

Automation tools in AV workflows

Building digital pipelines using standard IT tools.

Jonáš Svatoš Národní filmový archiv, CZ No Time To Wait! #3 London, UK. 25.-26.10.2018



Various steps when working with AV content

- Ingest
- Verification
- Storage
- Transcoding
- Packaging
- Delivery
- ..



Workflow management?

- DAMS
- Digipres systems like Archivematica, Preservica, Adlib
- Manual flow

Národní filmový archiv

Scripts are good, but..

- Individual scripts do not serve a generic purpose
- Complex workflows must chain many steps together
- Logging anyone?
- Auditing the logs anyone?



Is this deterministic?

(venv) [digilab@dll-render-1 ~]\$ cd /mnt/vault/INGEST/UPP-TRANSFER/ZBOROV/ (venv) [digilab@dl1-render-1 ZBOROV]\$ ll total 204140708 -rwxr--r-. 1 1028 users 963133204 22. říj 16.09 Zborov 20181022 original scan.wav -rw-r--r-. 1 1024 users 68012932294 16. říj 18.29 ZBOROV ALL REELS GRADED WIP 181015 1080 24p LT.mov -rw-r--r-. 1 1024 users 68969349901 17. říj 20.52 ZBOROV ALL REELS GRADED WIP 181016 1080 24p LT.mov -rw-r--r-. 1 1024 users 69167624548 23. říj 13.08 ZBOROV bez orezu 1080 24p 422LT MONOfinal 181023.mov -rwxr--r-. 1 1028 users 961585926 22. říj 16.10 Zborov restoration 20181022 bez sumu na zacatku.wav -rw-rw-r--. 1 1024 users 965447656 16. říj 10.12 Zborov restoration WIP 20181011.wav (venv) [digilab@dl1-render-1 ZBOROV]\$ bagit.py --md5 --sha1 . 2018-10-20 06:34:25,971 - INFO - Creating bag for directory /mnt/vault/INGEST/UPP-TRANSFER/ZBOROV 2018-10-20 06:34:25,975 - INFO - Creating data directory 2018-10-20 06:34:25.977 - INFO - Moving Zborov 20181022 original scan.way to /mnt/vault/INGEST/UPP-TRANSFER/ZBOROV/tmpnm4iw228/Zborov 20181022 original scan. wav 2018-10-20 06:34:25.977 - INFO - Moving ZBOROV ALL REELS GRADED WIP 181015 1080 24p LT.mov to /mnt/vault/INGEST/UPP-TRANSFER/ZBOROV/tmpnm4iw228/ZBOROV ALL RE ELS GRADED WIP 181015 1080 24p LT.mov 2018-10-20 06:34:25,978 - INFO - Moving ZBOROV ALL REELS GRADED WIP 181016 1080 24p LT.mov to /mnt/vault/INGEST/UPP-TRANSFER/ZBOROV/tmpnm4jw228/ZBOROV ALL RE ELS GRADED WIP 181016 1080 24p LT.mov 2018-10-20 06:34:25,980 - INFO - Moving ZBOROV bez orezu 1080 24p 422LT MONOfinal 181023.mov to /mnt/vault/INGEST/UPP-TRANSFER/ZBOROV/tmpnm4jw228/ZBOROV bez orezu 1080 24p 422LT MONOfinal 181023.mov 2018-10-20 06:34:25,981 - INFO - Moving Zborov restoration 20181022 bez sumu na zacatku.wav to /mnt/vault/INGEST/UPP-TRANSFER/ZBOROV/tmpnm4jw228/Zborov resto ration 20181022 bez sumu na zacatku.wav 2018-10-20 06:34:25,982 - INFO - Moving Zborov restoration WIP 20181011.wav to /mnt/vault/INGEST/UPP-TRANSFER/ZBOROV/tmpnm4jw228/Zborov restoration WIP 20181 011.wav 2018-10-20 06:34:25,983 - INFO - Moving /mnt/vault/INGEST/UPP-TRANSFER/ZBOROV/tmpnm4jw228 to data 2018-10-20 06:34:25,984 - INFO - Using 1 processes to generate manifests: md5, shal 2018-10-20 06:34:25,985 - INFO - Generating manifest lines for file data/ZBOROV ALL REELS GRADED WIP 181015 1080 24p LT.mov 2018-10-20 06:38:25,313 - INFO - Generating manifest lines for file data/ZBOROV ALL REELS GRADED WIP 181016 1080 24p LT.mov 2018-10-20 06:42:28,468 - INFO - Generating manifest lines for file data/ZBOROV bez orezu 1080 24p 422LT MONOfinal 181023.mov 2018-10-20 06:46:31,476 - INFO - Generating manifest lines for file data/Zborov 20181022 original scan.wav 2018-10-20 06:46:34.917 - INFO - Generating manifest lines for file data/Zborov restoration 20181022 bez sumu na zacatku.way 2018-10-20 06:46:38,264 - INFO - Generating manifest lines for file data/Zborov restoration WIP 20181011.wav 2018-10-20 06:46:41,848 - INFO - Creating bagit.txt 2018-10-20 06:46:41,896 - INFO - Creating bag-info.txt 2018-10-20 06:46:41,955 - INFO - Creating /mnt/vault/INGEST/UPP-TRANSFER/ZBOROV/tagmanifest-md5.txt 2018-10-20 06:46:42,013 - INFO - Creating /mnt/vault/INGEST/UPP-TRANSFER/ZBOROV/tagmanifest-sha1.txt (venv) [digilab@dl1-render-1 ZBOROV]\$ echo \$? (venv) [digilab@dl1-render-1 ZBOROV]\$ ls bag-info.txt bagit.txt data manifest-md5.txt manifest-sha1.txt tagmanifest-md5.txt tagmanifest-sha1.txt

