

# Eficiencia energética en un centro de supercomputación

### Energy efficiency in a supercomputing centre

◆ J. López Cacheiro, C. Fernández Sánchez, J.I. López, P. Rey, A. Feijoo, S. Díaz

#### Resumen

La eficiencia energética es un aspecto cada vez más importante dentro de un centro de supercomputación ya que los consumos eléctricos de los equipos (servidores, almacenamiento y red) se han disparado. Esto se pone de manifiesto en el hecho de que en la actualidad un centro de supercomputación como el CESGA tiene un consumo medio de unos 600KW, lo que implica un elevado coste en el suministro eléctrico. Con consumos tan elevados resulta muy importante evaluar mecanismos para mejorar la eficiencia energética ya que una reducción en el consumo o un mejor aprovechamiento de los recursos supone un gran beneficio tanto económico como para el medio ambiente. En este artículo se muestran cinco mecanismos de mejora que se han aplicado para mejorar la eficiencia energética en nuestro centro. Muchas de estas medidas pueden ser aplicadas a otros centros mientras que otras son más específicas y no servirán para todos los casos.

**Palabras clave:** eficiencia energética, consumo, medio ambiente, recursos.

#### Summary

Energy efficiency is becoming an increasingly important aspect within supercomputing centres as the power consumption levels of equipment (servers, storage and network) have shot up. This is demonstrated by the fact that, today, a supercomputing centre such as CESGA has an average consumption of around 600 kW, which means high energy costs. With such high consumption levels it is extremely important to evaluate mechanisms to improve energy efficiency, since a reduction in consumption, or better use of resources, would result in both economic and environmental benefits. This article outlines five improvement mechanisms that have been applied to improve energy efficiency in our centre. Many of these measures can be applied to other centres while others are more specific and will not be of use in all cases.

**Keywords:** energy efficiency, consumption, environment, resources.

## 1. Introducción

El CESGA cuenta actualmente con un CPD de 300 m<sup>2</sup> que dispone de 2 transformadores capaces de suministrar más de 1.5 MW y está respaldado con 4 SAIs con una capacidad total de 980 KVA y un grupo electrógeno de 1125 KVA. El mecanismo utilizado para refrigerar el CPD es mediante dos enfriadoras, con una capacidad refrigeradora de 580 KW cada una, que suministran agua fría a 8 climatizadoras interiores que a su vez enfrían el aire de la sala. El consumo medio de toda la infraestructura de supercomputación y de soporte ronda los 600KW por lo que resulta muy importante un correcto análisis de la eficiencia energética de los distintos componentes.

El hecho de que la eficiencia energética es uno de los temas de moda en el mundo de la supercomputación se ve claramente reflejado en la creación de la lista del Green500[1] una lista análoga al Top500[2], pero en donde se tiene en cuenta no sólo la potencia de cálculo sino también la eficiencia energética de los superordenadores.

Los superordenadores requieren de condiciones específicas de temperatura y humedad para operar que obligan a los centros de supercomputación a instalar costosos sistemas de climatización para garantizar su funcionamiento. Estos sistemas además consumen cada año una parte muy importante de la energía



El consumo medio de un centro de supercomputación como el CESGA es de 600 KW



Dado el elevado coste del suministro eléctrico, es muy importante un correcto análisis de la eficiencia energética de los distintos componentes



La ASHRAE aprobó cambios en sus recomendaciones para la temperatura en los CPD, pasando de los 25° a los 27°

Los procesadores Montvale incorporan mecanismos de ahorro energético como escalado de frecuencia y voltaje

eléctrica total del centro. Sin embargo, cabe destacar que los rangos de temperaturas en los que pueden operar los servidores han mejorado notablemente en los últimos años, por este motivo a principios de este año la American Society for Heating, Refrigerating and Air-conditioning Engineers (ASHRAE) aprobó cambios en sus recomendaciones para la temperatura en los CPD, pasando su recomendación de los 25°C a los 27°C, aunque alertan de las consecuencias que este aumento puede provocar en el aumento de ruido y en la aparición de puntos calientes. Este pequeño cambio consistente en aumentar la temperatura del CPD de 25 a 27°C puede suponer un importante ahorro energético y económico a final de año. Eso sí, siempre es importante a la hora de aplicar este tipo de cambios evaluar cuidadosamente sus consecuencias, tratando de prevenir posibles problemas.

