

JUNTA DE ANDALUCÍA

**Centro Informático
Científico de Andalucía
(CICA)**

Anuarios 2010 y 2011

Apoiando la Investigación en Andalucía



JUNTA DE ANDALUCÍA

INDICE

Carta del director	3
COMUNICACIONES	5
RED INFORMÁTICA CIENTÍFICA DE ANDALUCÍA (RICA+)	6
1. Descripción de RICA	6
2. Centros conectados a RICA	8
3. Novedades en RICA en 2010	8
4. Estadísticas de tráfico de red	10
5. Mejoras en infraestructuras	16
5.1. Equipamiento	16
5.1.1. Instalación de las fuentes de alimentación y tarjetas procesadoras redundantes en los nodos provinciales	16
5.1.2. Instalación de las nuevas tarjetas procesadoras en RT1 y RT2	19
5.1.3. Instalación de nuevos nodos de RICA	20
5.2. Configuraciones	20
5.2.1. VSS	20
5.2.2. BGP full-routing	21
5.3. Servicios Multicast e IPv6	21
5.4. Local Internet Registry	21
6. La nueva red de fibra en Andalucía	22
6.1. Inconvenientes de la actual RICA	22
6.2. Red Nacional RedIRIS-NOVA	23
6.3. RedIRIS-NOVA en Andalucía	23
7. Características de la nueva red RICA	23
7.1. Especificaciones de la red	23
7.2. catálogo de servicios	23
7.2.1. Ancho de banda bajo demanda	24
7.2.2. Conectividad de alta capacidad IPv4 e IPv6	24
7.2.3. Conexiones punto a punto	24
7.2.4. Difusión multicast sobre IPv4 e IPv6	24
7.2.5. Servicio de DNS	24
7.2.6. Movilidad: acceso federado a la red para usuarios móviles	24
7.2.7. MPLS (Multi-Protocol Label Switching)	24

7.2.8. NOC - equipo de gestión de red	25
7.2.9. Telefonía IP	25
7.2.10. Videoconferencias: H.323 Global Dialing Scheme (GDS)	25
7.2.11. VPN-1.2 (punto a punto) o VPLS (multipunto a multipunto)	25
7.2.12. VPN-1.3	26
7.3. Diseño de la nueva red RICA	26
7.3.1. Nodos centrales de RICA	27
7.3.2. Nodos provinciales de Almería, Cádiz, Córdoba, Granada, Huelva y Jaén	28
7.4. Estado actual del despliegue	28
SEGURIDAD	31
1. Gestión centralizada de incidentes de seguridad	32
2. Análisis de vulnerabilidades	33
3. Despliegue de IPv6 en DNS	34
4. Esquema Nacional de Seguridad	34
5. Nuevos sistemas LDAP	35
6. Repositorios internos de software	36
7. Informes mensuales de seguridad	36
8. Contactos con otros CERT	37
9. Implantación de certificados de TERENA / COMODO	37
10. Investigación sobre sistemas IDS	38
11. Test de herramientas para auditoría de seguridad	38
SISTEMAS / HPC	41
1. Actuaciones en la infraestructura dedicada a SICA	42
1.1. Base de datos standby	42
1.2. Direct routing	43
1.3. Implantación del nuevo Sistema de Información Científica de Andalucía (SICA2)	43
2. Novedades en el área de sistemas	45
2.1. Cluster de virtualización	45
2.2. Infraestructura de backup	46
2.2.1. Servidores	46
2.2.2. PCs de escritorio	47
2.3. Implantación del gestor de proyectos REDMINE	47
2.4. Implantación del gestor de versiones GIT	48

2.5. Sistema de autoencendido de ordenadores	48
2.6. Configuración de IPv6 en servidores	49
3. CICA y Europa	50
3.1. Proyecto EDGI	50
4. Proyecto de adecuación para la sostenibilidad y eficiencia energética en el CICA	50
4.1. Cambio de suelo técnico	52
4.2. Adquisición de máquinas de aire acondicionado	52
COMPUTACIÓN CIENTÍFICA DE ALTAS PRESTACIONES Y e-CIENCIA	54
1. Usuarios y proyectos de investigación	54
1.1. Usuarios y áreas de conocimiento	54
1.2. Grupos de investigación registrados	56
1.3. Publicaciones de resultados	57
2. Asistencia a eventos	57
2.1. II Jornadas Ibéricas de Supercomputación	57
2.2. III y IV Jornadas de Usuarios de la Red Española de Supercomputación	58
2.3. IV Reunión Plenaria de la Red Española de e-Ciencia	59
3. Hardware y software	59
3.1. Introducción a la tecnología hardware	59
3.2. Hardware para cálculo científico	59
3.3. Software	60
3.4. Pruebas de rendimiento	61
3.4.1. Introducción	61
3.4.2. Descripción de los servidores	61
3.5. Nuevos servidores	62
4. Estadísticas anuales	63
4.1. Consumo de CPU y memoria RAM	63
4.2. Estadísticas del cluster NOVA BULL	65
4.3. Ejecución de tareas, tiempo medio de espera	66
4.4. Colas de ejecución. Descripción y uso	69
4.5. Evolución en el periodo 2008-2010	70
4.5.1. Consumo de recursos	70
4.5.2. Tiempos de espera en las colas	71

4.6. Almacenamiento en disco compartido	72
5. Acceso a los recursos	74
5.1. ¿Cómo hacerse usuario del CICA?	74
5.2. Descripción del portal de e-Ciencia	74
APLICACIONES	79
1. Catálogo Colectivo del Consorcio de Bibliotecas Universitarias de Andalucía	80
2. Proyecto CONFIA	82
2.1. Introducción	82
2.2. Infraestructura hardware y servicios aportados	83
2.3. Descripción del estado actual de los elementos comunes	83
3. Forja RedIRIS - CICA	83
4. Portal de difusión de la ciencia Alnova	84
4.1. Introducción	84
4.2. Desarrollo inicial	86
4.3. Funcionalidad	86
4.4. Pruebas y mantenimiento	87
5. El CICA en las redes sociales	87
FORMACIÓN Y EVENTOS	89
1. Asistencia a eventos y talleres	90
1.1. Talleres para proyectos europeos	90
1.2. Reuniones de coordinación de EDGI	90
1.3. Otros eventos	90
2. Visitas al CICA	94
ANEXO: PRUEBAS HPC	97
1. Descripción de las pruebas	98
2. Realización de las pruebas y resultados	99
2.1. Batería de pruebas 1: Servidores con 2 procesadores	100
2.2. Batería de pruebas 2: Análisis de la interconexión	101
2.3. Batería de pruebas 3: Servidores de 4 CPUs con uso simultáneo de número distinto de cores	101
2.4. Batería de pruebas 4: Análisis del uso de Hyperthreading	102
2.5. Batería de pruebas 5: Virtualización	103

Carta del Director

Os doy de nuevo la bienvenida y sirva el presente documento para mostrar las principales actividades que el CICA ha realizado durante los años 2010 y 2011. Estos años caracterizados por una difícil coyuntura económica internacional de bastante calado, desde el Centro nos hemos preparado apropiadamente para los diferentes proyectos que se presentaban, y entre ellos destacar el proyecto de despliegue de la fibra oscura y que durante estos dos años hemos convertido en una realidad. Es importante indicar que desde este centro nos sumamos a la idea de que “crisis es a la vez una oportunidad”.

El despliegue físico de la nueva red RICA de fibra oscura comenzó a finales de 2010 y ha finalizado durante el 2011, quedando solo por instalar la electrónica de red en 2012 es el año de su puesta en marcha, previéndose un importante ahorro de coste de alquiler de líneas, e igualmente importante, un avance sin precedentes a nivel de telecomunicaciones y servicios sobre la red, que iremos poniendo en valor para la comunidad universitaria e investigadora de Andalucía a lo largo de los próximos meses y años, con posibilidades que ahora mismo incluso se nos escapa de la mente, pero que cambiarán de manera notoria nuestra manera de relacionarnos y colaborar con otros docentes e investigadores.

En el apartado de la computación de altas prestaciones, durante estos dos años hemos profundizado en el modelo de “housing óptimo y gratuito con contraprestaciones razonables” para las infraestructuras de cálculo y almacenamiento masivo que consiguen los grupos de investigación universitarios. Durante 2010 y 2011 se han aumentado sensiblemente las capacidades de cálculo y almacenamiento, no solo con equipos propios sino fundamentalmente con equipos que los grupos de investigación alojan en el Data Center del CICA a cambio de ceder al resto de la comunidad investigadora los tiempos muertos en los que no necesiten dichos recursos. Este modelo, además de aplicar economías de escala, tiene otras ventajas, tales como un ahorro energético de forma global, un control profesionalizado de las infraestructuras de cálculo y almacenamiento, y una tolerancia a fallos de los sistemas tanto a nivel físico (eléctrico) como lógico (backups).

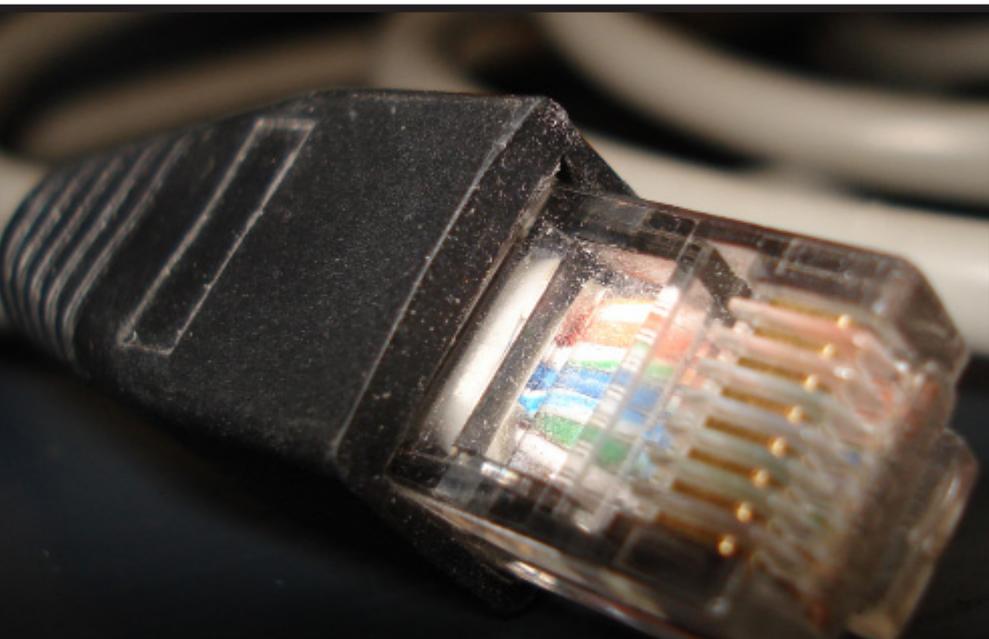
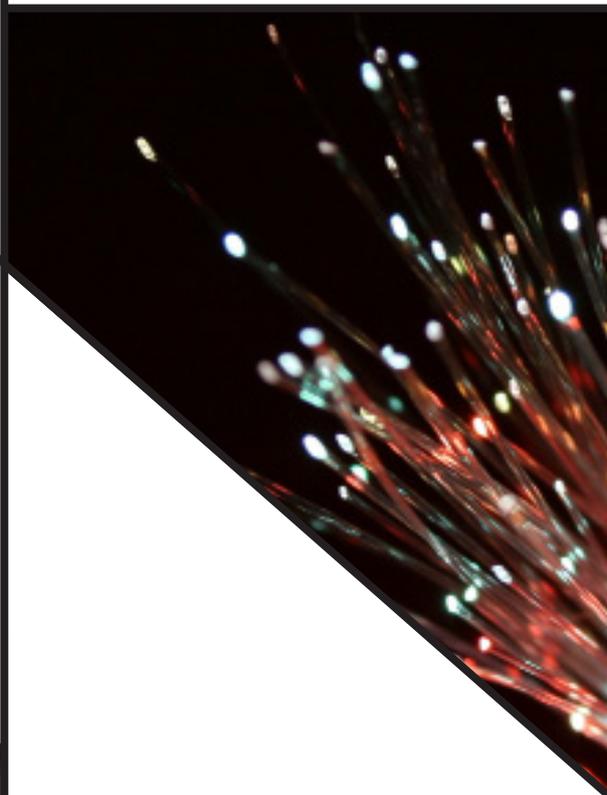
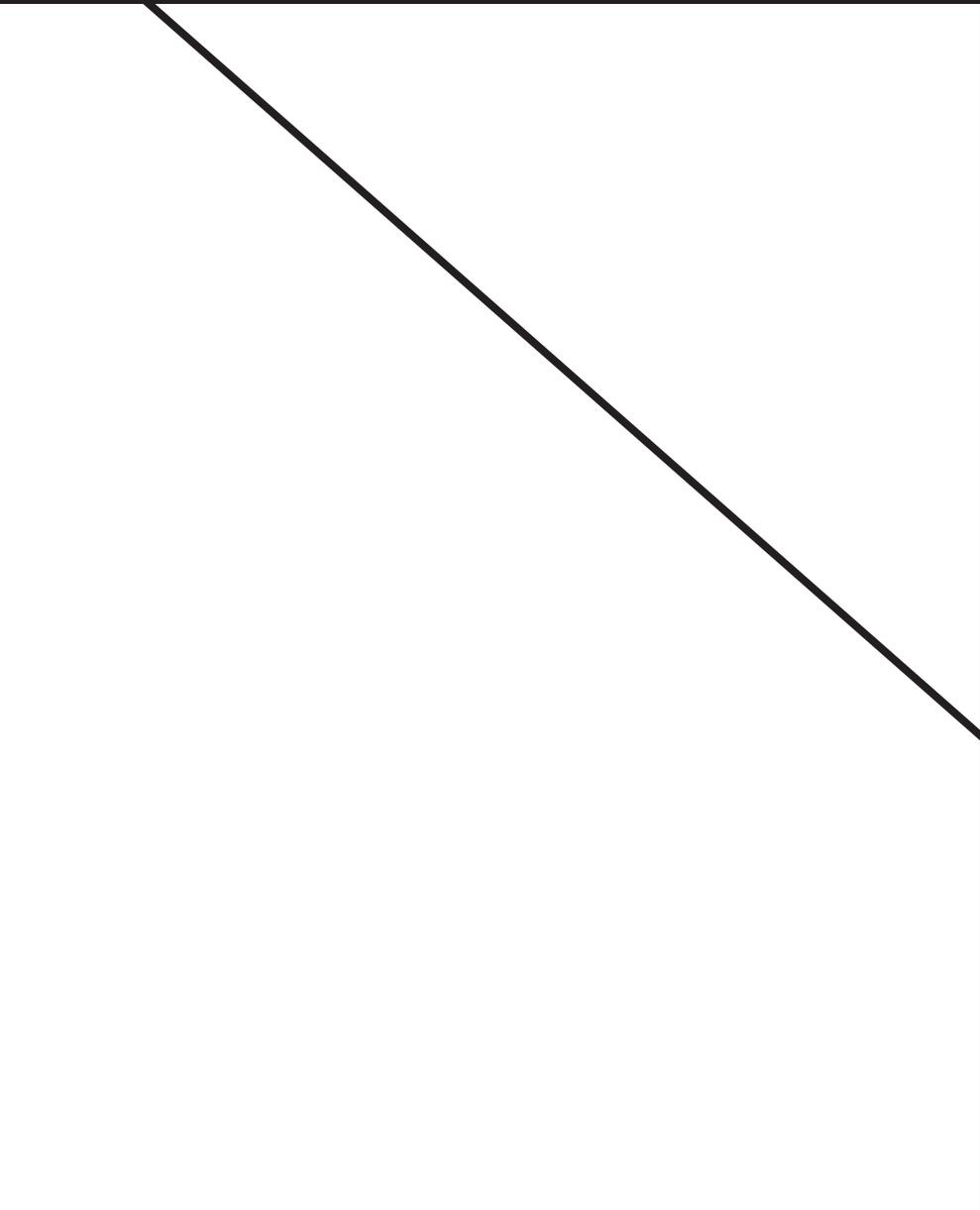
Otro aspecto a destacar y centrándonos en el propio edificio del CICA, hemos acogido con bastante entusiasmo los nodos primarios de otras Consejerías. En particular, y durante este año 2011 se ha trasladado y alojado en el CICA todos los servidores de la Consejería de Hacienda y Administración Pública de la Junta de Andalucía, estando previsto a lo largo de los próximos meses el traslado de otras Consejerías, convirtiéndose nuestro Centro en un nodo estratégico de la infraestructura en informática de Andalucía.

Por último y siguiendo con las infraestructuras del CICA, hay que decir que aun teniendo pendiente grandes cambios en las infraestructuras eléctricas, en 2010 y 2011 se han añadido nuevos equipos de acondicionamiento de la refrigeración al Data Center del CICA, con rendimientos energéticos muy superiores y gases no contaminantes, y se ha completado el proyecto de mejora de infraestructura eléctrica y facilities (grupo electrógeno, sistemas de alimentación ininterrumpida, etc.) que permitirá continuar con la hoja de ruta para las actuaciones a acometer en el inmediato futuro para que el data Center del CICA siga siendo un referente de calidad y buenas prácticas energéticas.

Por tanto terminamos 2011 con el objetivo económico cumplido y una buena carga de trabajo pendiente para 2012, pues hay que poner en marcha una infraestructura de comunicaciones de altísima capacidad, pensada y diseñada para un contexto temporal de 20 años. Ya no se habla de decenas de megabits por segundo sino de centenares de gigabits por segundo. Y es que hasta 2010 solo había un punto de conexión entre la Red Académica Española y Andalucía, mientras que en 2011 la infraestructura física une todas las provincias a través de instalaciones localizadas en los campus universitarios andaluces. Y también se han puesto los mimbres sobre los que basar el funcionamiento óptimo del centro en cuanto a supercomputación y rendimiento energético.

A lo largo del presente documento describimos de manera detallada los aspectos que anteriormente hemos descrito. Antes de finalizar quiero agradecer a todos los investigadores y docentes andaluces que colaboran con el CICA, su apuesta decidida por servirles, razón por la que diariamente justificamos nuestra razón de ser, la de servir a la enseñanza e investigación en Andalucía.

Juan Antonio Ortega



Comunicaciones

La evolución de la Red Informática Científica de Andalucía (RICA) hacia una red con más y mejores servicios forma parte del plan estratégico de CICA. Poner tecnologías de vanguardia a disposición de todas las Universidades y Centros de Investigación de Andalucía, ha sido uno de los objetivos fundamentales en el periodo 2010-2011.

Las necesidades de los científicos aumentan cada año y requieren de nuevas tecnologías como equipos de trabajo remotos con herramientas colaborativas de eCiencia, explotación de grandes bases de datos, sistemas de computación de alto rendimiento, minería de datos, sistemas de información geográfica, lingüística computacional, etc.