Automation to the rescue!



- Define inputs
- Implement steps by reusing existing tools and scripts
- Set validations and conditions for eventual branching
- Assign notifications
- Read the notifications :-)
- Audit the logs (GUI!)

There are generic tools available



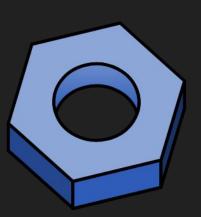
- Jenkins
- Ansible
- Rundeck
- Buildbot

••











Jenkins

- Build automation server
- Web GUI, written in Java
- As it is generic, we can use it for whatever purpose
- Multi-tenant
- Plugin ecosystem

Jenkins jobs overview

🧕 AAA [Jenkins] 🛛 🗙 🕂							×
← → C ^u ⁽¹⁾ ⁽²⁾ ⁽²⁾ ⁽²⁾ ⁽²⁾ ⁽²⁾ ⁽²⁾						Work 🔒 ••• ∞ ☆ ∞ 🛓	S ≫ Ξ
没 Jenkins					4 🔍 search	Jonáš Svatoš	log out
Jenkins >						ENABLE AUT	O REFRESH
쯜 New Item						Zadd	description
🌯 People	AAA	INGEST	TAPE-SETS TOOLS ZALOHY	ZZ-TAPE ZZZ-OLD +			
Build History	s	w	Name ↓	Last Success	Last Failure	Last Duration	
🚳 Edit View		*	CHECK-DCP	7 days 3 hr - <u>#19</u>	16 days - <u>#14</u>	21 sec	\bigotimes
🔍 Project Relationship		*	EMAIL_LINKS	1 hr 59 min - <u>#3884</u>	N/A	6 ms	
E Check File Fingerprint	0	63	MAKE-BAGIT	4 min 58 sec - <u>#12</u>	11 min - <u>#10</u>	56 sec	\bigotimes
🐡 Manage Jenkins	•	*	MAKE-BLURAY	21 hr - <u>#429</u>	6 days 21 hr - <u>#420</u>	16 min	
🍇 My Views	ā	-	MAKE-DCP	1 day 5 hr - <u>#191</u>	5 days 0 hr - #184	32 min	2
JQS Monitoring	-		MAKE-DCP-SCREENER			1 hr 27 min	
Credentials				1 day 0 hr - <u>#177</u>	1 day 2 hr - <u>#176</u>		\mathbf{D}
hew View	•	6	MAKE-DVD	2 days 3 hr - <u>#836</u>	2 days 3 hr - <u>#835</u>	2 min 5 sec	\mathbf{D}
Build Queue (1)		*	MAKE-EXCERPTS	53 min - <u>#547</u>	N/A	29 min	\bigotimes
MAKE-CHECKSUM		*	MAKE-HD-MASTER	11 days - <u>#42</u>	1 mo 10 days - <u>#36</u>	4 hr 57 min	\bigotimes
	•	*	MAKE-HD-PREVIEW	2 days 4 hr - <u>#14</u>	18 days - <u>#8</u>	7 min 5 sec	\bigotimes
Build Executor Status master + 12 computers (1 of 10 executors)	0	*	MAKE-LOSSLESS	11 days - <u>#6</u>	N/A	1 day 2 hr	Ø
	0	*	MAKE-SCREENER	2 hr 2 min - <u>#4932</u>	22 hr - <u>#4911</u>	3 min 6 sec	ø
	•	*	MAKE-STILLS	2 hr 29 min - <u>#635</u>	1 day 23 hr - <u>#616</u>	16 sec	Ø
	0	*	PRINT-STITEK	2 mo 29 days - <u>#158</u>	3 mo 8 days - <u>#153</u>	2,7 sec	ø
	Icon: <u>S</u> M	<u>I</u> L			Legend 🔝 RSS for all	RSS for failures 🔊 RSS for just late	st builds

