



Italian National Agency for New Technologies,
Energy and Sustainable Economic Development

Presentation of the CRESCO annual report: Research activities and results 2020 Accounting 2021 (reduced version)

Online, 07/04/2022

Davide De Chiara e Simone Giusepponi: TERIN-ICT-HPC



Contents

The 2020 Report

Accounting 2021

2020: some users and research groups

CRESCO Report 2020: statistics

CRESCO usage: trends in recent years

Cresco report 2020 overview



High Performance Computing on CRESCO infrastructure: research activities and results 2020



December 2021

Number of papers: 52

Total number of authors: 207

Number of pages: 284

Document format: PDF

PDF size (print version): 14.2 MB

PDF size (web version): 14,2 MB

Contributions provided by a selection of users of the CRESCO infrastructure

Scientific Editor: Davide De Chiara e Simone Giusepponi
ENEA, TERIN-ICT-HPC

Cover: ENEA, TERIN-ICT-HPC

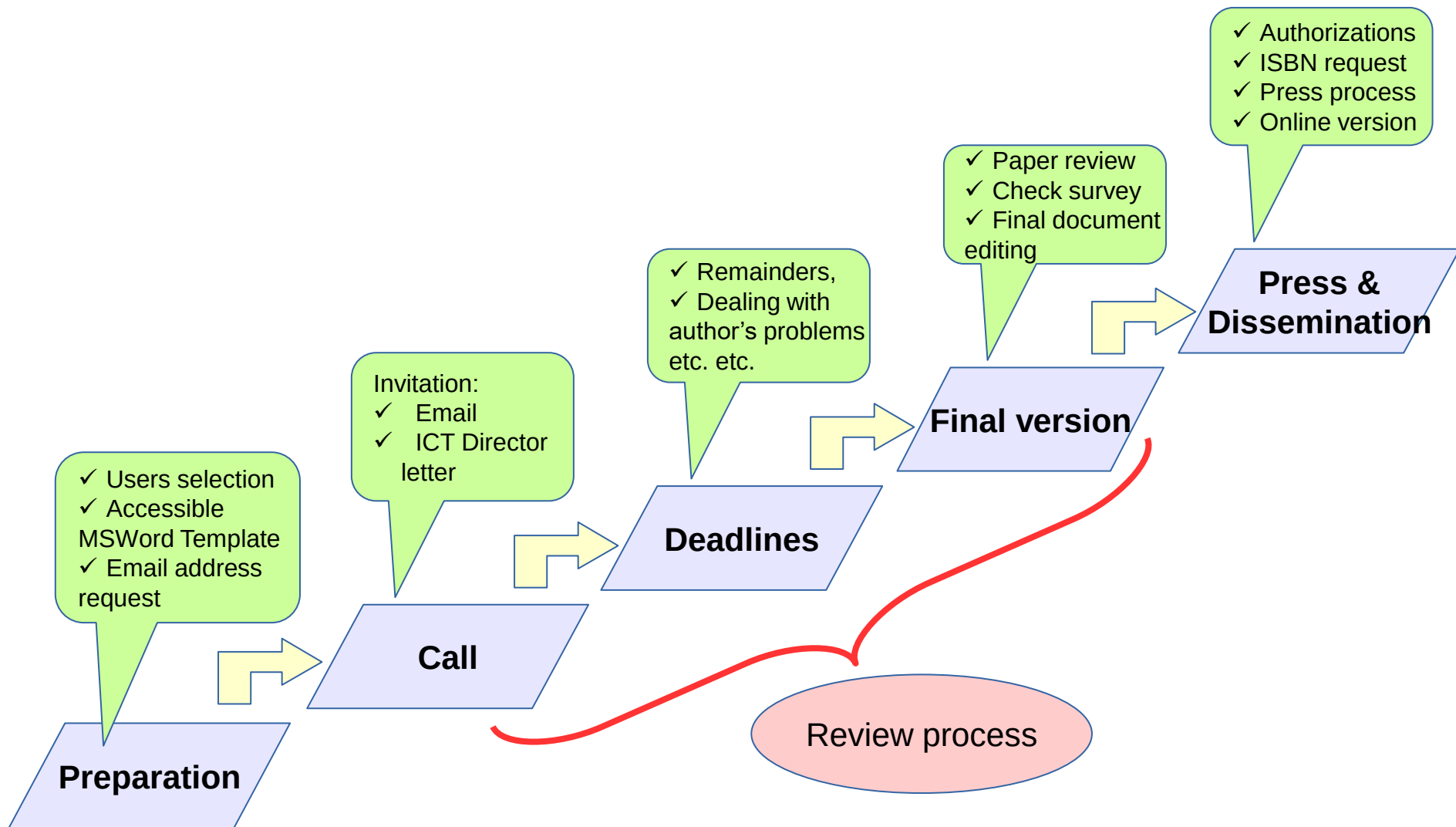
ISBN per edizione digitale: 978-88-8286-429-3

ISBN per edizione stampa: 978-88-8286-430-9

See also: [/afs/enea.it/project/eneagrid/html/SystemAdministration/CRESCO_Reports](https://afs.enea.it/project/eneagrid/html/SystemAdministration/CRESCO_Reports)
https://ict.enea.it/wp-content/uploads/2020/05/report_cresco_2020-ISBN.pdf



Report Preparation



Report preparation history

- 2021-06:
 - user selection from: Isf accounting (info from kerberos lists and mailing lists);
 - survey tool kit (limesurvey: <http://sondaggi.enea.it/index.php/>): user list, mails;
- 2021-07-30: invitation mail (~100 invited users);
- 2021-09-24: first deadline (~25 contributions only, despite the many reminders);
- 2021-10-16: deadline extensions, reminders, solicitations, mails, review;
- 2022-02: final check, author index, samples for the cover;
- 2022-03: authorization process for the publication by REL-PROM;
- 2022-03: press & dissemination;
- 2022-04: presentation prepared (this presentation).

◆ **Compressed PDF version online:**

https://ict.enea.it/wp-content/uploads/2020/05/report_cresco_2020-ISBN.pdf

Invited Users

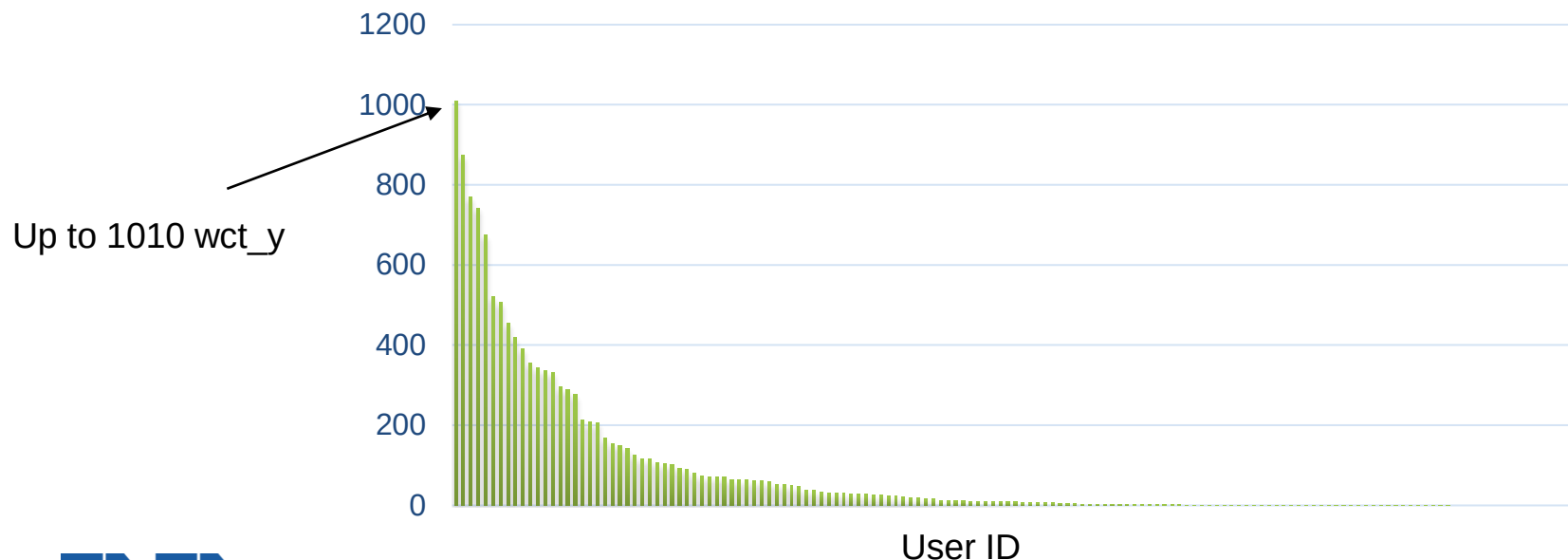
100/152 users invited

Selection criterium is:

$wct_y > 1.0 \parallel (av_totcore * > 5000)$

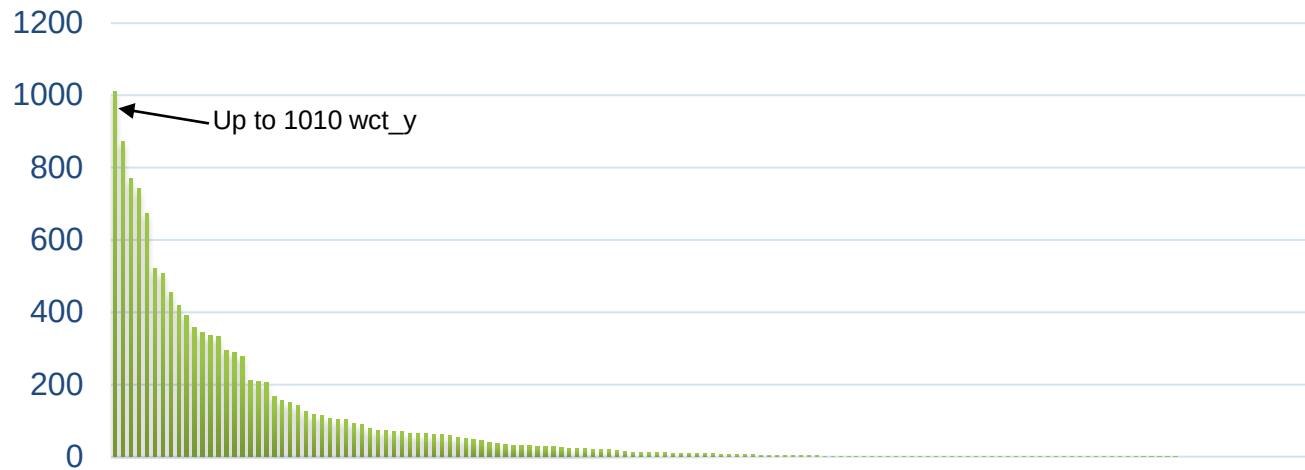
wct_y: total wall clock time (years)
njobs: number of submitted jobs
av_nc: average number of used cores
av_totcore: njobs*av_nc

Total run time by user 2020 wct_y

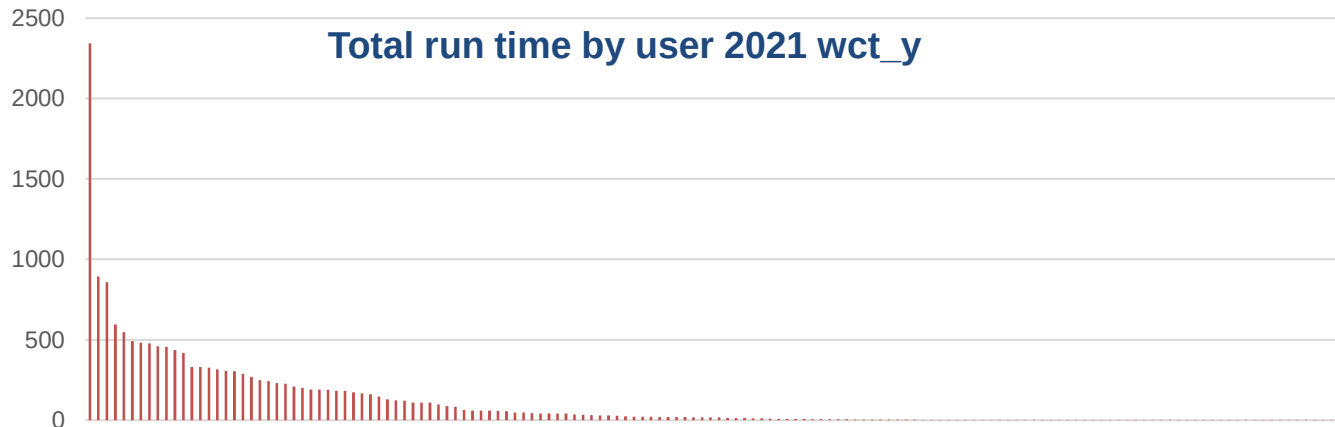


Wall clock time by user 2020-2021

Total run time by user 2020 wct_y



Total run time by user 2021 wct_y



In 2020

**~10 users are:
52%
of wct_y**

**~20 users are
~75%
of wct_y**

**Statistics
lead by
few power
users**

This has been reported
since the 2013
presentation
by Funel

Accounting

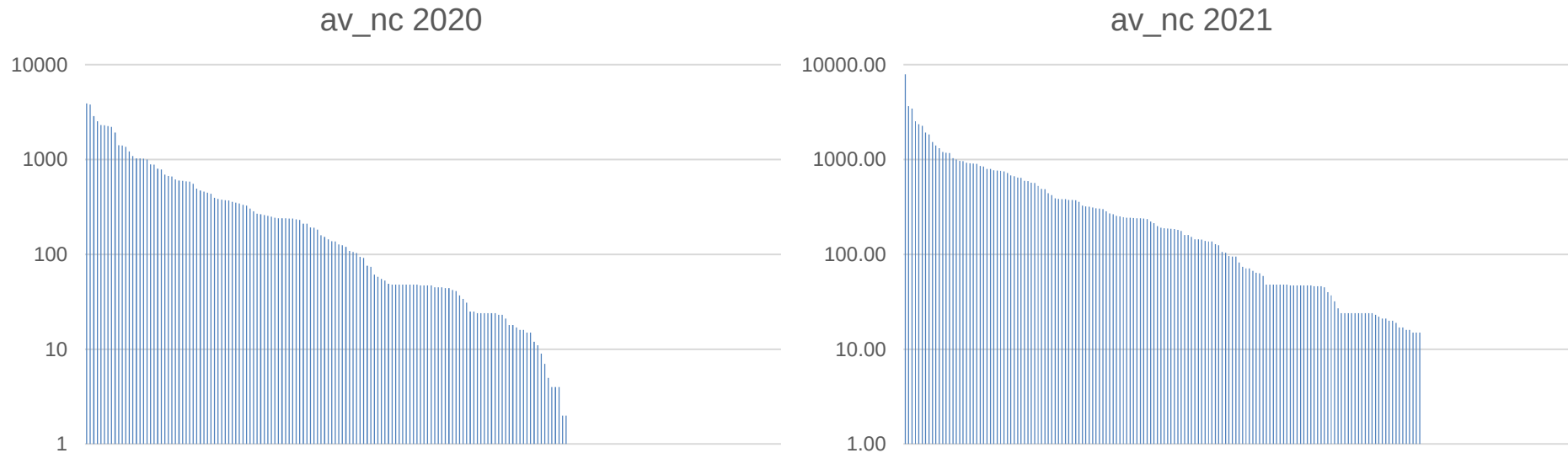
Accounting details 2020 - 2021

Cresco in 2020

Year	jobs	wct_y
2017	571258	3895
2018	1321623	4853
2019	1675438	9992
2020	1637704	12277
2021	1026449	16079

In 2022 something like 2021?

Average number of cores by user 2020-2021



78 (2020), 96(2021) users use more than 100 cores

Several (~20) users with more than 1000 cores in 2020-2021

Average number of cores is 933 (2020) and 969 (2021)

CRESCO6 has ~20000 cores.

Were there users who used more than 10000 cores in 2021?

No! There is a limit on LSF: up to 6144 cores (unless system queue for admin).

