

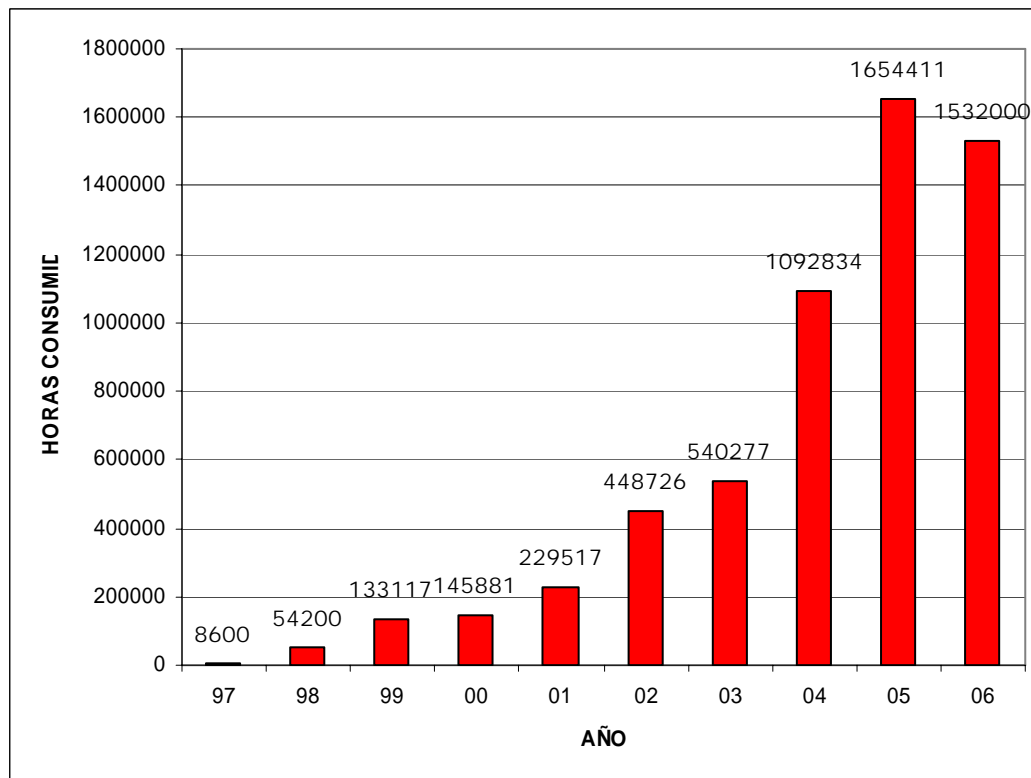
1.1 Servidores de cálculo

La fundación CESGA cuenta con diferentes arquitecturas de sistemas de cálculo de altas prestaciones disponibles para la comunidad de usuarios. Dependiendo de los algoritmos de cálculo, el usuario opta por la arquitectura de computación más adecuada.