Para cubrir todas estas necesidades, el Centro Informático Científico de Andalucía está inmerso en pleno despliegue de la nueva red de fibra oscura, con la que se aumenta considerablemente la capacidad de la red, con mejores recursos y nuevos servicios.

Otra de las mejoras destacables es la adopción de IPv6, siendo CICA una de las primeras instituciones en adoptar dicho protocolo en España





RED INFORMÁTICA CIENTÍFICA DE ANDALUCÍA (RICA+)

La red andaluza para la enseñanza superior y la investigación RICA ha seguido evolucionando en 2010 y 2011 para continuar proporcionando los servicios de comunicaciones avanzadas a la comunidad investigadora, y garantizar que sus necesidades actuales y futuras queden cubiertas.

Los retos de 2010 han sido: Dar servicio a mayor número de centros, continuar con la renovación tecnológica y dado el aumento de las necesidades de los usuarios en ancho de banda y servicios, aumentar la capacidad de las líneas de las provincias andaluzas que lo han requerido por el aumento de caudal.

1. DESCRIPCIÓN DE RICA

La Comunidad Autónoma de Andalucía dispone de una red de comunicaciones académica y de investigación denominada RICA (Red Informática Científica de Andalucía), conectada a la red académica española RedIRIS (Red de Interconexión de Recursos Informáticos). Esta red tiene una importancia crucial para los grupos de investigación y la comunidad educativa ya que les permite acceder a recursos distribuidos en red, así como participar e interactuar en lo que se conoce como e-ciencia. Esto es fundamental para el avance y el crecimiento de Andalucía en I+D+I.

RICA proporciona una serie de puntos de acceso a través de los cuales se conectan los centros adscritos a la red. Éstos son todas las universidades, centros públicos dedicados a la investigación, centros públicos dedicados a actividades académicas, y otros Agentes del Conocimiento andaluces. Los nodos de acceso a la red RICA se encuentran ubi-

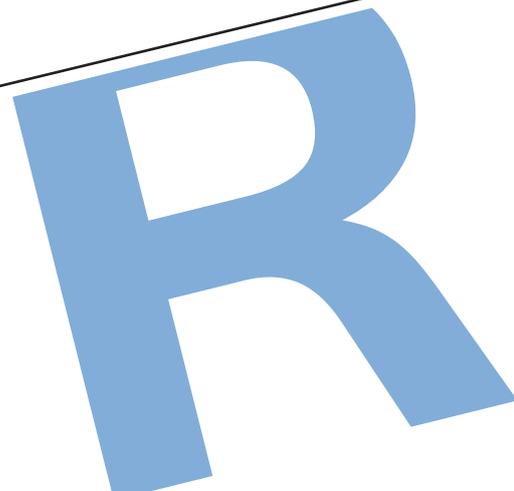
cados en cada una de las diez universidades públicas andaluzas. Los centros de investigación de cada provincia, se conectan al nodo de acceso de su provincia. El CICA gestiona y coordina esta red, así como los servicios que se prestan a través de ella.

Los servicios de comunicaciones que RICA ofrece a la comunidad académica y científica andaluza requieren como soporte una infraestructura básica de comunicaciones adaptada tecnológicamente a las necesidades de las instituciones y centros adscritos. Debido al carácter innovador de las actividades de estas instituciones, la red está expuesta a continuos cambios y avances tecnológicos. Las necesidades de los investigadores obligan a desplegar redes de muy alta capacidad y a integrar continuamente nuevos protocolos y servicios.

Las redes académicas deben trabajar principalmente en tres direcciones:

- Adaptación de protocolos.
- Definición de nuevos modelos de gestión.
- Estandarización de un catálogo de servicios.

Teniendo en cuenta las nuevas necesidades de los centros conectados y las características de la red es necesaria una evolución constante de la misma, para poder seguir ofreciendo servicios de calidad a la comunidad académica e investigadora. Con esta política se persigue ofrecer una red de última generación y alta capacidad, a la vez que flexible, robusta, fiable y adaptada a las actividades propias de los centros adscritos a RICA.



La red actual está basada en la tecnología MacroLAN/VPN IP, ofrecida por Telefónica. Esta tecnología es la empleada para interconectar los Nodos de RICA, a los que se conectan las universidades y centros de investigación. MacroLAN/VPN IP es un servicio de interconexión de redes locales de banda ancha sobre infraestructura IP basada en tecnología MPLS. Como consecuencia de esta tecnología, to-

Cada Nodo de acceso a RICA (ver la topología en figura 1) dispone inicialmente de un acceso principal de nivel 3 a la red MacroLAN de 400 Mb/s, y otro de respaldo de 100 Mb/s. El caudal de la línea principal del Nodo de acceso a RICA de Granada se amplió posteriormente a 800Mb/s, por la creciente demanda de ancho de banda.

Las líneas de respaldo siguen un trazado físico distinto del

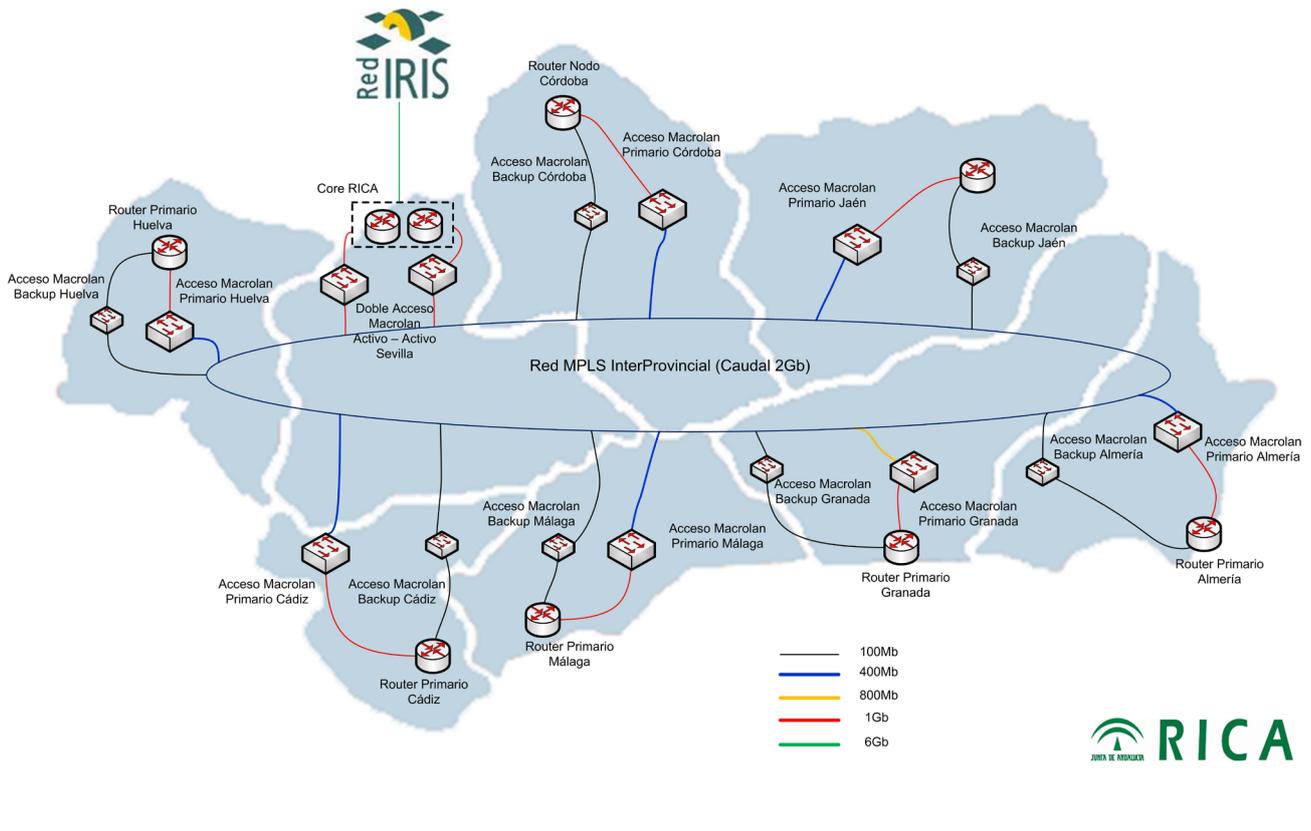


Figura 1: Topología de RICA en 2010

dos los recursos de RICA son accesibles con las mismas prestaciones desde cualquier ubicación.

Actualmente, la red de acceso a RICA es homogénea a nivel de equipamiento. Esta homogeneidad simplifica las labores de gestión y operación de la red.

Por otra parte, el equipamiento utilizado permite contar con las interfaces necesarias para conectar con la red de transporte MacroLAN de Telefónica y con los equipos de las diversas universidades e instituciones afiliadas a RICA en cada provincia. Cada una de las sedes está equipada con un router/switch que se encarga del intercambio de tráfico entre RICA y la red de cada centro en cuestión.

que sigue la línea principal, de forma que en caso de que se produzca un corte físico en una de las líneas no se vea afectada la otra.

El Nodo central de RICA, ubicado en CICA dispone de dos accesos diversificados MacroLAN de 1Gb/s, en modo balanceado. De este modo, si uno de los enlaces cae, el tráfico se redirecciona por el otro. La conexión del Nodo central de RICA con la red MacroLAN de Telefónica a 2Gb/s hace que toda la red RICA tenga ese ancho de banda de salida hacia RedIRIS (conexión con Internet).

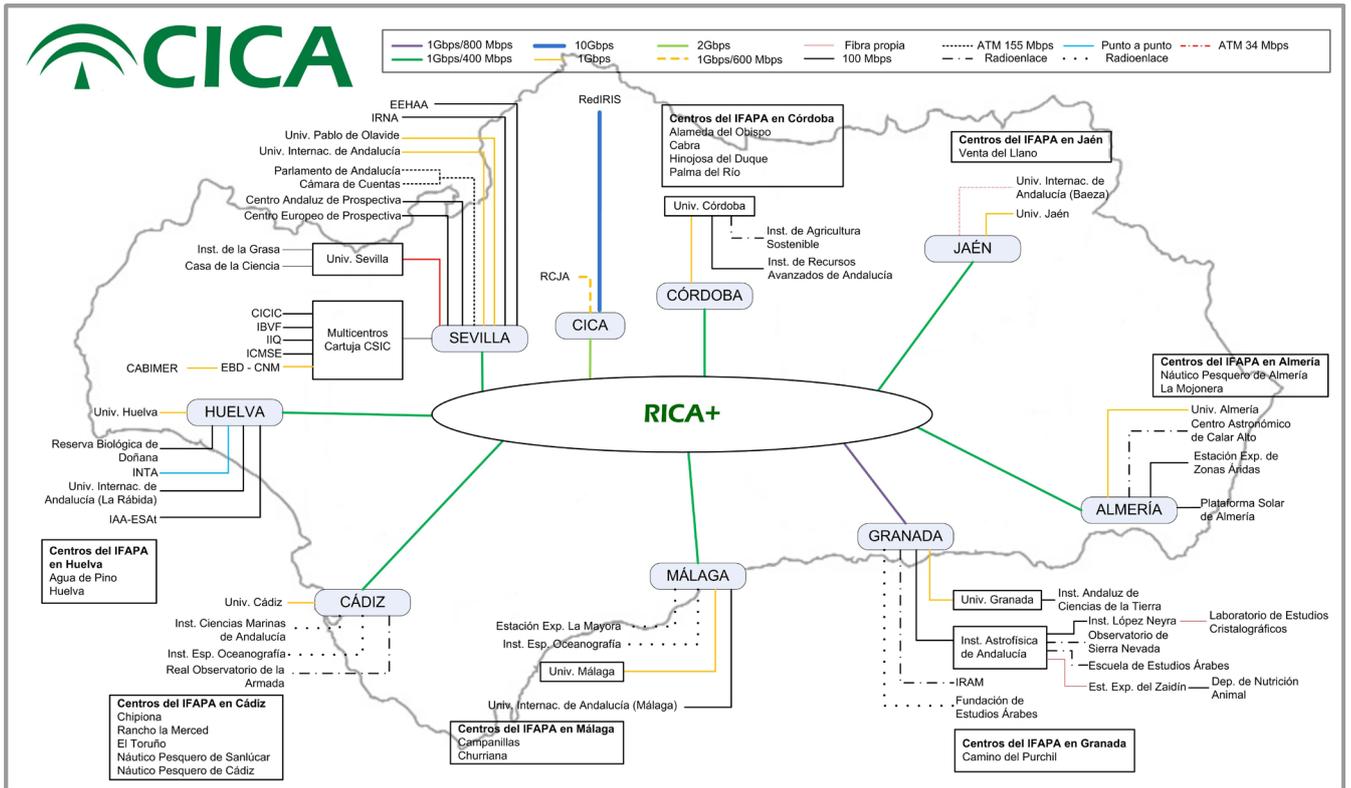


Figura 2: Mapa de centros de RICA en 2010

2. CENTROS CONECTADOS A RICA

La red RICA aumenta cada año, incorporando nuevos centros. Actualmente la red conecta los centros que aparecen en la figura 2 y cuyos datos más relevantes se detallan en la tabla de las páginas siguientes.

3. NOVEDADES EN RICA EN 2010

Durante los años 2010 y 2011, y en línea con la política de evolución constante de la red RICA, se han realizado numerosas ampliaciones en las conexiones con los centros. Se describen a continuación el detalle de cada una de ellas.

Aumento de ancho de banda RICA-RedIRIS

En 2011 se ha aumentado el ancho de banda de la conexión entre los routers principales del troncal del CICA y el de RedIRIS, ubicados en el CPD del CICA, pasando de una conexión mediante 6 cables de 1 Gb cada uno a una única conexión de 10Gb.

Universidad Pablo de Olavide

La Universidad Pablo de Olavide, ha aumentado su capacidad de conexión, migrando desde una conex-

ión ATM de 155 Mbps a una conexión MacroLan de 1 Gbps. Además, se ha instalado un nodo de RICA en las dependencias de la UPO, que proporciona conexión para la propia Universidad y los centros conectados anteriormente a través de ella. El nuevo nodo consiste en un equipo CISCO 3750.

Universidad Internacional de Andalucía

La sede de Sevilla de la Universidad Internacional de Andalucía (UNIA) pasa de una conexión ATM de 155 Mbps a MacroLan de 1Gbps.

Escuela Universitaria de Ingeniería Técnica Agrícola

El centro pasa de conectarse a través de la infraestructura de la UPO a conectarse directamente a infraestructura de RICA, más concretamente al nuevo nodo de RICA en la UPO.

Plataforma Solar de Almería

Se ha migrado la infraestructura de conexión instalándose nuevas antenas propias, que conectan el centro con el nodo de RICA en Almería. La velocidad de conexión ha aumentado de 7 Mbps a 36 Mbps (2x18Mbps).

Instituto de Astrofísica de Andalucía (CSIC)

Este centro, dependiente del Consejo Superior de Investigaciones Científicas y ubicado en la capital granadina, pasa a conectarse a través de MacroLan, con un ancho de banda de 300 Mbps.

Multicentros Cartuja

En este edificio dependiente del CSIC y ubicado en la capital sevillana, se ha instalado un nuevo nodo propio

de RICA, de forma que los diferentes centros que se conectan desde allí ahora lo hacen a través de este nodo.

Dichos centros, todos del CSIC, son los siguientes:

- Centro de Investigaciones Científicas Isla de la Cartuja, que a su vez aloja a los centros: Instituto de Bioquímica Vegetal y Fotosíntesis (IBVF), Instituto de Ciencia de Materiales de Sevilla (ICMS) y del Instituto de Investigaciones Químicas (IIQ).

Provincia	Centros	Conexión	Número Usuarios
Almería	Centro Astronómico Hispano-Alemán de Calar Alto	200 Mbps Radioenlace	60
	Estación Experimental de Zonas Áridas (CSIC)	100 Mbps Radioenlace	120
	Plataforma Solar de Almería (CIEMAT)	36 Mbps Radioenlace	200
	Universidad de Almería	1Gb/400 Mbps MacroLan	11400
Cádiz	Instituto de Ciencias Marinas de Andalucía (CSIC)	1 Gbps Fibra Óptica	100
	Instituto Español de Oceanografía	10 Mbps MacroLan	30
	Real Observatorio de la Armada	5 Mbps Radioenlace	70
	Universidad de Cádiz	1Gb/400 Mbps MacroLan	20000
Córdoba	Instituto de Agricultura Sostenible (CSIC)	1Gbps Fibra propia	200
	Instituto de Estudios Sociales Avanzados de Andalucía (CSIC)	100Mbps	100
	Universidad de Córdoba	1Gb/400 Mbps MacroLan	18200
Granada	Estación Experimental del Zaidín (CSIC)	100 Mbps	350
	Instituto Andaluz de Ciencias de la Tierra (CSIC)	100 Mbps	200
	Instituto de Astrofísica de Andalucía (CSIC)	300 Mbps MacroLan	200
	Instituto de Parasitología y Biomedicina "López Neyra" (CSIC)	100 Mbps	170
	Unidad de Nutrición Animal (CSIC)	100 Mbps	60
	Instituto de Radioastronomía Milimétrica (IRAM)	18 Mbps Radioenlace	50
	Universidad de Granada	1Gb/800 Mbps MacroLan	60000
Huelva	Reserva Biológica de Doñana	100 Mbps MacroLan	30
	Universidad de Huelva	1Gb/400 Mbps MacroLan	11000
	Universidad Internacional de Andalucía. Sede La Rábida	100 Mbps	2000
	Instituto de Técnica Aeroespacial	2 Mbps Serial	60
Jaén	Universidad de Jaén	1Gb/400 Mbps	16000
	Universidad Internacional de Andalucía. Sede de Baeza	34 Mbps	3000
Málaga	Estación Experimental "La Mayora" (CSIC)	10 Mbps MacroLan	100
	Instituto Español de Oceanografía	10 Mbps Radioenlace	90
	Universidad de Málaga	1Gb/400 Mbps MacroLan	36500
	Universidad Internacional de Andalucía Sede de Málaga	100 Mbps	2000
Sevilla	Centro Nacional de Aceleradores (CSIC)	100 Mbps	100
	Centro Nacional de Microelectrónica (CSIC)	1 Gbps Fibra	120

Centro de Investigaciones Científicas Isla de la Cartuja (CSIC)	1 Gbps Fibra	400
Centro Andaluz de Biología Molecular y Medicina Regenerativa (CABIMER)	1 Gbps	200
Instituto de Bioquímica Vegetal y Fotosíntesis (IBVF)	1 Gbps Fibra	100
Instituto de Investigaciones Químicas (IIQ)	1 Gbps Fibra	110
Instituto de Ciencia de Materiales de Sevilla (ICMS)	1 Gbps Fibra	60
Escuela de Estudios Hispanoamericanos (CSIC)	100 Mbps	80
Estación Biológica de Doñana (CSIC)	1 Gbps	200
Instituto de la Grasa y sus Derivados (CSIC)	18 Mbps Radioenlace	180
Instituto de Recursos Naturales y Agrobiología (CSIC)	1 Gbps	150
Centro Andaluz de Prospectiva (J. de Andalucía)	100 Mbps	15
Centro Informático Científico de Andalucía (J. Andalucía)	2Gbps Fibra	50
Centros TIC – Colegios e Institutos (J.Andalucía)	700 Mbps	
Cámara de Cuentas de Andalucía	155Mbps	150
Parlamento de Andalucía	155Mbps	150
Universidad de Sevilla	1Gb/400 Mbps MacroLan	60000
Universidad Internacional de Andalucía	1Gb/400 MacroLan	2500
Universidad "Pablo de Olavide"	1Gb/400 Mbps MacroLan	10000
Delegación del CSIC en Andalucía	36 Mbps Radioenlace	
TOTAL		256750

- El Centro Andaluz de Biología Molecular y Medicina Regenerativa (CABIMER)
- La Estación Biológica de Doñana (EBD)
- El Instituto de Microelectrónica de Sevilla del Centro Nacional de Microelectrónica (IMSE-CNM)

El nuevo nodo consiste en un equipo CISCO 3750.