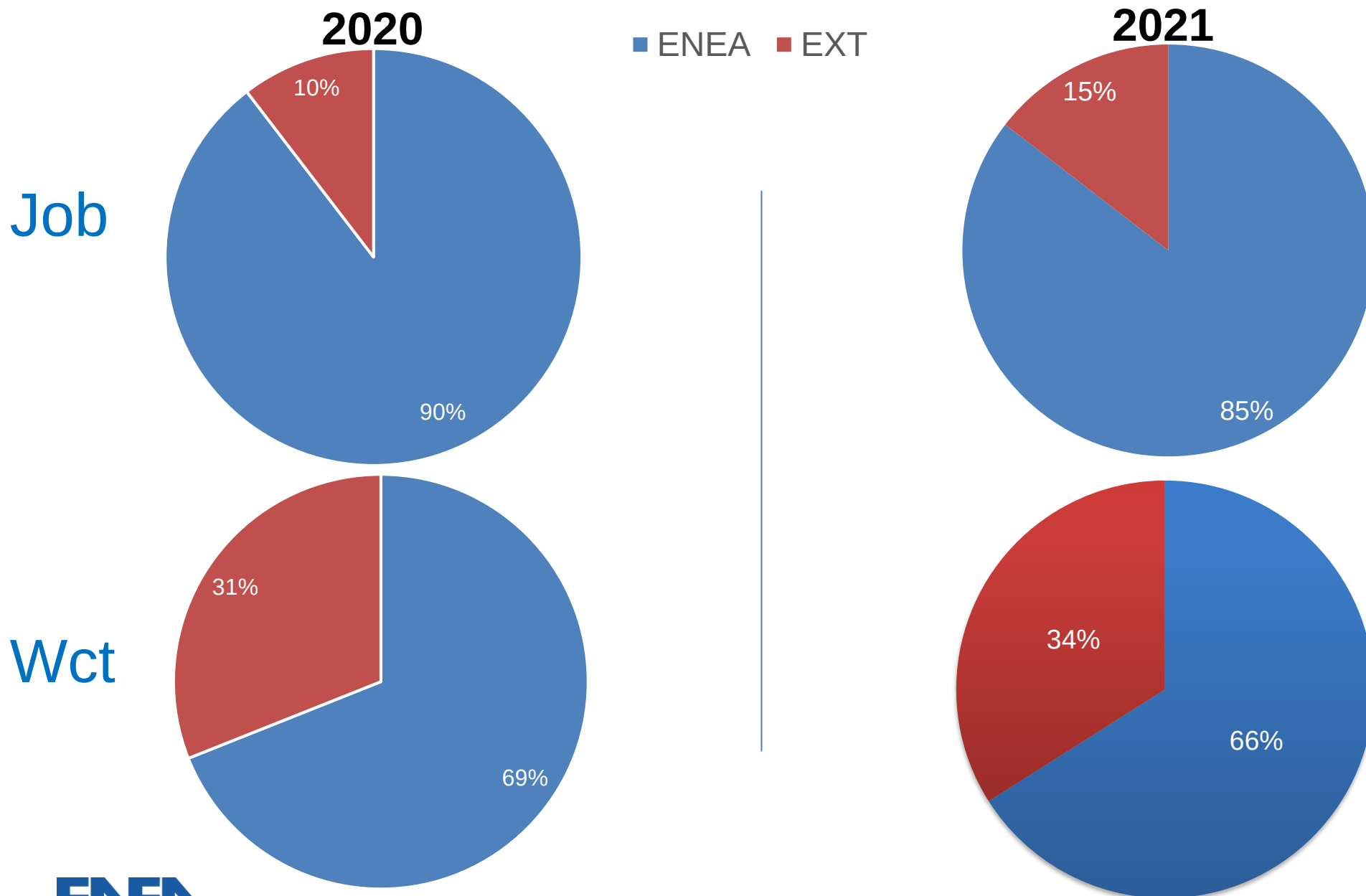
The 2020 CRESCO Users

Users 152

ENEA	108
TERIN	44
FSN	39
SSPT	21
other	4

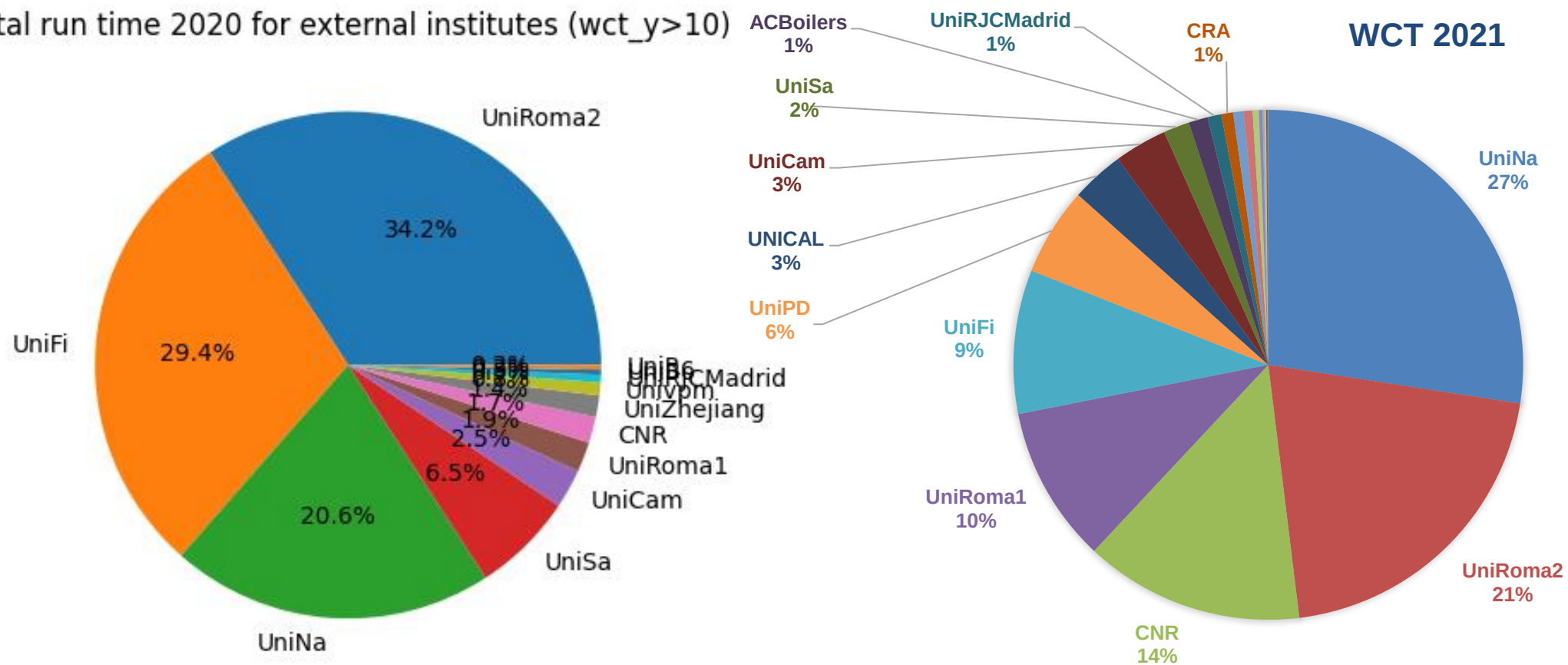
External	44
UniNa	9
UniRoma1	5
UniRoma2	4
UniFi	3
CNR	1
UniCal	1
UniSa	2
UniCam	2
UniRc	1
Univpm	1
UNIRCJMadrid	2
UniZhejiang	1
AVIO	1
AriaNet	2
Himet	1
Do	1
Simularia	2
ACBoilers	2
F4E	1
CRA	1
UniBo	1

ENEA and External usage 2020-2021



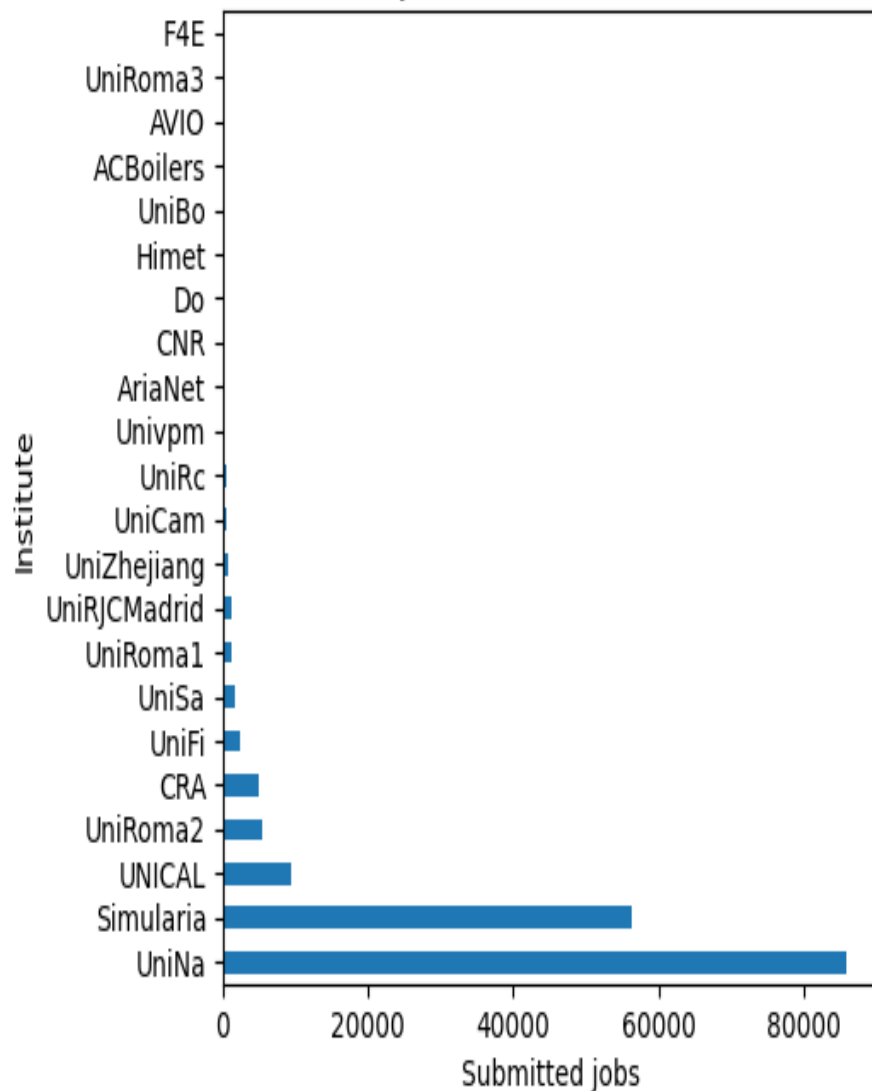
Usage by external institute – wct_y 2020-2021

Total run time 2020 for external institutes (wct_y>10)

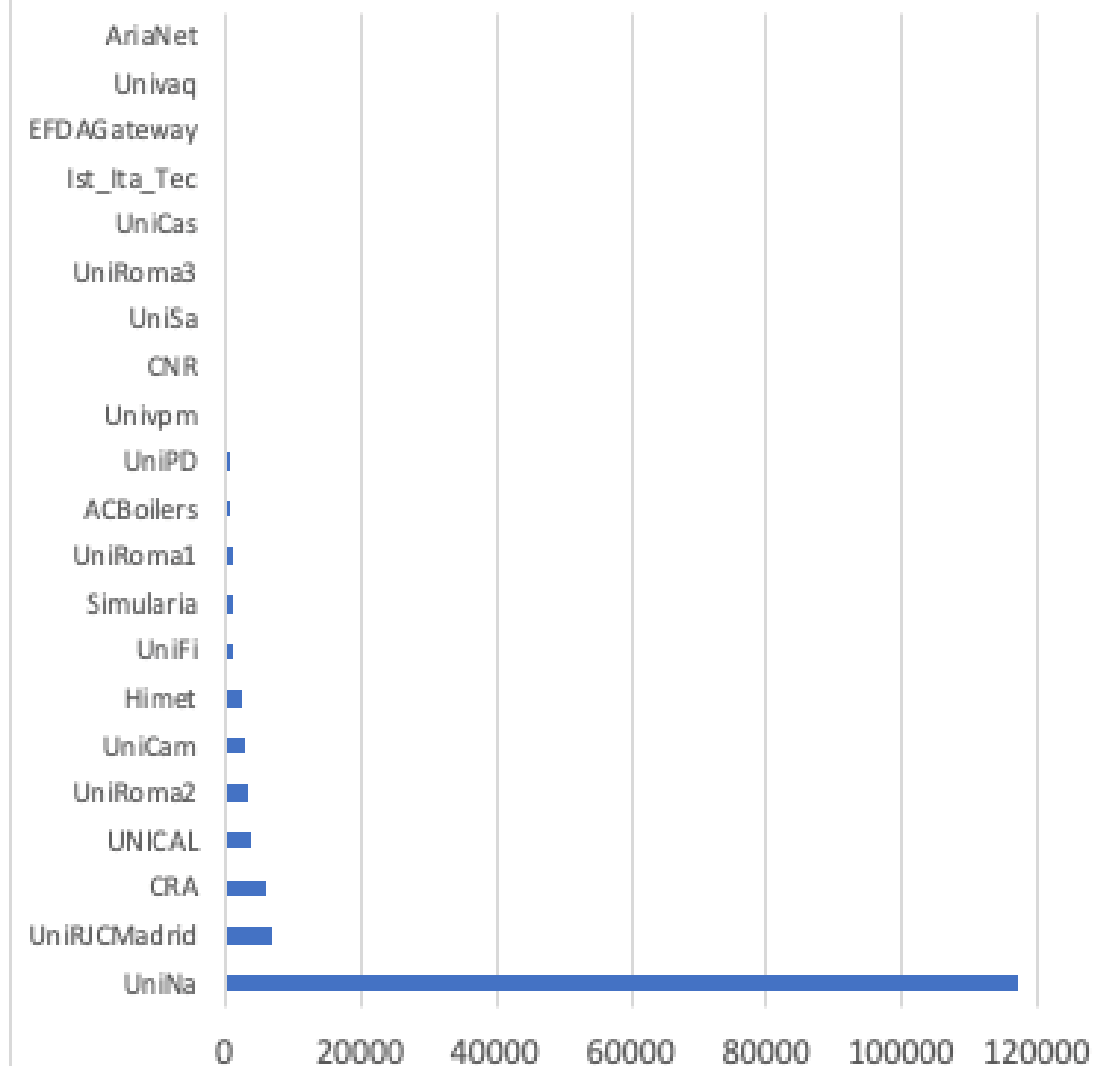


Usage by external institute – njobs 2020-2021

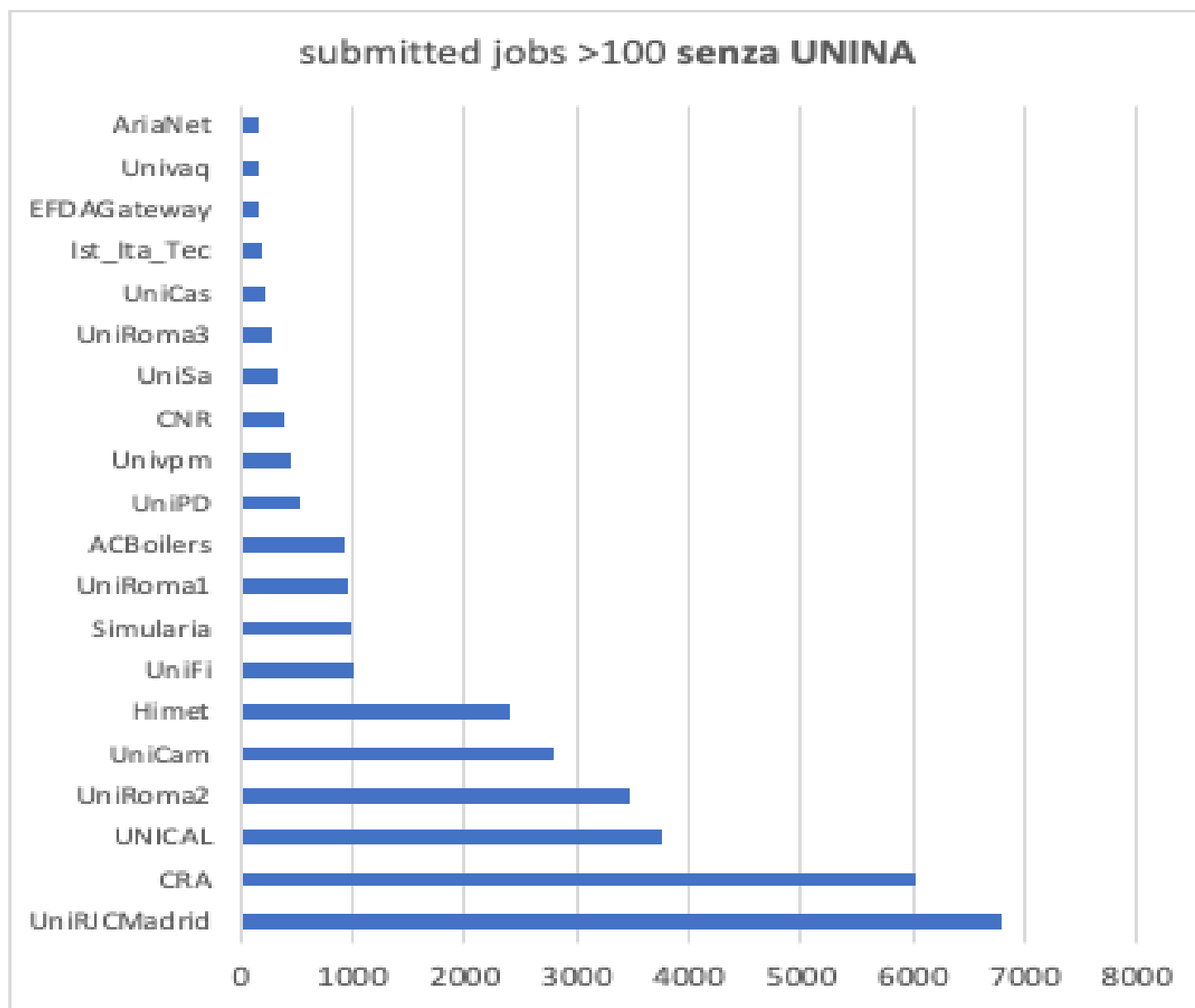
Submitted jobs 2020 for external institutes



2021 submitted jobs totali >100



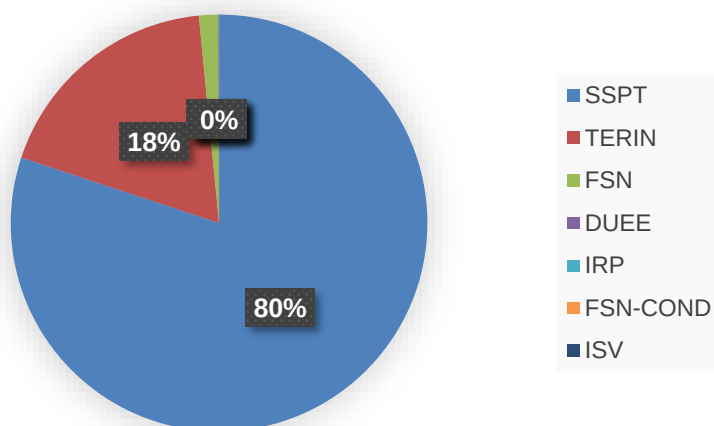
Usage by external institute – njobs 2021 senza UNINA



