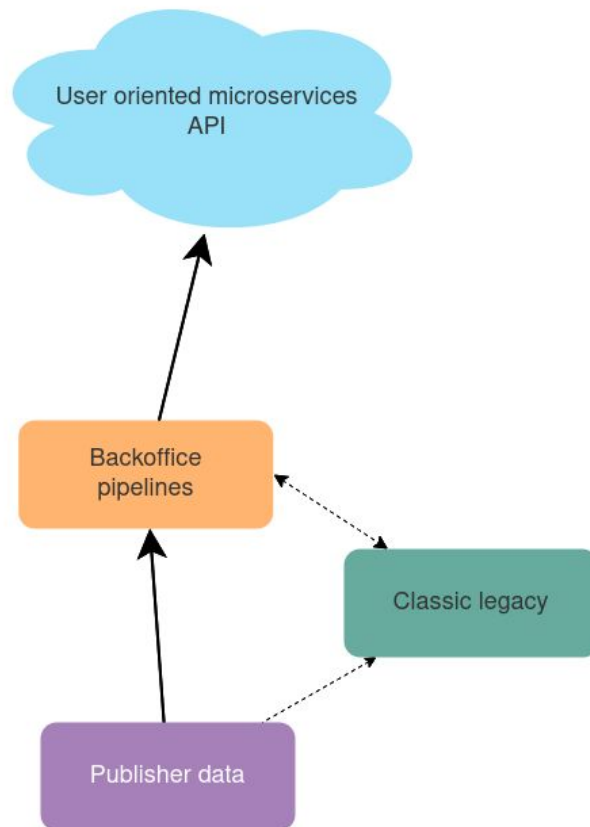# Infrastructure update

Nemanja Martinovic and the ADS Team

ADS Users Group Meeting, 15-16 Nov. 2021

Infrastructure overview

- **Backoffice**
  - Ingest and processing of data
    - Upgraded computing resources

- **Classic**
  - Legacy processes
    - Reducing dependency
    - Pruning away

- **Cloud**
  - API and user oriented microservices
    - Upgrades, new services

Backoffice - new computing resources

- **adsnest – new production environment**
  - 2x Intel(R) Xeon(R) Platinum 8268 CPU @ 2.90GHz
  - 96 threads total
  - 500 GB of RAM
  - RAID6 - 16x SSD
  - Backoffice transition to single docker-compose environment
    - Removed previously used VMs

- **adsnlp – home of machine learning projects**
  - 2x Intel(R) Xeon(R) Gold 6246 CPU @ 3.30GHz
  - 48 threads
  - 768 GB of RAM
  - 2x Tesla V100S PCIe 32GB

- **adsnull – home of backoffice dev server**
  - Full development environment
    - Made possible with the new computing resources

```
CONTAINER ID: NAMES                                                    STATUS
17f60b6066ab: backoffice_prod_data_pipeline_1                          Up 4 days
98a1e521b4b9: backoffice_prod_article_of_the_day_aod_1                 Up 5 days
7458d08c8633: backoffice_prod_classic_cronjobs_cronjobs_1             Up 5 days
d910ac46b5b1: backoffice_prod_montysolr_engine_1                      Up 5 days
ab3e0d2c000e: backoffice_system_monitoring_hub_grafana_1              Up 5 days
551db8cea023: backoffice_system_monitoring_hub_prometheus_1           Up 5 days
159f265ad318: backoffice_system_monitoring_hub_pushgateway_1          Up 5 days
467b36730c70: backoffice_system_monitoring_hub_alertmanager_1         Up 5 days
a107ae51d1ee: backoffice_prod_rasterizer_pipeline_1                   Up 5 days
9fcbbe273142: backoffice_prod_supervisor_pipeline_1                   Up 5 days
9c8497d48d0a: backoffice_prod_import_pipeline_1                       Up 5 days
77d8c7204ea1: backoffice_prod_fulltext_pipeline_1                     Up 5 days
57a054d08a93: backoffice_prod_ingest_report_svc_1                     Up 5 days
394d9f88520e: backoffice_prod_master_pipeline_1                       Up 5 days
1acb7965c5f1: backoffice_prod_rabbitmq_broker_1                       Up 5 days
415578a266a7: backoffice_prod_postgresql_db_1                         Up 5 days
b31582f9ce8f: backoffice_prod_stats_collector_svc_1                   Up 2 days
14c6cd708e79: backoffice_prod_orcid_pipeline_1                        Up 5 days
a5023102f834: backoffice_prod_myads_pipeline_1                        Up 5 days
d04baed8e613: backoffice_prod_doc_matching_pipeline_1                 Up 5 days
730df20fb56e: backoffice_prod_citation_capture_pipeline_1            Up 5 days
32a2cb8fda95: backoffice_prod_augment_pipeline_1                      Up 5 days
584baa736541: backoffice_system_logging_hub_graylog_1                Up 5 days (healthy)
be9d1f59f365: backoffice_system_logging_fluent-bit_1                 Up 5 days
abfa00074518: backoffice_prod_ads_articles_pipeline_1                Up 5 days
bb319a04ca9f: backoffice_system_logging_hub_elasticsearch_1          Up 5 days
0b0f616a7aae: backoffice_system_monitoring_node-exporter_1           Up 5 days
0363ee44a659: backoffice_system_monitoring_cadvisor_1                Up 5 days (healthy)
7aa1a8eaf992: backoffice_system_logging_hub_mongo_1                  Up 5 days
```

Cloud

- **New services:**
  - Tailor – continuous deployment platform (now in the cloud/availability)
  - Wordpress sites – migration from backoffice
    - Increased security
  - Nextcloud – own cloud solution / backup
  - BitWarden – password sharing platform