En este artículo nos centraremos en exponer algunos de los mecanismos de mejora de la eficiencia energética que hemos evaluado en el CESGA, compartiendo nuestras experiencias para que puedan ser de utilidad a otros centros. De cara a organizar el contenido hemos decidido centrarnos en cinco de los aspectos que consideramos más interesantes, y dedicar una sección a cada uno de ellos. En la última sección incluimos las principales conclusiones del estudio.

## 2. Mecanismos de ahorro energético de los servidores

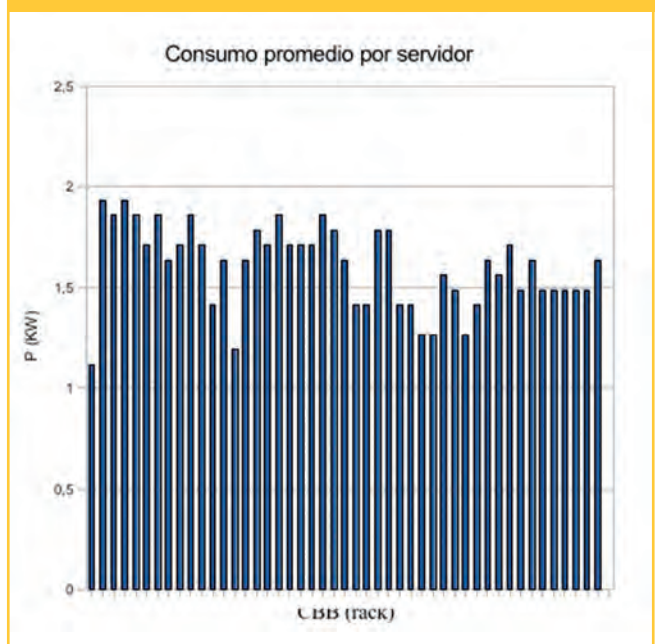
La mayor parte del CPD del CESGA está ocupado por el nuevo superordenador Finis Terrae (FT) que está formado por 142 nodos HP RX7640 y 3 HP Superdome, todos ellos con procesadores Itanium 2.

En cuanto a consumo eléctrico, los procesadores Itanium son conocidos por sus altos requerimientos de energía. En nuestro caso todos los RX7640 y el nuevo Superdome cuentan con procesadores Itanium 2 Montvale 9140N. Por lo tanto este será el procesador en el que centremos nuestro análisis. Estos procesadores emplean tecnología de 90 nm y tienen un TDP de 104 W. Los otros 2 Superdome cuentan con procesadores Itanium 2 más antiguos del modelo 2 Madison que emplean tecnología de 130 nm y tienen un TDP de 130W.

Los procesadores Montvale incorporan mecanismos de ahorro energético como escalado de frecuencia y voltaje (denominado tecnología Foxton) que permiten que tengan un consumo inferior a los Madison a pesar de tener dos cores en lugar de uno.

Para ver lo eficaces que son estos mecanismos de ahorro hemos realizado mediciones para comparar el consumo de los distintos servidores RX7640. Para ello se ha medido el consumo de cada rack y se ha tenido en cuenta que en cada uno de ellos hay 3 servidores para calcular el consumo promedio por servidor. Los resultados se presentan en la **figura 1**. Debe tenerse en cuenta que en estas mediciones influyen no sólo el

FIGURA 1. CONSUMO PROMEDIO POR SERVIDOR CALCULADO A PARTIR DEL CONSUMO TOTAL POR RACK



consumo de los procesadores sino también el consumo de otros componentes que puede variar con la carga del sistema como discos, memoria, etc. En las medidas también influye la gestión que el SO haga de los mecanismos de ahorro energético disponibles. En nuestro caso todos los nodos cuentan con la distribución Linux SLES10 SP1.