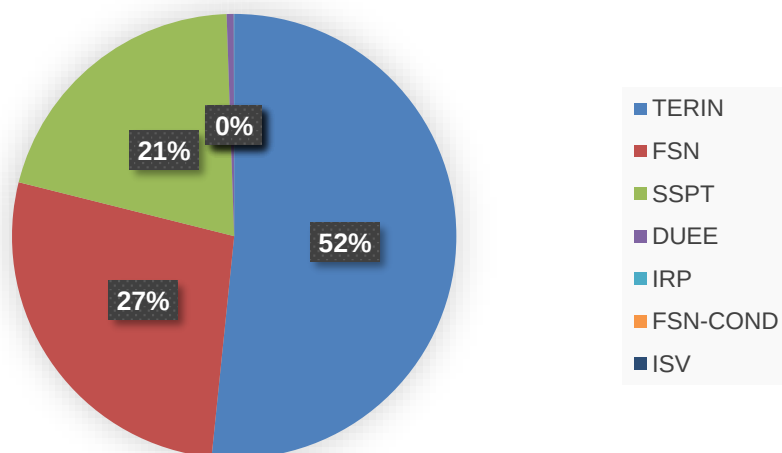
Usage by ENEA department 2020-2021

2020

njobs for ENEA departments

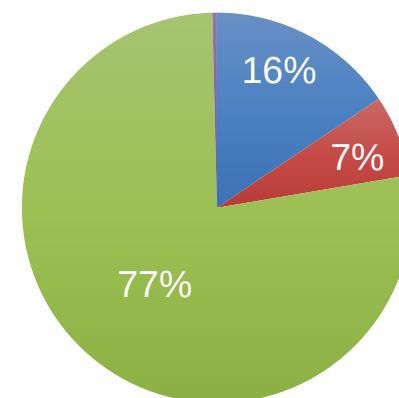


wct_y for ENEA departments



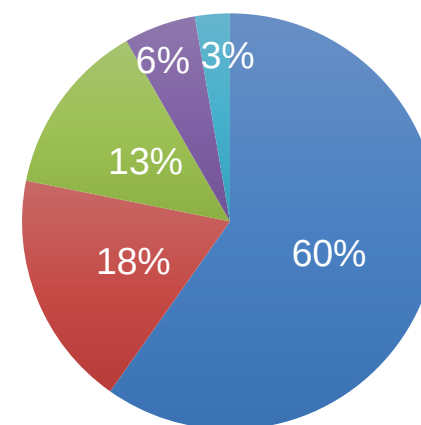
2021

njobs for ENEA departments 2021



■ TERIN ■ FSN ■ SSPT ■ DUEE ■ IRP ■ UTA ■ ISV

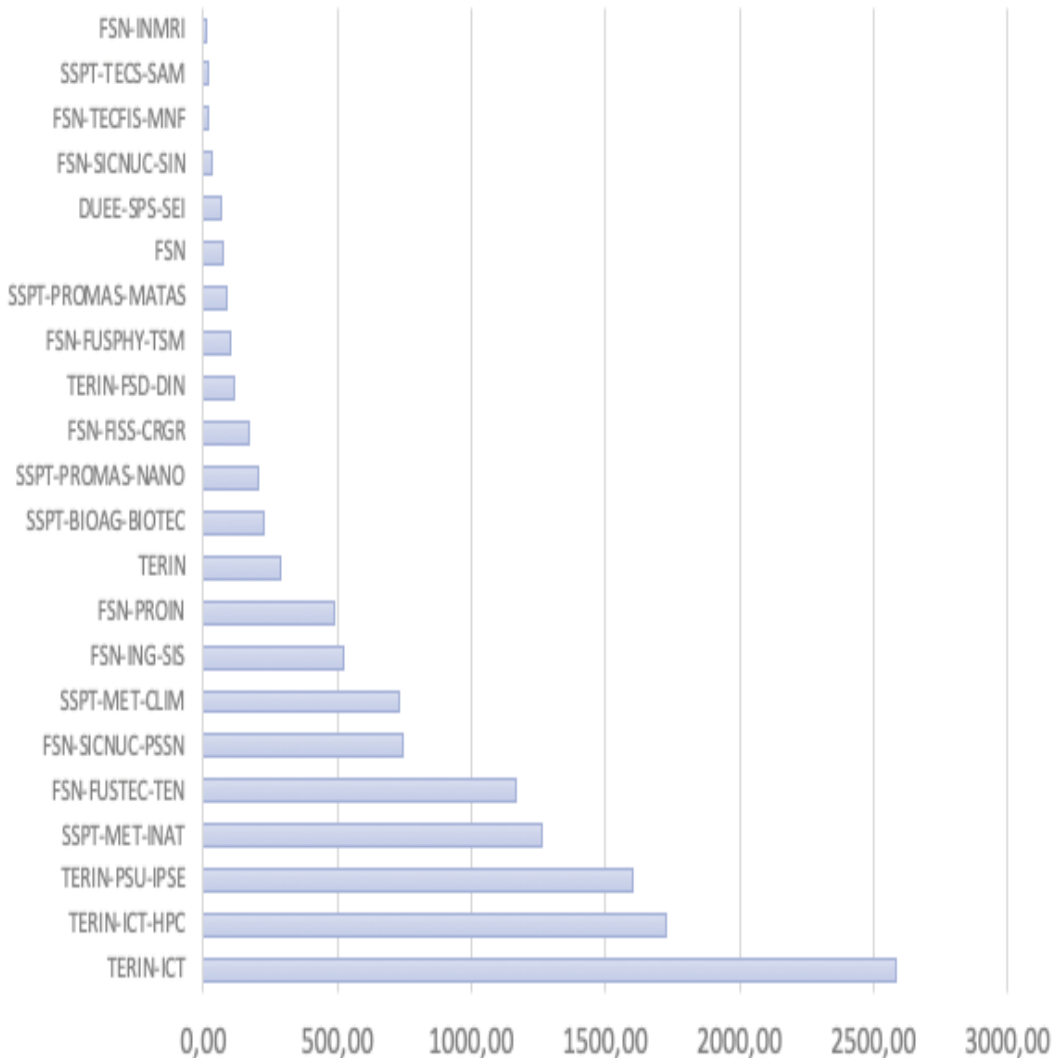
wct_y for ENEA departments 2021



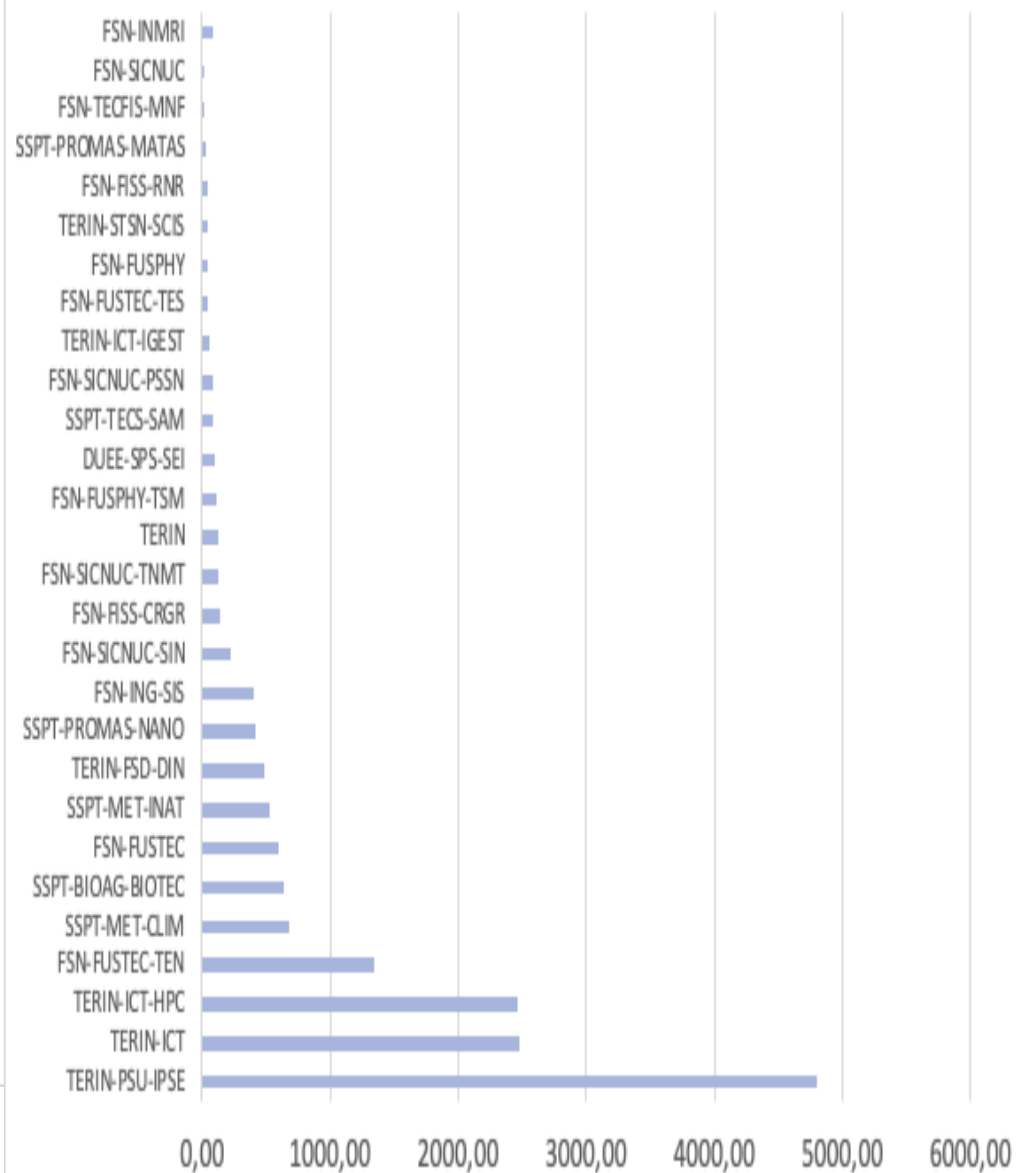
■ TERIN ■ FSN ■ SSPT ■ DUEE ■ IRP ■ UTA ■ ISV

ENEA usage laboratories wct_y 2020-2021

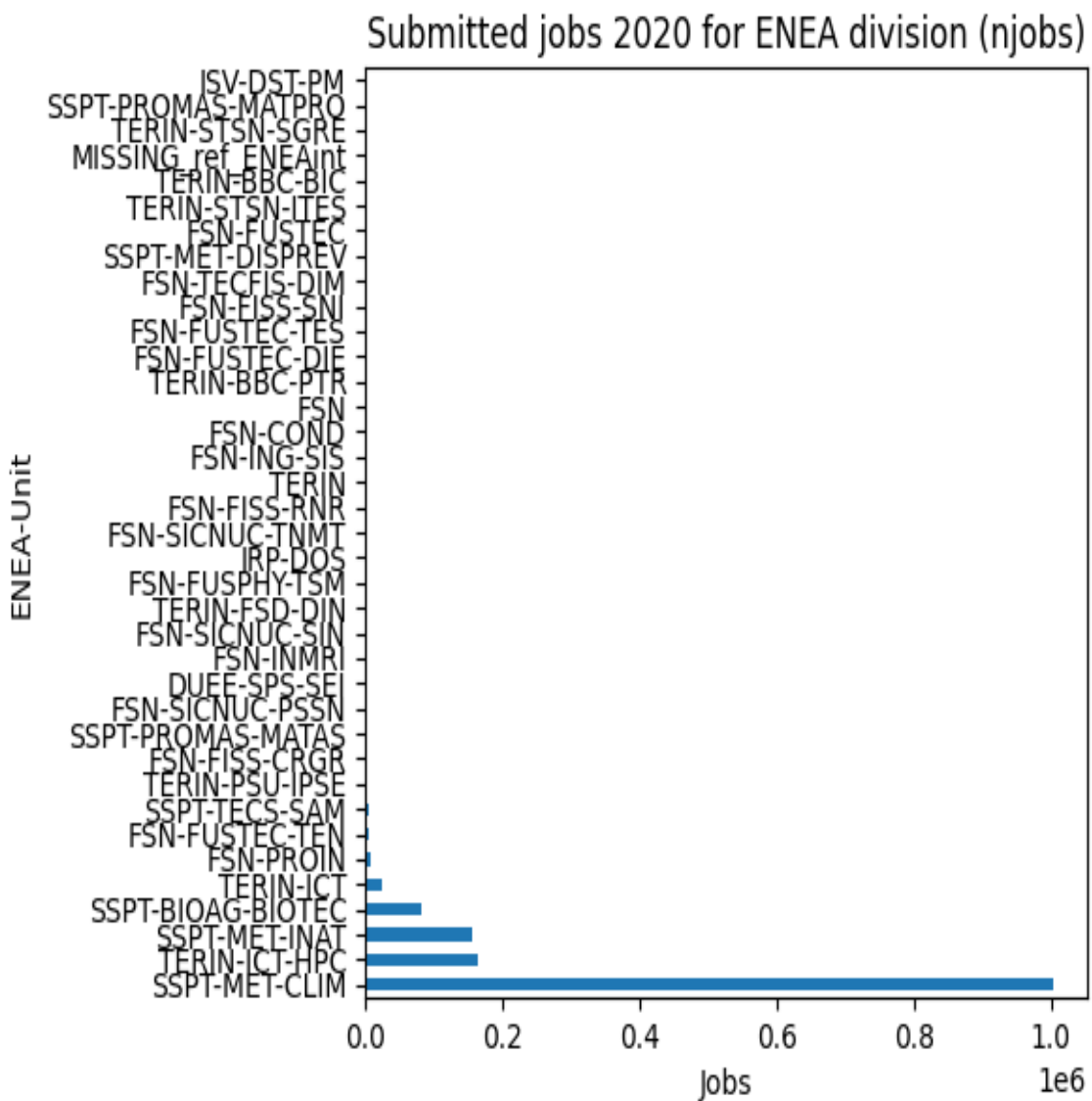
2020 wct_y



2021 wct_y



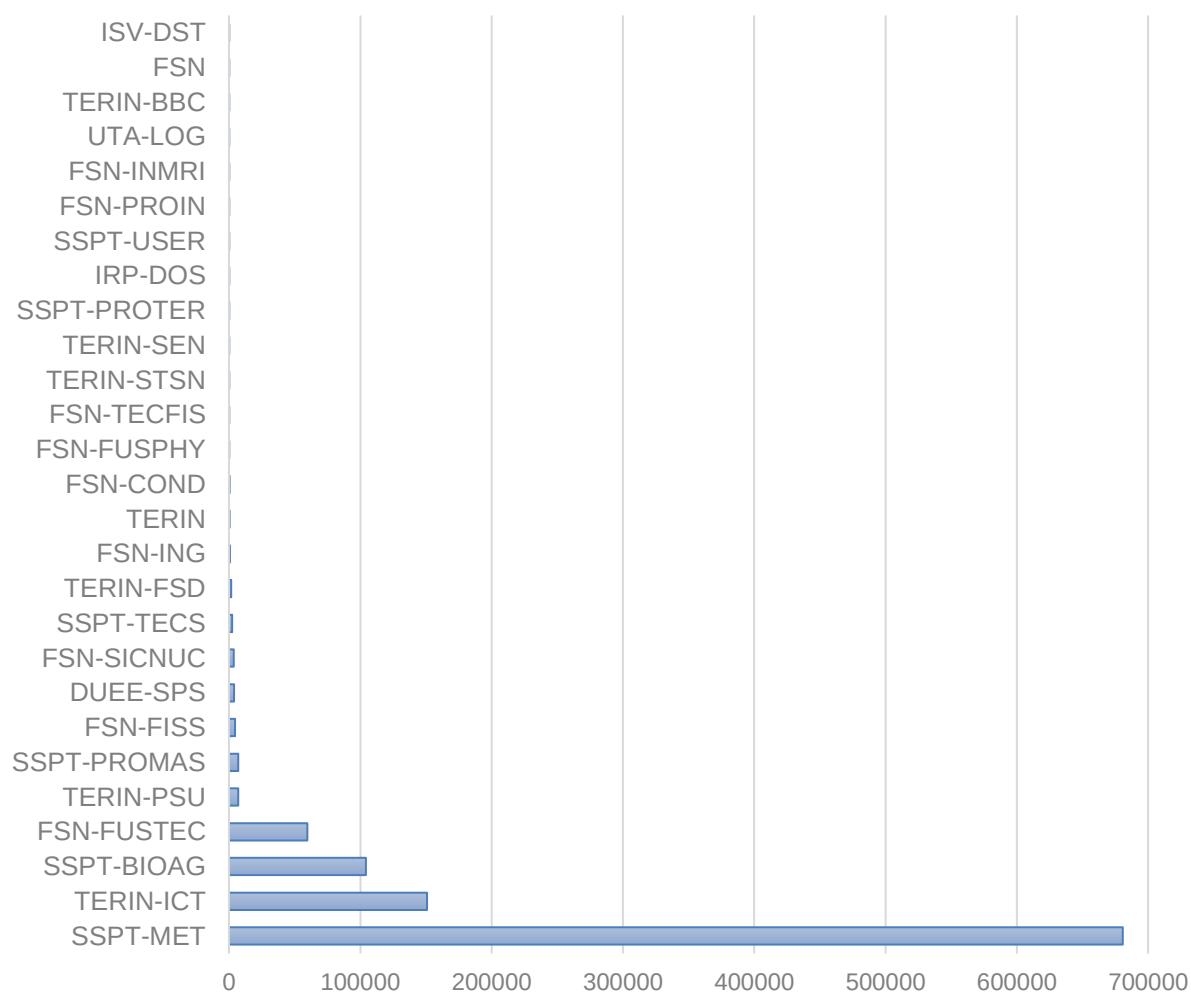
Usage by ENEA division - jobs 2020



DIV	njobs
SSPT-MET	1214808
TERIN-ICT	296510
SSPT-BIOAG	88462
TERIN-???	87479
FSN-PROIN	8780
FSN-FUSTEC	6971
SSPT-TECS	4342
SSPT-PROMAS	3632
TERIN-PSU	3208
FSN-FISS	2371
FSN-SICNUC	2256
FSN-FUSPHY	1164
DUEE-SPS	1138
FSN-INMRI	1056
TERIN-FSD	637
FSN-TECFIS	535
IRP-DOS	397
TERIN	320
FSN-ING	287
FSN-COND	247
FSN	208
SSPT-PROTER	194
TERIN-BBC	170
TERIN-STSN	5
MESSINA	3
ISV-DST	1

Usage by ENEA division - jobs 2021

Submitted jobs Division 2021



DIV	njobs
SSPT-MET	680694
TERIN-ICT	150703
SSPT-BIOAG	104206
FSN-FUSTEC	59587
TERIN-PSU	6939
SSPT-PROMAS	6936
FSN-FISS	4563
DUEE-SPS	3828
FSN-SICNUC	3643
SSPT-TECS	2410
TERIN-FSD	1561
FSN-ING	716
TERIN	433
FSN-COND	314
FSN-FUSPHY	292
FSN-TECFIS	272
TERIN-STSN	208
TERIN-SEN	128
SSPT-PROTER	101
IRP-DOS	67
SSPT-USER	34
FSN-PROIN	31

