

Maiatzeko estatistikak

Edu Ogando eta Txema Mercero

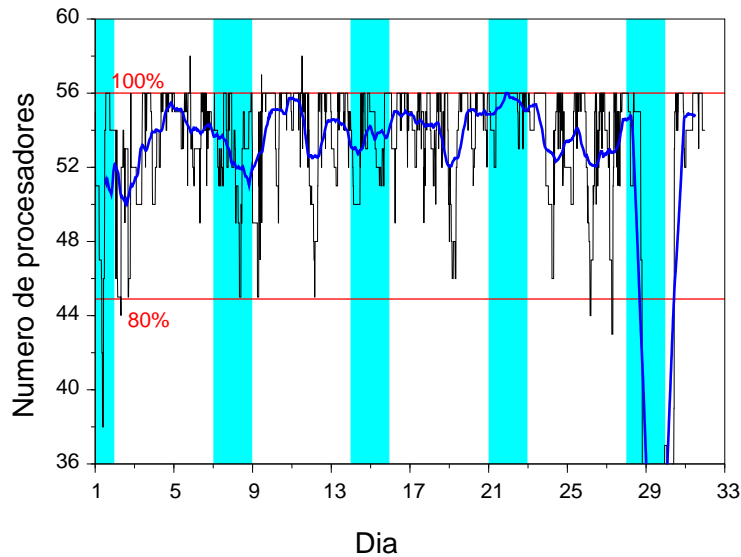
2005.eko ekainaren 23

Laburpena

Arina-ri maiatzean eman zaion erabilerari buruzko txostena aurkezten dugu. Hil honetako erabilera aurreko hilekoaren antzekoa izan da, ilareta-ko itxarondenborak pixka bat luzatu badira ere. Okupazio maila handia eta konstantea izan da, azkeneko asteburuan izan ezik, iraulketa handia gertatu zelako. Aipagarria da minutu gutxi kontsumitzen dituzten bidalitako lan kopuru handia; iruditzen zaigu horietatik batzuk amaitu egiten direla sarrera edo input fitxategiak prestatzerakoan gertatutako errore bategatik. Ehuneko hori txikiagoa izan da, aurreko hilekoarekin alderatuz. HTC-arentzako clusterra erabiltzen da gehien (iraupen txikiko kalkulu asko).

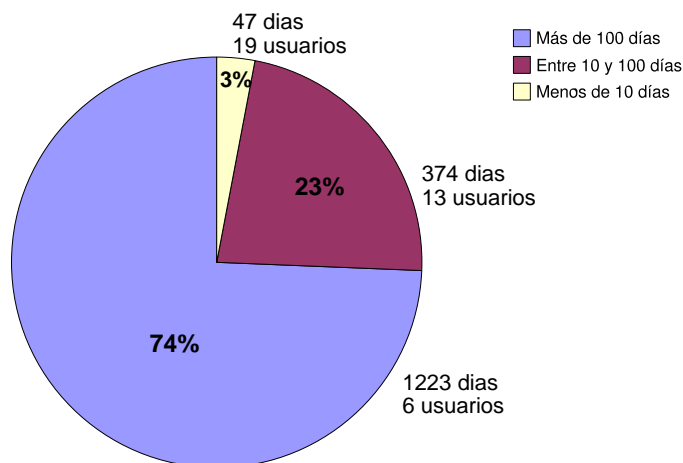
1 Erabilera eta estatistikak

Maiatzean 1264 lan atera ziren eta lan horiek 1736 egun erabilgarrietatik 1645 kontsumitu zituzten. Horrek esan nahi du clusterraren batez besteko okupazio



Irudia 1: Clusterraren okupazioa. Lerro urdinak batez besteko bat adierazten du, 24 orduren gainean. Zona itzaldunak jaiegunak dira.