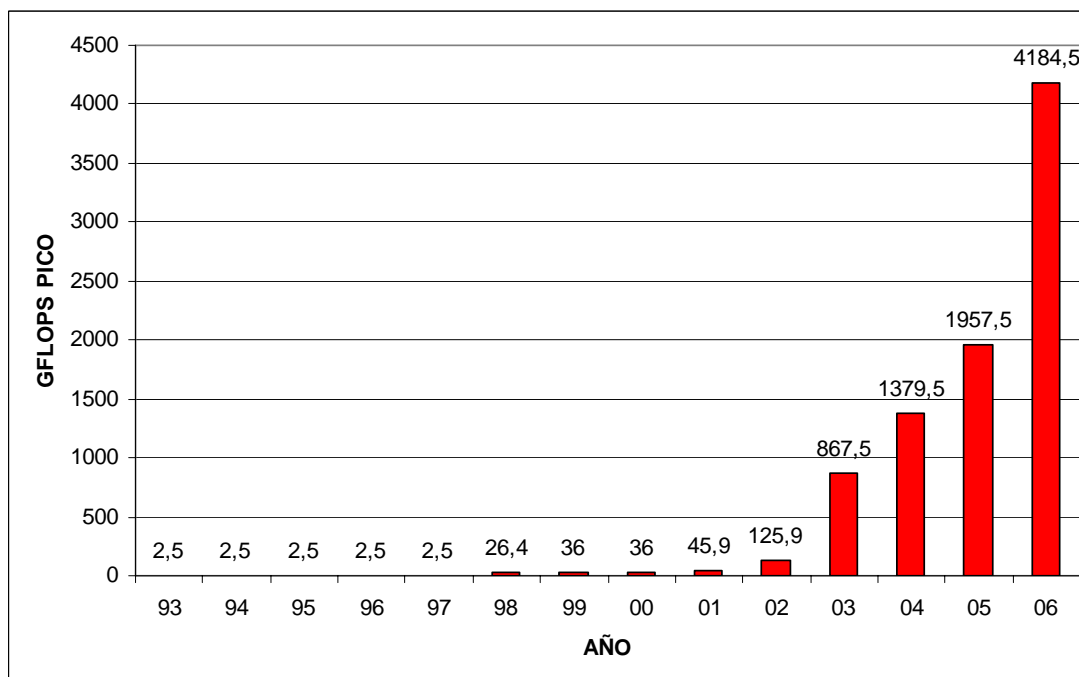
Los hechos más relevantes del año 2006 e cuanto a servidores de cálculo fueron:

- Firma del convenio FINISTERRAE
- Adquisición cluster grid a finales del mes de Diciembre

En los gráficos siguientes se reflejan los datos más significativos correspondientes a la actividad de cálculo en el CESGA:



EVOLUCIÓN DE LA CAPACIDAD DE CÁLCULO INSTALADA EN EL CESGA



En el año 2006 el CESGA no llevó a cabo nuevas adquisiciones de equipamiento de cálculo hasta finales de Diciembre en que se recepcionó un nuevo cluster de computación grid formado por 36 servidores DELL Poweredge 1955 y 1 servidor DELL Poweredge 2950 con un total de 1843 GFLOPS. Este sistema no entró en producción hasta el mes de Enero del 2007.