CRESCO Report

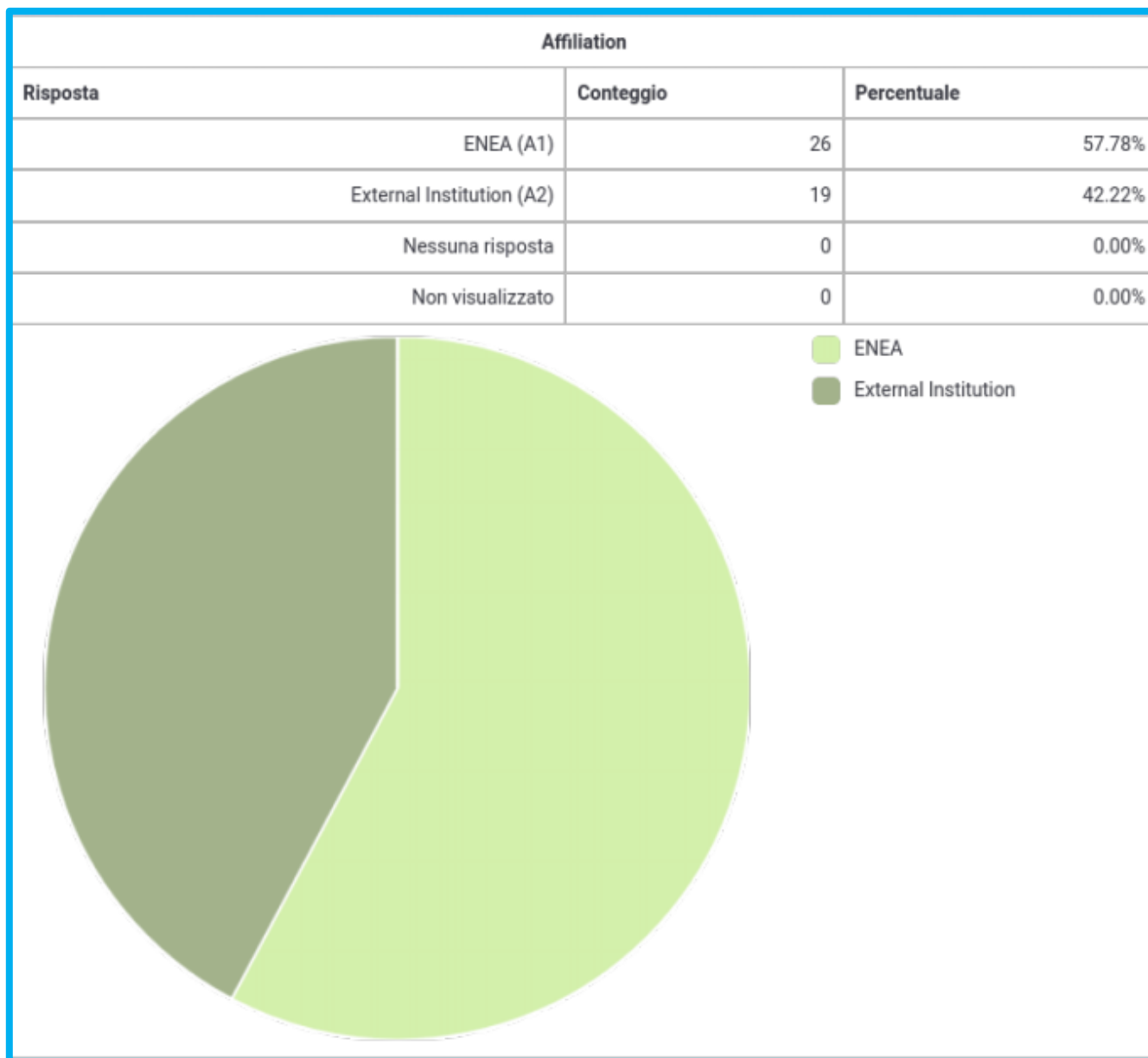
2020 report statistics

Contributed papers: statistics

- ✓ Invited users: 100/152
 - ✓ $wct_y > 1.0 \parallel (njobs * av_nc > 5000)$
- ✓ Survey responses: 45
- ✓ Papers : 45 (131 authors, ~3 authors/paper)
- ✓ Missing authors with $wct_y > 100$: 6 (on top30)
- ✓ Missing authors with $wct_y > 20$: 10 (on top60)

- ✓ 2020: we had 52 papers (D. De Chiara, S.Giusepponi)
- ✓ 2019: we had 42 papers (G.Santomauro, G.Guarnieri)
- ✓ 2018: we had 44 papers (G.Santomauro)
- ✓ 2017: we had 38 papers (M.Galli)
- ✓ 2016: we had 45 papers (M.Gusso)
- ✓ 2015: we had 46 papers (G.Ponti)
- ✓ 2014: we had 28 papers (F.Ambrosino)

ENEA and external Users



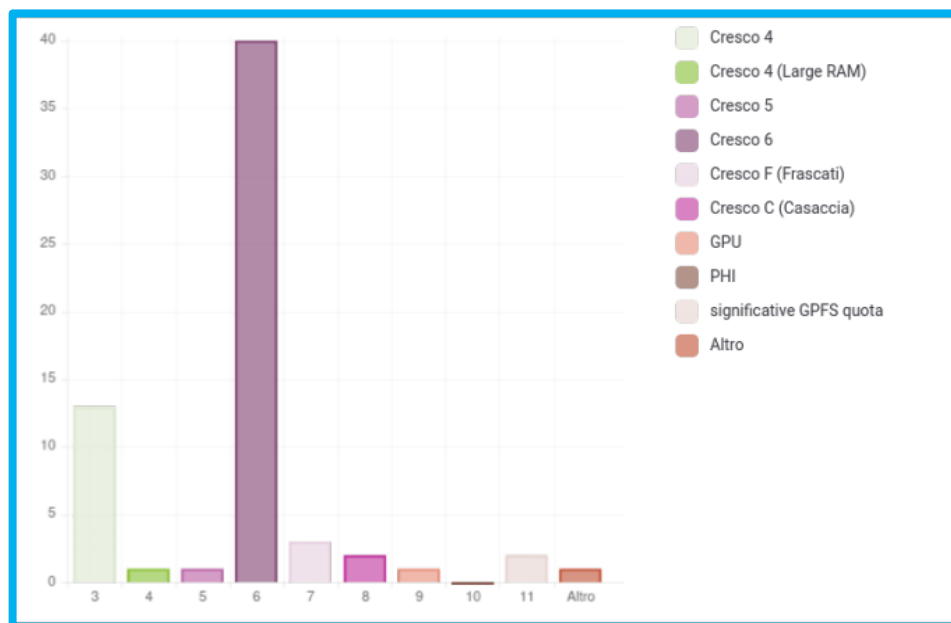
Contribution to a previous CRESCO Report



Risposta	Conteggio	Percentuale
Cresco Report 2008-2009 (1)	3	6.67%
Cresco Report 2009-2010 (2)	3	6.67%
Cresco Report 2010-2011 (3)	6	13.33%
Cresco Report 2012 (4)	7	15.56%
Cresco Report 2013 (5)	10	22.22%
Cresco Report 2014 (6)	9	20.00%
Cresco Report 2015 (7)	11	24.44%
Cresco Report 2016 (8)	13	28.89%
Cresco Report 2017 (9)	17	37.78%
Cresco Report 2018 (10)	23	51.11%
Cresco Report 2019 (11)	26	57.78%
Non visualizzato	14	31.11%

No contribution to previous reports: 14 (31%)

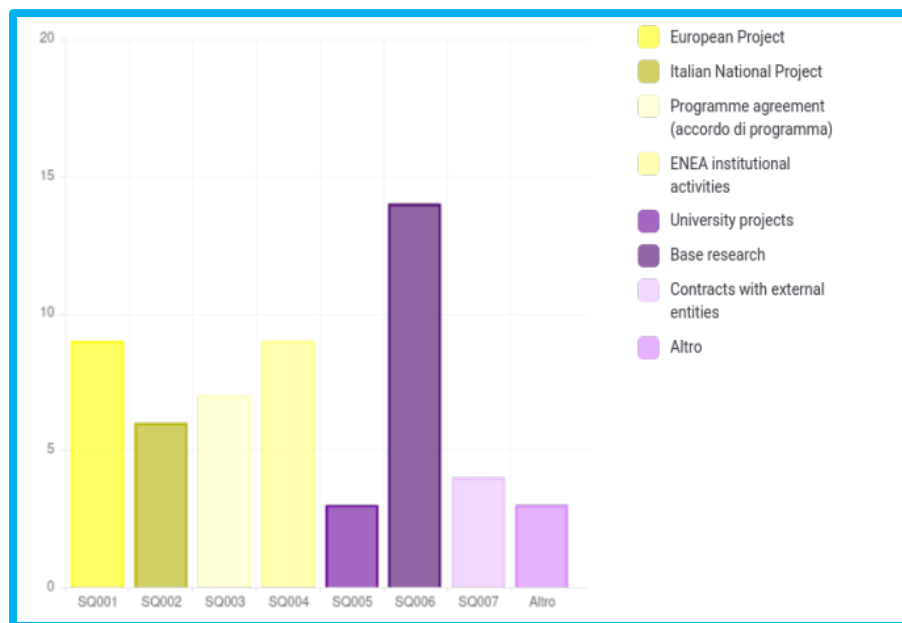
Clusters



Hardware		
	Conteggio	Percentuale
Cresco 4 (3)	13	28.89%
Cresco 4 (Large RAM) (4)	1	2.22%
Cresco 5 (5)	1	2.22%
Cresco 6 (6)	40	88.89%
Cresco F (Frascati) (7)	3	6.67%
Cresco C (Casaccia) (8)	2	4.44%
GPU (9)	1	2.22%
PHI (10)	0	0.00%
significative GPFS quota (11)	2	4.44%
Altro Sfogliare	1	2.22%

Cobranv1.frascati – desktop remoto

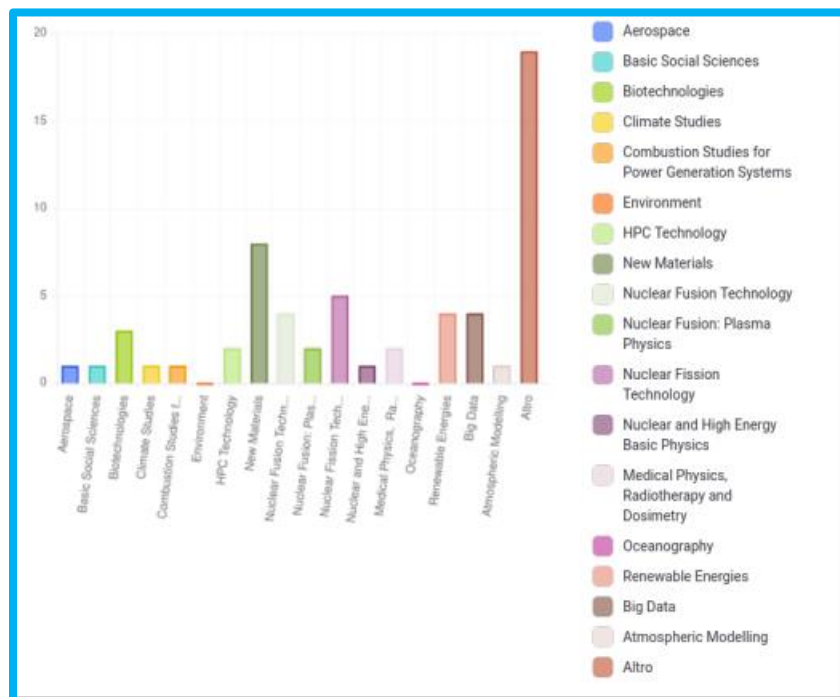
Context of the Activity



Context of the activity		
Risposta	Conteggio	Percentuale
European Project Sfogliare	9	20.00%
Italian National Project Sfogliare	6	13.33%
Programme agreement (accordo di programma) Sfogliare	7	15.56%
ENEA institutional activities Sfogliare	9	20.00%
University projects Sfogliare	3	6.67%
Base research Sfogliare	14	31.11%
Contracts with external entities Sfogliare	4	8.89%
Altro Sfogliare	3	6.67%

Many works on base research
 EU projects
 ENEA institutional activities
 Programme agreement

Research Domains



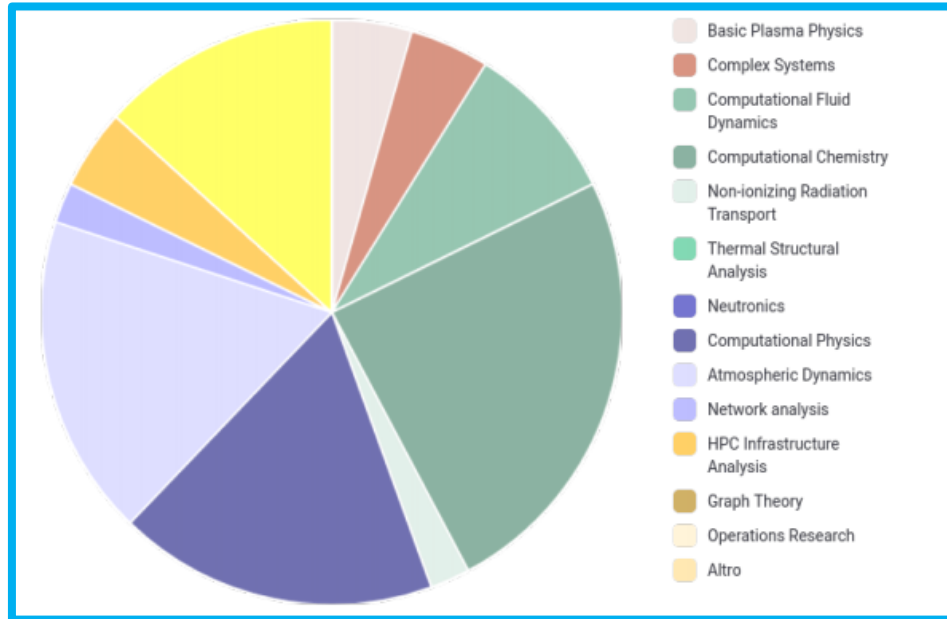
Other includes:

Radiation Protection	COVID-19
Molecular Dynamics	Statistics
Biophysics	Complex systems
Solid State Physics	Complex systems
Physics	Parallel Text Mining
Web Crawling - Big Data Analytics	Knowledge Extraction
Cultural Heritage	Metrology
Computational Drug Design	nuclear energy, MOX fuel
Energy Storage	Structural Bioinformatics
Computational Biology	computational nanoscience

Research & Development Domain		
	Conteggio	Percentuale
Aerospace (SQ001)	1	2.22%
Basic Social Sciences (SQ004)	1	2.22%
Biotechnologies (SQ003)	3	6.67%
Climate Studies (SQ002)	1	2.22%
Combustion Studies for Power Generation Systems (SQ005)	1	2.22%
Environment (SQ006)	0	0.00%
HPC Technology (SQ007)	2	4.44%
New Materials (SQ008)	8	17.78%
Nuclear Fusion Technology (SQ009)	4	8.89%
Nuclear Fusion: Plasma Physics (SQ010)	2	4.44%
Nuclear Fission Technology (SQ011)	5	11.11%
Nuclear and High Energy Basic Physics (SQ012)	1	2.22%
Medical Physics, Radiotherapy and Dosimetry (SQ013)	2	4.44%
Oceanography (SQ014)	0	0.00%
Renewable Energies (SQ015)	4	8.89%
Big Data (SQ016)	4	8.89%
Atmospheric Modelling (SQ017)	1	2.22%
Altro	19	42.22%



Modeling & Simulation Background

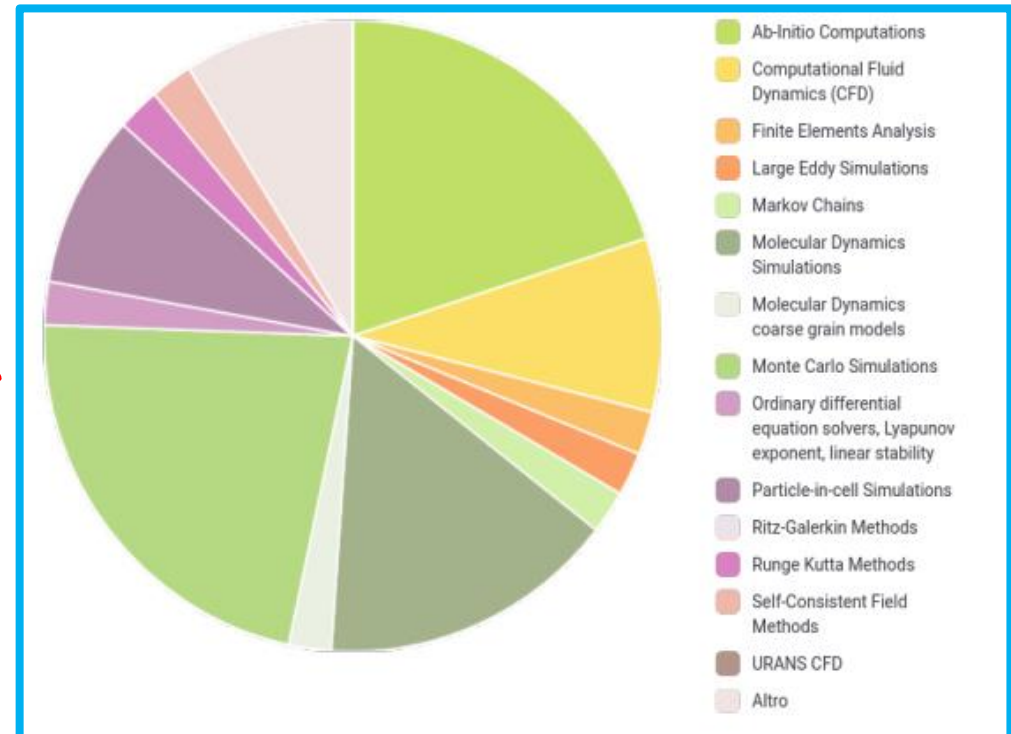


Computational Chemistry (11), Neutronics (8) and Computational Physics (8) are the main actors

Modeling & Simulation Background		
	Conteggio	Percentuale
Basic Plasma Physics (1)	2	4.44%
Complex Systems (2)	2	4.44%
Computational Fluid Dynamics (3)	4	8.89%
Computational Chemistry (4)	11	24.44%
Non-ionizing Radiation Transport (5)	1	2.22%
Non-ionizing Radiation Transport (6)	0	0.00%
Thermal Structural Analysis (7)	0	0.00%
Neutronics (8)	8	17.78%
Computational Physics (9)	8	17.78%
Atmospheric Dynamics (10)	1	2.22%
Network analysis (11)	2	4.44%
HPC Infrastructure Analysis (12)	0	0.00%
Graph Theory (13)	0	0.00%
Operations Research (14)	0	0.00%
Altro	6	13.33%

Computational Methodology

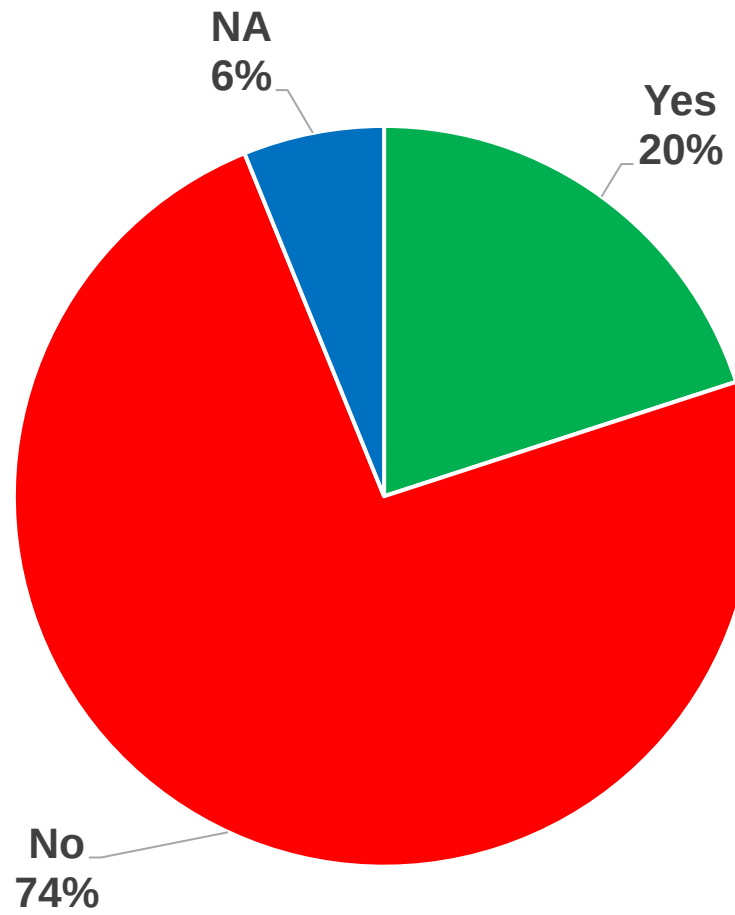
Computational Methodology:		
	Conteggio	Percentuale
Ab-Initio Computations (1)	9	20.00%
Computational Fluid Dynamics (CFD) (2)	4	8.89%
Finite Elements Analysis (3)	1	2.22%
Large Eddy Simulations (4)	1	2.22%
Markov Chains (5)	1	2.22%
Molecular Dynamics Simulations (6)	7	15.56%
Molecular Dynamics coarse grain models (7)	1	2.22%
Monte Carlo Simulations (8)	10	22.22%
Ordinary differential equation solvers, Lyapunov exponent, linear stability (9)	1	2.22%
Particle-in-cell Simulations (10)	4	8.89%
Ritz-Galerkin Methods (11)	0	0.00%
Runge Kutta Methods (12)	1	2.22%
Self-Consistent Field Methods (13)	1	2.22%
URANS CFD (14)	0	0.00%
Altro Sfogliare	4	8.89%



70% of the Montecarlo papers (10) use MCNP (7), also Ab-Initio (9) and Molecular Dynamics (7) are important.