Tiempo calculado por usuarios



Irudia 2: Denboraren kontsumoaren banaketa erabiltzaileen artean.

maila % 95 izan dela, apirilekoaren antzekoa (% 96). Oso okupazio handia da; esan daiteke makina bete-bete egon dela, makinaren aprobetxamendua erabatekoa izan dela. Ia lan guztiak 1 edo 2 nodoetara bidali dira, baina horregatik ez da clusterra zatitzen eta faktore horrek clusterra ia erabat betetzea ahalbidetzen du. 1. irudiak azaltzen du nola clusterraren okupazioa lehenengo eta azkeneko asteburuetan bakarrik jaitsi zela nabarmen; gainerako egunetan okupazio maila oso handia izan zen. Noizean behin 56 prozesadore baino gehiago egon dira, teknikariok probak egiten aritu garelako.

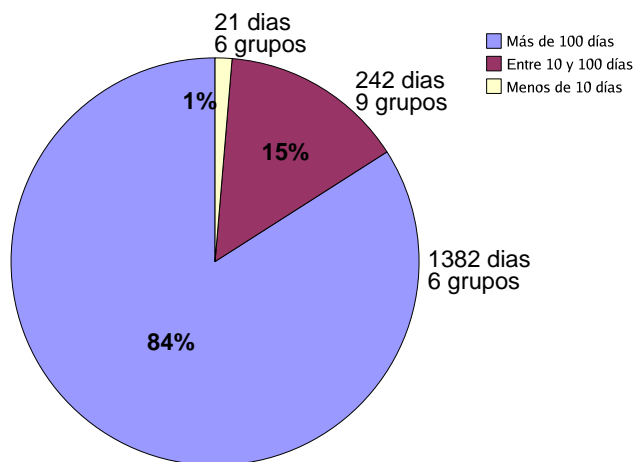
Une hauetan, 96 erabiltzaile daude Arina-n; maiatzetan 2 berri sartu dira. Erabiltzaile aktiboa egun bat baino gehiagoan kalkulatu duena bada, kalkuluak egin dituzten 38 erabiltzaileetatik 32 aktiboak izan dira. 2. irudiak kontsumitutako kalkuluaren denbora nola banatu den azaltzen du. Erabiltzaileen artean, 6k 10 egun baino gehiago eman dituzte kalkulatzeko; 13k 10-100 egun artean; eta gainerako 19k 10 egun baino gutxiago. Gehien kalkulatu duen erabiltzaileak 333 egun baino gehiago eman ditu kalkulatzeko.

Arina-n 32 ikerketa talde daude alta emanda, eta horietatik 18 aktiboak dira. 3. irudiak kalkulurako talde bakoitzak kontsumitu duen denbora azaltzen du. 6 taldek 100 egun baino gehiago kontsumitu dituzte eta 9 taldek 10-100 egun artean. Gehien kalkulatu zuen taldeak 585 egun erabili zituen. A eranskinean talde bakoitzak maiatzetan eguneko kontsumoa zehazten da.

Taldeek emandako denborarekiko kontsumitutakoari dagokionez, 4. taulak adierazten duen legez, zenbait taldek eman zitzaizen denbora baino askoz gutxiago kontsumitu dute. Beste talde batzuk, berriz, (clusterreko denbora erabilgarria dela eta) bere denbora ahitu dute edo ahitzeko zorian daude.

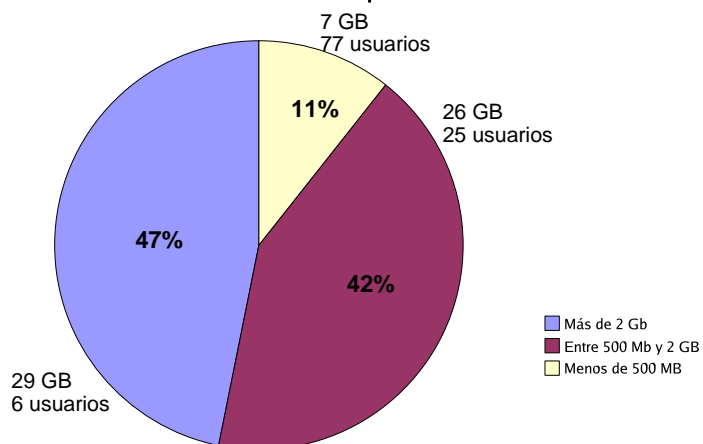
Hala, ikertzaileei eskatu diegu erabilgarri duten beren kalkulu denbora nola dagoen aztertzeko batetik, eta, datozen bi hilabeteetan denbora hori ez dutela kontsumituko pentsatzen badute, ordu batzuk liberatzeko bestetik.

