), for both inference and training on 16bits and 8bit (see this article).~~
        - ~~Another article that compares different GPUs for deep learning: https://oddity.ai/nl/blog/best-bang-for-buck-gpu/~~
        - ~~It is possible to limit power consumption, see https://timdettmers.com/2023/01/30/which-gpu-for-deep-learning/#Power_Limiting_An_Elegant_Solution_to_Solve_the_Power_Problem~~
        - ~~Power consumption around 250w (despites of its 450w TDP).~~
        - ~~This would be the following card: https://www.rueducommerce.fr/p-geforce-rtx-4090-24gb-verto-triple-fan-edition-pny-3510060-18279.html~~
    - NVIDIA GeForce RTX 4070 Ti
        - As of today, the most powerful GPU in 2 slots format.
        - https://www.inno3d.com/en/PRODUCT_INNO3D_GEFORCE_RTX_4070_Ti_X3_OC
        - https://www.alternate.fr/INNO3D/GeForce-RTX-4070-Ti-X3-OC-Carte-graphique/html/product/1891133

9 mai 2023 Monthly meeting

Attendees: Charles, Stephane, Pierre, Alex

News of the month

- re-install of off2 is progressing
    - opf install operational
    - still some work to be done on reverse proxy (wildcard certificates, etc)
    - we still have a problem on off2 NVMe disks, which is holding back migrations of images to ZFS
- moved 200 VM from ovh2 to ovh1
    - needed because staging use ZFS clone as nfs mounts on OVH3 and ovh2-> ovh3 latency is too high for such use
    - it requires to move around a lot of containers to gain space
        - it’s easy to move things around
        - we are quite short on storage
- incident on OVH3
    - a hard reboot fixed it
    - seems to be linked to a problem on sdc, where SMART was failing for a long time

- Lack of storage on robotoff machine
    - we can't really augment disk space because we are short on disk space on ovh1 / ovh2
    - the only reasonable option seems to be to move it to a new server (see points to discuss)

Issue review/triage/points of concern

Points to discuss

- new server at free, which configuration ?
    - goal:
        - put robotoff on it: needs CPU + GPU + memory
        - backup ZFS prod (complementary to OVH3, maybe not images, but OVH machines & data ?)
    - some points:
        - discussion: https://openfoodfacts.slack.com/archives/C1FPYCWM7/p1683272552984549
        - bare metal or proxmox?
            - some drop (overhead unclear, max ~15%) in performance when using proxmox
            - robotoff has not a strong need for HA (High availability) and have only some data to backup
        - it's a new machine at free
            - with really good bandwidth with the rest of the prod infrastructure
            - we can't go too often at free, so let's put a good server…
            - consider consumption efficiency
        - on GPU
            - memory is the most important point (to have all models in memory as loading/unloading is slow)

18 avr. 2023 Monthly meeting

Attendees: Stephane, Pierre, Alex

News of the month

- re-install of off2 is on the way
    - nginx reverse proxy ready (also with stunnel)
        - new ip address :)
    - creating and syncing datasets for opff
    - currently installing opff
- We have a problem with NVMe (sharing same PCI card) on off2 server

    - quite weird, the server boot, up to showing login prompt, but after some minutes it reboots
    - disabling pci slot 2 makes the server works
    - Christian changed the PCI card but it's still failing
    - we suspect it's a disk which is broken
    - documented on off2 reinstall
- another accident: off3 was down because of no space left on device, because of mongodb logs (no log rotate)
    - put log rotate

Issue review/triage/points of concern

- filebeat (logs in elasticsearch) is broken
    - WIP but I didn't find the time to fix it (seems my understanding is limited !)
    - Raphaël proposes a complete reboot of the install (we can loose past logs, not a big deal)
- slow staging environment for open food facts
    - the root problem is that we use nfs mount of zfs clones on ovh3 and ovh3 is 10 ms away ! That's far too much for nfs
    - solution is to have zclones of most data on ovh1 / 2
        - we can only keep images (which is big) as a nfs mount
- We will test metabase using staging mongodb this can help the team gather metrics and get insights