  - All of these – open licenses, free for the project

- **Cost saving policies**
  - Reserving instances

    Overview below is from March (when the purchase was made), until 10th of November (few days ago)

🔒 RI Utilization

| $26,754.35	$15,631.19	$11,123.15	$11,128.42
On-Demand Cost Equivalent ❶	Effective Reservation Cost ❶	Total Net Savings ❶	Total Potential Savings ❶

4

Service migrations

- Anonymous ftp → sftp;
- Final result:
  - Shutting down of anonymous ftp
  - Moving the system to the cloud
  - Creating:
    - S3 bucket backups
      - (s3 - simple storage service, efficient, cheap and secure way of storing objects)
    - user accounts for publishers
      - separate user structure
    - Manuals, user support
    - Pulling content, security
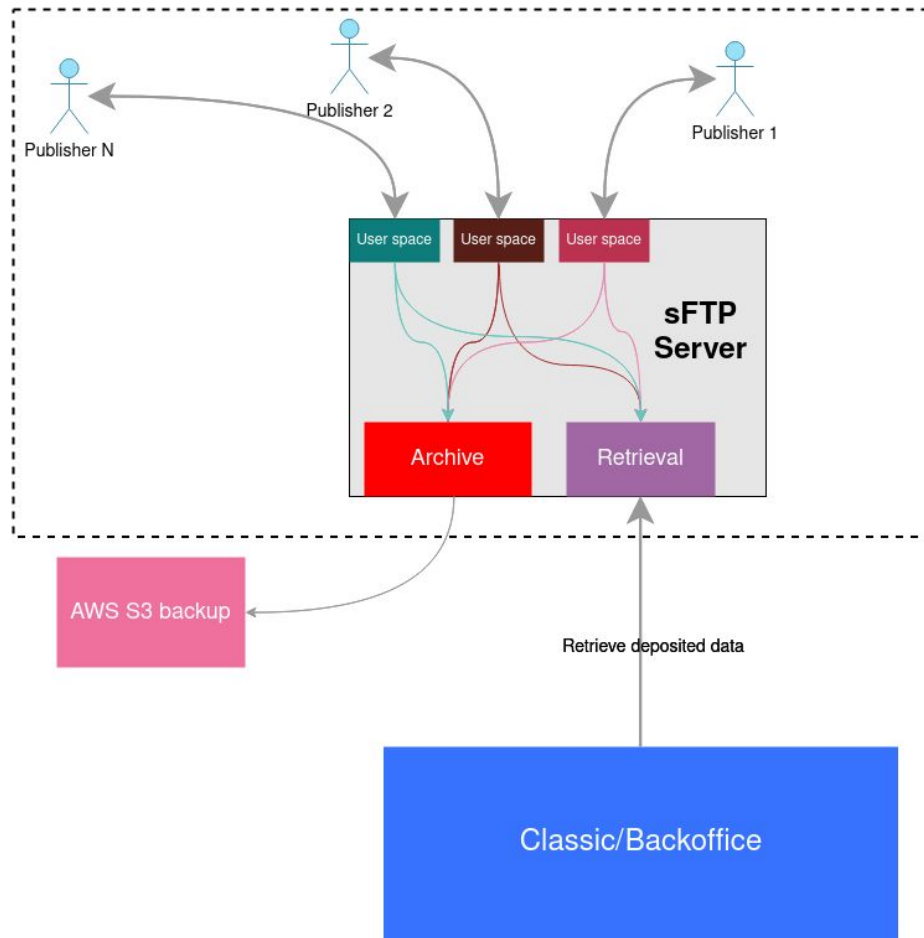
Service migrations - part 2

- **Moving from adsfe → adsdmz;**
  - Part of closing off external access to the backoffice and reducing dependency to the classic infrastructure
  - Working with CF and SysHelp people;
  - Optimization of serving content;
  - Switching to the https is in progress;

- **domain migrations -> to Route53**
  - All ADS domain names are centralized under AWS DNS service

- **Migration to the GitHub Actions**
  - CI environment, move from Travis
  - Cost savings (part of getting GitHub for free)
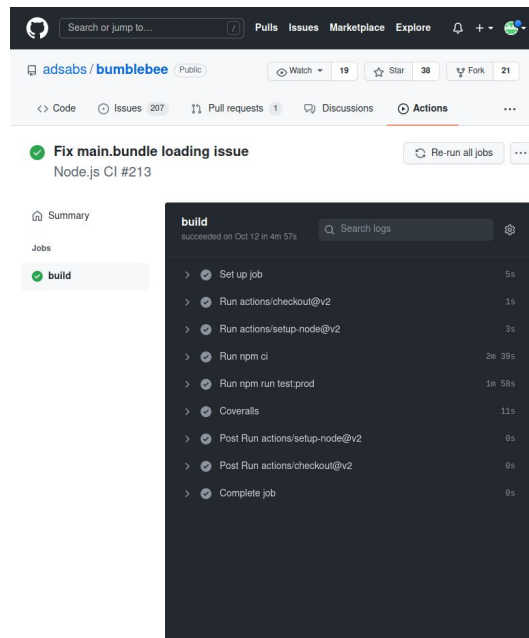
Overview

- Few hundred deployments, continuous development, high availability of the ADS (we had 2 power outages during last year)
- Cost savings:
  - reservation of instances
  - new open source services
  - utilization of available resources to the maximum (GitHub Actions, Route53)
  - Slack is now free!
- Evolution of existing services:
  - sFTP
  - PDF serving
  - New frontend servers for classic stuff

- New hardware, optimizations

- Plans for the future