Instituto de Recursos Naturales y Agrobiología (IRNA)

Cambio de conector para conseguir mayor velocidad, pasando de 100 Mbps a 1 Gbps. También se han eliminado los convertidores de medios que había entre los centros, disminuyendo los posibles puntos de fallo.

Centros TIC

Los Centros TIC de Andalucía, dependientes de la Consejería de Educación de la Junta de Andalucía y reparados por toda la geografía andaluza, se conectan a Internet a través de la infraestructura de RICA. En el año, se ha aumentado la capacidad de conexión para estos centros, pasando de 600 Mbps a 700 Mbps.

Cambios de ubicación

El nodo de Almería se ha trasladado desde el CPD de Almería al PDP de RedIRIS.

El nodo de RedIRIS ubicado en CICA se ha trasladado desde la sala general de servidores y comunicaciones del CPD de CICA a una zona específica de nueva instalación junto a la sala general mencionada, donde comparte instalaciones eléctricas y de seguridad con la infraestructura de comunicaciones de la parte administrativa de la Red Corporativa de la Junta de Andalucía.

4. ESTADÍSTICAS DE TRÁFICO DE RED

Se presentan a continuación las estadísticas de tráfico para la salida de la red RICA hacia RedIRIS. En las figuras 3 y 4 se puede observar el aumento progresivo de tráfico conforme avanza el tiempo, lo que justifica la necesidad constante de ampliación y mejora de la red.

Si se analiza el tráfico a nivel provincial, se puede observar la misma tendencia de crecimiento que en el enlace hacia RedIRIS. En las figuras 5 a 20 se muestran las estadísticas de tráfico para los enlaces provinciales. Como se puede observar si se comparan las gráficas pertenecientes al periodo anual de 2010, con las pertenecientes a 2011, el tráfico parece ligeramente superior en 2010.

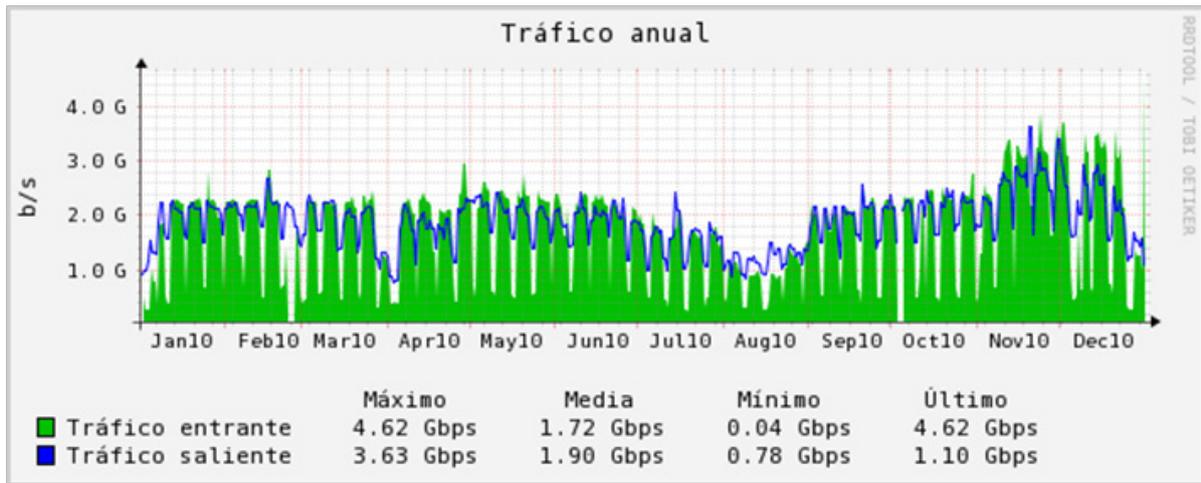


Figura 3: Tráfico RICA - RedIRIS (2010)

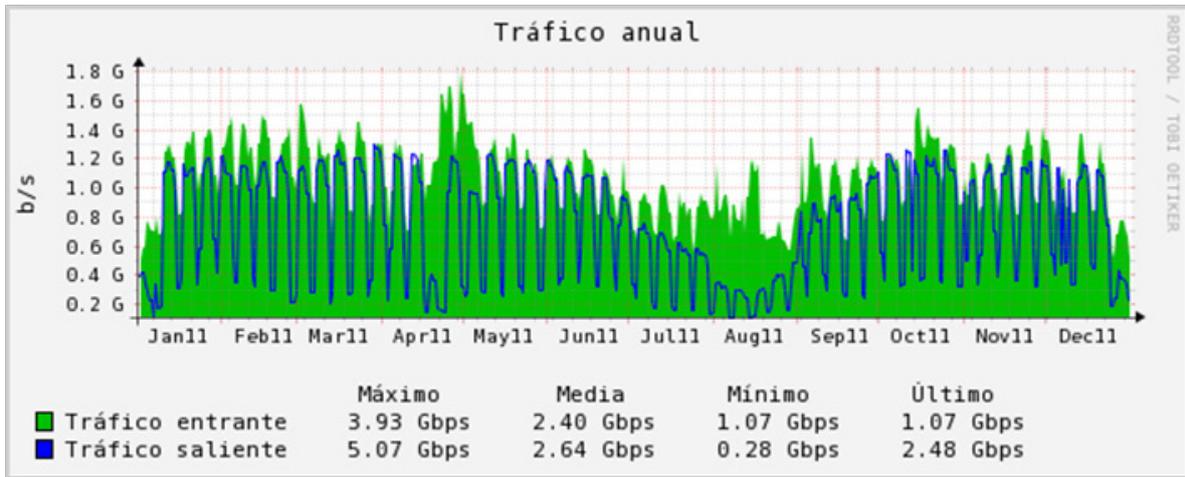


Figura 4: Tráfico RICA - RedIRIS (2011)

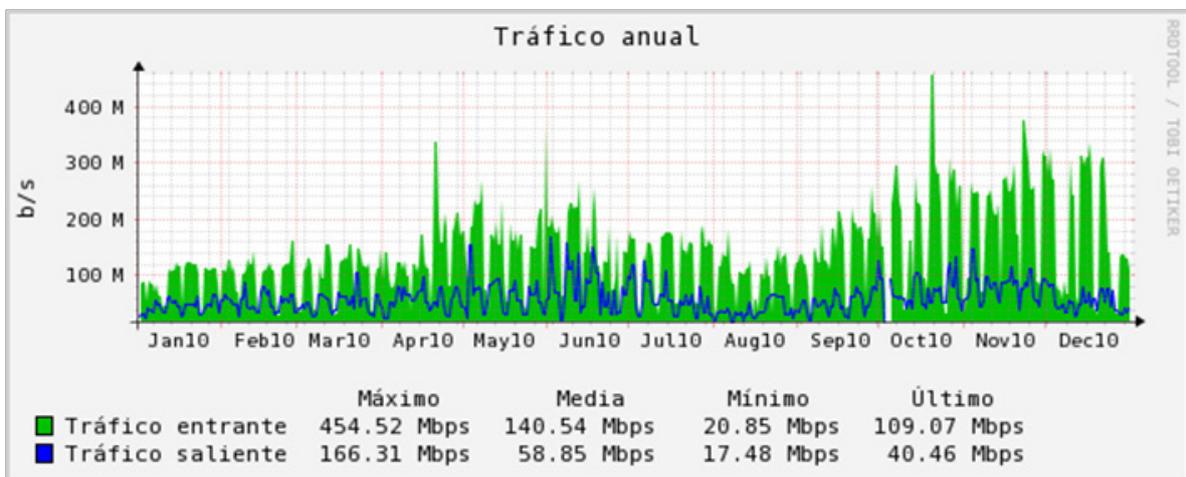


Figura 5: Provincia de Almería (2010)

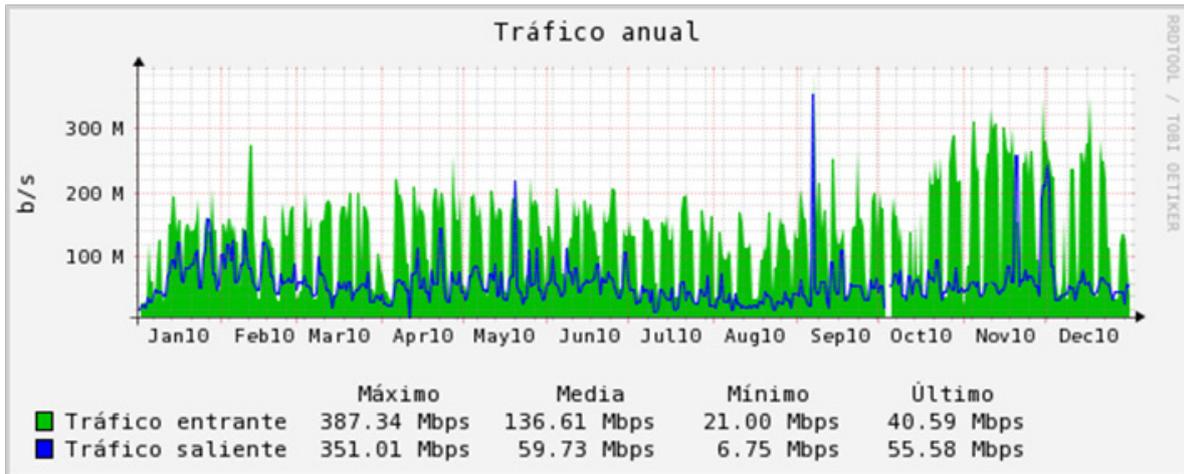


Figura 7: Provincia de Cádiz (2010)

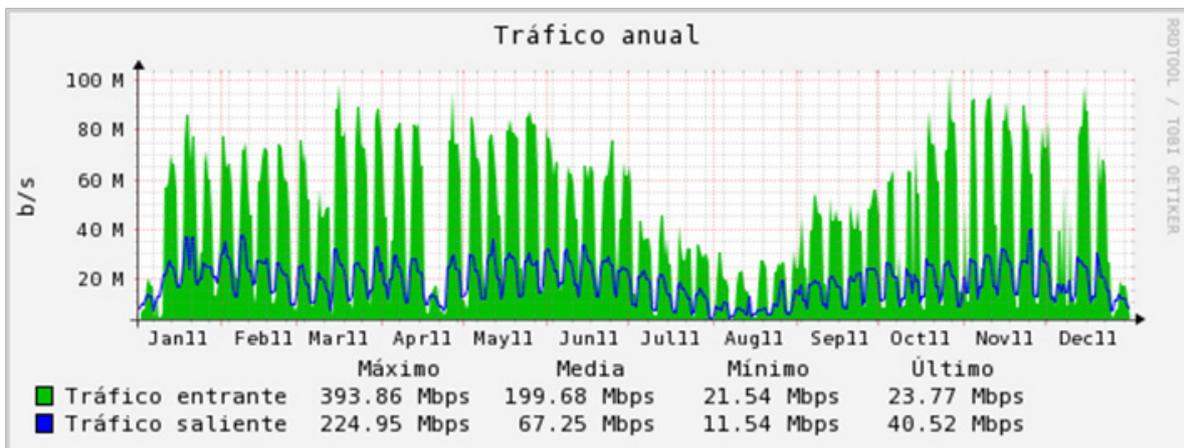


Figura 8: Provincia de Cádiz (2011)

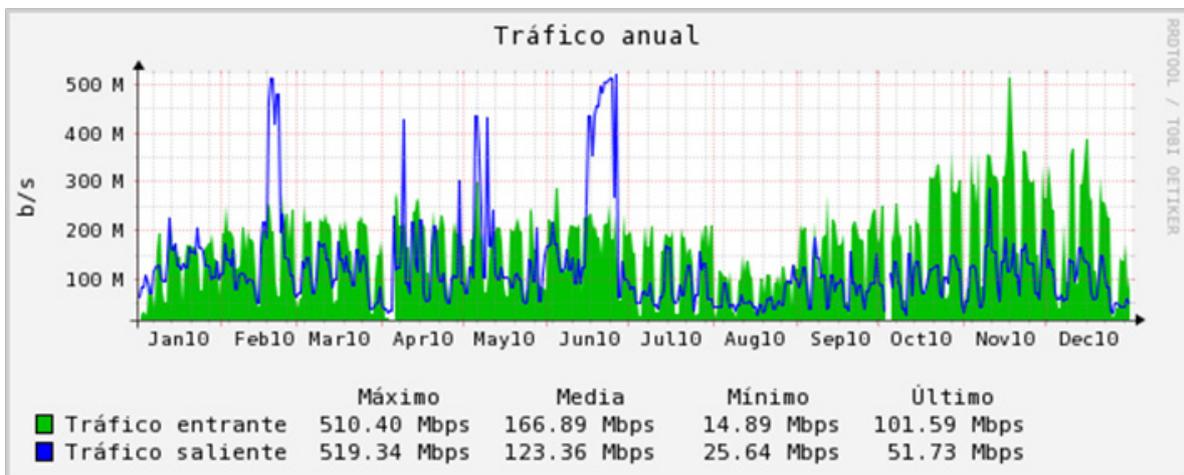


Figura 9: Provincia de Córdoba (2010)

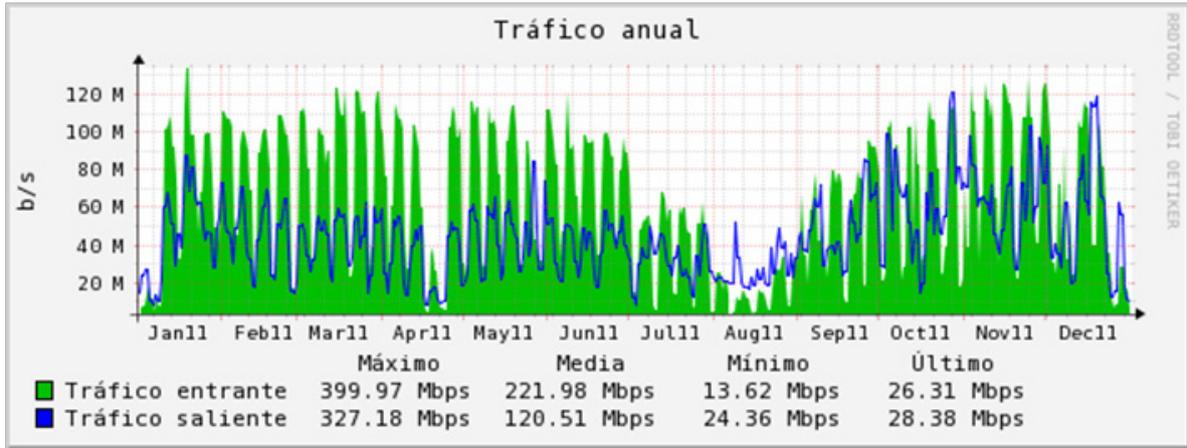


Figura 10: Provincia de Córdoba (2011)

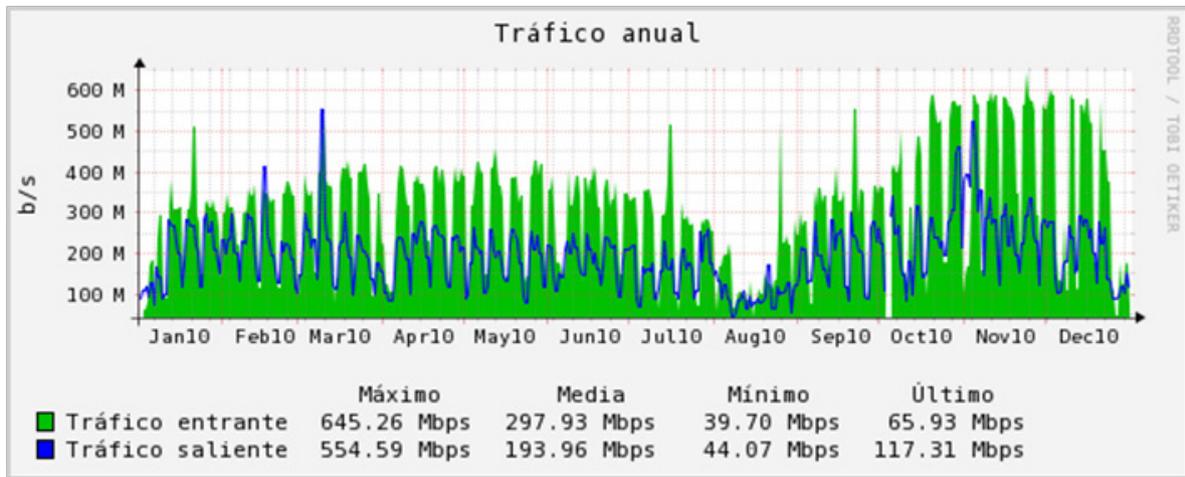


Figura 11: Provincia de Granada (2010)

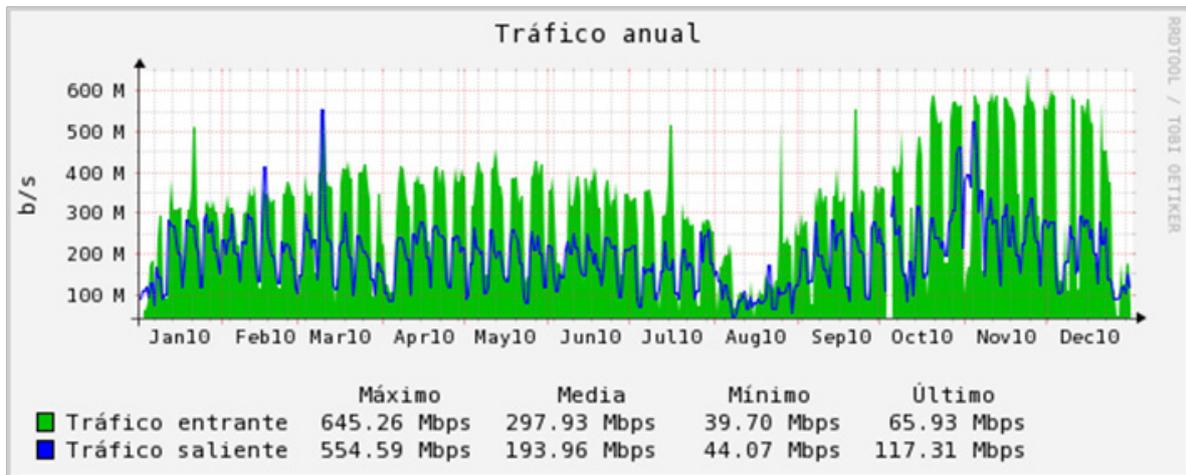


Figura 12: Provincia de Granada (2011)