Page generated: Oct 17, 2018 2:47:48 PM CEST REST API Jenkins ver. 2.121.3 SCM Sync status : 🔆 Last operation @ Tue May 08 22:59:52 CEST 2018

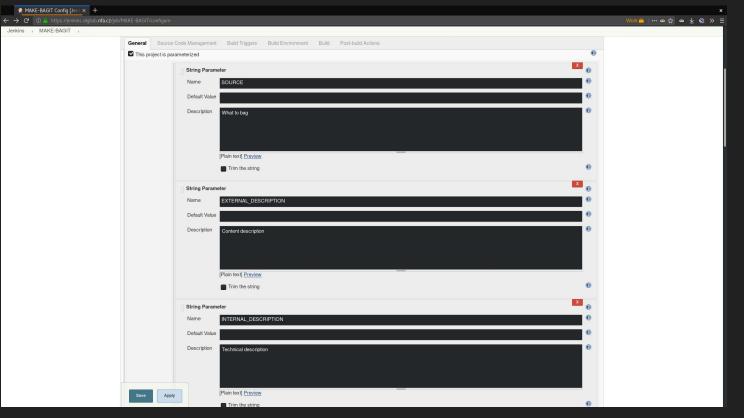


Jenkins job parameters

co <u>↓</u> š Svatoš	S ≫ ≡
š Svatoš	log out
	— I
	_

Národní filmový archiv

Jenkins job parameters config



Nárďdní

filmov \archi\

Jekins build log

MAKE-BAGIT #12 Consol × Parameters [Jenkins] × +					×
← → C ① ⓐ https://jenkins.digilab.nfa.cz/view/AAA/job/MAKE-BAGIT/12/console	Work 🧰	··· • 🖒 🕁	∞⊻	S >	» ≡
Jenkins > AAA > MAKE-BAGIT > #12					
2018-10-17 14:43:03,203 - INFO - Generating manifest these for file data/tengths.xts 2018-10-17 14:43:03,203 - INFO - Generating manifest lines for file data/DCP/SFC 24fps 6ch reel0.wav					
2010-10-17 14:43:03,512 - INFO - Generating manifest lines for file data/DCF/SFC_241p5_ocf_retto.wav 2018-10-17 14:43:03,512 - INFO - Generating manifest lines for file data/DCF/SFC_241p5 och retto.wav					
2010-10-17 14-05-05-2012 - INFO - Generating manifest lines for file data/DCF/SPC 24fpS_6ch reello.wav					
2018-10-17 14:43:06,605 - INFO - Generating manifest lines for file data/DCF/SPC 2445 sch restor.way					
2018-10-17 14:43:09,587 - INFO - Generating manifest lines for file data/DCP/SFC 24fps 6ch reel2.wav					
2018-10-17 14:43:12,000 - INFO - Generating manifest lines for file data/DCP/SPC 24156 Sch reel3.vav					
2018-10-17 14:43:14,007 - INFO - Generating manifest lines for file data/DCP/SPC 24fps 6ch reel4.wav					
2018-10-17 14:43:16,285 - INFO - Generating manifest lines for file data/DCP/SPC 24fps 6ch reel5.wav					
2018-10-17 14:43:19,017 - INFO - Generating manifest lines for file data/DCP/SPC 24fps 6ch reel6.wav					
2018-10-17 14:43:20,653 - INFO - Generating manifest lines for file data/DCP/SPC 24fps 6ch reel7.wav					
2018-10-17 14:43:22,925 - INFO - Generating manifest lines for file data/DCP/SPC 24fps 6ch reel8.wav					
2018-10-17 14:43:25,920 - INFO - Generating manifest lines for file data/DCP/SPC 24fps 6ch reel9.wav					
2018-10-17 14:43:29,251 - INFO - Generating manifest lines for file data/Decklick/SPC decklick reel1.wav					
2018-10-17 14:43:30,779 - INFO - Generating manifest lines for file data/Decklick/SPC_decklick_reel2.wav					
2018-10-17 14:43:33,525 - INFO - Generating manifest lines for file data/Decklick/SPC_decklick_reel3.wav					
2018-10-17 14:43:35,148 - INFO - Generating manifest lines for file data/Decklick/SPC_decklick_reel4.wav					
2018-10-17 14:43:36,898 - INFO - Generating manifest lines for file data/Decklick/SPC_decklick_reel5.wav					