Evolución Consumo de CPU por Sistema

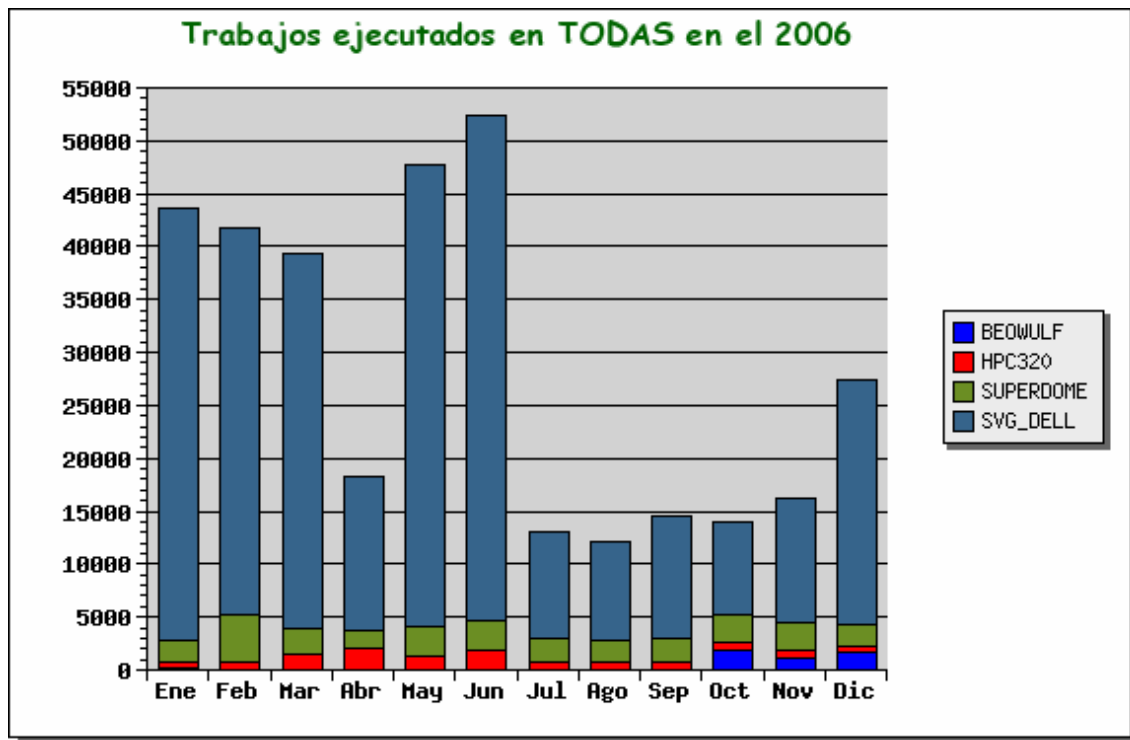


	Ene.	Feb.	Mar.	Abr.	May.	Jun.	Jul.	Ago.	Sep.	Oct.	Nov.	Dic.	Total
CESGA	3611:36	3537:06	4149:36	3870:23	3801:02	6195:21	4030:42	8854:11	3955:51	4603:42	5525:59	16880:34	69016:08
CSIC	14639:10	24147:42	13785:35	11497:47	13838:10	8431:45	14218:36	13447:00	10751:12	31879:41	30874:19	22886:28	210397:30
UDC	52570:34	39246:28	35224:48	26774:40	43111:37	34321:34	27925:36	32305:06	30322:28	25511:00	21113:32	36876:44	405304:13
USC	45897:35	40811:27	53966:21	57329:20	65712:48	55379:43	64559:54	70084:29	70411:59	56073:11	55756:57	57669:03	693652:52
UVI	15748:02	8560:29	10059:03	5645:27	8320:07	8325:22	4268:30	7782:10	9488:16	31273:15	30570:14	13889:17	153930:18
TOTAL	132466:59	116303:15	117185:25	105117:39	134783:47	112653:47	115003:21	132472:59	124929:48	149340:52	143841:03	148202:08	1532301:03

Durante el año 2006 los servidores más utilizados han sido el cluster SVG y el servidor HP Integrity Superdome. Durante este año no se realizó ninguna adquisición adicional (exceptuando la ampliación del cluster SVG con nodos grid pero que no se puso en producción en este año) mostrando todos los servidores unos niveles de ocupación cercanos al 100%. Diversas incidencias en el servidor de cálculo Superdome así como la priorización de trabajos paralelos (cuya eficiencia es peor que la de los secuenciales) han provocado una disminución en el número de horas consumidas por los investigadores desde 1,654,411 horas en el año 2005 hasta 1,532,301, lo que representa una disminución del 7.4%. Durante todo el año el servidor HPC320, instalado en el año 2002, también continuó con un nivel alto de utilización (entorno al 75%)

