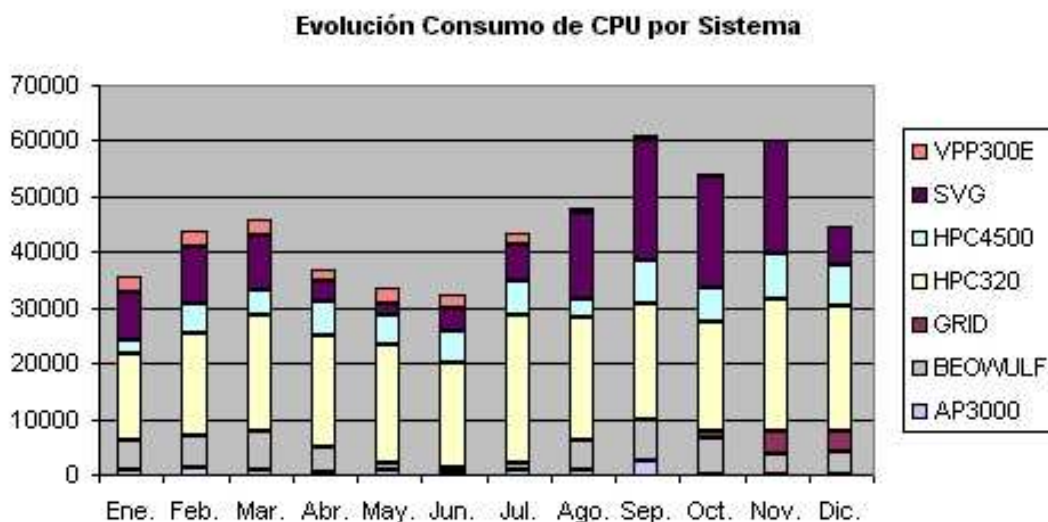
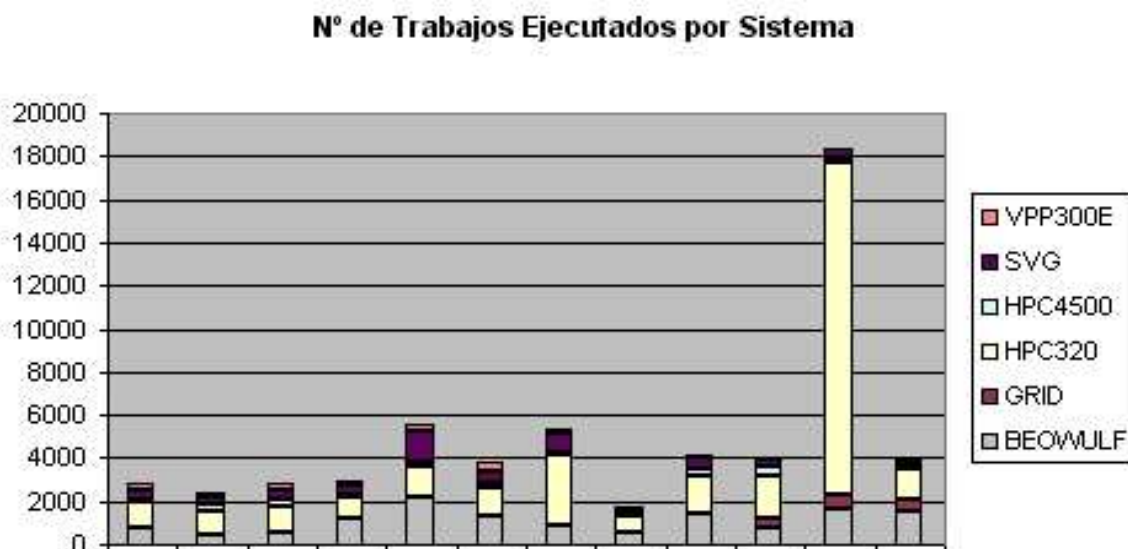