El consumo más bajo que se observa en el rack CBB1 con un consumo promedio por servidor de 1,12 KW. Esto se debe a que en el momento de la medición los 3 servidores no estaban siendo utilizados, presentando un uso de CPU inferior al 1%. En el otro extremo tenemos racks donde la potencia promedio por servidor llega a 1,93 KW. Por tanto, en caso de que el servidor no esté siendo utilizado, el consumo se reduce en 0,81 KW con respecto al consumo a carga completa. Suponiendo una situación ficticia en la que todos los nodos RX7640 no estén siendo utilizados esto implica una reducción de 115 KW, y en una situación de operación normal el ahorro promedio es de 43 KW.

### 3. Consolidación de servicios

En el clúster SVGD contamos con 80 nodos Dell PE750 monoprocesador con CPU Pentium IV Prescott de 3.2GHz fabricados con tecnología de 90nm y con un TDP de 84W. A la vez también contamos con varios chasis de blades. Uno de los más modernos es un chasis HP BladeSystem c-Class que cuenta actualmente con 8 blades con procesadores Xeon quad-core X5355 a 2.66GHz. Una característica muy interesante de los nuevos chasis de blades es que permiten monitorizar el consumo a través de una página web con lo que no es necesario realizar mediciones individuales.

Se han realizado medidas sobre ambos tipos de nodos de la potencia eléctrica consumida por core y se muestran los resultados obtenidos en la **tabla 1**. A la vista de estos resultados se puede ver que la diferencia de potencia eléctrica consumida es de aproximadamente 100W. Si consideramos el precio que nos cuesta actualmente la electricidad (0,09€/KWh) esto supone un ahorro anual de 631€ por cada blade (8 cores), en nuestro caso que contamos con 80 nodos el ahorro promedio anual sería de 6.310€.

Este es un ejemplo de cómo la consolidación de servidores que utilizan procesadores de generaciones anteriores usando nuevos modelos permite ahorrar energía eléctrica y, en este caso, también espacio. También demuestra el esfuerzo que están realizando los fabricantes para mejorar no sólo la capacidad de procesamiento de los nuevos equipos, sino también el ratio de eficiencia eléctrica.

Los nuevos chasis de blades permiten monitorizar el consumo a través de una página web

La diferencia de potencia eléctrica consumida es de 100 W, lo que supone un ahorro anual de 631 euros por cada blade

TABLA 1. CONSUMO COMPARATIVO POR CORE ENTRE DISTINTOS NODOS DEL CLÚSTER SVGDTOTAL

Tipo de nodo	CPU	Potencia promedio por core (W)
Dell PE750	Pentium IV Prescott 3.2 GHz (90nm)	130 ± 40
ProLiant BL460c G1*	Intel Xeon X5355 @ 2.66 GHz (65nm)	27 ± 5

\* Consumo promedio obtenido a partir del consumo total del chasis que cuenta con 8 blades



## 4. Almacenamiento

Entre otros el CESGA cuenta con un sistema de almacenamiento de HP paralelo basado en Lustre. Este sistema está formado por 20 servidores HP Proliant DL380 G5 y 76 cabinas de discos SF520, cada una de ellas con 12 discos SATA Seagate Barracuda ST3250620NS de 250GB cada uno. La capacidad útil de almacenamiento del sistema son 142 TB.

Este sistema de almacenamiento consume 19KW, lo cual corresponde a un 7% del consumo del FT, o lo que es lo mismo 12 nodos RX7640.

Si tenemos en cuenta los datos de la **tabla 2**, podemos ver que el consumo máximo de cada disco son 12,60W lo cual supone un consumo máximo teórico de 11,5 KW. Este consumo puede bajar hasta los 0,7 KW si el sistema pasa los discos al modo reposo.