Points to discuss

- OPFF install
    - For OPFF, I used "off" user for apache, this requires editing /etc/apache2/envars, should we install using www-data instead ?
    - no because we need to run scripts with off
    - Open issues
        - some things are logged to $data_root/logs (look at current logs)
        - make a symbolic link to /var/logs/apache2 ?
    - Product Opener install more based on git repository:
        - host configurations and file in infrastructure github
            - sync scripts
        - containers configurations in openfoodfacts-server github
        - git clone the repo somewhere and use symlinks to configure files (eg systemd units are symlinks to the repo files)
        - there are some exceptions of course (still some files to edit like /apache
- I would really like to have a step toward CD on open food facts production but I would like to have a simple system using pull instead of push (more secure)
    - thanks to chat GPT, I understand you can use a socket with systemd to lauch a script, that would do the deployment.
        - Potential Drawback though: if it fails, you might not have the logs (but the github action could monitor end of script and send back the logs)

        - old PR of Olivier on CD for different flavor
            - may have problem because of volume names (see)
- (question from Charles) Hard to know which version of Open Food Facts is in production
    - remove tests from release please
    - expose changelog in /files/ (so that it's accessible from the website)

14 mars 2023 Monthly meeting

Attendees: Alex, Christian, Charles, Pierre, Vova, Stéphane

News of the month

- Migration happens: see the report (and the PR) - off2
    - off2 mongodb moved temporarily to off3
    - off2 hardware upgraded, installed proxmox on it
- We are experiencing perf issues with temporary server for mongodb
    - complains from users
- memcached on off1 was down for a long time, not noticed
- (currently) off2 is down, constantly rebooting…
    - good news: ipmi works
    - Christian is trying to solve it
- We bought
    - 3 new 15Go disks and Chrisian tested them
        - 2 for off1 upgrade
        - 1 as spare for off1 / off2
    - some 16Go optane disks

Issue review/triage/points of concern

Points to discuss

- target architecture.
    - see https://github.com/openfoodfacts/openfoodfacts-infrastructure/pull/191

Plan for 2023

- server migration with a clean containerized install - and zfs syncs
- better encrypted two-way communication between data centers (stunnel + https)
- backups checks
    - testing backups through staging
        - with automated deployment of new clones
    - monitoring backups

- more active checks (monitoring - alerts)
- less false positives in alerts
- better monitoring
- GPU server for inference and possibly one for training (not hosted) - may
- [ ] move it to repository when agree

14 févr. 2023 Monthly meeting

Attendees: Alex, Bohdan, Stéphane

News of the month

- rsync (false?) problem on off1→ off2 → ovh3
  - use root to run the rsync (but still use off on ovh1 side)
  - launched a manual sync of images to ovh3 (will take time…)
  - removed old data (taking time, and maybe resources of mongodb)
- we have a date for first step in migration of off1/off2 servers : this friday
  2023-02 Upgrade for off1 and off2
- ZFS on boarding on friday 24th

Issue review/triage/points of concern

- Alex still have to understand IO issues on staging

Points to discuss

- will we move all the data from off2 to the OSM server ?
  - the influxdb docker can be shutdown (now metrics goes on another influxdb) and Alex backuped the data
  - we need more cleaning
  - how will data be copied ?
  - what will be wiped on off2 disks during the migration/upgrade?
  - at the minimum, we need to move MongoDB (server + database)
    - needs to be mongo 4.4.3
- new off1 architecture (see below)
  - use a general nginx proxy
  - try to have paths that are the same as docker ones (eventually change those)
- we met Bohdan who wants to contribute
  - docker / network / CICD

Migration discussion

TODO

- reconfiguration of ipmi ip addresses of off1 and off2
- migration to add disks to off1 and off2 and goes full proxmox + zfs
- documentation

Migration of server at free

- when installing hardware, the most important point is storage, as you can't go back easily as is
- 4 disk - let ZFS manage it totally
- 2 SSD - for the OS (mirror partitions) / as ZFS part or cache for ZFS disk
- we got a card fhor 4 SSD