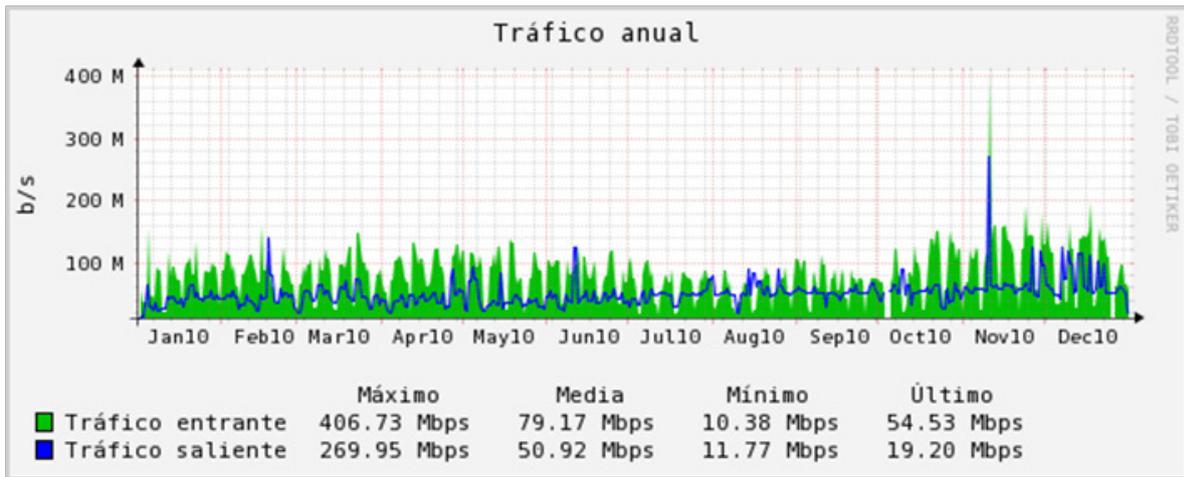


Figura 13: Provincia de Huelva (2010)

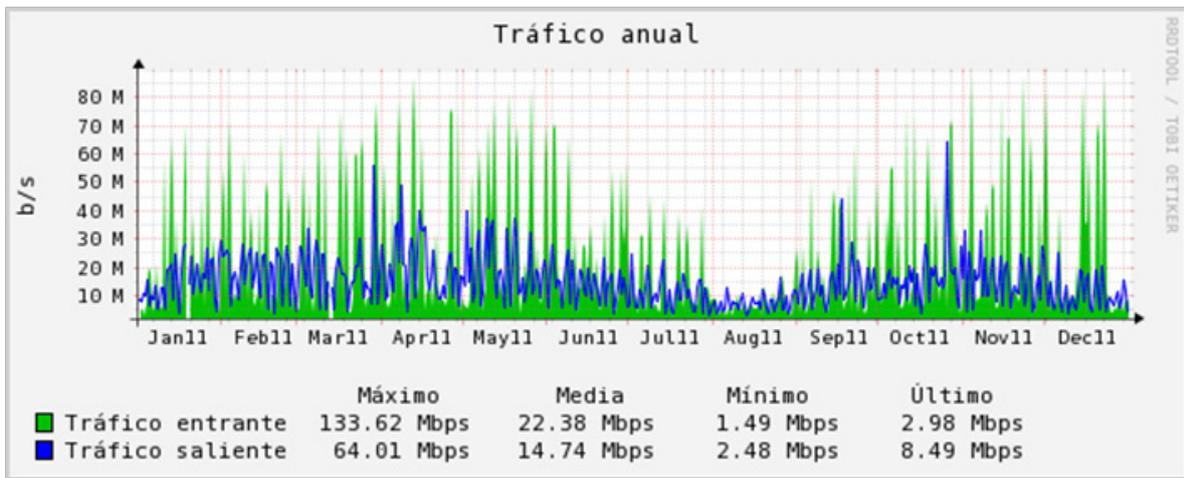


Figura 14: Provincia de Huelva (2011)

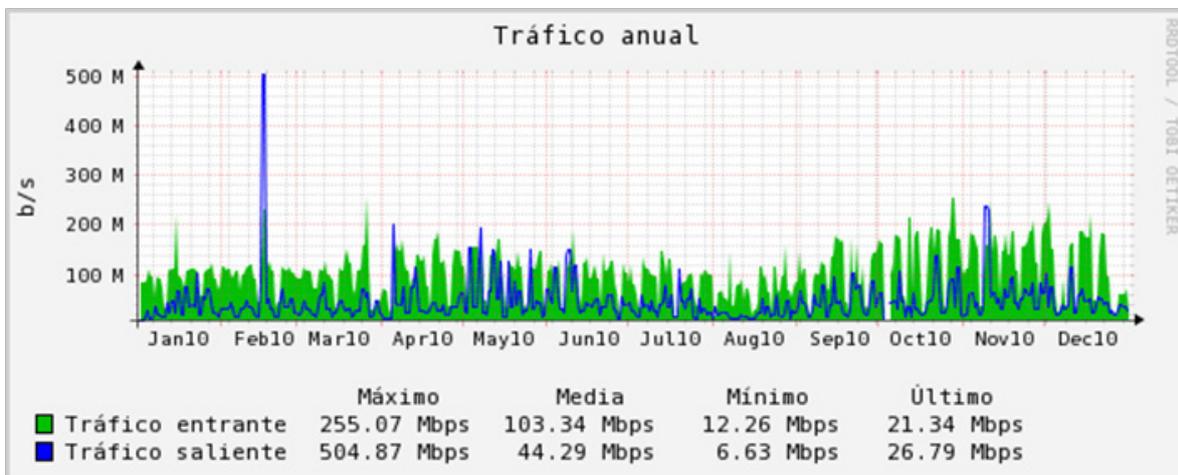


Figura 15: Provincia de Jaén (2010)

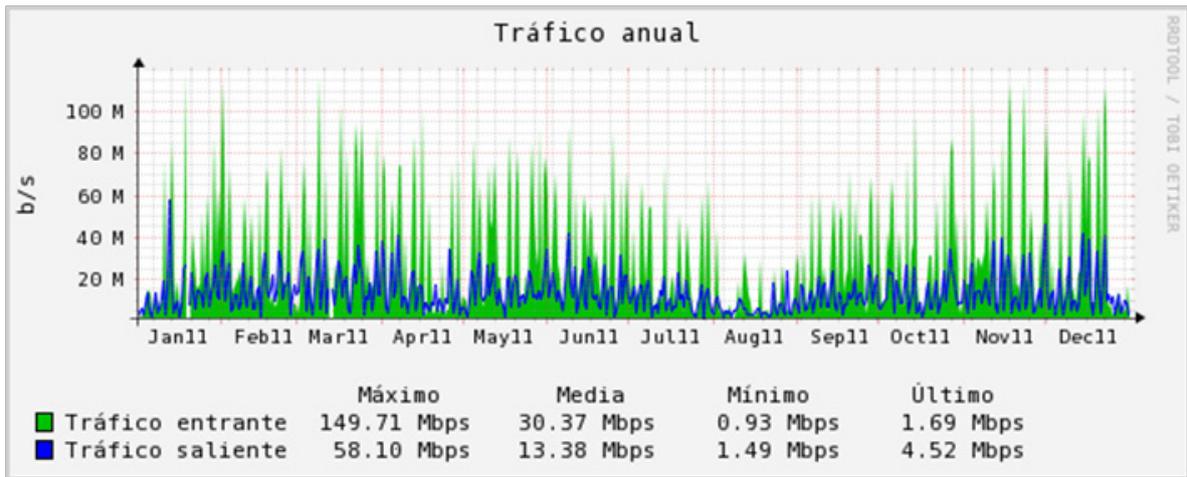


Figura 16: Provincia de Jaén (2011)

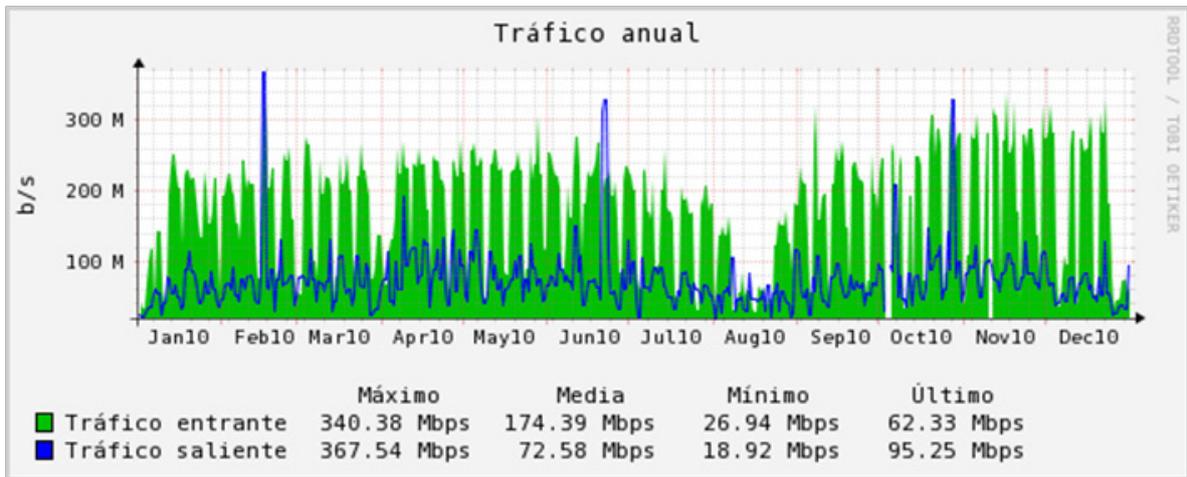


Figura 17: Provincia de Málaga (2010)

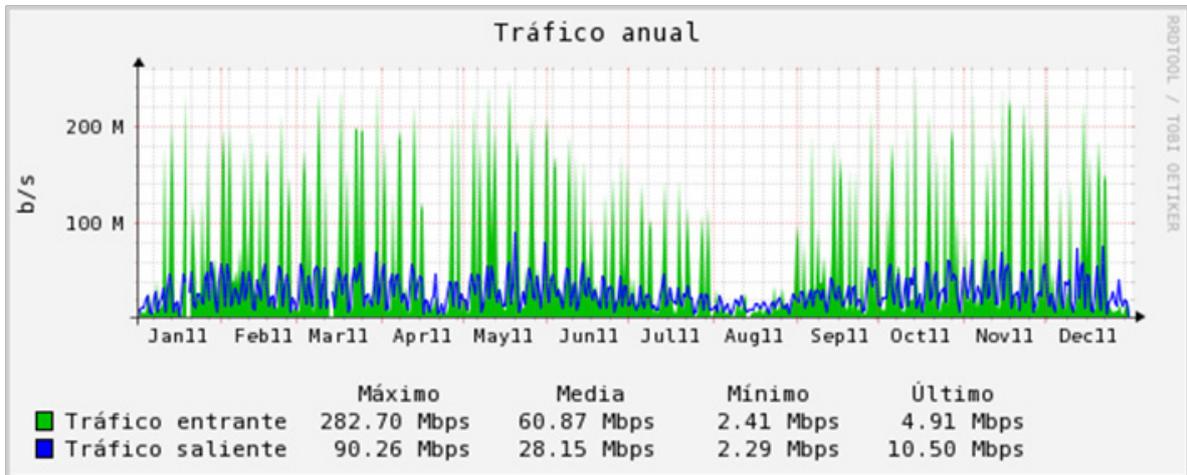


Figura 18: Provincia de Málaga (2011)

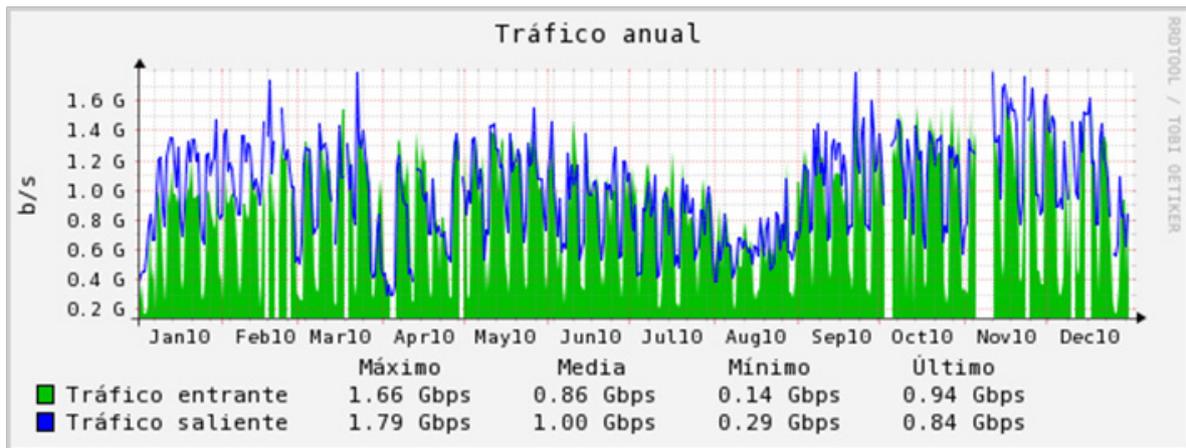


Figura 19: Provincia de Sevilla (2010)

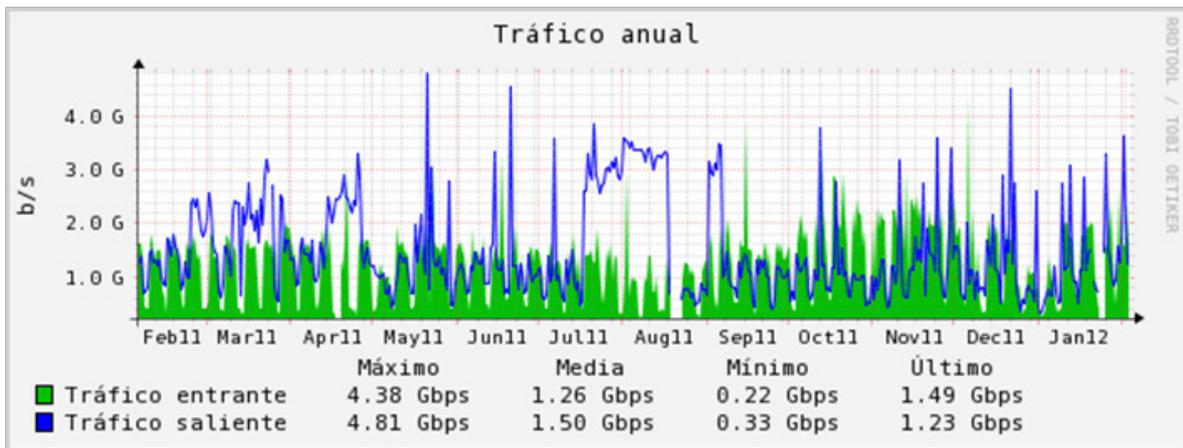


Figura 20: Provincia de Sevilla (2011)

Esto no es real, ya que la tendencia es a la inversa; el tráfico en todos los enlaces de la red experimenta un crecimiento anual, como consecuencia del mayor número de equipos conectados, la evolución de las aplicaciones y tecnología, aumento del tráfico en streaming, etc.

La razón de esta aparente disminución de tráfico es que los valores tomados para calcular las medias de tráfico en las gráficas de 2010 eran valores máximos de tráfico en las interfaces durante los intervalos de tiempo, mientras que los valores tomados para 2011, son valores medios en estos intervalos. De este modo, las gráficas representan un valor más próximo a la media de tráfico y en contraposición, muestra en menor medida la saturación de las líneas.

Por otro lado, desde 2011 y debido al aumento de tráfico

basado en el protocolo IPv6, desde CICA se empieza a monitorizar dicho tráfico, que se presenta en las gráficas 21 a 28.

5. MEJORAS EN INFRAESTRUCTURAS

5.1. Equipamiento

5.1.1. Instalación de las fuentes de alimentación y tarjetas procesadoras redundantes en los nodos provinciales.

Con el objetivo de garantizar la fiabilidad del servicio ofrecido por RICA en cada nodo provincial en 2010 se han instalado nuevas tarjetas procesadoras y fuentes de alimentación redundantes en todos los nodos.