2018-10-17 14:43:37,855 - INFO - Generating manifest lines for file data/Restored/SPC_full_restored_0424					
2018-10-17 14:43:37,922 - INFO - Generating manifest lines for file data/Restored/SPC_full_restored_0424	_reel1.w	av			
2018-10-17 14:43:38,118 - INFO - Generating manifest lines for file data/Restored/SPC_full_restored_0424					
2018-10-17 14:43:38,428 - INFO - Generating manifest lines for file data/Restored/SPC_full_restored_0424					
2018-10-17 14:43:38,918 - INFO - Generating manifest lines for file data/Restored/SPC_full_restored_0424					
2018-10-17 14:43:39,314 - INFO - Generating manifest lines for file data/Restored/SPC_full_restored_0424					
2018-10-17 14:43:39,883 - INFO - Generating manifest lines for file data/Restored/SPC_full_restored_0424					
2018-10-17 14:43:40,257 - INFO - Generating manifest lines for file data/Restored/SPC_full_restored_0424					
2018-10-17 14:43:40,746 - INFO - Generating manifest lines for file data/Restored/SPC_full_restored_0424					
2018-10-17 14:43:41,033 - INFO - Generating manifest lines for file data/Restored/SPC_full_restored_0424					
2018-10-17 14:43:41,366 - INFO - Generating manifest lines for file data/Restored/SPC_full_restored_0424					
2018-10-17 14:43:41,859 - INFO - Generating manifest lines for file data/Restored/SPC_full_restored_0424		av			
2018-10-17 14:43:42,506 - INFO - Generating manifest lines for file data/Scanned_source/SPC_poz_reell.wa					
2018-10-17 14:43:43,795 - INFO - Generating manifest lines for file data/Scanned_source/SPC_poz_reel2.wa					
2018-10-17 14:43:44,517 - INFO - Generating manifest lines for file data/Scanned_source/SPC_poz_reel3.wa					
2018-10-17 14:43:45,342 - INFO - Generating manifest lines for file data/Scanned_source/SPC_poz_reel.va 2018-10-17 14:43:46,197 - INFO - Generating manifest lines for file data/Scanned source/SPC poz_reelS.va					
	1				
2018-10-17 14:43:46,696 - INFO - Creating bagit.txt 2018-10-17 14:43:46,699 - INFO - Creating bagit.fotxt					
2018-10-17 14:43:40,097 - INFO - Creating Dag-InTO.txt 2018-10-17 14:43:46,700 - INFO - Creating /mnt/central/TEMP JONAS/SPC zvuk/tagmanifest-sha256.txt					
2018-10-17 14:43:46,719 - INFO - Creating /mmt/central/term_JUNAS/sFC_zuVuk/tagmanifest-sha2so.txt 2018-10-17 14:43:46,719 - INFO - Creating /mmt/central/termP_JONAS/sFC_zuVuk/tagmanifest-sha512.txt					
2010-10-17 14-43-40,739 - HWO - Cleating /mml/clean(ta/Tenr_JOHA/SPC_ZVH/Lagmanifest-Sha312.ttt [WS-CLEANUP] Deleting project workspace[WS-CLEANUP] done					
Finished: SUCCESS					
FIRING, JUCES					