En este punto resulta interesante comparar el coste comparativo que tendría almacenar backups en cinta con respecto al que supondría tenerlos almacenados en disco en el SF5. Para simplificar, en el caso del sistema de backup en disco tendremos en cuenta sólo el consumo de los discos y en el caso del sistema de backup a cinta despreciaremos el consumo de las librerías de cinta. Suponiendo que los discos estén en modo reposo, el consumo durante un año supone un coste de 575€. En general, sólo este coste no justifica la implantación de un sistema de almacenamiento jerárquico (HSM).

El sistema de almacenamiento de HP consume 19KW, lo cual corresponde a un 7% del consumo del FT

TABLA 2. CONSUMO ELÉCTRICO DE UN DISCO SEAGATE BARRACUDA ST3250620NS SEGÚN LOS DATOS FACILITADOS POR EL FABRICANTE

Estado	Potencia (W)
<i>Average idle current</i>	9,30
<i>Average seek power</i>	12,60
<i>Standby power</i>	0,80

El sistema de climatización (HVAC) es el lugar donde se presentan la mayor parte de las oportunidades de mejora de la eficiencia energética

## 5. Free Cooling y Recuperación de Calor

El sistema de climatización (HVAC) es el lugar donde se presentan la mayor parte de las oportunidades de mejora de la eficiencia energética. La eficiencia del mismo es la que determina en último lugar el grado de eficiencia global del CPD, la cual se suele medir a través de la métrica denominada PUE (*Power Usage Effectiveness*):

$$PUE = \text{Potencia Total} / \text{Potencia Equipos IT}$$

Actualmente el PUE del CESGA es ligeramente inferior a 2. Lo ideal sería que el PUE estuviese lo más próximo a 1 posible. Compañías como Google han conseguido que el PUE promedio de sus CPDs sea de 1.22, llegando en algunos de ellos a valores de 1.15.

En esta sección presentamos los dos primeros mecanismos que pueden contribuir a mejorar notablemente la eficiencia del sistema de climatización. En concreto a través del *Free Cooling* la compañía Megawatt ha anunciado la construcción de un centro con un PUE de tan sólo 1.04[3].

El CESGA cuenta con dos enfriadoras con una capacidad de 580 KW cada una que se encargan de enfriar el agua caliente procedente del CPD hasta 11°C. Esta agua llega a las climatizadoras interiores que la aprovechan para enfriar el aire del CPD. Posteriormente el agua caliente vuelve a las enfriadoras y se repite el ciclo.

El *Free Cooling* es un mecanismo que permite a las enfriadoras aprovecharse de las bajas temperaturas del aire exterior para enfriar el agua sin necesidad de utilizar los compresores. En general, la experiencia dice que si la diferencia de temperatura entre el aire exterior y la temperatura del agua es de más de 4 °C las enfriadoras empiezan a utilizar el *Free Cooling*.

Existe una herramienta web proporcionada por *The Green Grid*[4] que nos permite hacer un cálculo del ahorro que se obtendría utilizando *Free-Air Cooling* (una variante de *Free Cooling* en la que se usa directamente el aire del exterior para enfriar el CPD) en función de los parámetros del CPD y de su localización. En el caso del CESGA obtenemos un ahorro estimado de 260.000 € al año.

A pesar de que esta estimación parece un poco elevada, ya que estima que se podrá utilizar el *Free-Air Cooling* durante un 93% del año en Santiago de Compostela, lo importante es ver que se pueden conseguir grandes ahorros utilizando esta tecnología.

En el caso de la tecnología *Free Cooling* con la que cuentan nuestras enfriadoras, los ahorros son menores ya que el aire exterior se utiliza para enfriar el agua y no para enfriar directamente los equipos.