Today:

- much CPU / IO wait eaten by rsync on images

Christian's Plan

- buy hardware:
  - 2 x 2To SSD
  - 1 SSD in 3DX Point dedicated to logs - 16Go
  - RAM 256G
- ZFS: training - remote and in presence
  - framadate on friday to choose a date
- ask one more IP to free
- 1 we move mongodb there for some days (we already share local addresses) making a new VM
  - either we do only one cluster
  - OSM France server can host mongodb for some days
  - looking at https://www.computel.fr/munin/openfoodfacts/off2.openfoodfacts/mongo_ops.html it's stable enough we can do it when we want
- [**At the datacenter**] upgrade off2
  - with disk 4 x 14T (1 redudancy) all in ZFS
  - a bit more RAM
  - install system + proxmox
  - reboot checks
- configure off2:
  - migrate photo:
    - by copying ovh3 snapshot (so that zfs sync will retry easily after by)
    - then rsync off1 -> off2 and sync off2 -> ovh3 using zfs
    - Note when migrating sto: we did a double sync and then replace with symbolic link until all was symbolic links
    - Here :
      - mount off2 zfs of photo as NFS off1

            - rsync off1 -> off2 and little by little replace folders by symlink (like we did for sto)
    - migrate sto:
        - it's instantaneous
        - see if we pull or push zfs sync
    - migrate containers making things more modularity
        - proxy nginx - ssl - with it's own IP (**first**)
        - opf
        - obf
        - opff
        - off public
        - off pro
        - shared folders are by NFS
        - ! ssl certificates renewals with wildcards + DNS / OVH (acme.sh)
    -
- **[At the datacenter]** off1

Benchmark NFS - iotop - nettop (xxxperf aussi)

10 janv. 2023 | Infrastructure monthly meeting

Participants:

- Alex
- Pierre
- Raphaël
- Rory
- Stéphane
- Vova

News of the month

- Stephane did some cleaning on off2 /srv2 because disk usage was really high
    - still high usage
- Raphaël setup Label studio annotation tool https://annotate.openfoodfacts.org/projects/
- We had a hard time this week-end and this week: server down. See Infrastructure availability - January 2023
- Vova has begun some work on monitoring !
    - two low level dashboard: monitoring prometheus - monitoring machines
    - refactor of configuration

Issue review/triage/points of concern

- **Nginx -** there are two distribution (one with battery included) it could be interesting to use it https://openresty.org/en/
- **Disk usage on off1 off2 is dangerously high !**
    - a lot of things (e.g. old images from producers) could be moved elsewhere

        - where?
    - more disks coming
- **Infrastructure availability issues -** [Infrastructure availability - January 2023]
    - Actions already taken
        - banned bots and very heavy users of the API
        - reduced number of workers process in Apache configuration: avoid running out of memory
        - reduced monitoring query rate (black box)?
            - and test a single product or a text / status page, instead of getting the front page
    - To do:
        - investigate search queries that make apache processes grow to 8Gb
            - `35080 off 20 0 9665136 8.812g 2880 R 63.9 14.1 0:59.24 /srv/off/cgi/se`
            - `35235 off 20 0 9625372 8.779g 1976 R 63.6 14.0 0:55.88 /srv/off/cgi/se`
            - `35005 off 20 0 9664880 8.816g 2876 R 63.0 14.1 1:00.45 /srv/off/cgi/se`
            - `35307 off 20 0 9637484 8.791g 2812 S 60.0 14.0 0:55.18 /srv/off/cgi/se`
                - do more logging
                - investigate hard limits in apache (e.g. 2 min cpu / 2 gb per query)
        - investigate removing swap completely
            - it's not a problem generally but it would make regaining control faster
            - at vova company they are generally off
            - we could try to test it on a server (hammering it, seeing if stop apache is faster)
        - repair ipmi/drac on off1 and off2
        - implement nginx rate limiting
            - analyze bot traffic before
                - how/where
            - check with historic logs who would have been blocked
            - https://nginx.org/en/docs/http/ngx_http_limit_req_module.html#limit_req_dry_run
        - more cache ?
        - use an extra nginx reverse proxy on another machine or container?
            - then we can turn off what we want, reroute it, show a nice error message
        - adding exporters to monitor
            - better use standard process for exporters
            - we can get binaries for most exporters