Code Development

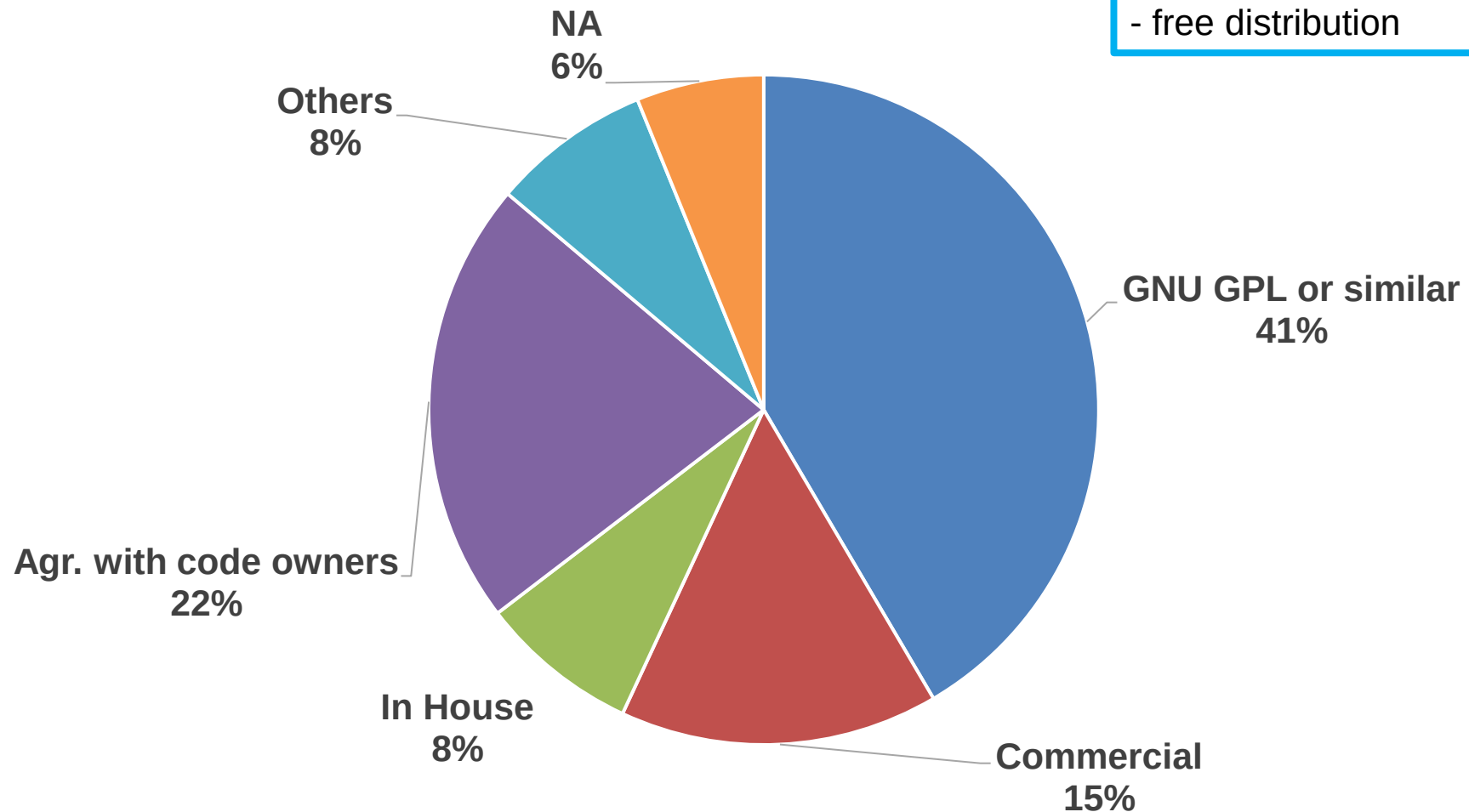
1/5 of contributors works on software codes



License type

- Most of code has a free-type license,
- 22% agreement with code owners
- 15% cases of commercial software use

- Others:
- free for research activity
 - controlled export license
 - open source
 - free distribution



Scalability tests?

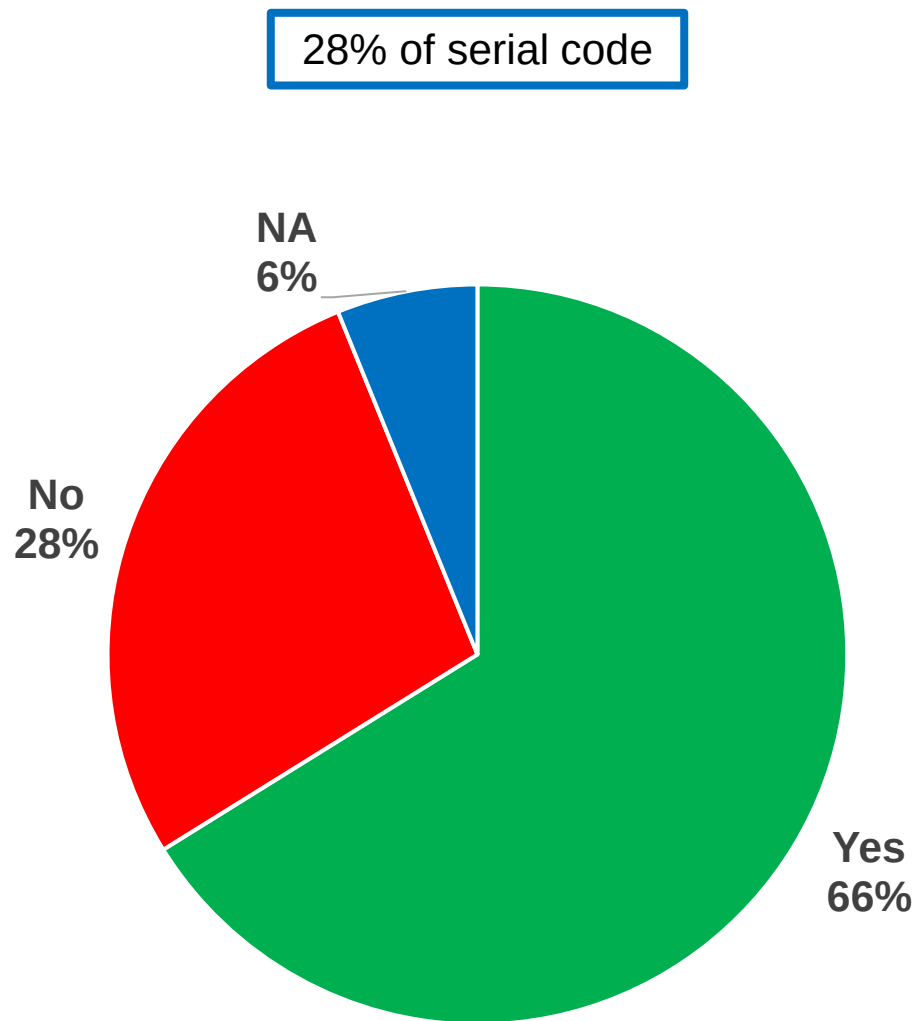
Scalability: Have you done scalability tests of the code?			
Risposta		Conteggio	Percentuale
Sì (Y)		19	42.22%
No (N)		26	57.78%
Nessuna risposta		0	0.00%
Non visualizzato		0	0.00%

~40% of users
do perform
scalability tests



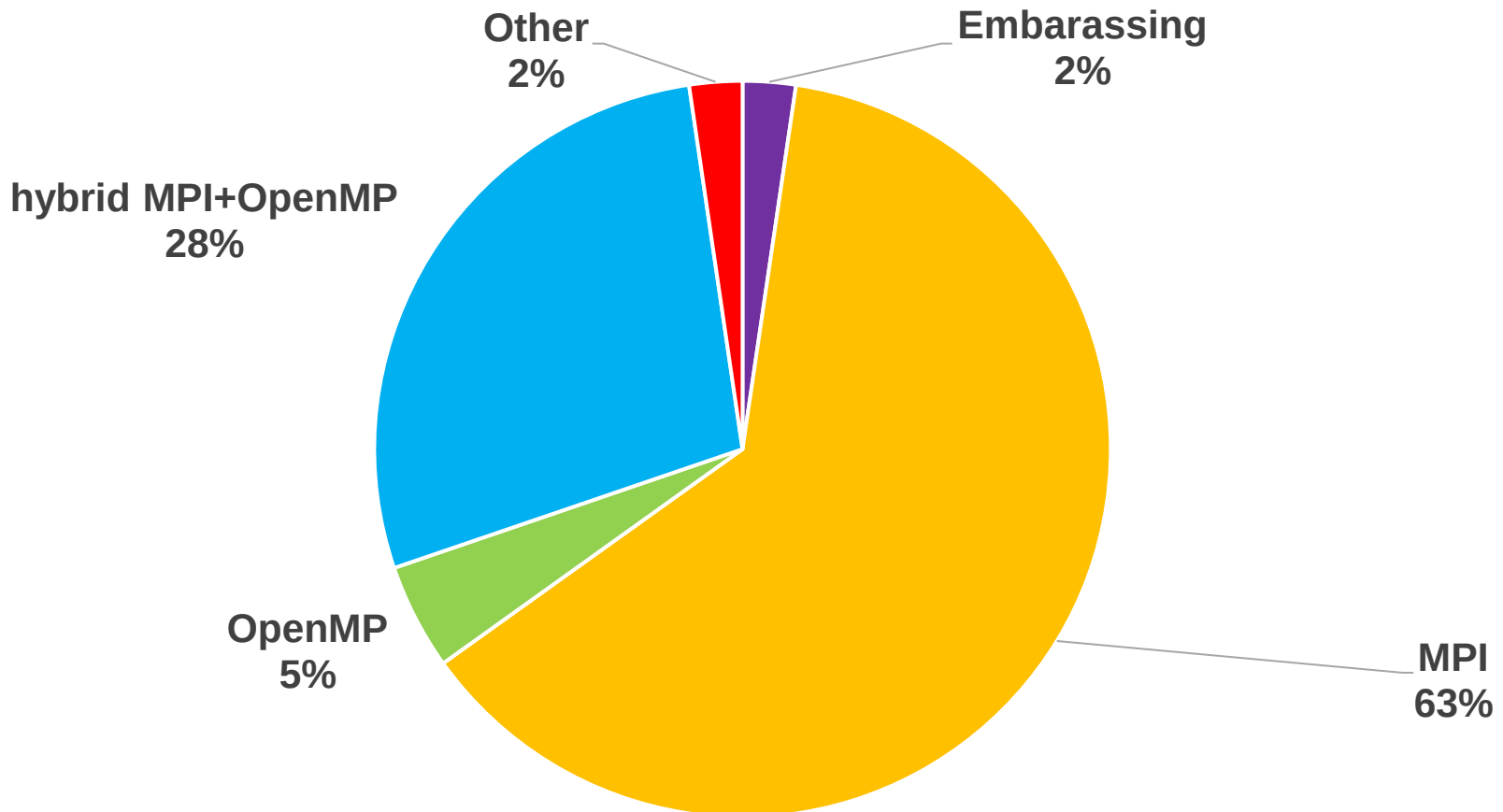
~60% of users
do not perform
scalability tests

Parallelism implemented

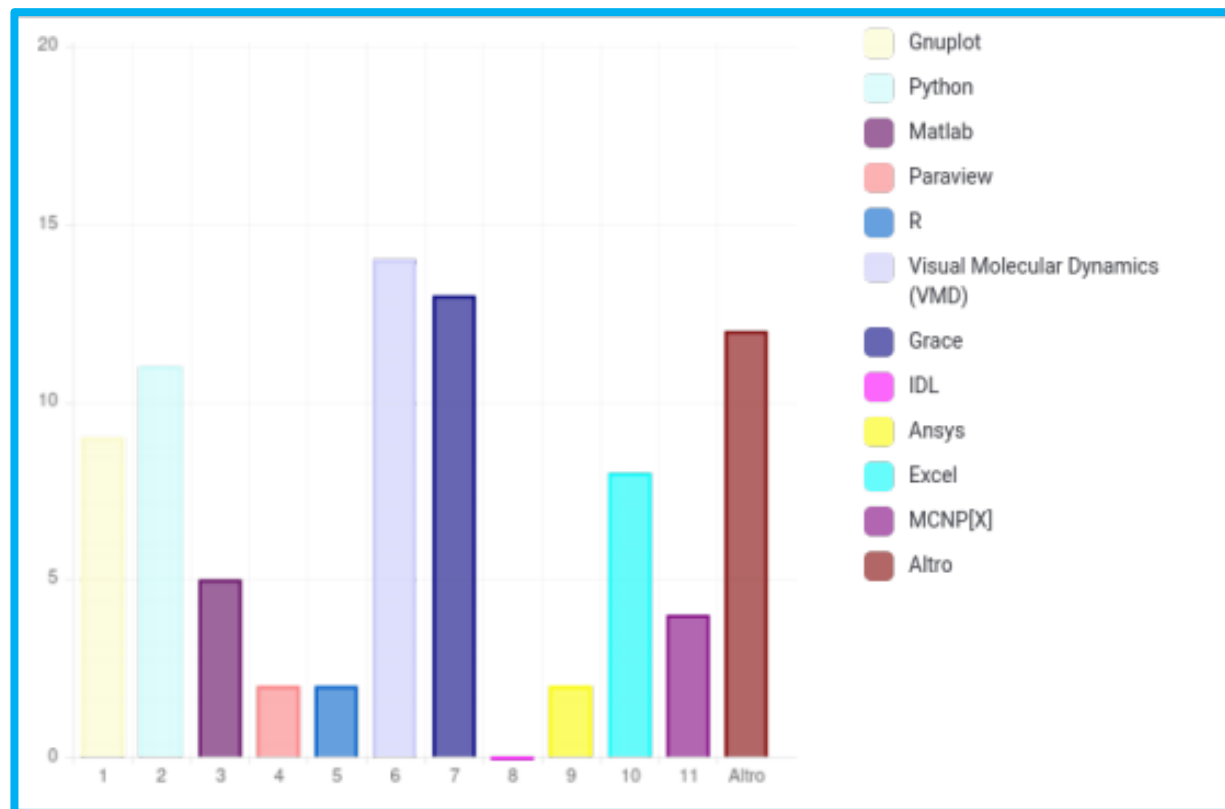


Type of parallelism implemented

Most users use MPI, but large fraction of hybrid code



Data Visualization software



Data Visualization: Specify the tools you used to view simulation results

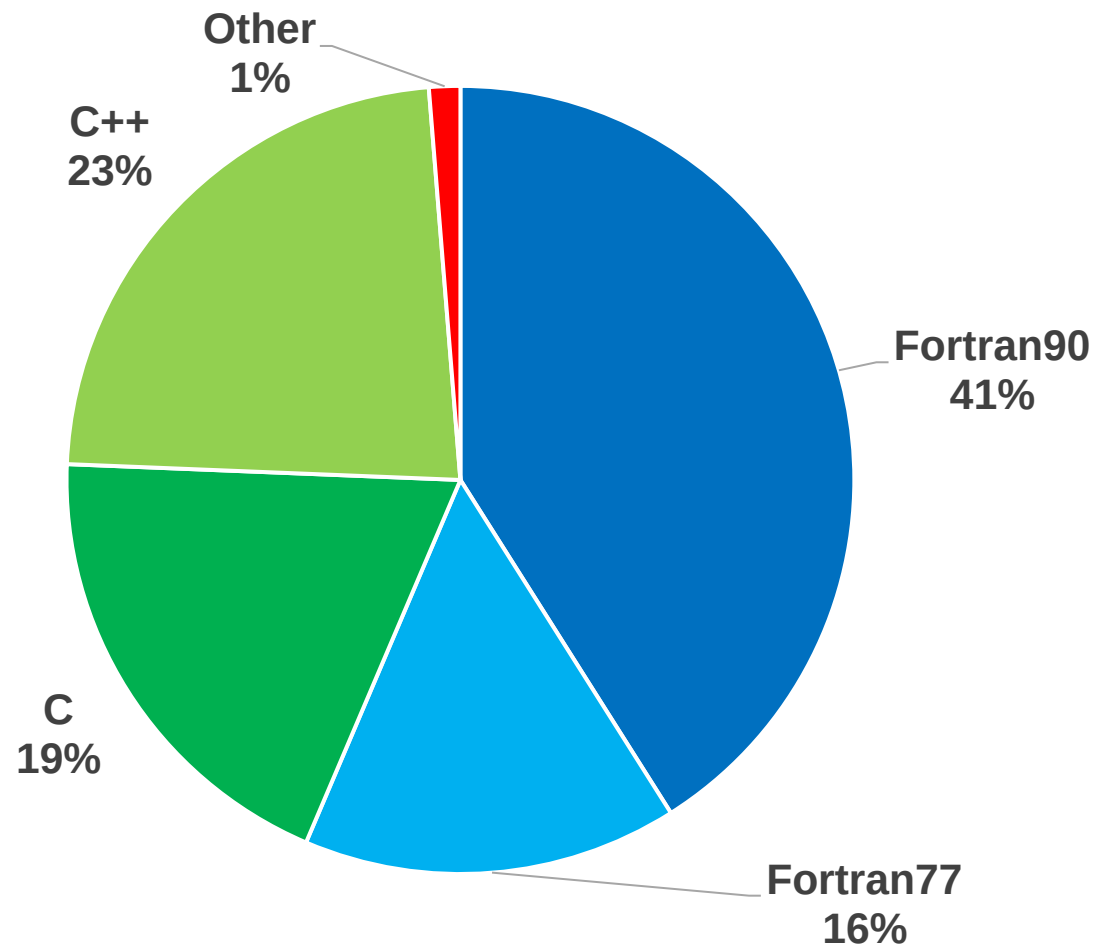
	Conteggio	Percentuale
Gnuplot (1)	9	20.00%
Python (2)	11	24.44%
Matlab (3)	5	11.11%
Paraview (4)	2	4.44%
R (5)	2	4.44%
Visual Molecular Dynamics (VMD) (6)	14	31.11%
Grace (7)	13	28.89%
IDL (8)	0	0.00%
Ansys (9)	2	4.44%
Excel (10)	8	17.78%
MCNP[X] (11)	4	8.89%
Altro	12	26.67%

12 other tools are cited:

Octave, Chimera, HeaRT, XCRYSDEN, TaLTaC2, NCL, Ferret, HIGZ.