Número de Trabajos Ejecutados por Sistema

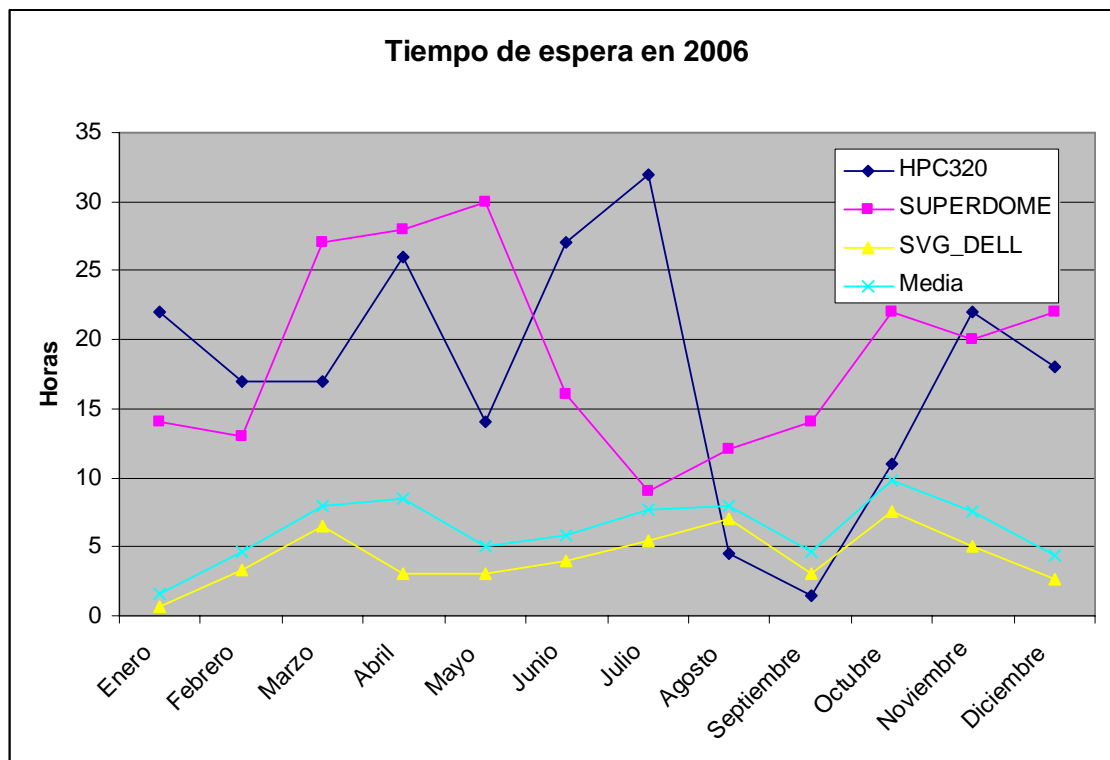
NUMERO DE TRABAJOS EJECUTADOS POR SISTEMA EN 2006



	Ene.	Feb.	Mar.	Abr.	May.	Jun.	Jul.	Ago.	Sep.	Oct.	Nov.	Dic.	Total
BEOWULF	126	64	33	21	14	11	25	8	37	1788	1153	1626	4906
HPC320	597	774	1448	1959	1248	1931	799	824	728	737	660	622	12327
SUPERDOME	2164	4305	2367	1827	2851	2660	2199	1978	2307	2670	2586	2056	29970
SVG_DELL	40769	36571	35474	14375	43655	47778	9998	9346	11438	8797	11890	23181	293272
Total	43656	41714	39322	18182	47768	52380	13021	12156	14510	13992	16289	27485	340475

El número de trabajos ejecutados por sistema representa la cantidad de simulaciones que los usuarios han realizado en cada uno de los servidores de cálculo. Este valor depende no sólo de la capacidad de cálculo disponible sino también de los recursos necesarios para la ejecución de las simulaciones. Los sistemas con mayor número de trabajos han sido el SVG y el Superdome. Especialmente el primero, debido a que es un servidor orientado a productividad, es decir, a la realización de un gran número de trabajos independientes con necesidades de cálculo relativamente bajas. Respecto al año 2005 se produjo una reducción del 2.4% en el número de trabajos al pasarse de 348,559 a 340,475