Tiempo consumido por grupos



Irudia 3: Denboraren kontsumoaren banaketa taldeen artean.

Disco usado por usuarios



Irudia 4: Diskoaren erabilera erabiltzaileen partetik.

Taula 1: Arina-n kalkulatu denboraren estatistikak, egunetan (uztailaren 1ean). Lehenengo zutabeak uztailean arte emandako kalkulatu denbora jasotzen du; bigarrenak apirilean eta maiatzean kontsumitutakoa; eta hirugarrenak, berriz, erabil daitekeen denbora. Azkeneko zutabea oraingo kontsumo erritmoa aintzat hartuz emandako denborarekiko ekainaren 30ean izango den desbiderapen bat da (% 0 batek esan nahi du eskatutakoa kontsumitu dela). Zenbakiak egunetan daude.

Taldea	Emandakoa	Kontsumoa	Erabilgarria	Desbiderapena (%)
wenzu	29	87.74	-58	+351.18
qfegp	468	574.77	-106	+83.93
waets	145	170.04	-24	+75.32
waaap	71	65.71	6	+37.12
pojuu	1462	1115.45	347	+14.41
poxlp	562	429.54	132	+14.54
qfla	50	37.83	12	+13.49
qocpn	62	34.47	28	+10.27
qofcm	468	309.74	159	-0.88
infl	41	20.80	20	-25.13
qfjfg [†]	132	62.94	69	-28.75
wuash	4	1.73	2	-37.77
waars	104	32.04	72	-38.46
wdaga [†]	448	156.55	291	-47.61
maaeg	35	12.56	23	-47.05
qpjag [†]	73	24.10	49	-50.65
qijgl	156	31.66	124	-69.61
wepmf	234	35.79	198	-77.09
ccjla	250	28.89	221	-82.67
oijbb [†]	12	0.63	12	-92.54
qfjmf	20	0.75	20	-94.60
qflp	8	0.11	8	-98.02

[†]Orduak bueltatu egin ditu (4320, 5000, 2000 eta 700 ordu hurrenez hurren), bere kodeekin arazoak izan dituelako eta aurreikusitako kalkulatu egiterik izan ez duelako; horiek ekainerako eskatzen dituzten taldeen artean banatuko dira.

4 irudian ikus daiteke diskoaren banaketa erabiltzaileen artean. Une honetan, 63.2 GB erabiltzen dira, disko erabilgarriaren % 25.

2 Arina-ren ilarak

Lehenik eta behin, aipatu behar da erabiltzaile gehienak 1 edo 2 nodo beteak eskatzen dituztela (4 edo 8 prozesadore). Horri esker, hobeto aprobetxatzen da clusterra, nodoei trinko eusten zaielako. 2 nodoekin eginiko lanak gutxienak dira eta, horregatik, denbora asko pasatzen dute ilaran sartu aurretik.

Ilaretara bidalitako lanei dagokienez, apirilean gertatu zen bezala, % 30ak lehenengo 15 minututan amaitu zuen, eta horietako asko ilaretara bidalitako

Ilara	Lan kopurua	Kontsumitutako denbora	Egikaritzeko batez besteko denbora (orduak)	Batez besteko nodoak
vfast	167	25	2.3	1.58
s_fast	207	104	12.0	1.00
s_medium	9	1	2.9	1.00
s_slow	1	1	23.4	1.00
p_fast	391	564	9.9	3.49
p_medium	65	466	45.6	3.77
p_slow	18	440	105.9	5.55
special	1	24	283.7	2.00

Taula 2: Ilaretako estatistikak. Egikaritzeko batez besteko denbora lan bat egikaritzen pasatzen den denbora erreala da.

Taula 3: Koletarako proposatutako denbora tarte berriak. Aldatutakoak **bel-tzez** nabarmentzen dira.

Ilara	Denbora tarteak
vfast	0-4 ordu
s_fast	4-24 ordu
s_medium	1-7 egun
s_slow	7-30 egun
p_fast	0-36 ordu
p_medium	36-96 ordu
p_slow	4-8 egun
Especial	Baldintza berezietako lanak soilik

input direlakoetan sortutako erroreengatik beharbada.