Národní filmový archiv

Jenkins job tasks



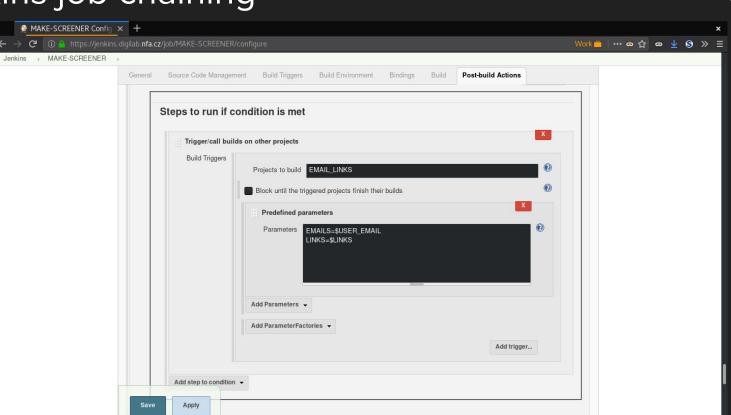
🤵 MAKE-BAGIT Config [Jen] 🗙	+	×
$\leftarrow ightarrow {f C}$ () ${f O}$ https://jenkins.dig	jilab nfa.cz /job/MAKE-BAGIT/configure	Work 💼 🚥 👁 🏠 👁 👱 🥱 » 🚍
Jenkins MAKE-BAGIT		
	General Source Code Management Build Triggers Build Environment Build Post-build Actions	
	Execute shell X Command #!/bin/bash set +x virtualenv-3 venv source venv/bin/activate pip installquiet bagit export PYTHONUNBUFFERED=1 See the list of available environment variables	
	Advanced Execute shell	0
	Command #!/bin/bash sot +x Source venv/bin/activate ORGANIZATION="Národní filmový archiv" ORGANIZATION ADDRESS="Malešická 12, Praha 3, 130 00, Czech Republic" CONTACT_EMAIL="digitallaboratory" CONTACT_EMAIL="digitalboratory" bagit.pysource-organization "\$ORGANIZATION"organization-address "\$ORGANIZATION_ADDRESS"	
	See the list of available environment variables Advanced Save the list of available environment variables	



Jenkins past build detail

🧶 Parameters [Jenkins] 🛛 🗙 🕂									×					
🕂 $ ightarrow$ C 🗍 🔒 https://jenkins.digilab.nfa.cz/job/MA	KE-BAGIT/16/parameters/		Work 🖻	••• @	☆	∞ ±	9	»	≡					
👰 Jenkins		4 Q search		2	Joná	iš Svatos	ś	log or	ut					
Jenkins → MAKE-BAGIT → #16 → Parameters						ENABLE A	UTO R	EFRES	đ					
🛧 Back to Project	Build #16													
Q Status	Balla #10													
📄 Changes	Parameters													
Console Output	SOURCE	/mnt/vault/V_RESENI/MASTER/FRANTA_TYMAL												
Edit Build Information		What to bag												
🚫 Delete Build	EXTERNAL_DESCRIPTION	Vyber krizeneckeho pro Frantiska Tymala												
Parameters		Content description						-						
Environment Variables	INTERNAL_DESCRIPTION	Format QuickTime, ProRes 422HQ, Full HD 24fps, LPCM 24bit 48kHz						Í.						
Rebuild		Technical description												
🜲 Previous Build	INTERNAL_IDENTIFIER	Trello #2914												
		AIS number						_						
	NODE	dl1-render-1												