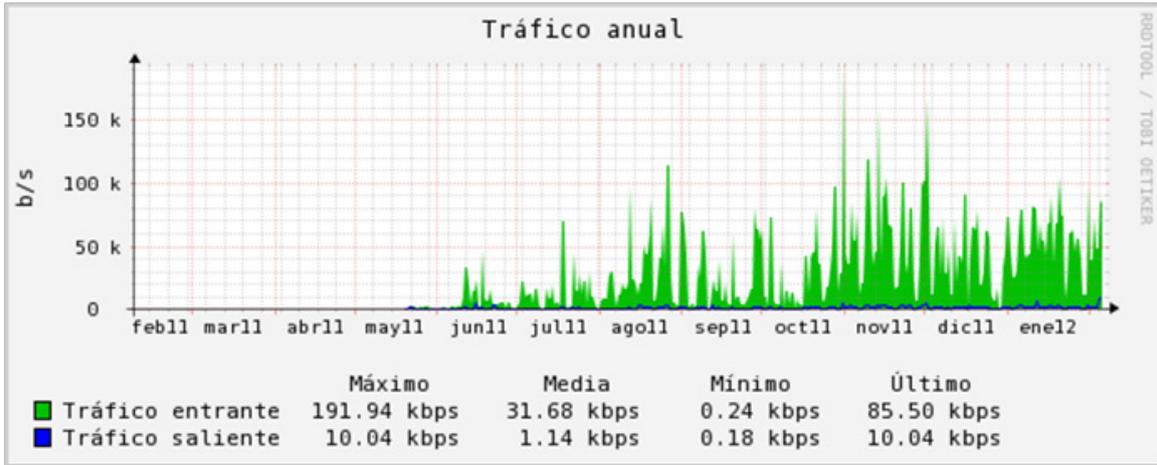


Figura 21: IPv6 - Provincia de Almería (2011)

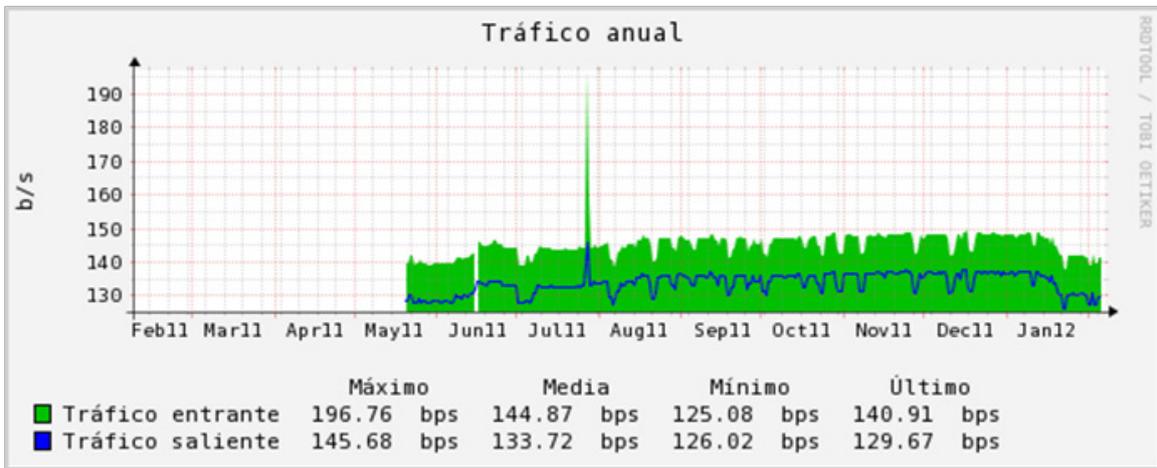


Figura 22: IPv6 - Provincia de Cádiz (2011)

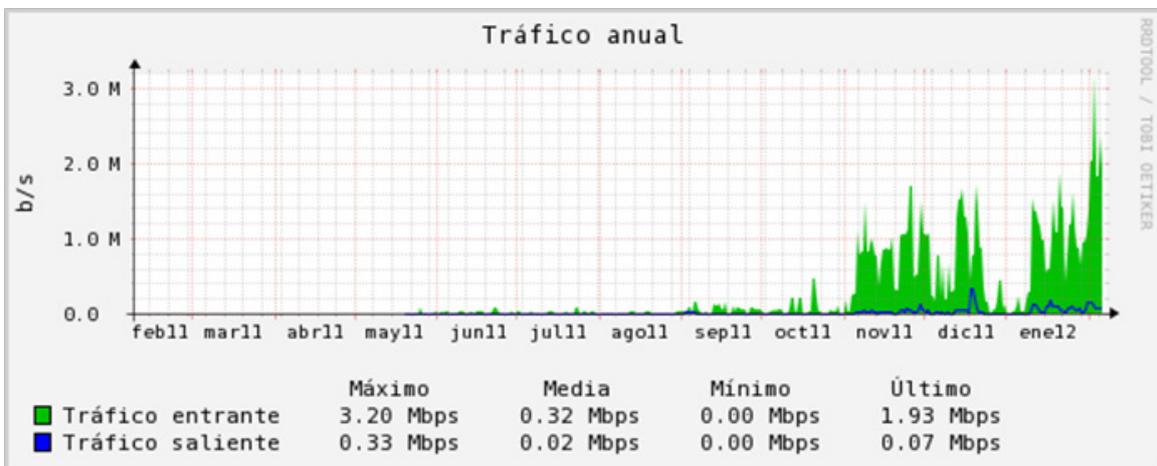


Figura 23: IPv6 - Provincia de Córdoba (2011)

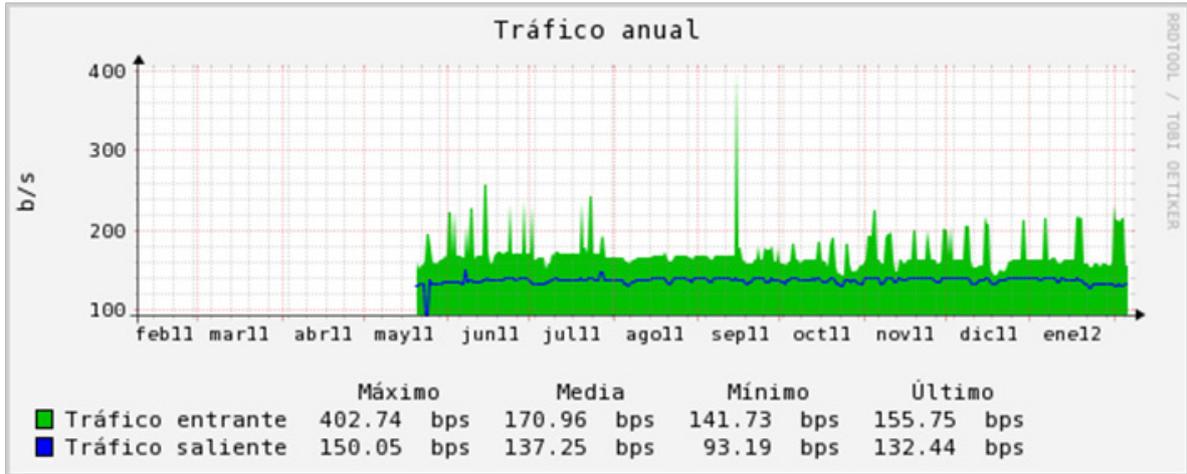


Figura 24: IPv6 - Provincia de Granada (2011)

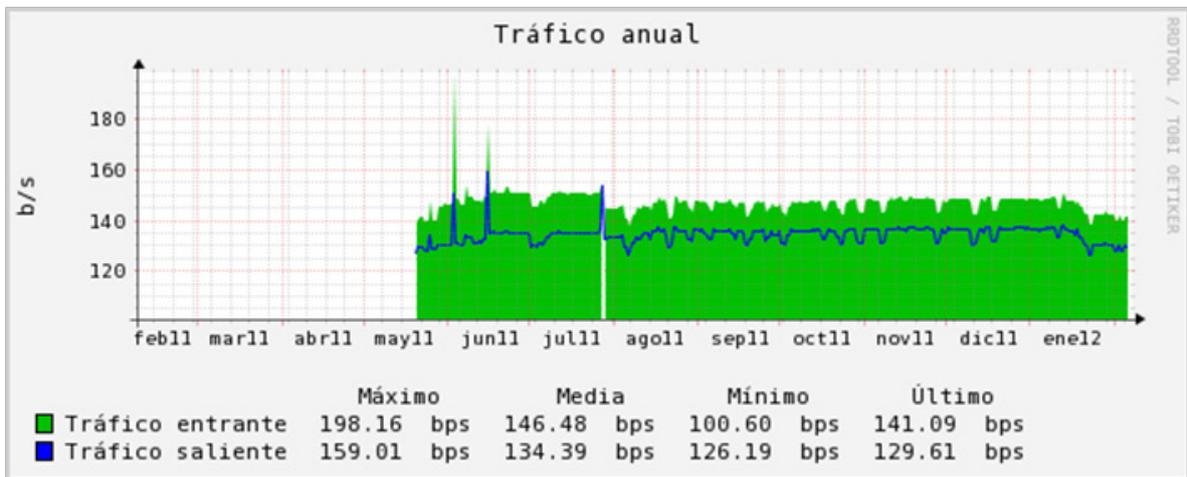


Figura 25: IPv6 - Provincia de Huelva (2011)

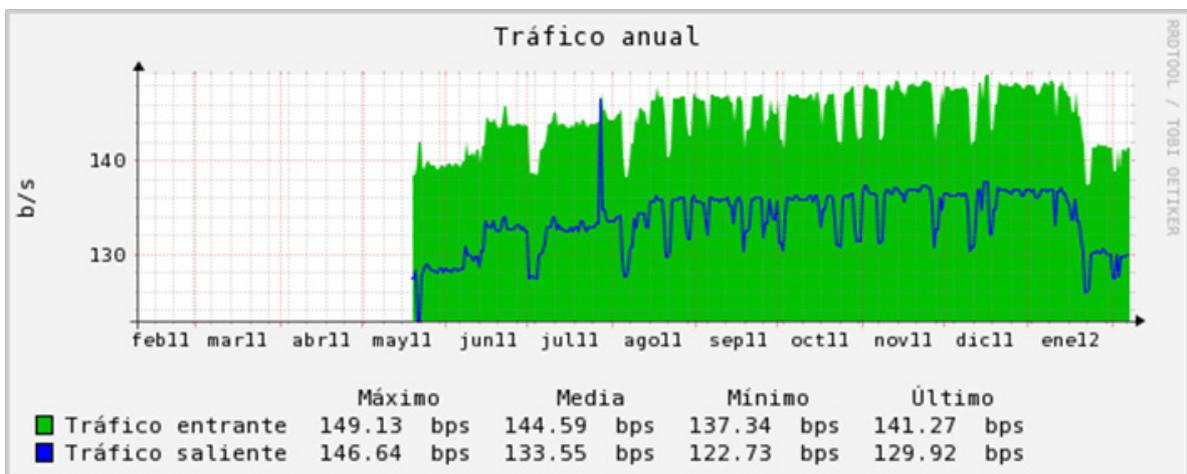


Figura 26: IPv6 - Provincia de Jaén (2011)

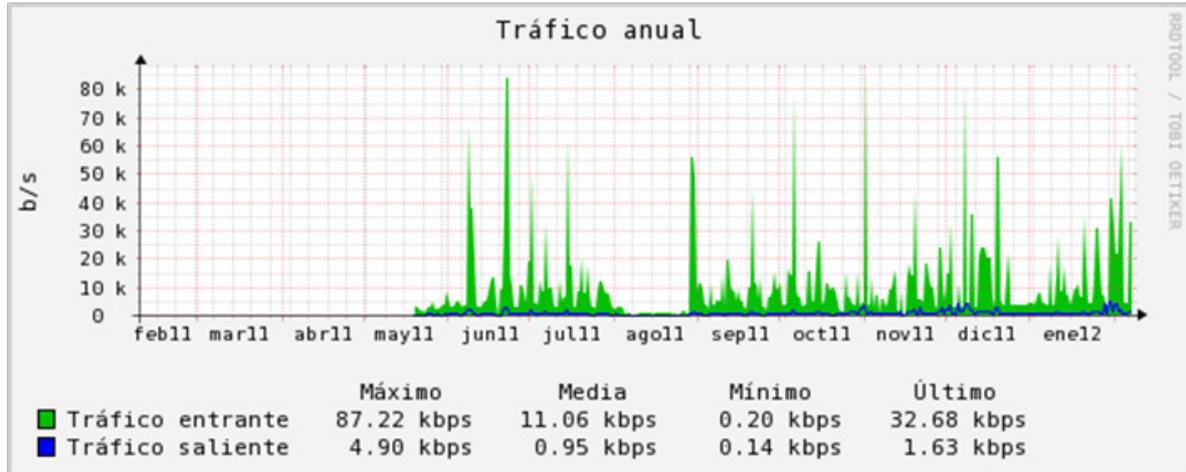


Figura 27: IPv6 - Provincia de Málaga (2011)

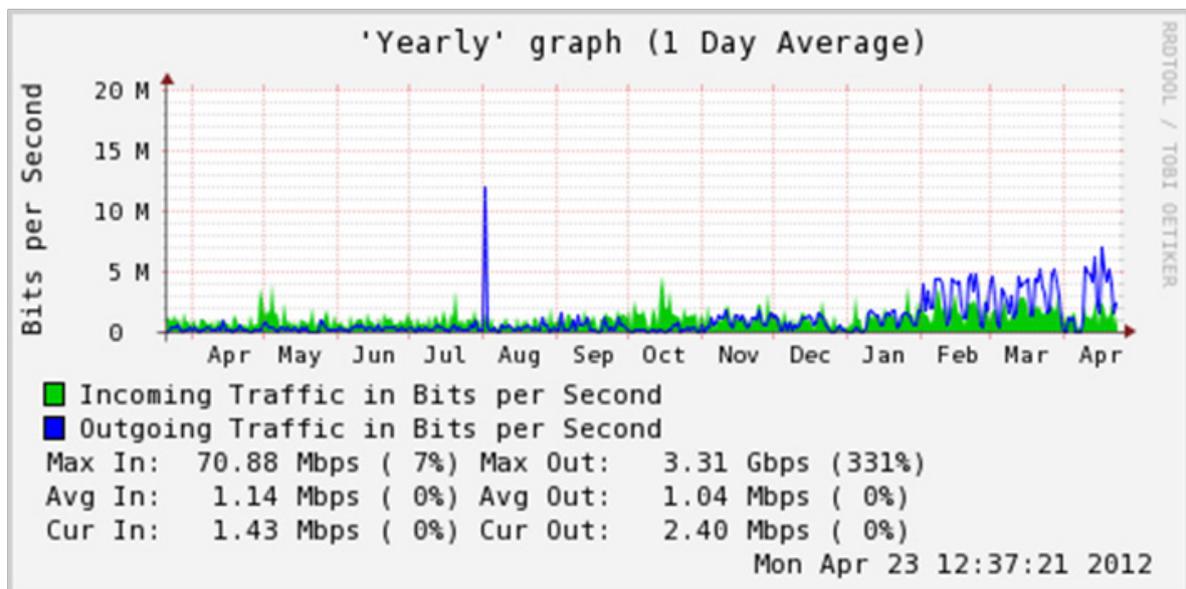


Figura 28: IPv6 - Provincia de Sevilla (2011)

En concreto han sido tarjetas procesadoras RE-S-1300-2048-S con 1300Mhz y 2Gb de memoria soportadas por los Juniper MX-240 instalados en cada nodo. Estas tarjetas están configuradas en modo “Hot-Standby o Espera Activa”, de forma que estén listas para tomar el control del equipo en caso de fallo de la tarjeta principal, sin que este hecho llegue a apreciarse en la calidad del servicio ofrecido.

De igual manera, se ha llevado a cabo la instalación de una segunda fuente de alimentación PS 1.2-1.7 kW; 100-240V AC in, que garantiza redundancia en la potencia que necesita cada Juniper MX-240 para su correcto funcionamiento.

5.1.2. Instalación de las nuevas tarjetas procesadoras en RT1 y RT2.

También en 2010 se han instalado dos nuevas tarjetas procesadoras WS-C6506 en los routers troncales. Estas tarjetas abren las puertas de VSS, una tecnología que permite tratar a los equipos troncales como un único “conmutador virtual”.

Adicionalmente, las nuevas tarjetas procesadoras capacitan a RICA para establecer BGP Full-Routing entre los routers troncales de RICA y los de RedIRIS, razón a la que, si unimos una mayor capacidad de procesamiento de rutas, garantiza una mejora ostensible del servicio ofrecido por la red RICA.

5.1.3. Instalación de nuevos nodos de RICA

Igualmente en 2010 se ha proporcionado conexión directa a RICA a dos nuevos nodos tras la instalación de equipos Cisco WS-C3750G-48TS-S en Multicentros Cartuja y en la Universidad Pablo de Olavide (UPO) respectivamente. De esta forma se equipara la conexión de estos 2 nodos a la de los nodos provinciales, estando conectados directamente a RICA.

Los equipos Cisco WS-C3750G-48TS-S cuentan con 48 puertos Ethernet 10/100/1000, capacidad para 4 SFP y los requisitos necesarios para conectarse a los equipos EDC de Telefónica que dan acceso a la Macrolan.

5.2. Configuraciones

5.2.1. VSS

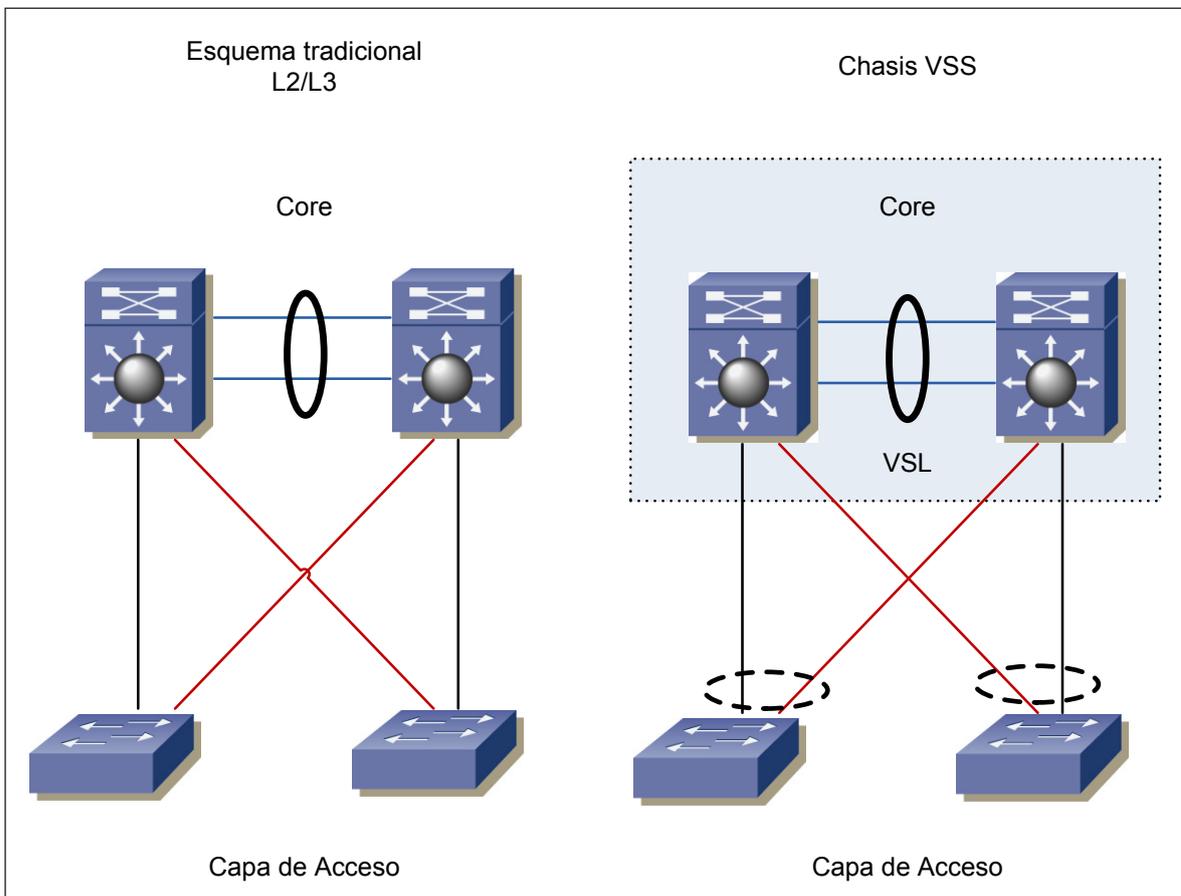
Se han virtualizado los Nodos centrales de RICA a través de la tecnología VSS (ver figura 29).