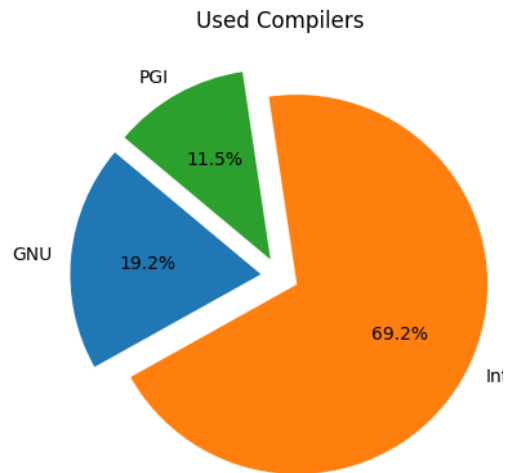
Computer Languages

The most used language is Fortran: 57%;
followed by C and C++: 42%
less than of 1/6 of Fortran is the ancient 77 version

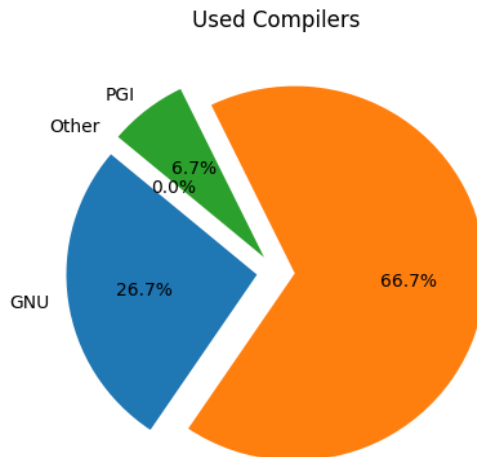


Compilers

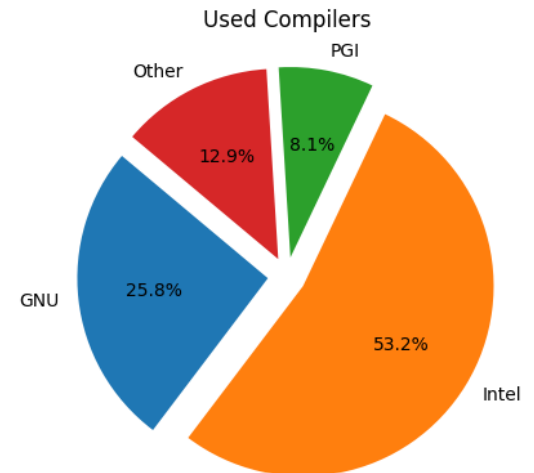
Report 2017



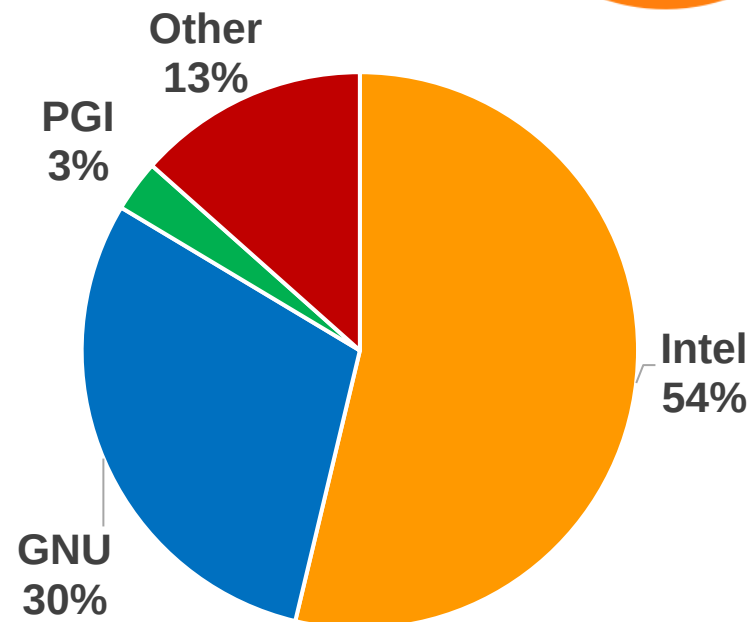
Report 2018



Report 2019



The Intel compiler is the most used, but it is decreasing



Some Used Software

~40 different software cited

Nuclear: MCNP (7), ERANOS, FISPACT, FLUKA, GEANT4.

MD: GROMACS (4), LAMMPS (4), NAMD (2), OCCAM (2), AMBER.

Ab-Initio: Quantum Espresso (5), VASP (3), ORCA (2), CPMD, CP2K, Siesta.

CFD: (Computational Fluid Dynamics) Ansys Fluent, OpenFOAM.

Environment: WRF (Weather Research and Forecasting).

Other: HeaRT, R, TaLTaC2, Photoscan, Python...

Most used libraries

Linear Algebra:

- BLAS (Basic Linear Algebra Subprograms) – 6
- LAPACK (Linear Algebra PACKage) – 6
- ScaLAPACK (Scalable Linear Algebra PACKage) – 2
- ELPA (Eigenvalue solvers for Petaflop Applications) – 1

MKL (Math Kernel Library) – 5

Fast Fourier transform

- FFTW3 – 3

Data Format

- NetCDF (Network Common Data Form) – 2
- HDF5 (Hierarchical Data Format) – 2

ESMF (Earth System Modeling Framework) - 1

JEFF (Joint Evaluated Fission and Fusion File) NEA Data Library – 1

CRESCO usage 2008-2021

Trends 2008-2021

Cresco Usage 2008-2021

Some major events			
YEAR	CLUSTER	PEAK TFlops	CORES
2008-15	CRESCO1 CRESCO2	25	672 2048->2720
2013-17	CRESCO3	20	2016
2014	CRESCO4	100	4864
2015-18	CRESCO5	24	640
2018	CRESCO6	700	10416->
2019	CRESCO6+	1.400	->20832

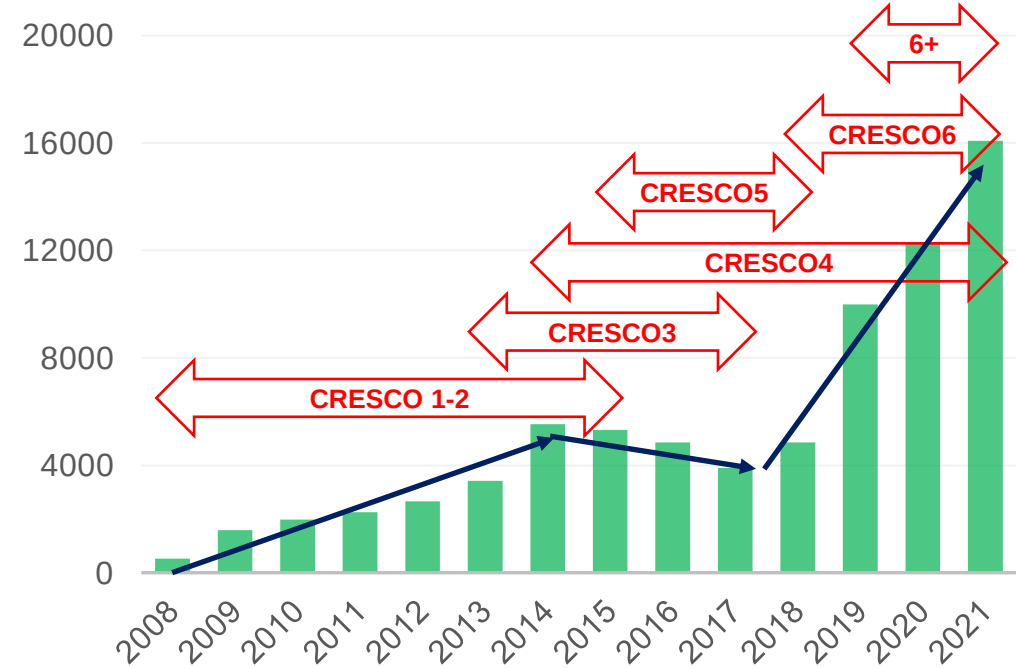
Wall Clock Time

- 2008 → 2014
- 2014 → 2017
- 2017 → 2021

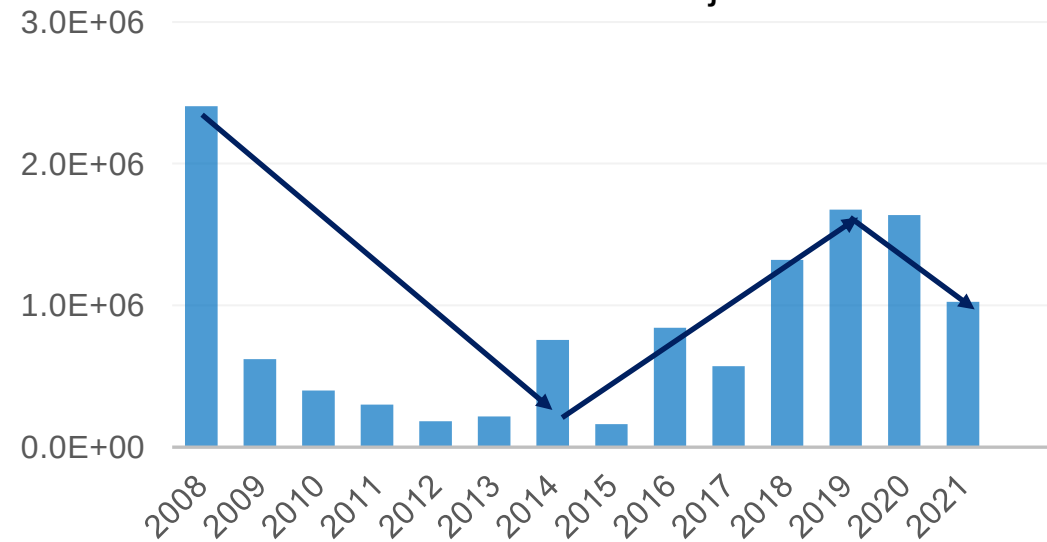
Number of jobs

- 2008 → 2014
- 2014 → 2019
- 2017 → 2021

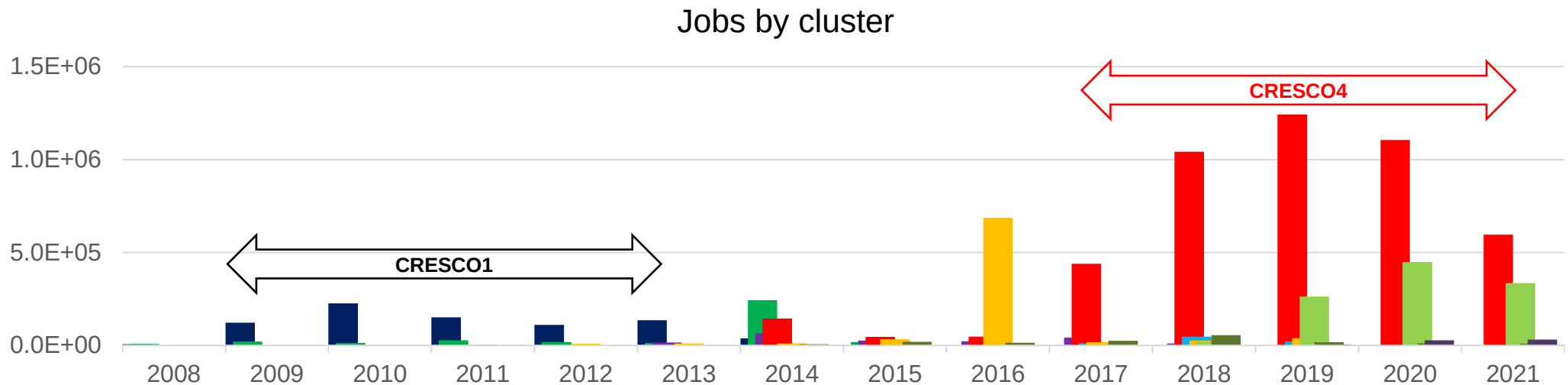
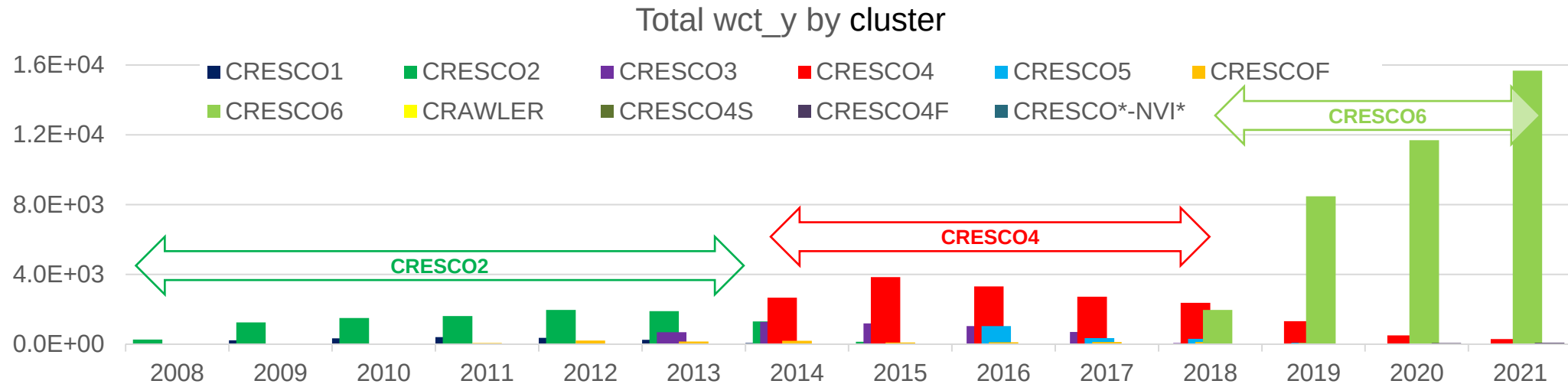
CRESCO Wall Clock Time (years)