Otra de las tecnologías de ahorro energético con la que cuenta una de las enfriadoras es la de recuperación de calor. Esta funcionalidad permite a la enfriadora aprovechar el calor generado por los compresores para calentar agua, que es utilizada en el sistema de calefacción del centro. Uno de los problemas que se nos puede plantear es que, en algunas situaciones, este sistema puede ser incompatible con el *Free Cooling*, ya que si se está aprovechando el aire exterior para enfriar el agua y no se están utilizando los compresores entonces no tendremos calefacción. En el CESGA en estos casos contamos con la posibilidad de utilizar una bomba de calor que funciona de forma independiente a las enfriadoras.

## 6. Cierre de pasillos

Es importante una correcta distribución de los racks en el CPD para que se establezca una buena circulación de aire. Hoy en día la configuración que se considera óptima es colocar los racks en filas con las partes frontales enfrentadas entre sí formando pasillos calientes y fríos. En algunos casos en tipos de máquinas específicas como los superdome, que no siguen el esquema habitual de coger aire frío por delante y expulsarlo por detrás, es necesario hacer reajustes y establecer pasillos fríos adicionales.

En esta sección veremos que un aspecto muy importante es asegurar que no se produce una re-circulación del aire caliente. Una solución que se suele adoptar es cerrar una clase de pasillos, o los fríos o los calientes. A priori, la solución en la que se cierran los pasillos calientes es la más eficiente energéticamente pero plantea ciertas complicaciones técnicas ya que es necesario canalizar el aire caliente de vuelta a las climatizadoras. Para evitar este problema en el CESGA hemos decidido optar por el cierre de los pasillos fríos. Adicionalmente existe otra limitación debida al sistema de extinción mediante gas HFC227, ya que éste está dimensionado para actuar con la configuración actual del CPD (sin cierre de pasillos). El cierre de los mismos obligaría a una reforma de este sistema. Para evitarlo la solución adoptada consiste en dejar sin cerrar el techo del pasillo de modo que el cierre no afecte a la propagación del gas HFC227.



El *Free Cooling* permite a las enfriadoras aprovecharse de las bajas temperaturas del aire exterior para enfriar el agua sin utilizar compresores



Es muy importante asegurar que no se produce una recirculación del aire caliente. Para eso se cierra una clase de pasillos: los fríos o los calientes



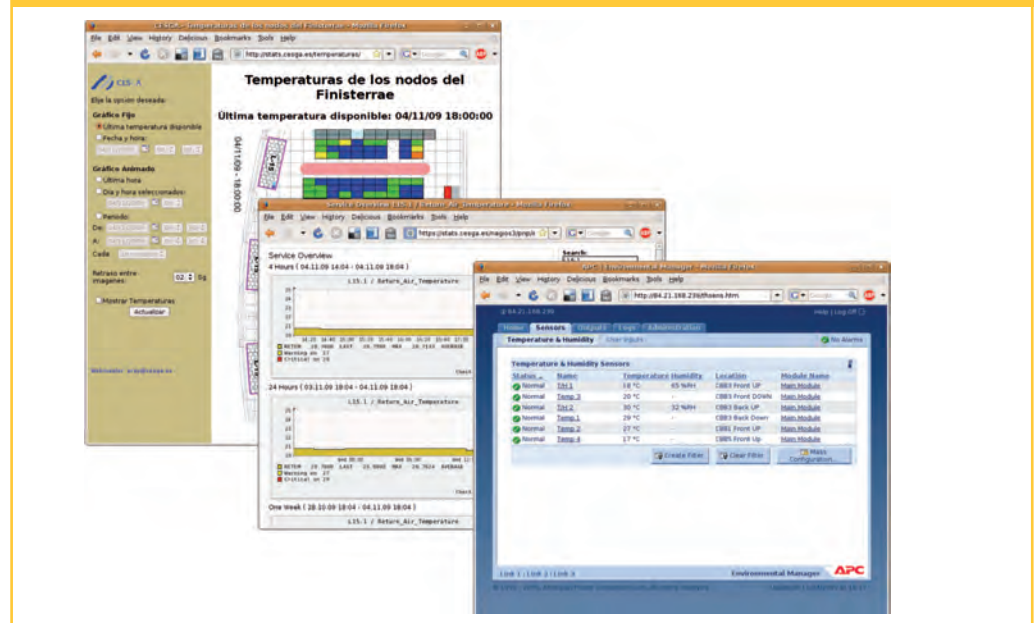
Adicionalmente, se han colocado paneles ciegos para evitar la recirculación del aire en los huecos entre racks que quedan debido a la presencia de columnas en el CPD.