Centro de Supercomputación de Galicia

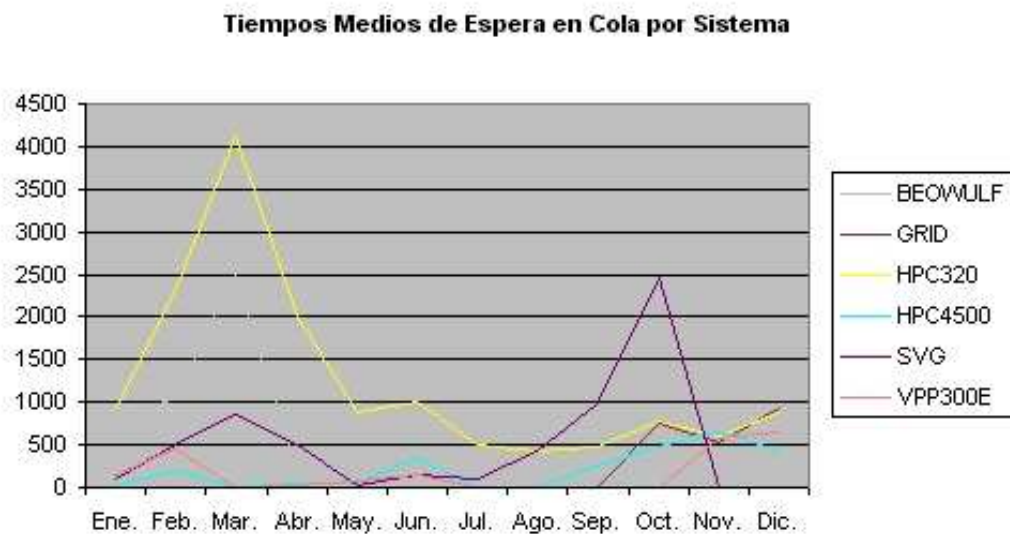
Estadísticas 2003



A lo largo del año 2003 se han consolidado los consumos de horas en el HPC320, con un nivel de utilización cercano al 100% de su capacidad a lo largo de todo el año y en los cluster linux Beowulf y SVG, a los que se han añadido otros nodos (que denominados nodos de computación GRID) para apoyar a estos clusters en momentos de alta demanda de recursos. Al mismo tiempo, disminuyó muy notablemente el consumo de horas de CPU en los servidores de cálculo paralelo Fujitsu AP3000 y el servidor vectorial VPP300, debido fundamentalmente al bajo rendimiento que estos sistemas, los más veteranos del CESGA, ofrecían en comparación con las más recientes incorporaciones. Finalmente, estos servidores dejaron de prestar servicio en el último mes del 2003, para ser reemplazado por el HP Integrity Superdome (que entra en producción en enero del 2004). El número total de horas de CPU consumidas por los usuarios del CESGA en el 2003 aumentó hasta 540.000 desde las 437.000 horas del año 2002, lo que representa un incremento superior al 20%.

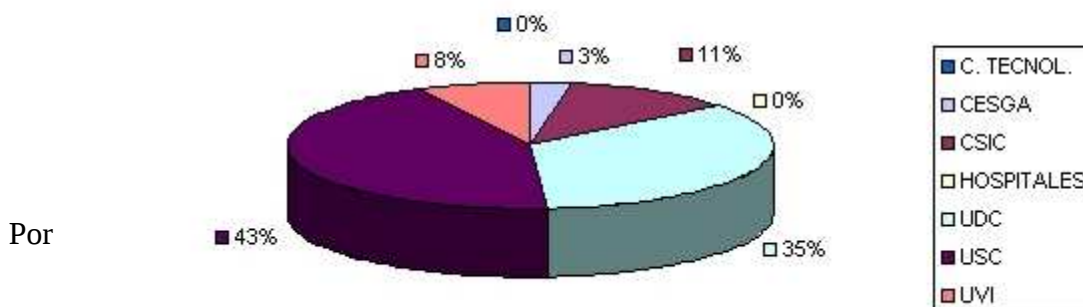


El número de trabajos ejecutados por sistema representa la cantidad de simulaciones que los usuarios han realizado en cada uno de los servidores de cálculo. Debido al carácter altamente variable en cuanto a consumo de recursos como es el tiempo de cálculo, el número de simulaciones que se realizan depende fuertemente de la cantidad de recursos que solicitan los usuarios del centro. En general, el promedio ha aumentado al aumentar el número de recursos disponibles, disminuyendo progresivamente la cantidad de trabajos ejecutados en los sistemas VPP300 y AP3000, mientras que aumentaban el número de trabajos en los sistemas más potentes, especialmente el HPC320. En el mes Noviembre, la simulación de procesos en astrofísica en el servidor HPC320 elevó el número de trabajos en este sistema de forma considerable al ser estos de una duración relativamente corta.



