

2020-01-07 Statusmeeting

- [Update on NAS latest tests and developments](#)
- [Status of the production sites](#)
- [Next meetings](#)
- [Any other business?](#)

Agenda for the joint NetarchiveSuite tele-conference 2020-01-07, 13:00-14:00.

Participants

- BNF: Clara, Sara, Géraldine
- ONB: Andreas
- KB/DK - Copenhagen: Tue, Stephen, Anders, Kristian
- KB/DK - Aarhus: Colin, Sabine
- BNE: -
- KB/Sweden: Par, Thomas, Peter

Join from PC, Mac, Linux, iOS or Android:

<https://kjdk.zoom.us/j/104443571>

Or an H.323/SIP room system:

H.323: 109.105.112.236
Meeting ID: 104 443 571

SIP: 104443571@109.105.112.236

Or Skype for Business (Lync):

<https://kjdk.zoom.us/skype/104443571>

Or Telephone:

Denmark: +45 89 88 37 88 or +45 32 71 31 57
United Kingdom: +44 203 051 2874 or +44 203 481 5237 or +44 203 966 3809 or +44 131 460 1196
Finland: +358 9 4245 1488 or +358 3 4109 2129
Sweden: +46 850 539 728 or +46 8 4468 2488
Norway: +47 7349 4877 or +47 2396 0588
US: +1 669 900 6833 or +1 646 558 8656
Meeting ID: 104 443 571

International numbers available: <https://zoom.us/u/acRu0MV3xJ>

You can join a meeting by using apps from a pc, a tablet or a smartphone, but you can also use the browser based version (it works with newer versions of Chrome or Firefox)

Please click OK if you see the system dialog.

Launching...

If nothing prompts from browser, [download & run Zoom](#).

If you cannot download or run the application, [click here](#) to help your browser.

Update on NAS latest tests and developments

Status of the production sites

Netarkivet

Our biggest issue during the last weeks of 2019 was a harvest problem between Christmas and New Year, that is to say, there were no uploads (neither from the last broad crawl for 2019 nor from the selective crawls) to the Copenhagen bit archive from 24 December 2019 to 1 January 2020. 453 jobs failed with upload errors. About 2400 files are missing on the server, but the system did not "tell" us, that there was no server space left. The minSpaceleft setup implemented before Christmas apparently did not work.

We will work on uploading the missing 2400 files when the broad crawl is finished. We started this broad crawl on 8 November 2019.

An electrical cut in Copenhagen last night probably has produced some loss, as the broad crawl was not finished.

News from our test environment: the newest snapshot NAS 5.7 IIPC Heritrix bundle is able to run on NAS 5.5 without problems.

BnF

A happy New Year and best wishes to all for 2020 from the BnF web archiving team!

Our broad crawl, which started on October 12th, finished on December 23rd. It represents 2.2 billion URLs and 118,17 TB of compressed data. Despite technical problems related to our infrastructure (25% of the jobs were killed by their HarvestController because Heritrix needed too much time to initialise), it took less time than last year (11 weeks in 2019). Its size exceeds our initial budget of 110 TB due to an average weight per URL which is higher than our estimates (from 55 421 bytes to 57 044 bytes). We'll analyse the reports to understand this increase : it probably comes from an evolution of the websites.

We welcome a newcomer in our team : Alexandre Faye. He will be in charge of cooperation with researchers and international cooperation.

ONB

BNE

KB-Sweden

Next meetings

- February 4, 2020
- March 3, 2020
- April 7, 2020
- May 5, 2020
- June 9, 2020
- July 7, 2020
- September 8, 2020
- October 6, 2020
- November 3, 2020
- December 8, 2020
- January 5, 2021

Any other business?

€€€€€€€€