En el CESGA contamos con distintos mecanismos para monitorizar la temperatura en el CPD (ver **figura 2**):

- Página web propia para monitorizar las temperaturas de los equipos (temperatura del aire de entrada al servidor, CPU, MB)
- Sensores de APC para monitorizar temperaturas y humedades
- Monitorización de la temperatura de las climatizadoras a través de Nagios.

**FIGURA 2. SISTEMAS DE MONITORIZACIÓN DE TEMPERATURA DISPONIBLES EN EL CESGA: MAPA DE TEMPERATURAS, NAGIOS Y APC**

El CESGA cuenta con sensores de APC para monitorizar temperaturas y humedades



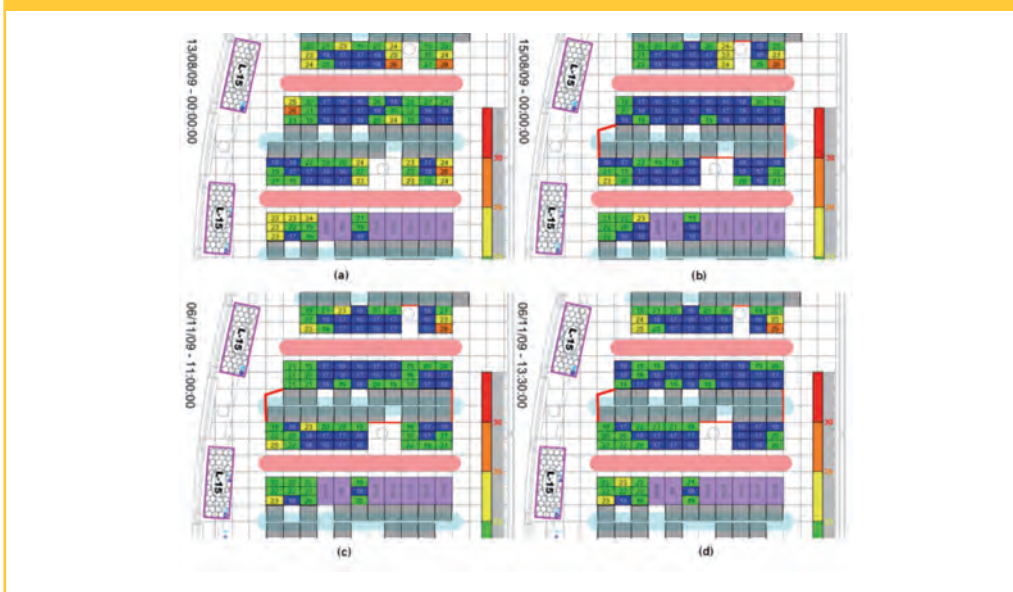
Los únicos nodos que resultan perjudicados por el cierre son los que están en el rack

En especial el primero de ellos nos ayudará a ver la eficiencia de las medidas adoptadas en uno de los pasillos.

En la **figura 3** se muestra el efecto de la colocación de las mamparas y paneles ciegos en uno de los pasillos. Como se puede ver comparando las gráficas 3a y 3b la temperatura a la que reciben el aire los nodos desciende en general, especialmente en el caso de los tres nodos que anteriormente estaban a 26 °C y que ahora pasan a 22 y 20 °C, no quedando ningún nodo en la zona naranja. Los únicos nodos que resultan perjudicados por el cierre son los nodos que están en el rack de la izquierda donde se localiza el único nodo del pasillo en la franja amarilla, a 23 °C. La temperatura de dos de los nodos en este rack aumentó en 2 °C tras el cierre.