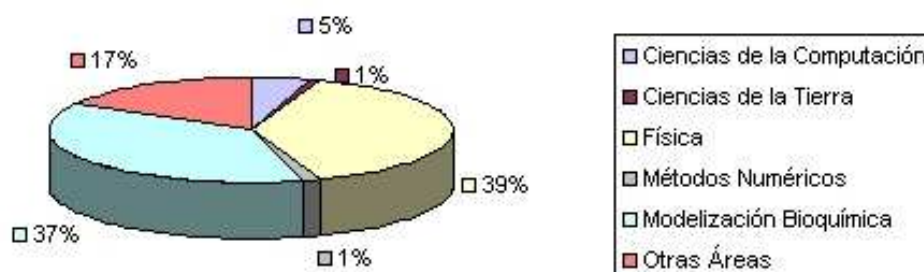
Los tiempos medios de espera en cola representan, de media, el tiempo que los usuarios tiene que esperar desde que solicitan los recursos del CESGA hasta que sus simulaciones comienzan a utilizarlos. Estos tiempos varían en función de la cantidad de simulaciones que se estén realizando e idealmente debería ser próximo a cero para evitar demoras. Sin embargo, el alto nivel de ocupación de los sistemas de cálculo provoca que las simulaciones deban esperar por orden de llegada su turno hasta que existan recursos suficientes. Por tanto, a mayor tiempo de espera en cola, mayor nivel de saturación presentan los recursos de computación. Como se observa en la gráfica, estos tiempos de esperas disminuyen en los períodos de vacaciones de verano y de navidad (sin llegar a ser nulos generalmente) y aumentan significativamente a medida que se van incorporando nuevos usuarios. También es necesario tener en cuenta que este tiempo de espera no aumenta todavía más debido a que los propios investigadores deciden no utilizar los recursos cuando los tiempos de espera en cola aumenta.

Distribución de CPU por Institución



instituciones, el mayor consumo de horas lo registró en el año 2003 la Universidad de Santiago con un 43% de las horas, seguida de la Universidad de A Coruña con un 35% del consumo total. En su conjunto, las tres universidades gallegas representan el 86% del consumo total de horas del centro, el CSIC el 11% de las horas consumidas y sólo un 3% de las horas se han destinado a proyectos participados por el CESGA. El principal cambio respecto al año pasado se produce al pasar la Universidad de Santiago a ser la institución con mayor consumo, frente al año 2002 en que fue la Universidad de A Coruña.

Distribución de CPU por Área de Trabajo



Por áreas de trabajo, los cálculos relacionados con el estudio de la física consumieron en el año 2003 el 39% de las horas de cálculo, mientras que el 37% de las horas se dedicaron a investigaciones relacionadas con la modelización bioquímica, representando estas dos áreas el 76% del consumo de horas. Frente al año 2002, en el cual la modelización bioquímica ocupó la primera posición con un 36% de las horas de cálculo, este año los cálculos relacionados con la física han adquirido mayor protagonismo al pasar del 24% de las horas consumidas en el año 2002 al 39% en el año 2003.

En el número de cuentas activas, que representa aquellas cuentas de usuario con un consumo significativo de horas de cálculo a lo largo del año, destaca el sistema HPC320 con el mayor número de cuentas activas, 154, debido a que también es el sistema con mayores prestaciones del centro en el año 2003. En el año 2002 fue el servidor vectorial Fujitsu VPP300 el que mayor

número de cuentas activas tenía con 106. El número de cuentas activas ha disminuído respecto al año anterior en los sistemas que están cercanos a su período de amortización y que por tanto ofrecen las menores prestaciones del centro: el AP3000, el VPP300 y el HPC4500, y han aumentado en los clusters Linux Beowulf (de 17 a 51, triplicando el número de cuentas), SVG (de 31 a 45, 45% de incremento) y también en el servidor HPC320 (de 96 a 154, 60% de incremento). Globalmente, el número total de cuentas activas ha crecido desde las 397 del año 2002 a 446 en el año 2003, lo que representa un crecimiento del 12%, a pesar de que han dejado de prestar servicio 2 servidores de cálculo.