VSS es una tecnología, propietaria de CISCO, que convierte varios Cisco Catalyst 6500 en un único “conmutador virtual”.

tador virtual”. Esto conlleva un aumento de la eficiencia operativa aumentando el *throughput* del equipo. En la fase inicial de esta tecnología son dos equipos los que podrán operar como un único “conmutador virtual”. Una de las características necesarias para poder implantar esta tecnología es el modelo de supervisora de los equipos así como un módulo de transmisión. Concretamente necesita el modelo Virtual Switching Supervisor 720 10GE. Estas supervisoras hacen que el tanto el plano de datos como la matriz de conmutación con capacidad de 720 Gbps estén activas al mismo tiempo, por lo que se combinan pudiendo alcanzar un *throughput* de 1400 Gbps. Los dos chasis están sincronizados para que cualquiera de los dos pueda tomar el control del “chasis virtual” en caso de fallo.

Los beneficios del uso de VSS

- Aumenta la eficiencia operativa mediante la simplificación de red y reduce los gastos de administración del switch (un único punto de gestión, elimina la necesidad del uso de otros protocolos de “Hot Standby” como Hot Standby Router Protocol (HSRP) o Virtual Router Redundancy Protocol (VRRP).



Gráfica 29: Comparación entre el modelo tradicional y el modelo VSS

- Permite la creación de Etherchannels (802.3ad o Port Aggregation Protocol (PAgP)) multichassis, es decir, permite crear Etherchannels usando interfaces de dos equipos distintos físicamente, pero dentro del mismo “chasis virtual”.
- Maximiza la utilización de todos los puertos 10 Gigabit Ethernet en el “Virtual Chassis”.
- Escala el throughput del equipo hasta 1400 Gbps.

5.2.2. BGP full-routing

Durante el año 2010 se ha modificado la conexión del router troncal de RICA con RedIRIS a BGP full-routing, que ha sido posible gracias a las mejoras introducidas en el hardware de los routers troncales de RICA.

5.3. Servicios Multicast e IPv6

Todos los centros conectados a RICA disponen de conectividad Multicast e IPv6 y pueden disfrutar de todos los servicios ofrecidos a través de ellos. En el caso de multicast, destacamos el proyecto AccessGrid y las emisiones de Opera Oberta y Flamenco en Red.

En el año 2010 también se han realizado pruebas de multicast sobre IPv6 satisfactorias entre el CESCA y los Nodos de RICA en Jaén, Almería y CICA. Estas pruebas van a permitir la recepción de futuras emisiones de contenido multimedia utilizando el protocolo IPv6.

5.4. Local Internet Registry

A finales del año 2011, CICA ha terminado su registro como LIR (Local Internet Registry) en RIPE NCC (Réseaux IP Européens Network Coordination Centre). RIPE es la organización sin ánimo de lucro que proporciona la coordinación técnica para la infraestructura de Internet. Esta organización actúa como RIR (Regional Internet Registry) proporcionando recursos de Internet globales y servicios relacionados (IPv4, IPv6 y números de AS) a los miembros de la región del ámbito de RIPE NCC.

Un número de AS es un número de Sistema Autónomo (Autonomous System) que permite el enrutado con los grandes proveedores de Internet.

La comunidad de miembros de RIPE está formada principalmente por Internet Service Providers (ISPs), operadores de comunicaciones, entidades gubernamentales y grandes entidades localizadas en Europa, Oriente Medio y Asia Central (ver figura 30).

Como miembro de RIPE NCC, CICA puede:

- Solicitar espacios de direcciones IPv4, IPv6 y números AS (Autonomous System).
- Gestionar la información de registro de la RIPE DB (base de datos de RIPE).
- Certificar digitalmente los números de recursos de Internet.
- Asistir a los meetings de RIPE NCC, proporcionar feedback y votar las decisiones que se tomen a nivel de comunidad.
- Asistir a los cursos de formación de RIPE NCC.

Además, se ha solicitado y puesto en funcionamiento un AS público para la red RICA, agrupando el espacio de direcciones que RICA tiene asignado en un único bloque para que sea enrutable desde Internet.

elevado precio del alquiler de líneas de altas prestaciones condiciona la capacidad de transmisión de la red. Esto es un inconveniente para proyectos que requieren gran capacidad de transporte de datos, evitando que se lleven a cabo o encareciendo desmesuradamente su coste.

Por otro lado, RICA está condicionada por el catálogo de servicios que ofrece el proveedor, dependiendo de las características de su red y de su equipamiento.

Por ejemplo, en la red MacroLAN de Telefónica no existe soporte nativo para IPv6 ni para multicast, teniendo que recurrir a soluciones poco eficientes y difíciles de gestionar para poder ofrecer estos servicios a la comunidad científica y universitaria que los demandan.

6.2 Red Nacional RedIRIS-NOVA.

Actualmente se está desplegando una red de comunicaciones avanzada para la comunidad académica y de investigación española basada en tecnología de fibra óptica de alta capacidad (figura C13). La mencionada red sustituirá a la actual RedIRIS-10, dotándola de mayor velocidad y alcance. Al contrario que con la actual red, la nueva RICA extenderá sus nodos ópticos hasta muchos de los centros afiliados.

La nueva red mejorará la capacidad colaboradora de los investigadores andaluces y la colaboración con otros investigadores de fuera de Andalucía, y permitirá la investigación en campos con elevados requerimientos de ancho de banda. Además, la nueva red andaluza mejorará las posibilidades de localizar en Andalucía ciertas infraestructuras científico-tecnológicas europeas que requieren conexiones fiables de alta velocidad. Permitirá igualmente una conexión rápida y eficaz con las redes de otras comunidades y países vecinos.

6.3. RedIRIS-NOVA en Andalucía

En Andalucía se utilizará como medio físico la red de fibra nacional RedIRIS-NOVA. Como consecuencia de esto, resultará la nueva red académica y de investigación de Andalucía, RICA, formando parte de la infraestructura nacional de RedIRIS-NOVA. RICA tendrá conectividad con todas las demás redes autonómicas de España así como a las redes académicas y científicas de otros países, e Internet comercial.

7. CARACTERÍSTICAS DE LA NUEVA RED RICA

7.1. Especificaciones de la red.

Las necesidades intrínsecas de los centros adscritos a RICA hacen que no se comporte exclusivamente como una red de transporte, sino como una plataforma de servicios. La actividad investigadora propia de dichos centros, exige que el equipamiento de red de RICA ofrezca:

- Soporte de los protocolos de red que permitan la implementación de entornos adecuados para la puesta en marcha de proyectos de investigación.
- Rendimiento de conmutación y encaminamiento suficiente para las tasas de transmisión en los interfaces de red de los equipos.

Los objetivos de la red propuesta son:

- Mayor ancho de banda.
- Mayor fiabilidad.
- Fácil escalabilidad, tanto en ancho de banda como en servicios.
- Mayor simplicidad en la gestión y operación de la red.

7.2. Catálogo de servicios

Los servicios a ofrecer dentro de la nueva red RICA requieren de la adquisición del nuevo equipamiento. Estos servicios que se proporcionarán son:

- Ancho de banda bajo demanda
- Conectividad de alta capacidad IPv4 e IPv6
- Conexiones punto a punto
- Difusión multicast sobre IPv4 e IPv6
- Servicio de DNS
- Movilidad: acceso federado a la red para usuarios móviles
- MPLS (Multi-Protocol Label Switching)
- NOC – equipo de gestión de red
- Telefonía IP
- Videoconferencias: H.323 Global Dialing Scheme (GDS)
- VPN-L2 (punto a punto) o VPLS (multipunto a multipunto)
- VPN-L3

Todos ellos se describen en los siguientes subapartados.

7.2.1. Ancho de banda bajo demanda.

Las instituciones pueden solicitar enlaces físicos dedicados bajo demanda, de hasta 10Gb/s con otra institución que esté a su vez conectada bien a RICA, RedIRIS o a alguna de las redes europeas situadas en la huella de fibra oscura de GÉANT2. Se utiliza la tecnología 10GE WAN-PHY como tecnología de transporte.

Este hecho soluciona el problema de las redes congestionadas por el alto volumen de tráfico cursado, haciendo la operación del circuito más rápida y más flexible. Controlando y balanceando el uso de la red por aplicaciones de alta capacidad se puede categorizar flujos de tráfico, consiguiendo así la calidad de servicio requerida.

El usuario puede reservar previamente la capacidad necesitada para transferir grandes cantidades de información, de forma rápida y efectiva.

7.2.2. Conectividad de alta capacidad IPv4 e IPv6.

El CICA pone a disposición de las Instituciones afiliadas una red de comunicaciones de alta capacidad, que permite a la comunidad académica e investigadora andaluza participar en proyectos de investigación nacionales e internacionales donde se requieran servicios avanzados de comunicaciones.

El servicio de conectividad proporciona a las instituciones afiliadas un acceso básico a Internet. Es necesario hacer uso de este servicio de conectividad para poder tener acceso al resto de servicios.

La red troncal soporta, de forma nativa, los protocolos IPv4 e IPv6 con una configuración en "dual-stack". El servicio de conectividad incluye esta tecnología y para poder hacer uso de ella es necesario que la institución disponga de direccionamiento IPv4 e Ipv6.

7.2.3. Conexiones punto a punto.

La red ofrece la posibilidad de reservar conexiones virtuales punto a punto sobre ella, para facilitar el trabajo del usuario, así como evitar la congestión en el enlace por la acción de otros usuarios. Este hecho permite crear conexiones de alta capacidad y calidad de servicio diferenciada.

7.2.4. Difusión multicast sobre IPv4 e IPv6.

El servicio de multicast optimiza el uso del ancho de ban-

da en aplicaciones multimedia, ya que permite el envío de tráfico hacia muchos destinos sin tener que replicar la información para cada receptor. La difusión multicast se realiza de modo nativo, lo que beneficia la configuración de las conexiones.

Entre las transmisiones que se han realizado utilizando la red, destacan los proyectos Òpera Oberta, Flamencoenred y Access Grid, que enlaza varios grupos geográficamente distantes a través de una red de gran ancho de banda, mediante recursos multimedia de audio y vídeo de gran calidad.

7.2.5. Servicio de DNS.

El CICA ofrece a sus instituciones el servicio de DNS o servidor de nombres de dominios, a través de la infraestructura de red existente y empleando para ello sus propios servidores.

7.2.6. Movilidad: acceso federado a la red para usuarios móviles.

Eduroam es el servicio mundial de movilidad segura desarrollado para la comunidad académica y de investigación. Este servicio permite que estudiantes, investigadores y personal de las instituciones participantes tengan conectividad Internet a través de su propio campus y cuando visitan otras instituciones participantes.

RICA, como red académica y de investigación participante en la iniciativa, permite el acceso federado eduroam a través de su infraestructura.

7.2.7. MPLS (Multi-Protocol Label Switching).

La conmutación de etiquetas multiprotocolo (Multiprotocol Label Switching), o MPLS es un tecnología usada para establecer caminos conmutados de etiquetas (Label Switching Paths) sobre redes IP, ya sea de forma manual o automática. Las etiquetas añadidas a cada paquete permiten una transmisión más rápida y fluida en la red, comparándolo con el método tradicional IP.

MPLS permite implementar los siguientes servicios sobre una red IP:

- Redes privadas virtuales de nivel 3.
- Redes privadas virtuales de nivel 2, ya sean punto a punto o conexiones multipunto a nivel 2.

- Ingeniería de tráfico (MPLS-TE), usada para controlar las rutas tomadas por los paquetes de datos, orientada a optimizar el uso de la red y evitar puntos de congestión en la red.

7.2.8. NOC – equipo de gestión de red.

El equipo de gestión de red ofrece un servicio de atención de incidencias y solicitudes de red a las instituciones afiliadas a RICA. Además, es el encargado de implementar, gestionar y monitorizar los servicios de la red.

Para interactuar con los usuarios, se ha creado un sistema de tickets, cuyo objetivo es el de notificar averías y tareas programadas producidas en el servicio en el caso de que éstas afecten a partes significativas del mismo (nodos de la red, líneas troncales, o conectividad internacional).

7.2.9. Telefonía IP.

La telefonía IP entre las universidades y centros de investigación andaluces será también otro de los servicios que va a proporcionar la nueva red de fibra.

La Voz sobre IP (VoIP, Voice over IP) es una tecnología que permite la transmisión de la voz a través de redes IP en forma de paquetes de datos.

La Telefonía IP es una aplicación inmediata de esta tecnología, de forma que permita la realización de llamadas telefónicas ordinarias sobre redes IP u otras redes de paquetes utilizando un PC, gateways y teléfonos estándares. En general, servicios de comunicación (voz, fax, aplicaciones de mensajes de voz, etc) que son transportados vía redes IP, Internet normalmente, en lugar de ser transportados por la red telefónica convencional.

Esta tecnología utiliza la red de datos existente para curar el tráfico de voz, lo que permite la comunicación entre los usuarios de la red directamente, sin utilizar los servicios de ningún proveedor de telefonía.

7.2.10. Videoconferencias: H.323 Global Dialing Scheme (GDS).

El servicio de videoconferencia posibilita la celebración de reuniones punto a punto o reuniones multipunto de diferentes grupos de trabajo distantes geográficamente, proporcionando un ahorro de tiempo y de costes de desplazamiento a los investigadores de las instituciones conectadas a la red. Las herramientas que se pueden usar van desde el audio y el vídeo a las pizarras electrónicas y

los editores de texto compartidos.

La tecnología soportada es H.323, utilizando la red existente. Dentro de este marco, el servicio de numeración GDS permite:

- La implementación de un plan de numeración universal que permita la interconexión de zonas H.323 adoptando por ello el plan que se utiliza en las Redes Nacionales de I+D del resto del mundo.
- Simplificar la realización de videoconferencias dentro de la comunidad.
- Facilitar la convergencia tecnológica entre redes de conmutación de paquetes y circuitos al permitir la integración con planes de numeración de telefonía convencional.
- Proporcionar la infraestructura que dé soporte al plan de numeración así como los recursos para la gestión de las nuevas incorporaciones.
- Facilitar la realización y organización de videoconferencias y VoIP.

7.2.11. VPN-L2 (punto a punto) o VPLS (multipunto a multipunto).

Este servicio permite interconectar a usuarios localizados en distintos puntos geográficos como si se encontraran en la misma LAN, permitiendo así realizar una gestión global de su red de usuarios pasando por alto la capa IP introducida por la red del proveedor a través de la cual se comunican los usuarios finales.

El servicio VPLS, Virtual Private LAN Switching, permite que todos los elementos conectados a la red en las diferentes sedes funcionen como si estuvieran conectados en la misma LAN.

Pueden prestarse servicios de VPNs de nivel 2 haciendo uso de la conexión de RICA con RedIRIS, de manera que los servicios de RICA pueden extenderse al resto de la geografía española, incluso europea y mundial. Así pues, sería factible una extensión de LAN mediante VPLS entre las redes de una universidad y cualquier otra o la dedicación de enlaces de gran ancho de banda para proyectos de colaboración europeos entre centros tecnológicos ubicados en distintos países.

La tecnología utilizada para implementar este servicio son túneles MPLS (LSPs). La funcionalidad de MPLS traffic-engineering (MPLS-TE) se utiliza para controlar la ruta que siguen los túneles de forma que se optimice la utilización de la red.

7.2.12. VPN-L3.

Los servicios VPN L3 permiten por un lado provisionar de forma sencilla VPNs sobre entornos heterogéneos en cuanto a tecnologías de acceso. Y por otra parte, posibilita una mayor simplicidad de gestión al delegar en la red

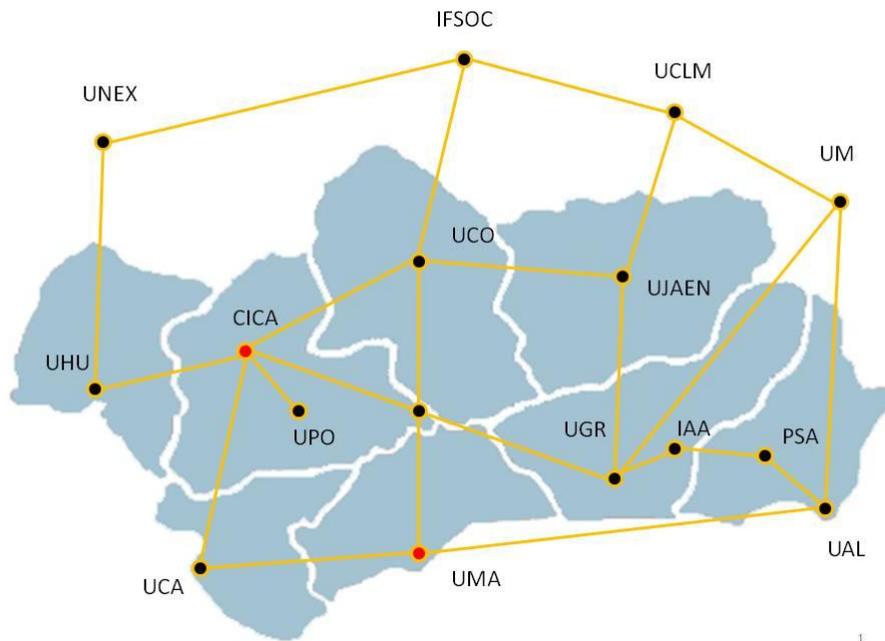


Figura 32: Topología física de la red

el routing entre distintas sedes y departamentos así como la gestión de seguridad (firewalls centralizados), NAT, etc. Igualmente permiten ofrecer servicios diferenciados y soportar mecanismos de garantía de ancho de banda por servicio.

Las VPN-L3 RFC2547bis se basan totalmente en estándares y constituyen la solución ideal en redes con topologías variadas u orientadas a conexiones “todos-con-todos”.

7.3. Diseño de la nueva red RICA.

En la figura 32 se muestra la huella de fibra óptica desplegada por RedIRIS en el proyecto RedIRIS-NOVA en la comunidad autónoma de Andalucía, así como los “Puntos de Presencia” o PdP de RedIRIS, donde se instalan los equipos de transmisión ópticos.