Points to discuss

- on availability
    - nginx is a good place to control service degradation
    - we can condition cache on existence cookie -> enable caching whole pages
- should we use cloud-init and templates on Proxmox for QEMU VMs ? (and for containers ?).
- when we deploy a new project, remember to open an issue for monitoring so we can add prometheus exporters, blackbox ping etc…

13 déc. 2022 | Infrastructure monthly meeting

Participants: Alex Garel Christian Quest Stéphane Gigandet rorypowis@gmail.com

News of the month

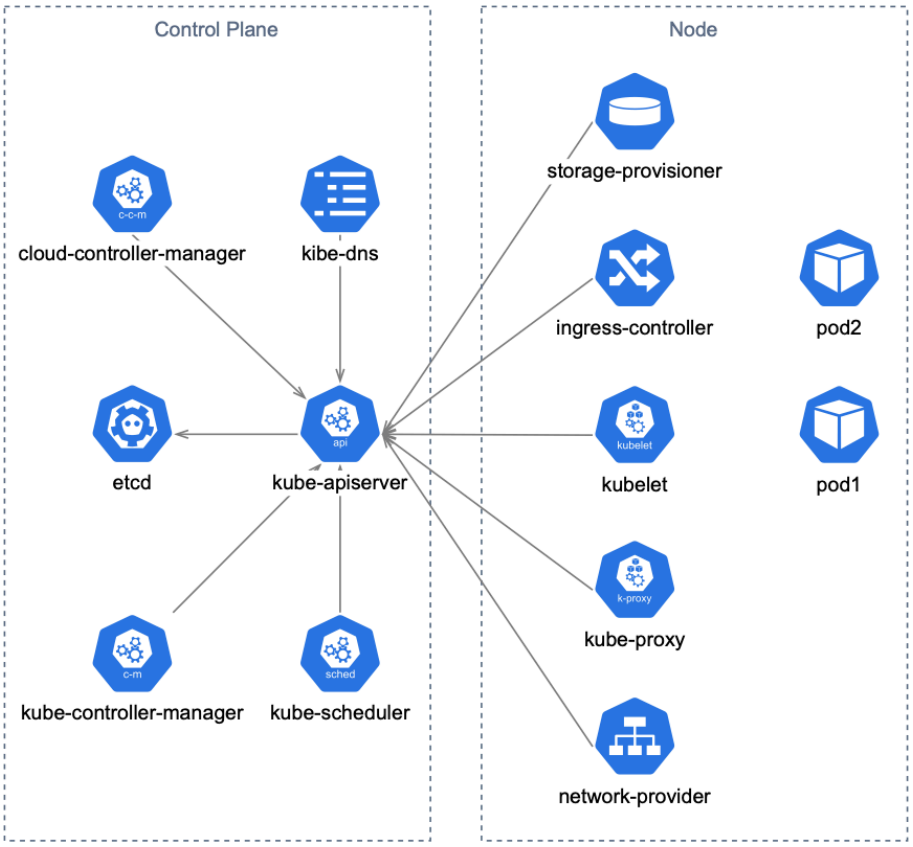
- moved monitoring to its own VM (see also report)
    - I used a direct NFS volume instead of a bind mount on NFS mount
        - Alex suspects it's far better than a bind mount which has edge cases with LXC, especially on a volume with loads of files, it takes 3 min to spawn a Docker container
- add more monitoring:
    - exporters where not deployed on every node (in our case, virtual machines containing containers)
    - still a lot remains to be done (missing exporters, etc.)
    - Vova has some examples for postgres exporter + dashboard (from perf table and monitoring stats)
- new deployments:
    - (staging) openfoodfacts-search (integration, was already deployed)
    - metrics (for business metrics) : https://github.com/openfoodfacts/openfoodfacts-metrics
    - taxonomy-editor (and a fix) https://ui.taxonomy.openfoodfacts.net
    - facets knowledge panels