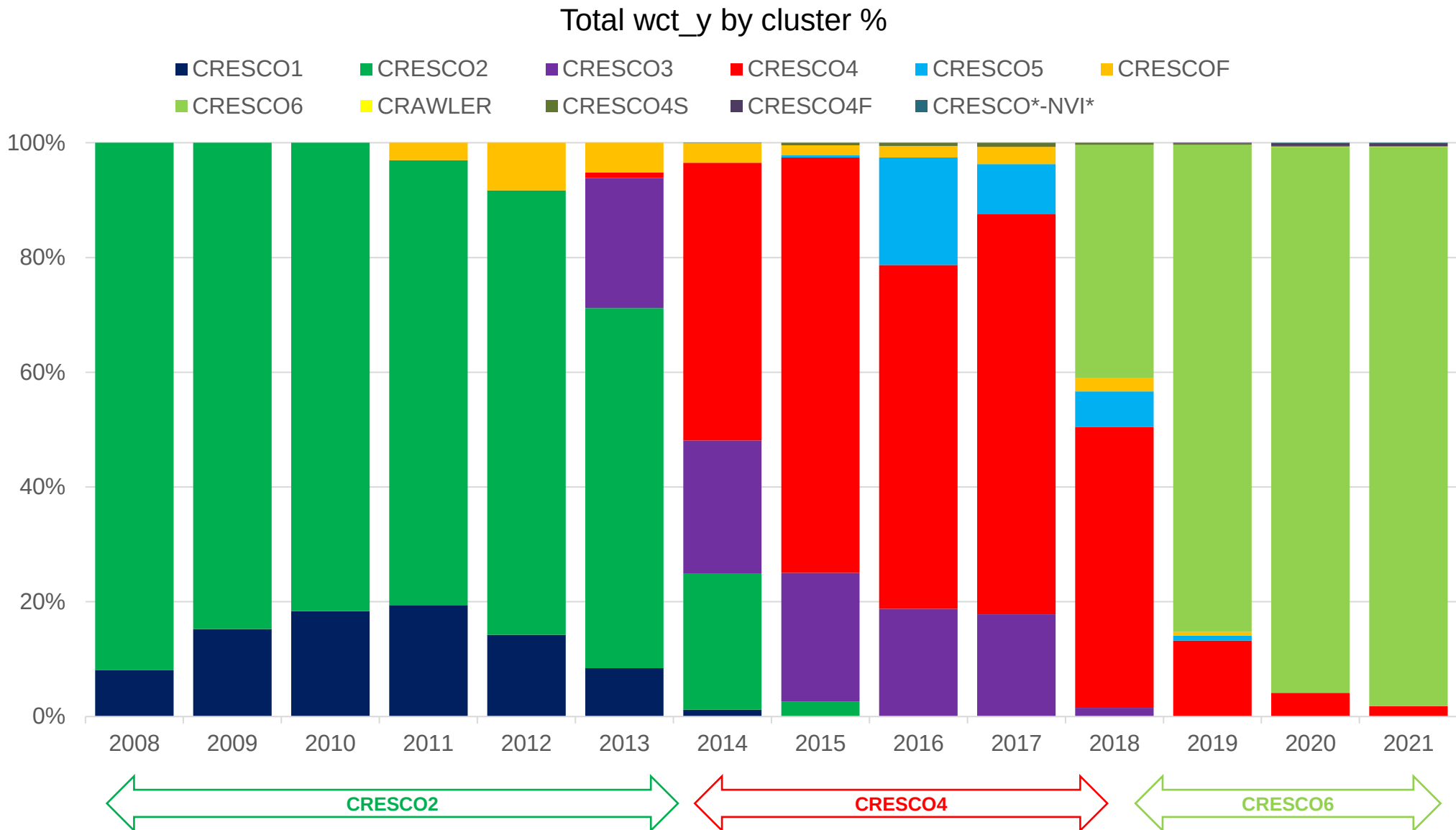
CRESCO Number of jobs



Usage by clusters in recent years

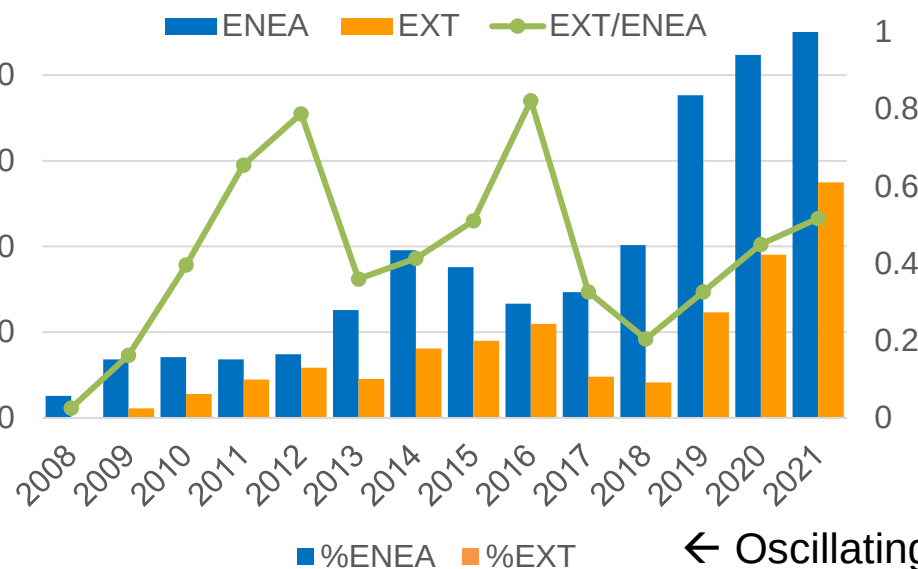


Usage by clusters in recent years

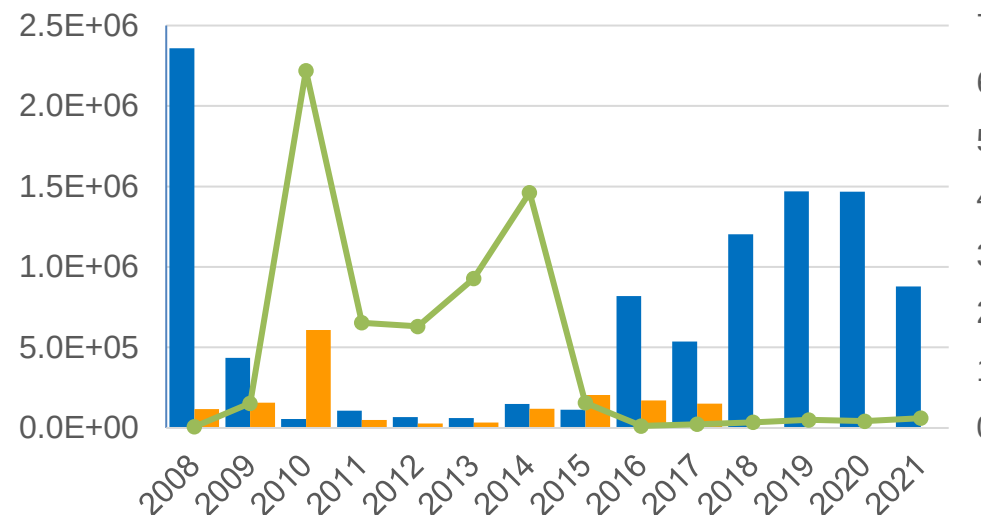


ENEA and External Users

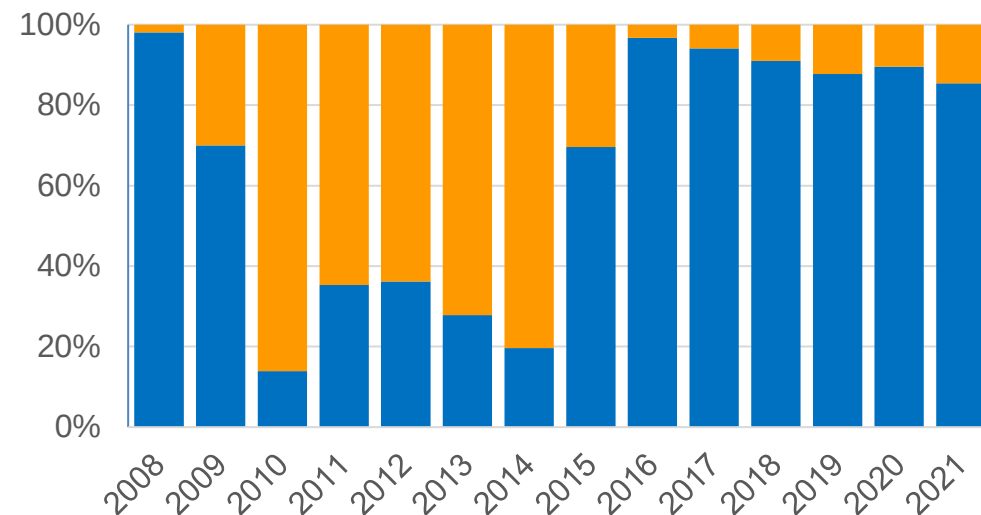
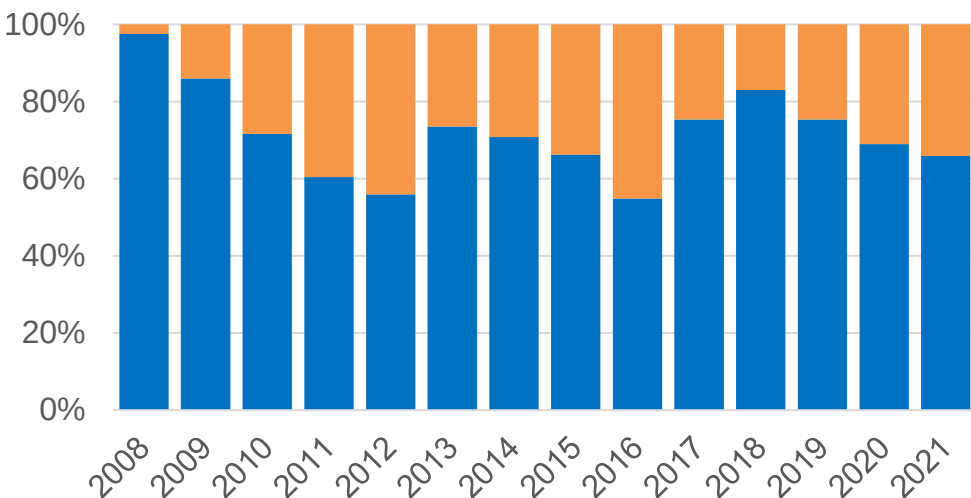
wct_y ENEA and EXTERNAL users



Jobs ENEA and EXTERNAL users



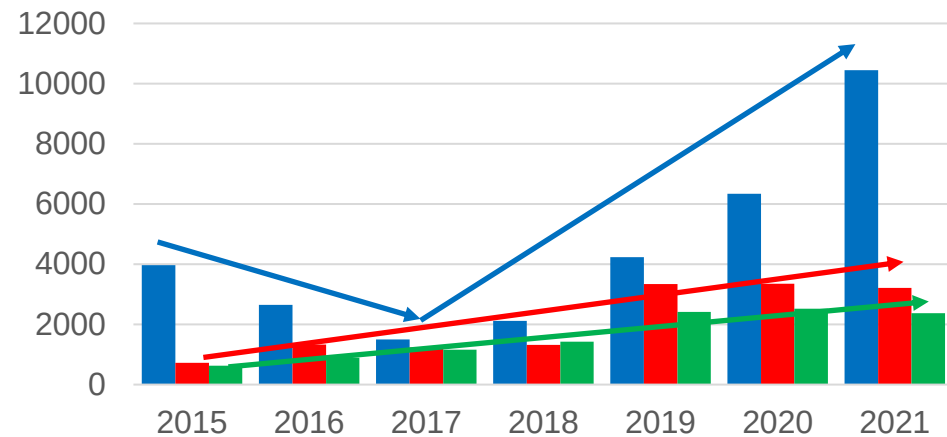
← Oscillating usage by users



ENEA Department Usage

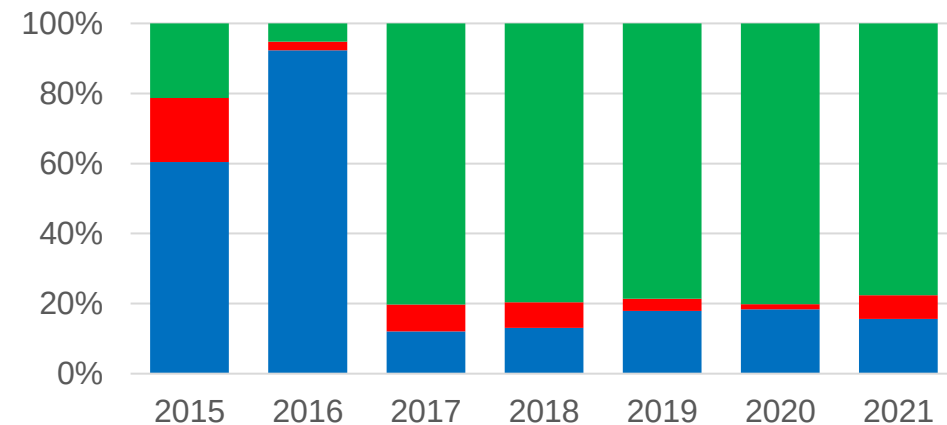
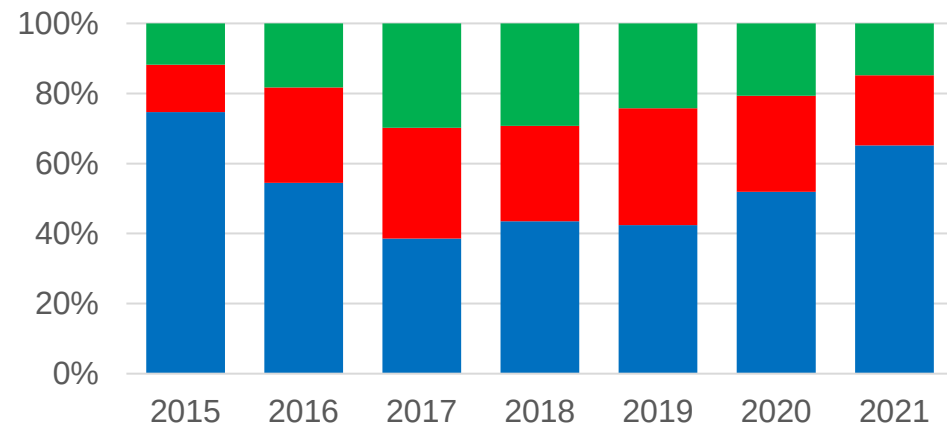
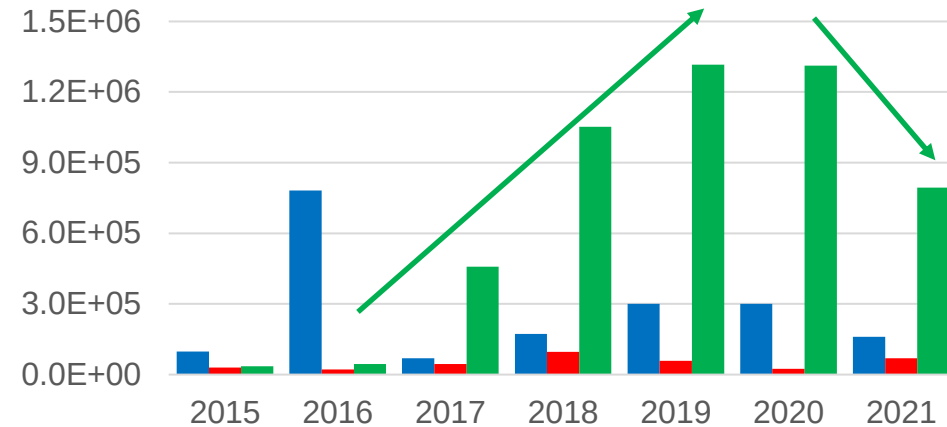
Wct_y ENEA departments

■ DTE-TERIN ■ FSN ■ SSPT



njobs ENEA departments

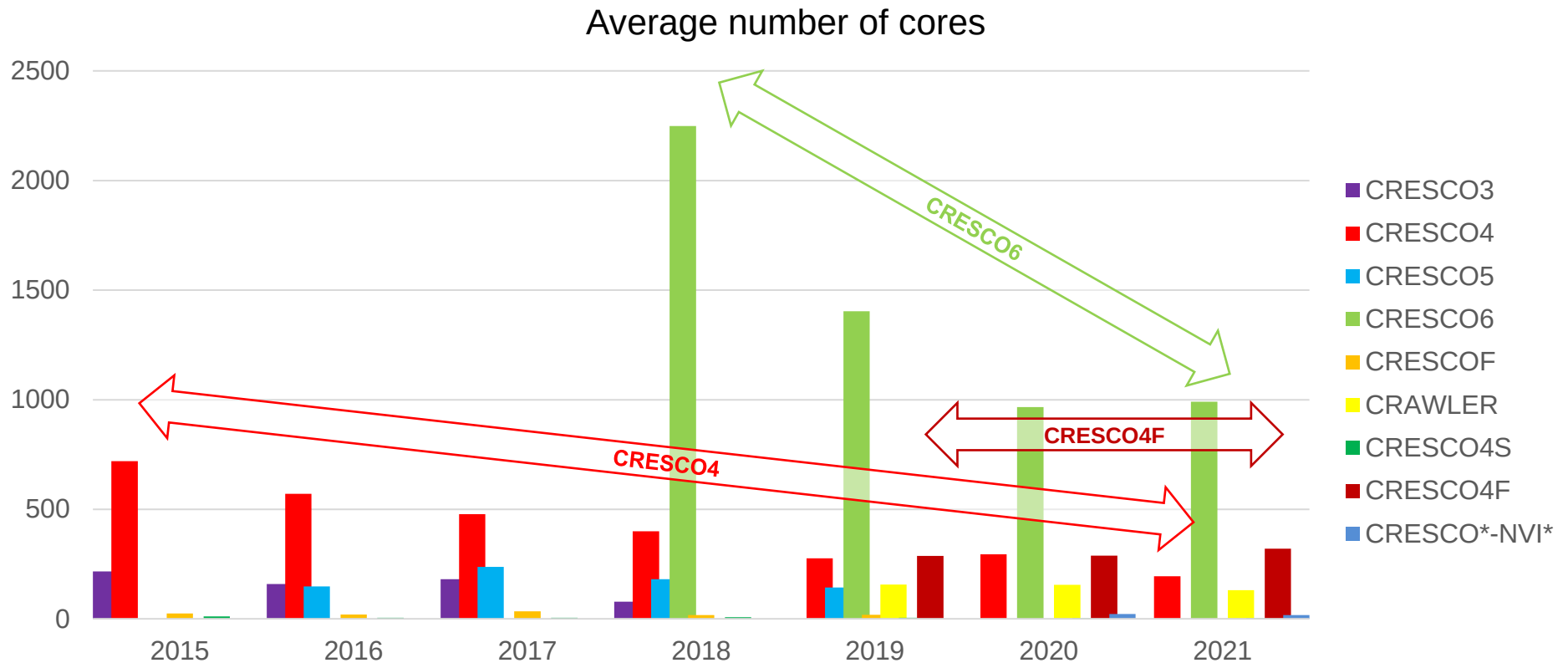
■ DTE-TERIN ■ FSN ■ SSPT



DTE-TERIN main user

Increasing number of jobs by SSPT in 2016-2019

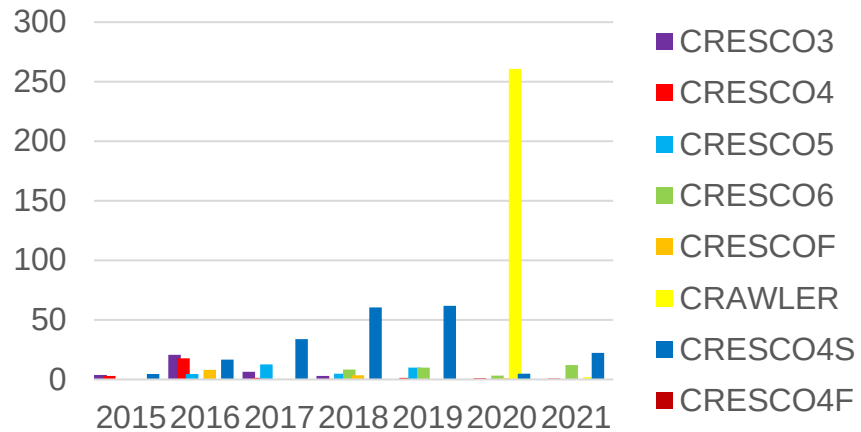
Parallelism



less usage of parallelism in recent years
In 2018 first CRESO6 tests used many cores
In 2019 first CRESO4F runs