Sobre esta red óptica se desplegará la red IP de RICA, cuya topología se muestra en la figura 33. En cada uno de los PdP de RedIRIS-NOVA en Andalucía irá alojado un rou-



Imagen 34: Bastidor rectificador

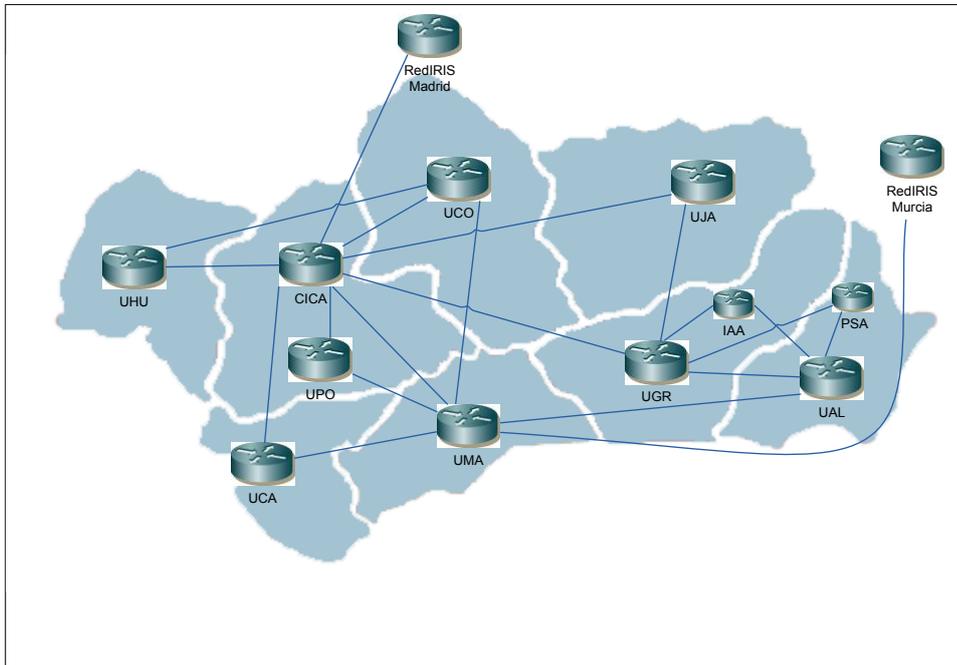


Figura 33: Topología IP de RICAnova

ter de RICA, dando lugar a la red IP que interconectará a todas las Universidades y centros de investigación de nuestra comunidad.

7.3.1. Nodos centrales de RICA,

Los equipos ubicados en estos nodos cursan el tráfico agregado de la red completa. A través de estos puntos pasa el tráfico destinado a Internet y a otras redes académicas nacionales e internacionales, por lo que son un punto crítico en la red y deben estar correctamente redundados. Además de las conexiones del troncal de la red, estos equipos también deben soportar las conexiones de los centros afiliados a RICA en esa provincia.

Por otra parte, el equipamiento de estos nodos (ver imágenes de la 34 a la 37) debe ser capaz de soportar todo el tráfico de la red autonómica, ya que en el caso de que se produzca un fallo en su homólogo, el tráfico total de RICA se encaminará a través de él. Esto es posible porque cada Nodo tiene conexión con cada uno de los dos nodos centrales, sea de manera directa o a través de otro nodo IP, pudiendo intercambiar el punto de salida dinámicamente mediante protocolos de enrutamiento. También se puede realizar de forma administrativa para balancear tráfico y repartir la carga entre ambos nodos.

Por ello, cada uno de estos nodos debe permitir la agre-



Imagen 35: Bastidor óptico



Imagen 36: Bastidor de fibra óptica

gación de líneas de alta velocidad. Por un lado las provenientes desde los Nodos provinciales de RICA, de interconexión con el otro NPR y dos la conexión con la red nacional. Se estima que para conectar los nodos centrales con RedIRIS IP es suficiente con dos enlaces de 10Gb/s, ampliables bajo demanda si el tráfico troncal lo requiriese.

7.3.2. Nodos provinciales de Almería, Cádiz, Córdoba, Granada, Huelva y Jaén.

Estos nodos se conectan a los nodos centrales de RICA mediante una línea directa con cada uno, actuando una de principal y otra de respaldo. Estas conexiones serán de 10Gb/s, aunque deben poder ampliarse en un futuro hasta los 40Gb/s o 100Gb/s, formando troncal de RICA de 40Gb/s – 100Gb/s.

7.4. Estado actual del despliegue.

El despliegue de la nueva Red está basado en la realización de una serie de hitos que se están llevando a cabo

desde comienzos del año 2010 y se espera finalizarlos a mediados de 2012. Estos hitos consisten en realizar los replanteos en las sedes donde se ubicarán los PdPs, la entrada de la fibra óptica diversificada, el acondicionamiento de los CPD que no cumplan con las especificaciones, la designación de la ubicación definitiva del equipamiento de red, elaboración de la documentación de las condiciones técnicas, la firma del acuerdo de alojamiento, la instalación de los bastidores (fibras, rectificador y equipos ópticos) y las pruebas de la infraestructura una vez desplegada.

Desde el CICA se están completando muchos de estos hitos y en los que no dependen exclusivamente de l equipo de comunicaciones del centro, se actúa como coordinadores de las diferentes partes involucradas en el proyecto para que dichos hitos se completen de la manera más rápida y satisfactoria posible.

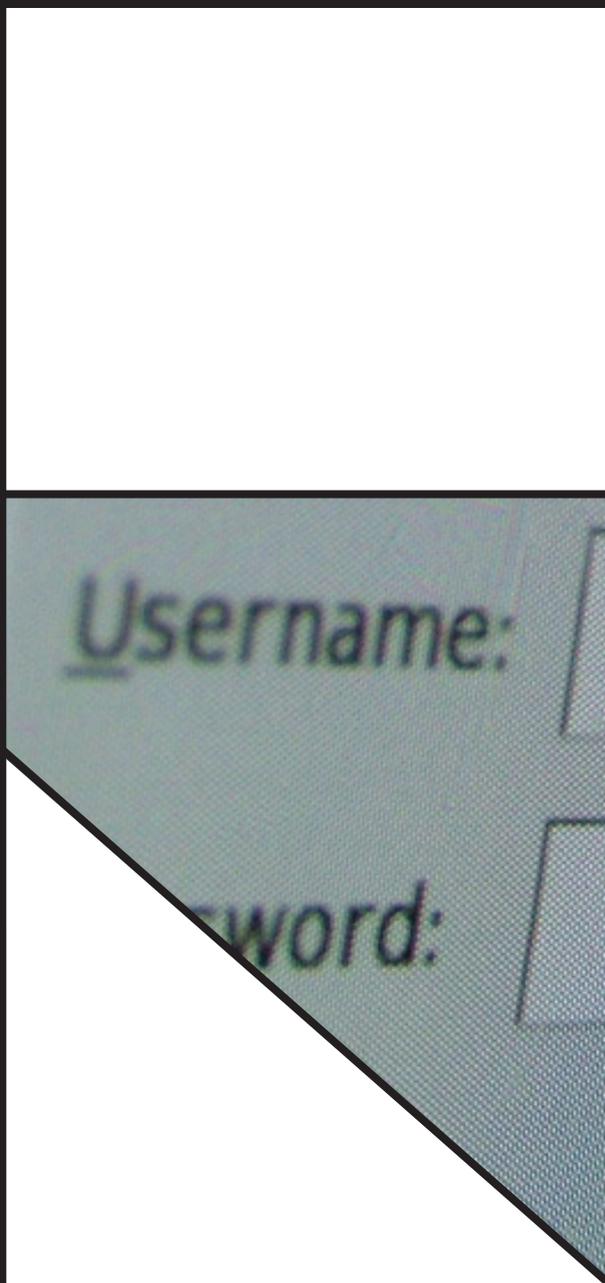
En las dos tablas de la página siguiente se detalla el nivel de ejecución del despliegue tanto a finales de 2010 como a finales de 2011.

PdP	Replanteo	Entrada de fibra al CPD	Acondicionamiento CPD			Designación del sitio de huella	Documentación condiciones del PdP	Firma acuerdo alojamiento	Instalación de bastidores		
			Aire	Suelo	Incendio				Óptico	Rectificador	Equipos
UAL	Realizado	2 caminos de fibra	Realizado	Realizado	Realizado	Realizado	Validada por institución	Pendiente	Pendiente	Pendiente	Pendiente
UCA	Realizado	En proceso	Realizado	Realizado	Realizado	Realizado	Pendiente obra	Pendiente	Pendiente	Pendiente	Pendiente
CICA	Realizado	2 caminos de fibra	Realizado	Realizado	Realizado	Realizado	Enviada a RedIRIS	Pendiente	Pendiente	Pendiente	Pendiente
UCO	Realizado	2 caminos de fibra	Realizado	Realizado	Realizado	Realizado	Validada por institución	Pendiente	Realizado	Realizado	Realizado
UGR	Realizado	En proceso	Realizado	Realizado	En proceso	Realizado	Pendiente obra	Pendiente	Pendiente	Pendiente	Pendiente
UHU	Realizado	2 caminos de fibra	Realizado	Realizado	Realizado	Realizado	Validada por institución	Pendiente	Realizado	Realizado	Realizado
IAA	Realizado	2 caminos de fibra	Realizado	Realizado	Realizado	Realizado	Validada por institución	Pendiente	Realizado	Pendiente	Pendiente
UJAEN	Realizado	En proceso	Realizado	Realizado	Realizado	Realizado	Validada por institución	Pendiente	Pendiente	Pendiente	Pendiente
UMA	Realizado	2 caminos de fibra	Realizado	Realizado	Realizado	Realizado	Validada por institución	Pendiente	Realizado	Realizado	Realizado
PSA	Realizado	1 camino de fibra, pendiente diversificación	Realizado	Realizado	Realizado	Realizado	Validada por institución	Pendiente	Realizado	Realizado	Realizado
UPO	Realizado	En proceso	Pendiente	Pendiente	Pendiente	Realizado	Pendiente obra	Pendiente	Pendiente	Pendiente	Pendiente

Datos de final de 2010

PdP	Replanteo	Entrada de fibra al CPD	Acondicionamiento CPD			Designación del sitio de huella	Documentación condiciones del PdP	Firma acuerdo alojamiento	Instalación de bastidores		
			Aire	Suelo	Incendio				Óptico	Rectificador	Equipos
UAL	Realizado	2 caminos de fibra	Realizado	Realizado	Realizado	Realizado	Validada por institución	Pendiente	Realizado	Realizado	Realizado
UCA	Realizado	2 caminos de fibra	Realizado	Realizado	Realizado	Realizado	Validada por institución	Pendiente	Realizado	Realizado	Realizado
CICA	Realizado	2 caminos de fibra	Realizado	Realizado	Realizado	Realizado	Enviada a RedIRIS	Pendiente	Realizado	Realizado	Realizado
UCO	Realizado	2 caminos de fibra	Realizado	Realizado	Realizado	Realizado	Validada por institución	Pendiente	Realizado	Realizado	Realizado
UGR	Realizado	2 caminos de fibra	Realizado	Realizado	Realizado	Realizado	Validada por institución	Pendiente	Realizado	Realizado	Realizado
UHU	Realizado	2 caminos de fibra	Realizado	Realizado	Realizado	Realizado	Validada por institución	Pendiente	Realizado	Realizado	Realizado
IAA	Realizado	2 caminos de fibra	Realizado	Realizado	Realizado	Realizado	Validada por institución	Pendiente	Realizado	Realizado	Realizado
UJAEN	Realizado	2 caminos de fibra	Realizado	Realizado	Realizado	Realizado	Validada por institución	Pendiente	Realizado	Realizado	Realizado
UMA	Realizado	2 caminos de fibra	Realizado	Realizado	Realizado	Realizado	Validada por institución	Pendiente	Realizado	Realizado	Realizado
PSA	Realizado	2 caminos de fibra	Realizado	Realizado	Realizado	Realizado	Validada por institución	Pendiente	Realizado	Realizado	Realizado
UPO	Realizado	2 caminos de fibra	Realizado	Realizado	Realizado	Realizado	Validada por institución	Pendiente	Realizado	Realizado	Realizado

Datos de final de 2011



Seguridad

El Centro Informático Científico de Andalucía ha seguido prestando durante 2010 y 2011 servicios de asesoría en materia de seguridad a los centros de investigación y universidades conectados a la red RICA+. Siguiendo las recomendaciones de las normativas internacionales, se ha implantado un servicio de tracking de incidencias y base de datos de conocimiento, a fin de disponer de un aprendizaje posterior al incidente con el fin de evitar futuros problemas que ya hayan sido solventados.

En materia de despliegue de IPv6, en 2010 el servidor de nombres del Centro Informático Científico de Andalucía responde de forma nativa a peticiones del nuevo protocolo siendo uno de los centros pioneros de España. En este mismo ámbito se están investigando las implicaciones de seguridad que conlleva así como la implantación de firmas en las zonas distribuidas a los centros.

Otra de las iniciativas llevadas a cabo gracias al Servicio de Identidad de RedIRIS, es el Servicio de Certificados Personales, a través del cual los usuarios del CICA dispondrán de certificados de usuario firmados por una entidad reconocida con los que podrán firmar y cifrar todas sus transacciones electrónicas.



GESTIÓN DE INCIDENTES

El Centro Informático Científico de Andalucía desempeña el rol de nexo de unión entre los distintos centros educacionales y/o de investigación de la comunidad andaluza. Todos ellos dependen en direccionamiento de CICA, debido a la estructura jerárquica del sistema de nombres de dominio (DNS), y por lo tanto somos el primer filtro por el que pasan las notificaciones de incidentes de seguridad u otro tipo, antes de acabar en su destino final.

1. GESTIÓN CENTRALIZADA DE INCIDENTES DE SEGURIDAD.

La ingente tarea de la gestión de incidentes de seguridad puede convertirse en todo un problema si se hace de manera manual y sin automatismos, ya que son cientos los incidentes recibidos cada semana debido al gran número de usuarios existentes bajo el paraguas de la red RICA+.

El departamento de Seguridad de CICA además de detectar problemas de seguridad informática y enviarlos a las entidades correspondientes recibe a su vez notificaciones sobre posibles incidentes de cinco fuentes distintas principalmente:

- RedIris
- Detectores de spam
- Universidades
- Gestoras defensoras de derechos de autor
- ISPs externos

- Órganos judiciales y fuerzas de seguridad del Estado

La solución a este problema se haya en un sistema estable de gestión centralizada de incidencias, así se puede contar con un repositorio común de problemas y sus soluciones para casos futuros, y todo un equipo de seguridad puede supervisar de manera coordinada las alertas recibidas.

De entre todo el software existente para estas labores se eligió Request Tracker (utilizado ampliamente por RedIris) debido a las características que pone a nuestra disposición:

- Panel de control vía web
- Basado en gestión a través de correo electrónico
- Visualización rápida del estado de los casos
- Potente y rápido sistema de búsquedas
- Fácil de interconectar a otros sistemas
- Altamente configurable y modificable al estar desarrollado en Perl
- Permite gestión de usuarios y grupos
- Permite establecer dependencias entre incidentes
- Posibilidad de añadir módulos externos (Perl)

Durante el año 2010 se han recibido 1173 incidencias sobre infracción de derechos de autor y 189 de incidentes de seguridad relativos a Iris-Cert. En la figura 37 se observa la progresión.



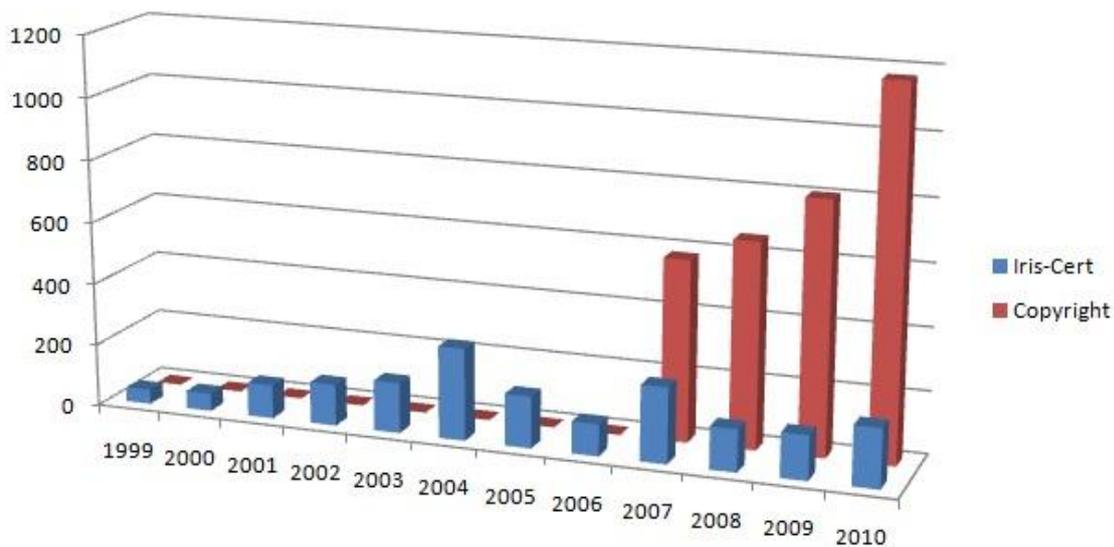


Figura 37: Incidentes de Seguridad en 2010

La razón del aumento de incidencias recibidas a lo largo de los años viene dada por la mejora en los mecanismos de detección en los motores de los sistemas de detección de intrusos (IDS) y también por el aumento de la actividad de los usuarios malintencionados a través de las redes sociales, botnets, etc.

Asimismo, CICA colaboran con otros agentes externos como INTECO, en su despliegue nacional de sondas de detección de nuevas alertas y patrones de ataque.

2. ANÁLISIS DE VULNERABILIDADES.

La introducción de nuevos servicios y funcionalidades en el mercado de las tecnologías de información es una realidad que sucede a un ritmo vertiginoso, y los responsables

en el campo de la seguridad informática somos responsables de auditar dichas novedades a fin de dotar a las organizaciones de un nivel de seguridad óptimo de acuerdo a las necesidades del negocio. Aunque la información nos desborda hemos de ser capaces de seguir ciertos mecanismos automatizados para llevar a cabo una buena labor de monitorización.

Siempre se debe intentar alcanzar un punto medio entre usabilidad y seguridad, y un desequilibrio no debería ser admisible en ninguno de los dos casos. Para hallar el punto de equilibrio entre ambos es necesario conocer los riesgos que implican la introducción de nuevas funcionalidades y establecer una serie de controles acordes al nivel del riesgo. Las vulnerabilidades son debilidades de seguridad asociadas a los sistemas de información de la organización. No constituyen por sí solas un daño, debe existir una amenaza (usuario malintencionado, error humano,

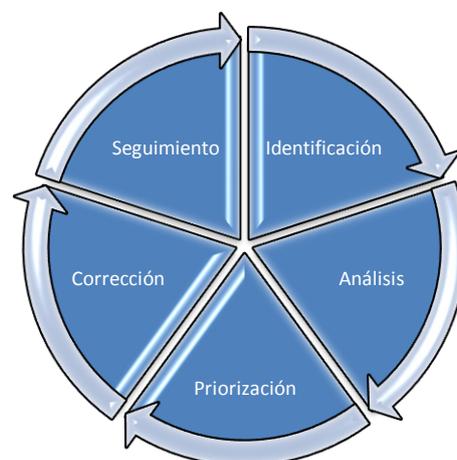


Figura 38: Método de trabajo frente a ataques

etc.) que las aproveche para afectar al funcionamiento de un activo. Es de vital importancia el trabajo proactivo en identificación, análisis, priorización, corrección y seguimiento de vulnerabilidades como método de trabajo para ir un paso por delante de los posibles atacantes (ver figura 38).