TIEMPO MEDIO DE ESPERA PARA PROCESO POR SERVIDORES EN 2006



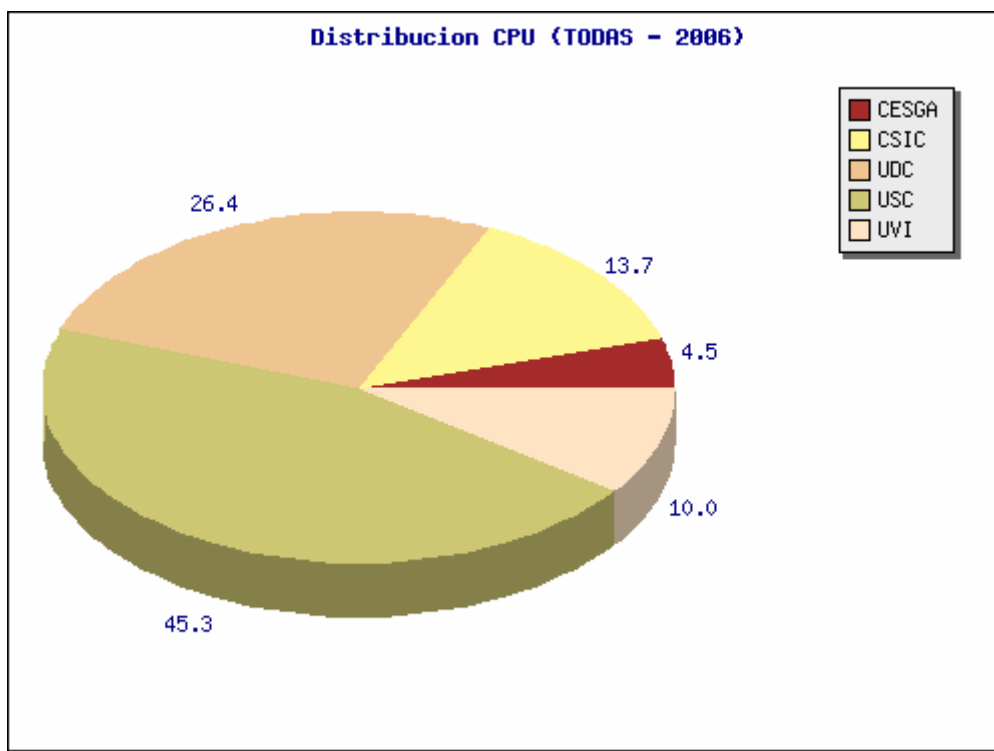
	Ene.	Feb.	Mar.	Abr.	May.	Jun.	Jul.	Ago.	Sep.	Oct.	Nov.	Dic.	Media
HPC320	0d-22:41:52	0d-16:50:30	0d-17:04:47	1d-02:42:30	0d-14:14:47	1d-03:08:41	1d-08:06:20	0d-04:30:18	0d-01:38:04	0d-11:49:37	0d-22:13:31	0d-18:23:56	0d-19:24:22
SUPERDOME	0d-14:14:54	0d-13:03:42	1d-03:30:01	1d-04:27:34	1d-05:59:43	0d-16:38:48	0d-08:54:38	0d-12:45:17	0d-14:47:22	0d-22:13:55	0d-20:05:50	0d-21:51:17	0d-18:59:27
SVG_DELL	0d-00:44:12	0d-03:22:00	0d-06:30:03	0d-03:28:44	0d-03:02:30	0d-04:22:12	0d-05:30:47	0d-07:22:32	0d-02:51:01	0d-07:46:20	0d-05:01:03	0d-02:46:06	0d-03:50:00
Media	0d-01:42:24	0d-04:38:40	0d-08:10:17	0d-08:31:02	0d-04:57:13	0d-05:51:35	0d-07:46:58	0d-08:14:26	0d-04:40:49	0d-09:52:25	0d-07:50:12	0d-04:25:07	0d-05:42:42

Tiempo Medio de Espera en cola por Sistema

Los tiempos medios de espera en cola representan el tiempo que por término medio deben esperar las simulaciones de los usuarios desde que solicitan los recursos del CESGA hasta que sus simulaciones comienzan a utilizarlos. Estos tiempos varían en función de la cantidad de simulaciones que se están realizando e idealmente debería ser lo más próximo a cero para evitar las esperas hasta obtener los resultados de las simulaciones. Sin embargo, cuanto más alto es el nivel de ocupación de los sistemas de cálculo, es necesario esperar más cantidad de tiempo hasta que existan los recursos suficientes. Por tanto, esta medida también es un buen indicativo de cuál es el nivel de saturación existente en los recursos de computación. Normalmente los tiempos de espera disminuyen en los períodos estivales de verano y navideños y aumentan significativamente a medida que se van incorporando nuevos usuarios. En la gráfica se aprecia que los sistemas con mayores tiempos de espera son el Superdome (por encima de las 20 horas) y el HPC320, con menores tiempos de espera durante los meses de Julio y Agosto. Mientras que el tiempo medio de espera bajó en el Superdome de 25 a 19 horas, en el HPC320 subió de 9 a 19 horas y en el SVG no varió. Los tiempos de espera del Beowulf (utilizado para investigación en paralelización) se debieron a incidencias en las tarjetas de red Mirynet que impidieron la realización de algunas simulaciones que requerían el 100% de los nodos disponibles.