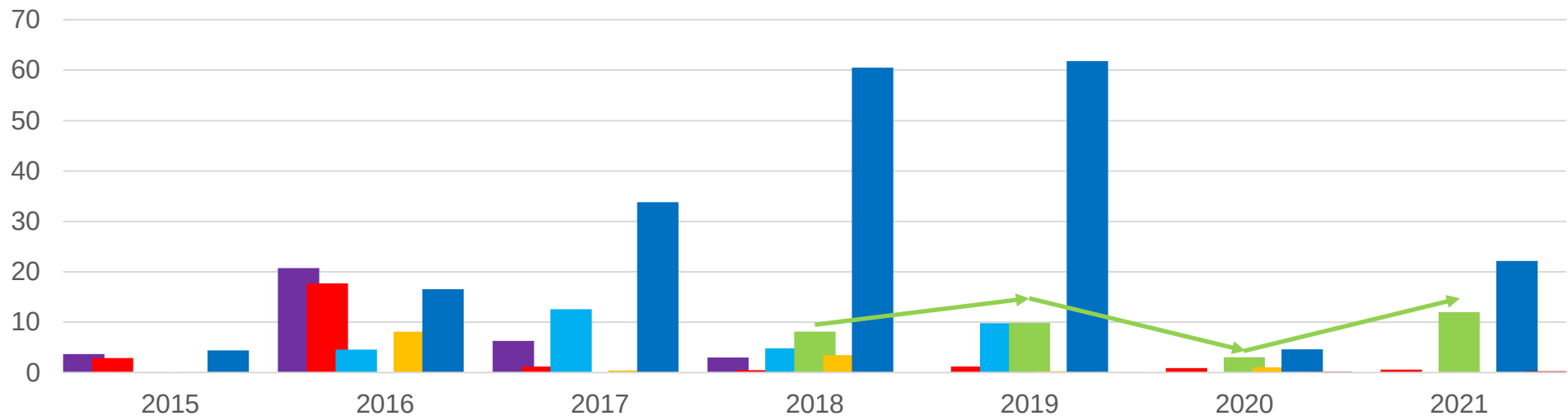
Wait time

Average wait time (hours)

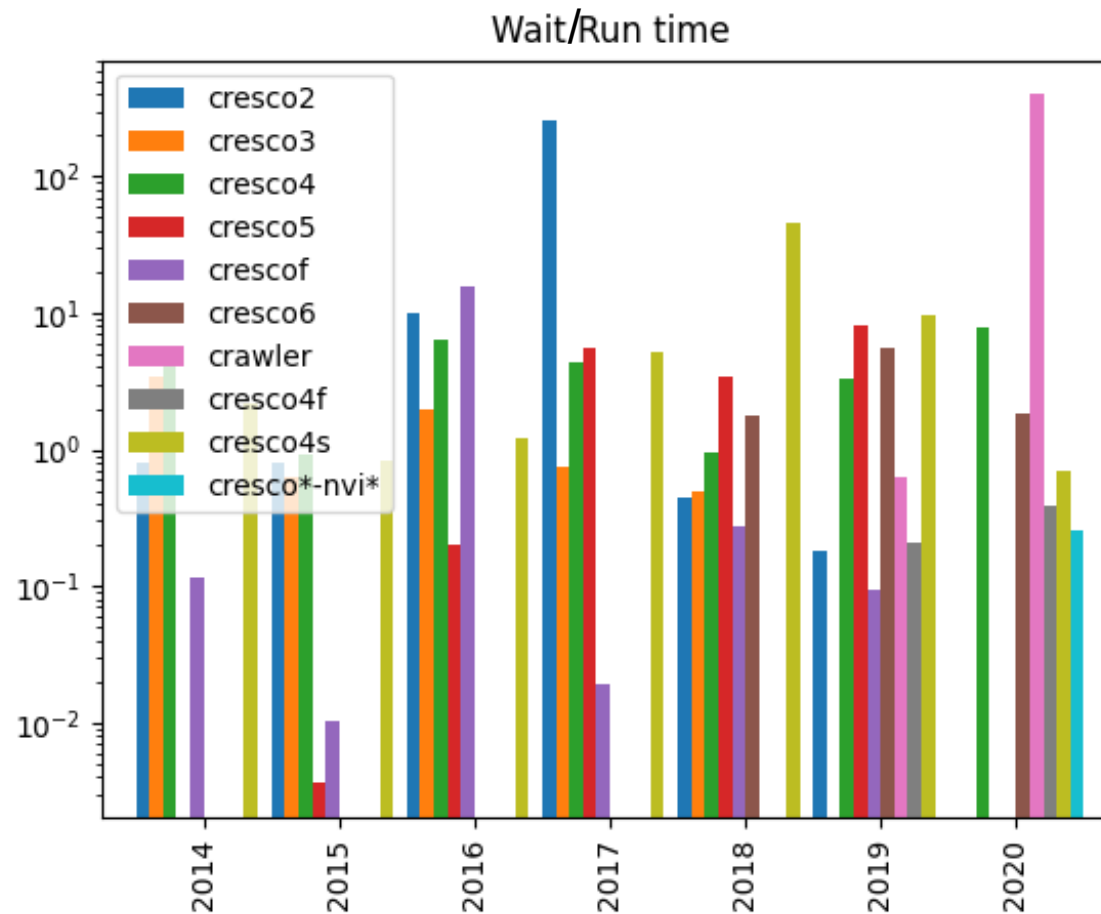


CRESCO6: increasing wait time in 2021.

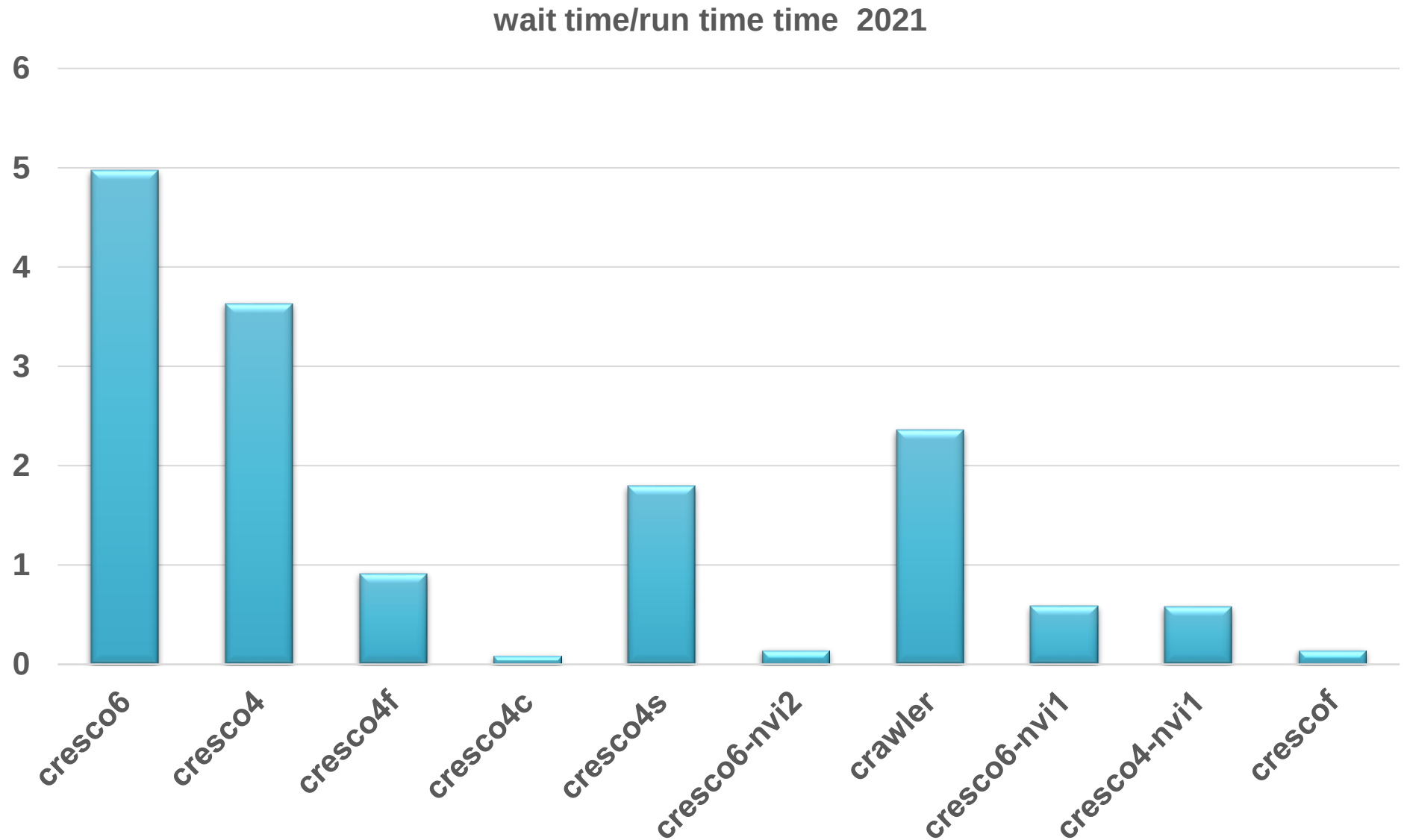
No CRAWLER



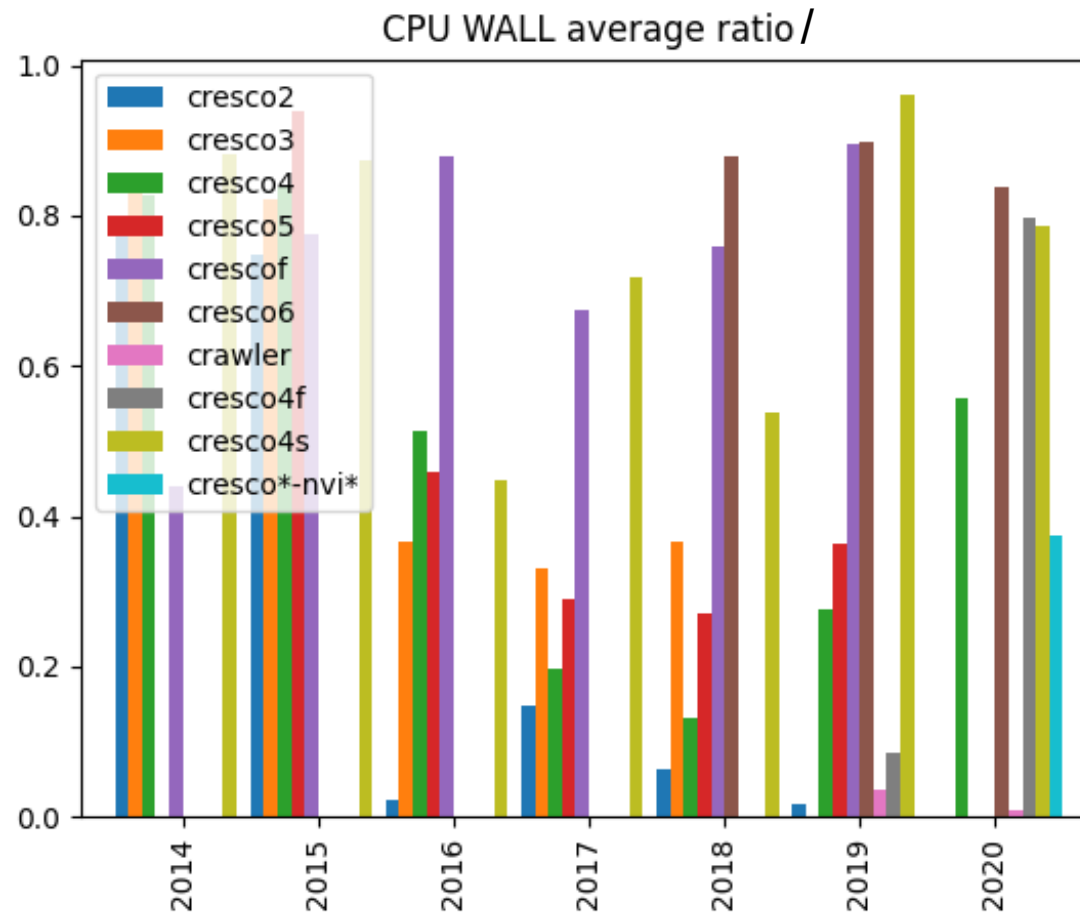
Wait/Run time



Wait/Run time 2021

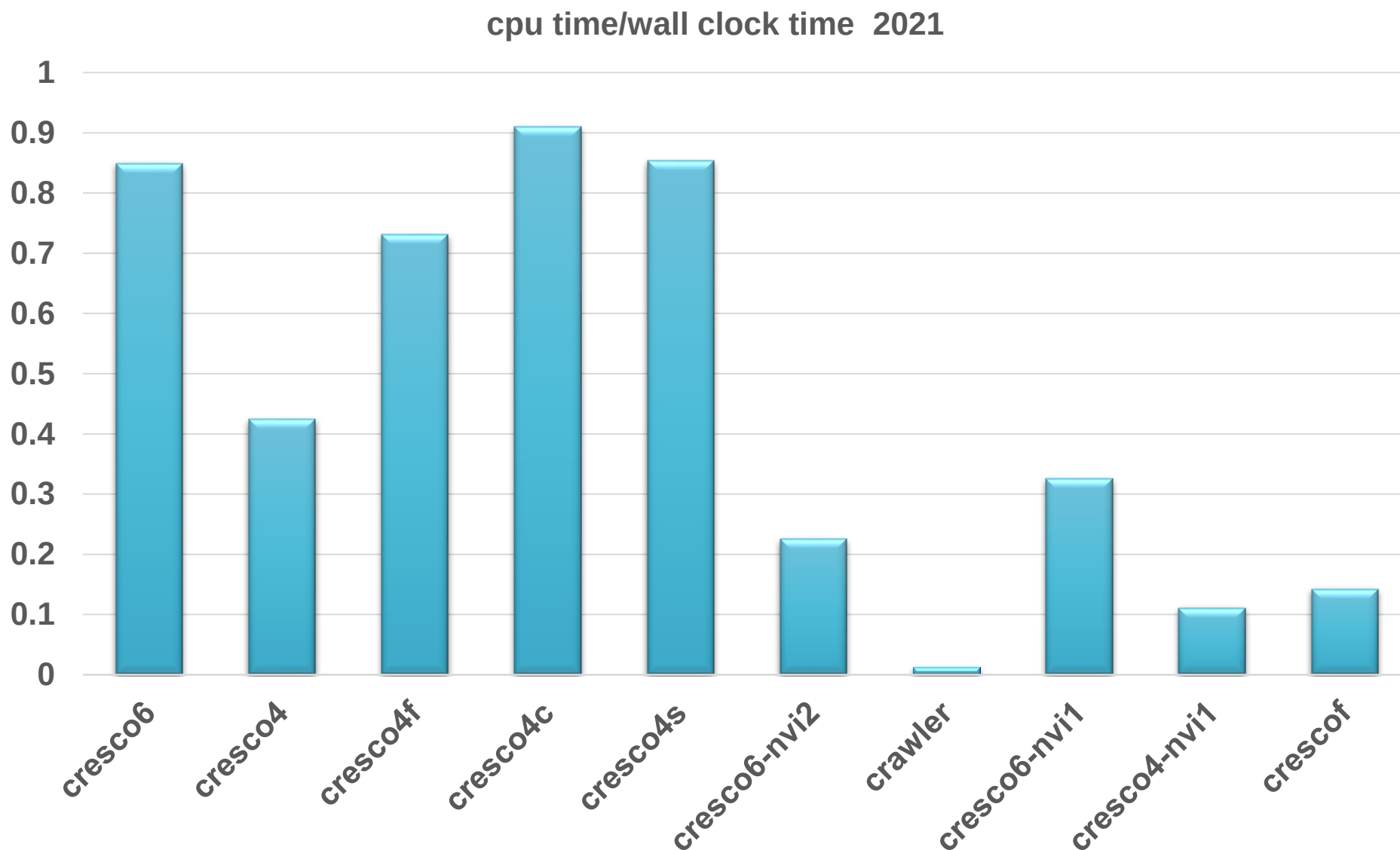


CPU time / Wall Clock time



Ratio at ~90% for CRESO6

CPU time/Wall clock time 2021



Cresco6 consumption 2020

ALL nodes	cresco6x226	cresco6x073
MEDIA KWh	MINIMO KWh	MASSIMO KWh
1644,423037	531,05272	2087,04325
Consumo anno 2020 in kwh		
697235,3676		

Cresco6 single node consumption & thermal growth 2020

sys_power in Watt	sample(sys_power)	Temp output- Temp input	Percentage utilization %
100	26026	9,809613	-
110	112399	9,349567	3
120	45004	11,968358	6
130	369971	12,896784	9
140	269650	12,308374	11
150	2171754	14,526158	14
160	439459	17,968541	17
170	83129	17,214005	20
180	77066	17,671204	23
190	79418	17,795147	26
200	97880	17,764702	29
210	152490	18,767604	31
220	275887	19,508719	34
230	390230	19,870661	37
240	356973	20,522958	40
250	495728	21,090366	43
260	527183	21,821673	46
270	822425	22,083209	49
280	1005577	22,719683	51
290	697265	23,274881	54
300	332593	23,482037	57
310	274125	23,244195	60
320	187388	24,852162	63
330	98650	25,609752	66
340	64856	26,213951	69
350	54265	27,339298	71
360	42520	27,692145	74
370	30983	28,534551	77
380	25711	28,727626	80
390	13365	28,314927	83
400	6641	26,809667	86
410	5342	26,009173	89
420	6988	26,081568	91
430	10015	26,390714	94
440	2666	27,501125	97
450	25	20.24	100

Cresco6 single node CPU power consumption & thermal growth 2020

cpu power	N° sample	percentage cpu	DeltaTemperature (Tout - Tin)
20	10	5,714286	4,2
30	7	8,571429	3,428571
40	589646	11,428571	8,156921
50	1772765	14,285714	9,334432
60	113287	17,142857	10,16144
70	90079	20	9,987467
80	53047	22,857143	8,722265
90	171917	25,714286	11,64981
100	374964	28,571429	13,50758
110	2013613	31,428571	14,02938
120	725662	34,285714	17,47129
130	98393	37,142857	17,12486
140	88545	40	17,71216
150	89738	42,857143	17,9347
160	128480	45,714286	18,334
170	218871	48,571429	19,23358
180	395337	51,428571	20,02045
190	420974	54,285714	20,17224
200	463206	57,142857	21,04882
210	572588	60	21,79385
220	922911	62,857143	22,38865
230	1278352	65,714286	22,73392
240	868832	68,571429	23,13479
250	302898	71,428571	24,21684
260	117423	74,285714	25,80029
270	66092	77,142857	26,96736
280	58671	80	27,46457
290	35203	82,857143	28,25947
300	32403	85,714286	27,52079
310	9395	88,571429	26,81245
320	6530	91,428571	26,34227
330	6794	94,285714	26,33721
340	13845	97,142857	26,34316
350	311	100	25,40836

Acknowledgments

Special thanks to:

The whole CRESCO team:

always working hard to maintain the clusters alive

Giovanni Bracco:

for accounting data files

Catia Masella:

for assistance in the use of the “Survey” system

Antonella Andreini, Paola Giaquinto:

for the printed version

Giuseppe Santomauro, Guido Guarnieri

for ongoing technical support

The CRESCO users:

for their contributions to the 2020 annual report

**ANY
QUESTIONS?**

Personal comments

On Report Preparation:

- Time to begin: as soon as possible? (For 2021-2022 Report start was on June)
- Only one template: Word? (For new accessibility rules)

On CRESCO usage:

- How to arrange High Availability activities?