Issue review/triage/points of concern

- Todo: add more monitoring (recent projects, more alerts)
    - dashboards
    - more alerts
    - alert-manager / blackbox exporter
- Try to use direct NFS mount on Open Food Facts in staging environment to see if it avoids very long startup times (to be tested)
- start implementing stunnel usage
- Plan a date for off1/2 servers disk migration? (Christian needed)
    - we still need to buy 2 disks (2 were already bought)

Points to discuss

- Urgency of clarifying deployment using best practices (prefix every volume and network with project and environment)
    - skip for now
- Should we install something like Portainer on each docker node ? (but limit it to visualization) - any other solution ?
    - UI for the running containers
    - Kubernetes Overview



- Should we deploy prometheus exporters outside current docker-compose deployments (services deployed in lxc's) ? Can we use docker for those ?
    - Vova propose to stick to prometheus
    - node exporter on every VM (note: there is a package in debian) + host
    - one prometheus instance should be enough
    - munin is already deployed
- Should we avoid port publishing even on VM ? (or is it exaggerated) - network can be a vector of attack for an intruder that passes first line (access to a VM with whatever user)
    - to connect docker networks between nodes
        - either we use stunnel manually
        - possible docker network overlays: flannel, cilium, https://www.nomadproject.io/, https://k3s.io/
- discussion on grafana for metrics or https://redash.io/
    - Vova suggest to stick to grafana (?) because it's a good open source tool
        - do not hesitate to use the query explorer

    - maybe https://www.timescale.com/ instead of Influxdb if we want to stick to SQL
- FOSDEM: we submitted a talk on CI/CD
    - Let's meet there 🙂

8 nov. 2022 | Infrastructure monthly meeting

Participants: Alex Garel Christian Quest Stéphane Gigandet Charles Népote contact@openfoodfacts.org rorypowis@gmail.com

News of the month

- off1 and off2 migrations:
    - Christian bought two 14TB disks and is testing them
- deployed ll (alias for ls -al) command on new machines post-install script

Issue review/triage/points of concern

- snapshot is locked problems (preventing VM backups): why does this happens ?

Points to discuss

- **Alex started an overview documentation**, can we discuss it ?
    - https://github.com/openfoodfacts/openfoodfacts-infrastructure/blob/docs-servers-overview/docs/overview.md
        - We have servers in two data centers => We have servers by two providers
    - contains info from https://wiki.openfoodfacts.org/Infrastructure ?
- **list/inventory of VM and servers**
    - We have two docs for current list of VM:
        - an excel: https://docs.google.com/spreadsheets/d/19RePmE_-1V_He73fpMJYukEMjPWNmav1610yjpGKfD0/edit#gid=0
        - a generated table on the README tracking tickets
    - Can we keep only one ?
        - the generated one have the advantage that we can track each service with a ticket
        - third option:
            - harvest data from machines - maybe using proxmox API
        - keep the idea of creating one ticket per service
            - we keep one open ticket for each machine / vm / container as long as it runs, to log all important changes
- doing stunnel like reverse proxy
    - on 101 VM
    - check that 101 VM is capable to use SIMD instructions for AES
    - one instance for all OVH side, using proxy1 ip address

Tasks

- [ ] start auto inventory project (Rory / Alex)
- [ ] continue with overview documentation
- [ ] move on monitoring soon to happen !
- [ ] Standup integration in #infrastructure