Distribución de CPU por Institución

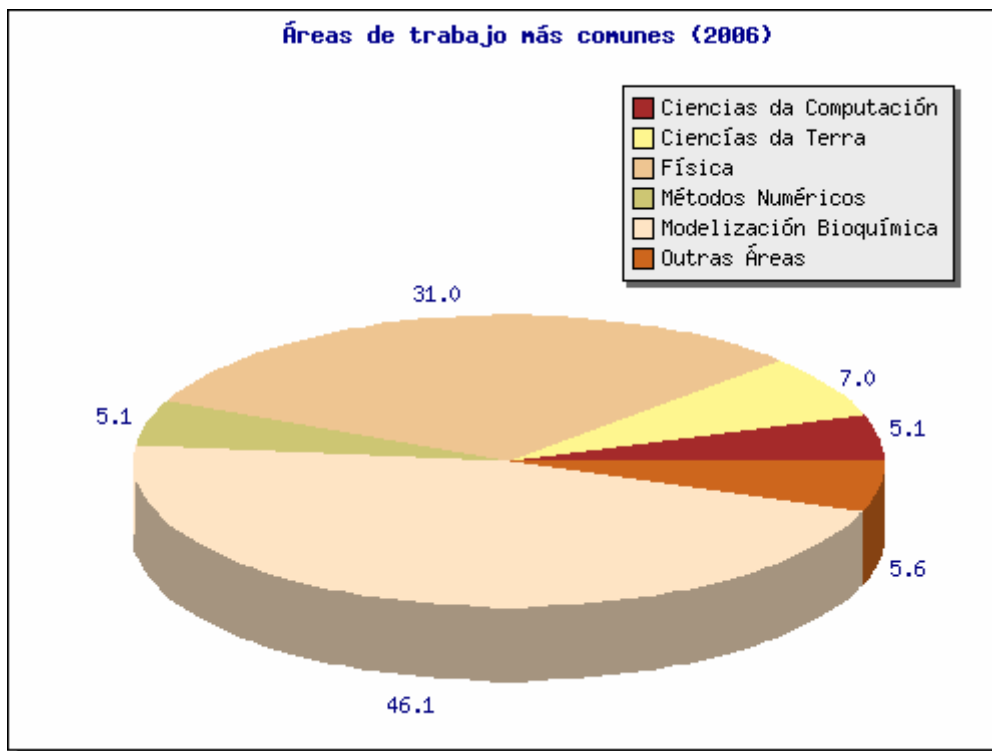
DISTRIBUCIÓN POR INSTITUCIONES DE LAS HORAS DE CPU CONSUMIDAS 2006



Institución	Consumo	Consumo (%)
CESGA	69204:50	4.5
CSIC	210058:50	13.7
UDC	404278:53	26.4
USC	694789:23	45.3
UVI	153969:6	10.0
TOTAL	1532301:3	100

Por instituciones, el mayor consumo de horas de computación lo registró en el año 2006 la Universidad de Santiago de Compostela con un 45.3% de las horas, seguida de la Universidad de A Coruña con un 26.4% del consumo total. En su conjunto, las tres universidades gallegas representan el 81.7% del consumo, el CSIC el 13.7% de las horas consumidas y un 4.5% de las horas se han destinado a proyectos participados por el CESGA. Respecto al año 2005, en el 2006 todas las instituciones a excepción de la Universidad de Santiago y de Vigo han incrementado su consumo,