Para revisar este efecto realizamos una prueba en donde extendimos la mampara que cierra el pasillo por la izquierda hasta el techo para evitar la recirculación de aire frío directamente a las climatizadoras (L15). El efecto de esta extensión se puede observar comparando las **figuras 3c y 3d**. Como se puede ver tras esta extensión la temperatura de los nodos del rack que había aumentado de temperatura baja y todos los nodos del pasillo reciben aire por debajo de 22 °C.

FIGURA 3. MAPA DE TEMPERATURAS DEL CPD: (A) ANTES DE COLOCAR MAMPARAS Y PANELES (B) TRAS LA COLOCACIÓN (C) ANTES DE EXTENDER LAS MAMPARAS HASTA EL TECHO (D) TRAS EXTENDER LAS MAMPARAS HASTA EL TECHO



El FT permite ahorrar en torno a 43 KW

## 7. Conclusiones

Se han presentado las experiencias del CESGA con cinco mecanismos de mejora de la eficiencia energética con los que hemos estado trabajando. Estas experiencias podrían ser de utilidad para otros centros de supercomputación y para otras instituciones que cuenten con CPDs. A continuación se resumen las principales conclusiones del estudio.

Se ha visto la importancia de los mecanismos de ahorro energético que incorporan los nuevos modelos de servidores que, en el caso del FT, nos permiten ahorrar en torno a 43 KW.

Se ha analizado la posibilidad de consolidar servidores en el SVGD utilizando blades, lo cual conllevaría un ahorro energético de 8KW.

Se ha comparado el consumo entre un sistema de almacenamiento en disco con un sistema en cinta para almacenar backups y hemos llegado a la conclusión de que la diferencia de consumo de los discos no justifica por sí sola el uso del sistema de cintas.

Se ha visto la importancia de utilizar un sistema de climatización basado en *Free Cooling*, el cual en zonas como Santiago de Compostela puede llevar a un ahorro de hasta 260.000 € anuales.

Se han analizado distintos mecanismos de cierre de pasillos y se ha presentado la solución adoptada en el CESGA.

Estos son sólo algunos de los mecanismos de mejora que hemos evaluado en el CESGA, pero existen muchos otros que no se presentan en este trabajo pero que también sería conveniente considerar como: eficiencia de los transformadores, eficiencia de las SAIs, corrección del factor de potencia, eficiencia de las fuentes de alimentación, etc.

En zonas como Santiago de Compostela se pueden ahorrar hasta 260.000 euros anuales utilizando el sistema de climatización basado en Free Cooling



Nos gustaría acabar este trabajo destacando que todos estos aspectos se están teniendo en cuenta en el diseño de un nuevo edificio que servirá para alojar el nuevo CPD del CESGA. Esperamos que gracias a los nuevos mecanismos de ahorro energético se pueda mejorar notablemente la eficiencia energética del CPD actual.

## Referencias

- [1] The Green500 List, <http://www.green500.org>
- [2] Top500 Supercomputing Sites, <http://www.top500.org>
- [3] Megawatt blog, <http://megawattconsulting.com/blog/?p=36>
- [4] Free Cooling Estimated Savings, [http://cooling.thegreengrid.org/europe/WEB\\_APP/calc\\_index\\_EU.html](http://cooling.thegreengrid.org/europe/WEB_APP/calc_index_EU.html)

Los aspectos señalados se están teniendo en cuenta en el diseño de un nuevo edificio que servirá para alojar el nuevo CPD del CESGA

J. López Cacheiro  
C. Fernández Sánchez  
J.I. López  
P. Rey  
A. Feijoo  
S. Díaz  
(sistemas@cesga.es)

Centro de Supercomputación de Galicia (CESGA)