2. taulak ilaretako estatistikak azaltzen ditu. Oker ateratako, berehala amaitzen diren edo dituzten lanak ezabatzeke, ilaretan 15, 30 eta 45 minutu baino gehiago kontsumitu zituzten lanak bakarrik izan ditugu kontuan, hurrenez hurren fast, medium eta slow ilaretan.

Lan gehienak fast ilaretara joaten dira. Dena den, kalkulu denboraren banaketa ilara paraleloen artean nahiko ekitatiboa da; fast ilara izan zen serieko lan gehien hartu zituen. Serieko ilaretan egikaritzen kontsumitzen zen batez besteko denborak aurrez definitutakoak baino dezente laburragoak dira; bien bitartean, ilara paraleloetan lanen iraupena ilaren definizioarekin bat dator.

Datu hauek aditzera ematen dute serieko ilaretako egikaritze denborak berdefinitu egin behar direla, banaketa hobea egon dadin eta serieko ilaretako denbora definitutakoarekin bat etor dadin. Ilara paraleloetan doikuntza txiki bat ere beharrezkoa izan liteke, ilara bakoitzean kontsumitzen den denbora hobeto banatua egon dadin. Honela, azterketa sakon baten ondoren, ilaren denborak zezor aldatzea erabaki dugu, jarraian adierazten den moduan:

4. taulan agertzen dira lanen batez besteko itxarondenborak eta batez beste kontsumitutako denborak. Kontsumitutako denborak kontuan hartzen du lanak pasatu dituen prozesadore kopurua. Daturik esanguratsuen hiru garren zuta-

Ilara	Kontsumitutako batez besteko denbora (orduak)	itxarondenbora (orduak)	denbora kalkulatzeko (%)
vfast	3.7	0.2	94
s_fast	12.0	6.2	66
s_medium	2.9	0.4	89
s_slow	23.3	0.0	100
p_fast	34.6	5.5	86
p_medium	171.9	15.1	92
p_slow	587.3	90.0	87
special	567.4	94.7	86

Taula 4: Ilara bakoitzeko lanen estatistika. Kontsumitutako denbora guztia da ilaran lan bakoitzak batez beste kontsumitutako denbora bider batez beste erabilitako prozesadore kopurua (1. taula). Azkenekoak adierazten du lana kalkulatzeko ari den denboraren ehunekoa; gainerakoa ilaran dago

bekoa da, non islatzen den kontsumitutako denboraren eta itxarondenboraren arteko ratioa. Itxarondenbora pixka bat luzatu egin da aurreko hilekoarekin alderatuz, fast ilaran izan ezik; hor gehiago luzatu da.

Beren kalkuluak egiteko arazoak izaten dituzten erabiltzaileak ditugu, biltegiak handiak behar baitituzte; izan ere, nodo lokaletako diskoek 36GB bakanrik dituztelako.

A Kalkulatzeko denboraren kontsumoa taldeka

Taldea	Kopurua lanak	Denbora guztia (egunak)	# batez bestekoa CPUak	Batez beste ilara (orduak)
TOTALA	1272	1645.11	3.26	0.23
pojuu	364	584.69	3.43	0.22
qfegp	8	284.30	7.05	6.10
qofcm	44	169.57	3.60	0.46
poxlp	93	134.11	3.89	0.31
wdaga	223	105.71	1.27	0.22
waets	26	103.44	3.99	0.16
wenzu	27	84.93	4.00	0.39
qocpn	132	34.47	3.76	0.17
waars	4	32.04	4.00	1.41
arinadm	79	22.99	2.49	0.08
qijgl	19	20.16	3.99	0.07
inffi	14	12.83	1.00	0.04
maaeg	56	12.56	3.95	0.10
wepmf	57	11.59	1.00	0.01
qpjag	68	10.75	1.00	0.09
waaap	2	9.71	4.00	0.21
ccjla	22	8.28	1.00	0.03
qfjfg	4	2.21	4.00	0.28
oijbb	5	0.58	1.00	0.06
qflp	23	0.11	1.00	0.01
qfjmf	2	0.07	1.00	0.55