DISTRIBUCION DE CONSUMO DE CPU POR AREAS DE TRABAJO

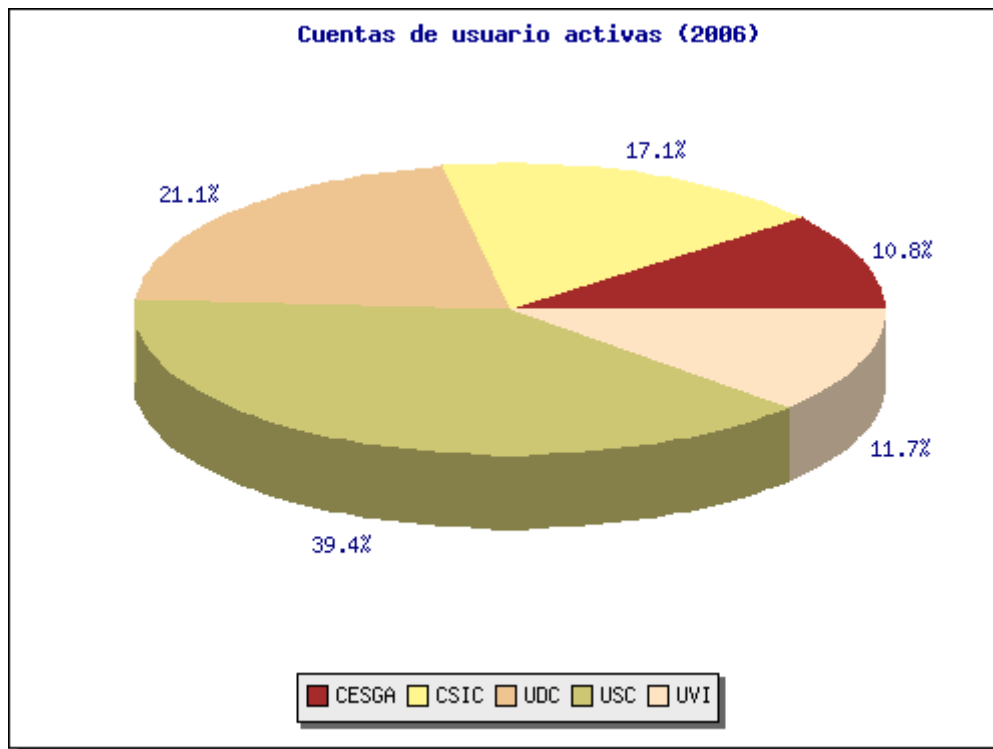


Área de trabajo	Consumo	Porcentaje
Ciencias da Computación	78068:24	5.1
Ciencias da Terra	106638:55	7.0
Física	471836:46	31.0
Métodos Numéricos	77680:16	5.1
Modelización Bioquímica	701666:10	46.1
Otras Áreas	85080:33	5.6
Total	1520971:4	100.0

Distribución de CPU por Área de Trabajo

Por áreas de trabajo, los cálculos relacionados con la modelización bioquímica consumieron en el año 2006 el 46.1% de las horas de cálculo, mientras que el 31% se dedicaron al estudio de la física, representando estas dos áreas el 77.1% del consumo de horas. La modelización bioquímica vuelve a ser, por tanto, el área de mayor demanda computacional, como sucedió en el año 2005. La modelización bioquímica ha pasado del 50.4% de utilización al 46.1%, mostrando una ligera disminución como ya sucedió el año anterior, frente a las otras áreas que han aumentado respecto al año 2005. Así, por ejemplo, se ha duplicado el porcentaje de utilización en el área de métodos numéricos