18 oct. 2022 | Grafana deployment

Participants: Alex Garel Raphaël Bournhonesque Stéphane Gigandet

Notes

- currently two grafana:
  - one non accessible on off2 (in a docker) with an influxdb
  - one for monitoring, which is right now on staging but should be moved with monitoring https://grafana.openfoodfacts.org/login and https://monitoring.openfoodfacts.org/login
- we have an ELK stack on monitoring
- proposition: have two different grafana:
  - one for openfoodfacts metrics → metrics.openfoodfacts.org
  - one for monitoring
- important metrics
  - analytics from robotoff
  - aggregation results (api on facets) sent by product opener in a script (robotoff metrics.py)
    - https://be.openfoodfacts.org/ingredients?stats=1&json=1
- what we want to send from free (off1 / off2) to “metrics” (and backward)
  - nginx logs analysis
    - we don’t want to use the grafana log analyzer directly
    - just analyze logs and send data to influxdb on a regular basis
  - would be interesting for off1 to be able to access influxdb

Tasks
https://github.com/openfoodfacts/openfoodfacts-infrastructure/issues/154

- [ ] create a docker-compose project with influxdb and grafana (for metrics) in openfoodfacts-infra [raphael]
- [ ] deployed on docker prod server 201 [raphael/alex]
  - [ ] ci/cd scripts
  - [ ] config of proxy1 (lxc 101 on ovh1)
  - [ ] dns config
- [ ] migrate influxdb data from off2 and monitoring [raphael]
- [ ] have a stunnel to talk to influx from free servers [alex]

11 oct. 2022 | Infrastructure monthly meeting

Participants:

- Alex
- Charles
- Christian
- Dariusz
- Pierre
- Rory
- Vova

Notes

- Alex: listed some priorities according to me
  https://github.com/openfoodfacts/openfoodfacts-infrastructure/issues?q=is%3Aissue+is%3Aopen+label%3AP1
  https://github.com/openfoodfacts/openfoodfacts-infrastructure/issues/88
  (deployments)
- Alex: some PRs on documentation
  - https://github.com/openfoodfacts/openfoodfacts-infrastructure/blob/develop/docs/reports/2022-07-11-infra-workshop.md
  - https://github.com/openfoodfacts/openfoodfacts-infrastructure/pulls
- Christian: project to add 14To x 2 to both off1 and off2
  https://github.com/openfoodfacts/openfoodfacts-infrastructure/issues/150
  - bought 2x14To disk - needs 2 more
  - off1 and off2 lacks disk space for images
  - switch 4+14TB mirrored disks on off1 and off2 to 4x14TB with RaidZ1 (3x14TB available = 42TB), that's 42-18 = 24TB new storage space.
    last year of images is about 2T
  - store pictures on zfs to avoid current rsyncs (as we do now for products and users)
  - at same rate should give us a couple of years before filling disks
  - side goals:
    - have proxmox on off1 / off2
    - have more focus VMs (proxmox CT aka LXC)
  - technically
    - we will use OSM France server to host mongodb while reconfiguring off2
- Pierre proposal on meeting structure:
  - bug triage
  - free expression on last month tasks
  - eventual deep dive
- How to best contribute

    - github issues / slack ?
- Instead of a spreadsheet for services
    - catalog of services
    - or in metrics (telegraf on every machines)
    - or through proxmox / docker apis
- Proxy documentation
- Mail gateway - https://github.com/openfoodfacts/openfoodfacts-infrastructure/blob/develop/docs/mail.md
- docker / proxmox discussion
    - shan't we use only one ? Advantage: simplify the stack
        - make VM images for proxmox instead of docker ?
            - not that easy because docker-compose manage more than one machine, also networks and volumes
    - Charles: debian / bare metal more easy
    - Christian: also proxmox offers the clustering of the bare metal and in a very easy way
    - Alex: docker easier for Continuous Deployment (because of idem potency)
    - proxmox and docker are complementary
- Alex is happy to document everything needed, so just ask :-)

Tasks

- ☑ ~~Decide whether this document should be public or not ?~~
    - ☐ → **NO**
- ☑ ~~create an issue for off1 off2 migration~~
- ☑ ~~give vova access to infrastructure and monitoring projects + assign issue on moving https://github.com/openfoodfacts/openfoodfacts-infrastructure/blob/develop/docs/reports/2022-07-kibana-down-es-circuit-breaking-exception.md~~