Jenkins job chaining



Nárợdní filmový archiv

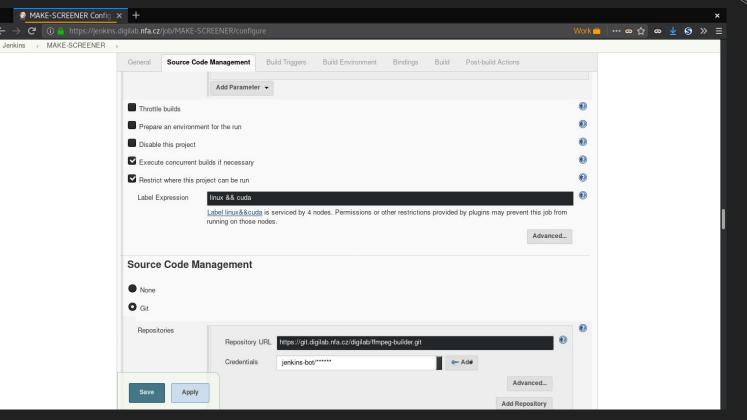
Jenkins slaves

6 Idle

									Nodes [Jenkins] ×
🔒 🚥 ෙ රු 💿 👱 😏	Work 🖴 🕴 🚥						mputer/	lab. nfa.cz /co	C 🛈 🤷 https://jenkins.digil
⑦ Jonáš Svatoš ∣ Io	0	🔍 search	4 🤇						Jenkins
ENABLE AUTO REP			· · · · · · · · · · · · · · · · · · ·						is ⊩ Nodes ⊩
Space Response Time	Free Temp Space	Free Swap Space	Free Disk Space	Clock Difference	Architecture	Name ↓	S		lack to Dashboard
N/A N/A	N/A	N/A	N/A	N/A		brazil			lanage Jenkins
N/A N/A	N/A	N/A	N/A	N/A		<u>dl1-pi-1</u>			lew Node
0,58 GB 33ms	850,58 GB	7,76 GB	945,11 GB	In sync	Linux (amd64)	dl1-projekce-1			Configure
3,80 GB 34ms	3,80 GB	860,30 MB	458,67 GB	In sync	Linux (amd64)	dl1-render-1		-	d Queue
N/A N/A	N/A	N/A	N/A	N/A		dl1-skener-1			ilds in the queue.
1,36 GB 32ms	11,36 GB	886,00 MB	11,36 GB	In sync	Linux (amd64)	dl1-storage-1-read		-	d Executor Status
9,16 GB 5ms 🚽	9,16 GB	23,70 GB	689,83 GB	In sync	Linux (amd64)	dl1-render-2			aster
N/A N/A	N/A	N/A	N/A	N/A		dl1-render-3			dle dle
N/A N/A	N/A	N/A	N/A	N/A		dl1-render-4			azil (offline)
6,24 GB 4ms 🚽	26,24 GB	7,25 GB	331,10 GB	2 sec behind	Windows 7 (amd64)	dl1-render-5			1-pi-1 (offline)
N/A N/A	N/A	N/A	N/A	N/A		dl1-render-6			1-projekce-1
N/A N/A	N/A	N/A	N/A	N/A		dl1-render-7			
6,20 GB 0ms 🚽	6,20 GB	393,74 MB	6,20 GB	In sync	Linux (amd64)	master			dle
31 min 31 min	31 min	n 31 min	n 31 min	a 31 min	31 min	Data obtained			dle
N/ 6,20 G	N/ 6,20 G	N/A 393,74 MB	N/A 6,20 GB	N/A In sync		dl1-render-7 master			dle 1 -render-1 dle

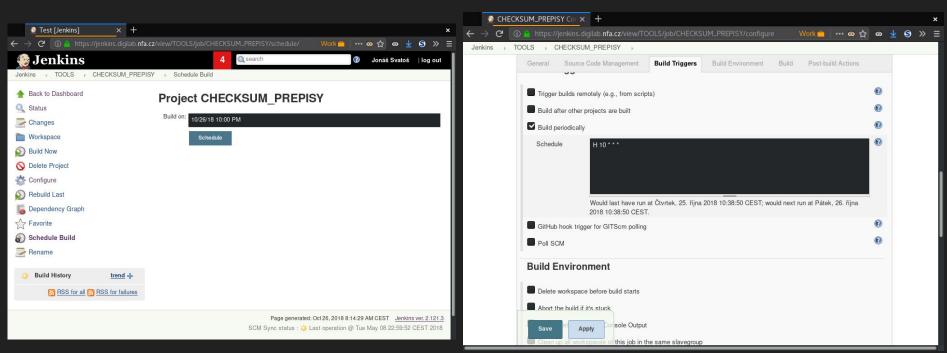
Nárợdní filmový archiv

Jenkins slave tag matching



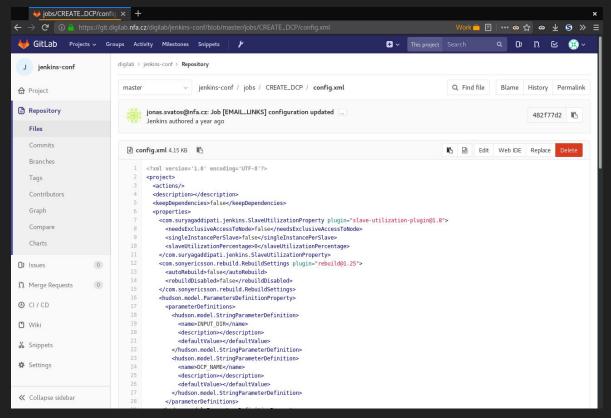
Jenkins as a cron







Jenkins job config in Git



Conclusions



- Standardized workflow needs deterministic steps
- As little human intervention as possible
- There are generic automation servers available
- Log retention!
- Jenkins is great



Thanks! Questions?

jonas.svatos@nfa.cz

github.com/NFAcz