EVOLUCION DE CUENTAS ACTIVAS POR SISTEMA Y AÑO



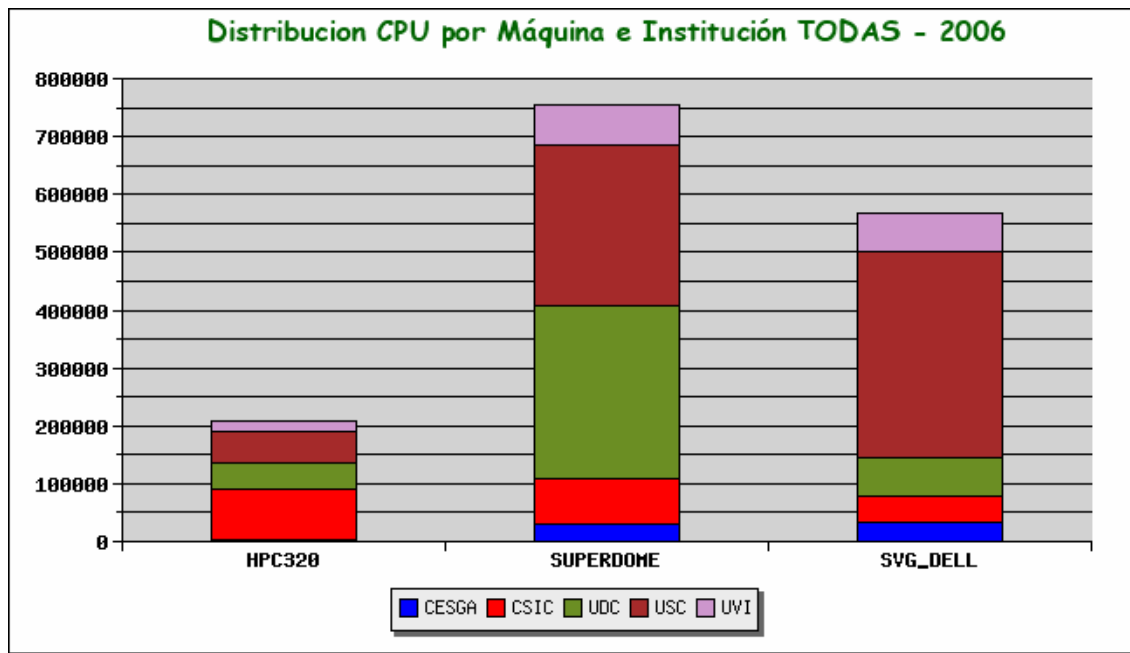
#	CESGA	CSIC	UDC	USC	UVI	Total
HPC320	17	41	35	76	20	189
SUPERDOME	21	41	54	105	35	256
SVG_DELL	22	16	32	45	12	127
Total	60	98	121	226	67	572

Número de Cuentas Activas por Sistema

El sistema con mayor número de cuentas activas (es decir, usuarios con un consumo significativo de horas de cálculo a lo largo del año) es el HP Superdome, con 256 usuarios con consumo (34 más que en el año anterior). El HPC320 ocupa la segunda posición con 189 cuentas activas y por último 127 del cluster SVG. En total, el número de cuentas activas ha crecido desde las 564 del año 2005 a 572 este año, lo que supone un crecimiento tan sólo del 2% (en el año 2005 el incremento en el número de cuentas activas fue del 10%), quizás motivado por la saturación de los sistemas actuales.

Porcentaje CPU por Institución y Máquina

PORCENTAJE DE CPU POR INSTITUCION Y MAQUINA



	HPC320	SUPERDOME	SVG_DELL
CESGA	1.8	4.1	6.1
CSIC	42.2	10.3	7.8
UDC	20.9	39.4	11.4
USC	26.9	36.8	63.2
UVI	8.2	9.4	11.6

En esta gráfica podemos ver cuáles son los sistemas más demandados por cada una de las instituciones que utilizan los servicios de cálculo del CESGA. Como se puede apreciar, los investigadores del CSIC utilizan principalmente el servidor HPC320, con más de un tercio de la utilización del sistema, mientras que el Superdome se reparte principalmente entre los investigadores de las universidades de Santiago y de Coruña y el SVG es utilizado mayoritariamente por los investigadores de la Universidad de Santiago.