Los análisis de vulnerabilidades deben ser siempre ejecutados de manera periódica con el fin de detectar inconsistencias entre los resultados que se esperan tener y lo que realmente existe. De manera excepcional, se deben realizar análisis por demanda cuando haya cambios en el diseño de la red o de los sistemas o se realicen actualizaciones en los dispositivos.

El Departamento de Seguridad del CICA trabaja siguiendo la metodología anterior y realiza escaneos periódicos sobre los diversos equipos que conforman su dominio de control y los reporta a los administradores de sistemas con el fin de detectar posibles vulnerabilidades en el software instalado, puertas traseras o servicios no autorizados, asesorando posteriormente a los responsables sobre las medidas que deberían tomar para minimizar el riesgo.

3. DESPLIEGUE DE IPv6 EN DNS.

En los últimos meses de 2010 se anuncia, incluso con alertas, del fin de las direcciones de Internet tal y como las conocemos hasta ahora. Desde los distintos agentes registradores de cada continente se ha avisado de la indisponibilidad de nuevas direcciones IPv4. Los motivos son básicamente dos: el auge de dispositivos móviles conectados a Internet de manera permanente y el uso ineficiente de las direcciones IPv4.

IPv6 llegará a todas las organizaciones conectadas a la red antes o después, lo que motivó que en el CICA se realizara un diseño de la red teniendo en cuenta la implantación de IPv6 en un futuro no muy lejano. Es por ello que hoy día el CICA puede jactarse de ser ejemplo de previsión e implementación temprana en España del nuevo sistema de direccionamiento que dominará durante las siguientes décadas en Internet, trabajando tanto a nivel interno, como en un sus servicios ofrecidos al público, con un modelo de doble pila IPv4/IPv6 en el que se elige preferentemente ésta última.

Al igual que en el antiguo sistema de direcciones, en la versión 6 de IP juega un papel fundamental para las organizaciones su servidor de nombres (DNS), que requerirá múltiples cambios y adaptaciones para su puesta en producción, además de la curva de aprendizaje que el



nuevo protocolo supone para el personal dedicado a gestionar los servidores de nombres.

En el CICA se dispone de un prefijo de direcciones heredado desde RedIris, el cual se subdivide y delega para dar conectividad dentro de la red RICA+ a los distintos centros de investigación y Universidades afiliadas. Esto quiere decir que ya se está en disposición de ofrecer asesoramiento en despliegue, conectividad, servicios y delegación de direccionamiento en IPv6 a los clientes de la red RICA+ cuando estos soliciten su adhesión al nuevo protocolo de red.

Desde el Departamento de Seguridad de CICA no se han olvidado las complicaciones que las nuevas tecnologías introducen en las organizaciones y se ha tomado parte de forma activa en el análisis de las nuevas vulnerabilidades de IPv6 y en el filtrado del tráfico no deseado.

4. ESQUEMA NACIONAL DE SEGURIDAD.

El Esquema Nacional de Seguridad (ENS) es un ambicioso plan del Ministerio de Administraciones Públicas que pretende dotar a todo el ente público nacional de un vocabulario y bases comunes en cuanto a la seguridad de la información relativa a los ciudadanos. A través de este plan, las administraciones públicas contarán con una serie de políticas, protocolos y procedimientos de actuación guiados por el nivel de criticidad de los datos manejados.

En el Departamento de Seguridad del CICA se iniciaron los trámites para la implantación en el centro de la normativa ISO 27001, relativa a la seguridad de la información en cualquier de sus formatos (física/digital), previamente a la aprobación y publicación en el BOE del ENS.

Ambas normativas convergen en una serie de factores que a día de hoy son de vital importancia en la seguridad informática:

- Seguridad integral: donde no caben acciones improvisadas o puntuales que suelen ser causadas por errores humanos.



Figura 39: Ciclo de Deming

- Proactividad: las actividades en materia de seguridad de las organizaciones no pueden limitarse a la espera de un evento.
- Seguridad gestionada: debe existir un grupo estratégico que gestione la seguridad a todos los niveles.

Para dar comienzo al trabajo basado en procedimientos se está realizando un estudio previo sobre los diferentes procesos de actuación y de negocio, a través de un consejo interdepartamental, con el objeto de lograr una visión lo más amplia posible. Es de vital importancia implicar a todos los departamentos de la organización ya que el cumplimiento de los controles a implantar dependerán del buen hacer de todo el personal.

La aproximación desde el CICA para una óptima gestión de la seguridad se basa en el estudio de los procesos de negocio. Son dichos procesos los que constituyen el día a día de las organizaciones y los que están sujetos a posibles fallos o defectos operativos, por lo tanto deben ser auditados de manera periódica para asegurar que las evoluciones que sufren los activos (información, personas, tecnología) no introducen nuevas vulnerabilidades o barreras en el trabajo diario. Las pautas de actuación que se siguen en el CICA constituyen el ciclo de mejora continua o ciclo de Deming (ver figura 39).

5. NUEVOS SISTEMAS LDAP.

El protocolo LDAP es un mecanismo estandarizado a la hora de afrontar sistemas de cuentas de usuarios centralizados en entornos bajo gran demanda como es el caso del CICA.

Durante 2011 el sistema LDAP de CICA para prevenir la obsolescencia y resolver la falta de escalabilidad, Departamento de Seguridad diseñó e implantó un nuevo cluster de servidores para alcanzar este propósito.

La nueva arquitectura implantada afronta la necesaria escalabilidad y proporciona una gran capacidad de respuesta. Para ello se dispuso un sistema que permite la distribución de carga (mediante balanceo de carga), es capaz de proporcionar un rápido aprovisionamiento de sistemas adicionales dotando al cluster de servidores con nuevos nodos, así como un método de replicación de datos entre los nodos del cluster destinados a proporcionar el servicio LDAP.

Después del análisis de los requisitos, se optó por trabajar al completo con software libre y gratuito, usando:

- Debian GNU/Linux Squeeze como sistema operativo base, debido a la gran robustez, estabilidad y seguridad del sistema, al importante trabajo de mantenimiento del software realizado por la comunidad, y a su facilidad de uso.
- OpenLDAP 2.4 como servidor LDAP, que incorpora mecanismos nativos para replicación de datos y es dinámico, característica fundamental en la implantación.
- Pacemaker y Corosync como sistemas de gestión de la disponibilidad del cluster. Dos proyectos de software que actualmente son el estándar en cuanto a la infraestructura de alta disponibilidad en entornos Linux.
- LinuxVirtualServer (LVS) como sistema de balanceo de carga, software muy maduro que además es altamente integrable con Pacemaker y Corosync.

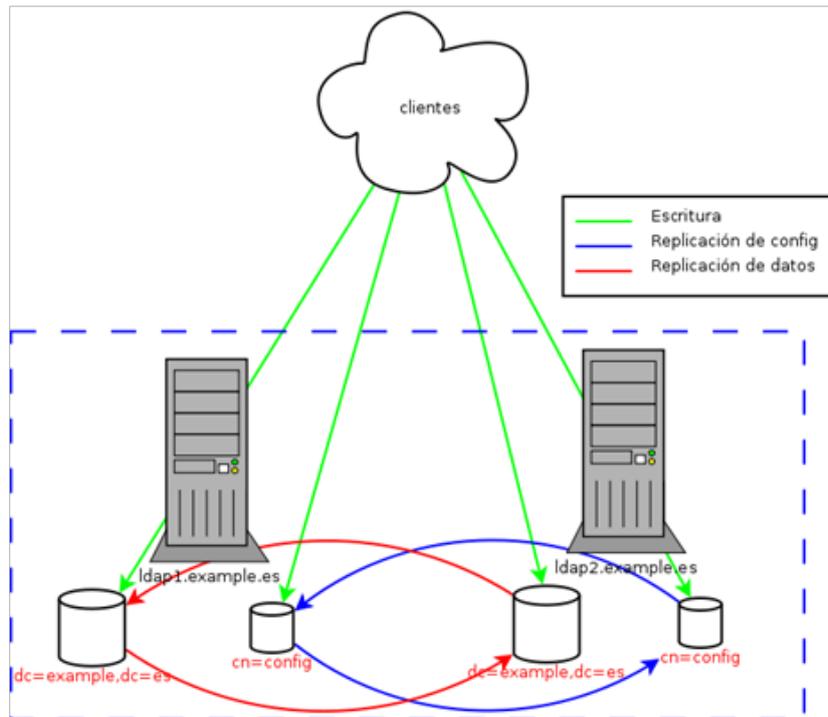


Figura 40: Esquema de nuevo LDAP

El cluster es de tipo activo-activo, con balanceo interno. La cantidad de nodos puede variar en función de las necesidades y es capaz de soportar gran cantidad de accesos simultáneos para afrontar una situación alta demanda.

El sistema soporta balanceo de carga en operaciones tanto de escritura como de lectura, como se muestra en el esquema de la figura 40.

Esta arquitectura es muy avanzada en cuanto a diseño de sistemas se refiere. Además, incorpora un despliegue total en IPv6, por lo que el funcionamiento es completamente transparente tanto en IPv4 como en IPv6.

Junto con esta implantación, se desarrollaron algunas herramientas para ayudar al administrador en la gestión del cluster.

Todo el software subido al repositorio es desarrollado, mantenido y empaquetado por el personal del Departamento de Seguridad.

Este repositorio ofrece paquetes en los formatos más usuales (DEB/RPM) para todas las máquinas de CICA, paquetes que se adaptan a las características de cada sistema y pueden instalarse en sistemas Linux de base, como Debian, Ubuntu, SUSE, CentOs y RHEL.

El software usado es completamente libre y gratuito: Debian GNU/Linux Squeeze (sistema operativo), apache2 (servidor web), reprepro (estructura de repositorio de paquetes deb) y createrepo (estructura de repositorio de paquetes rpm).

6. REPOSITARIOS INTERNOS DE SOFTWARE.

El Departamento de Seguridad de CICA ha desplegado un sistema de repositorios internos de software para sistemas operativos basados en Linux. El repositorio es solo para uso interno para CICA, y no es accesible desde otra red.

Desde este punto centralizado de distribución de software, es posible realizar un continuo trabajo de mejora sobre las configuraciones y utilidades de seguridad que se despliegan en cada máquina, así como un minucioso seguimiento de las actualizaciones.

7. INFORMES MENSUALES DE SEGURIDAD.

Desde mediados de 2011, todos los meses se redacta un informe donde se detallan los aspectos más importantes que se han tratado en el Departamento de Seguridad.

Algunas secciones que se pueden encontrar en dichos informes son:

- Incidencias: Muestra una tabla con las incidencias ocurridas, mostrando la fecha, afectados y la gravedad. Además, se describe detalladamente cada incidencia, explicando las acciones y contramedidas que

se han llevado a cabo para su solución y posibles recomendaciones para evitar que vuelva a ocurrir.

- Cambios en el firewall: Se muestran los cambios realizados en el firewall durante ese mes, resumiendo las reglas añadidas y borradas y a qué *vlanes* (redes locales virtuales) pertenecen.
- Repositorios internos: Muestra una tabla con los paquetes propios de nuestro repositorio interno, comentando si se ha hecho alguna actualización y los cambios realizados en la versión.
- Migraciones: En el caso de que haya habido una migración ese mes, se explica en este apartado.
- Fallos de servicios: Si se detectan fallos en algún servicio que son responsabilidad de nuestro departamento, aparecerá una descripción junto a la fecha de cuando se produjo.
- Gestión PKI: Se informa de la generación de nuevos certificados o revocaciones, ya sea desde la CA (Autoridad de Certificación) del CICA o desde la de Terena/Comodo.
- Actualidad: En este apartado se habla de las vulnerabilidades publicadas durante ese mes que pueden afectar a alguno de nuestros servicios.

Si alguna incidencia no llega a resolverse ese mes, se tiene en cuenta para el siguiente informe.

Una vez terminado, el informe se manda por correo a todos los usuarios del CICA y se sube al sistema interno de documentación Alfresco. Además, también se archiva una copia impresa junto a los demás informes.

8. CONTACTOS CON OTROS CERT.

El CICA se relaciona con varias entidades de gestión de incidencias de seguridad, como son (entre otras) AndalucíaCert, INTECO-CERT e IRIS-CERT.

Con cada una de estas entidades, el Departamento de Seguridad realiza un continuo intercambio de información, ya sea de forma activa (llamadas telefónicas directas, intercambio de correo electrónico, etc..) o de forma pasiva (suscripción a boletines, seguimiento en redes sociales y en la web, etc..)

AndalucíaCert (centro de seguridad TIC)

Es el centro experto para la gestión de la seguridad TIC de la Junta de Andalucía. Está formado por un equipo de

profesionales especializados en seguridad de la información, organizados según procesos definidos siguiendo modelos y estándares internacionalmente reconocidos y dotados de las herramientas específicas necesarias para el desarrollo de su actividad. Esta iniciativa de la Junta de Andalucía es un proyecto innovador, llamado a convertirse en el punto de referencia de la Seguridad de la Información en nuestra comunidad autónoma.

INTECO-CERT

La misión del INTECO (Instituto Nacional de Tecnologías de la Comunicación) es aportar valor e innovación a los ciudadanos, a las pymes, a las Administraciones Públicas y al sector de las tecnologías de la información, a través del desarrollo de proyectos que contribuyan a reforzar la confianza en los servicios de la sociedad de la información en España.

INTECO colabora con el sector TIC dentro de sus líneas de actuación, tanto de instituciones como de empresas privadas dedicadas principalmente al sector TIC.

IRIS-CERT

El Servicio de Seguridad de RedIRIS (IRIS-CERT) tiene como finalidad la detección de problemas que afecten a la seguridad de las redes de centros afiliados a RedIRIS, así como la actuación coordinada con dichos centros para poner solución a problemas de seguridad informática. También realiza una labor preventiva, avisando con tiempo de problemas potenciales, ofreciendo asesoramiento a los centros, organizando actividades de acuerdo con los mismos, y ofreciendo servicios complementarios.

IRIS-CERT también da soporte (coordinación de incidentes) a centros ajenos a RedIRIS, aunque de forma más limitada. Concretamente, IRIS-CERT actúa como último punto de contacto para los incidentes graves que puedan afectar al dominio “.es”, aceptando cualquier incidente de este tipo como evidencia para dar soporte a nuestra comunidad.

9. IMPLANTACIÓN DE CERTIFICADOS DE TERENA / COMODO.

Gracias a los acuerdos a los que CICA está suscrito siguiendo las directrices de RedIRIS para la comunidad, estamos implantando en todas nuestros servidores, servicios, máquinas y para personas certificados de TERENA/COMODO. Esta CA (Comodo) es aceptada por todos los navegadores, y es una autoridad certificadora reconocida internacionalmente que aportan globalización y prestigio a nuestros servicios ofrecidos.

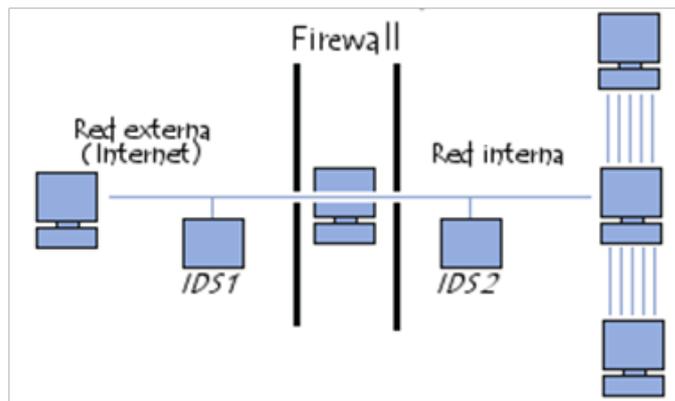


Figura 41: Esquema de IDS del CICA

En el Departamento de Seguridad generamos precisos informes sobre el uso, implantación y detalles de cada certificado generado. Cada vez que haya que generar un certificado se actualiza nuestro informe, y de esta manera aseguramos el funcionamiento de nuestro centro y su efectividad.

seguridad, mejorando no solo la seguridad, prevención y detección temprana, sino que además permite dedicar esos recursos humanos a tareas como el análisis de vulnerabilidades y desarrollo de políticas de seguridad entre otras muchas, mejorando así la productividad del departamento de seguridad.

10. INVESTIGACIÓN SOBRE SISTEMAS IDS.

Cada vez son más habituales los ataques a los grandes sistemas informáticos, sin embargo, la mayoría de los administradores siguen estando “ciegos” ante dichas intrusiones y ataques. Es por este motivo que desde hace más de 20 años se desarrollan nuevos campos dentro de la seguridad informática, entre ellos la detección de intrusos en los sistemas informáticos en grandes sistemas. En concreto la detección de intrusos en una red telemática sigue siendo, un gran problema de difícil solución. Dada la naturaleza del problema, no existe un método probado que permita la detección infalible.

Por estos motivos, desde CICA se ha decidido investigar y valorar la implantación de nuevos IDS que aseguren la calidad y el correcto funcionamiento de la seguridad de CICA. Todas esas detecciones son almacenadas en forma de registros de información que son de gran utilidad a la hora de reconstruir, mediante técnicas de análisis forense, qué es lo que le ha sucedido a un sistema informático.

Ante el gran crecimiento de servidores y de tráfico, el control de los sistemas de manera manual es totalmente inviable. Actualmente los sistemas están totalmente implantados y son funcionales, ayudando así a establecer medidas correctivas antes de que surjan incidentes de seguridad. Además de todo lo expuesto anteriormente, estos sistemas han permitido trasladar los recursos asignados a estas tareas y ser reaprovechados en otras áreas de la

11. TEST DE HERRAMIENTAS PARA AUDITORÍA DE SEGURIDAD.

El volumen de vulnerabilidades descubiertas en redes, sistemas y plataformas sigue creciendo exponencialmente, obligando a que las organizaciones de cualquier sector y tamaño, tengan que analizar frecuentemente las vulnerabilidades que puedan afectar su infraestructura tecnológica. Es imprescindible para toda organización satisfacer las siguientes necesidades:

- Poder determinar cuáles son las vulnerabilidades existentes en los activos de su red.
- Poder conocer cómo pueden estas vulnerabilidades comprometer a los activos de su red.
- Poder priorizar y remediar con efectividad las vulnerabilidades de los activos de su red.
- Poder garantizar el cumplimiento de regulaciones o auditorías respecto al estado de vulnerabilidades de su red.

En este ámbito, en el CICA se están implantando sistemas para realizar escaneos automáticos de vulnerabilidades en toda la red (ver figura 42). Se consigue así la detección de las vulnerabilidades de manera sistemática y desatendida. Tras elaborar nuestras propias bases de datos se realizan análisis exhaustivos para priorizar y subsanar las vulnerabilidades detectadas.

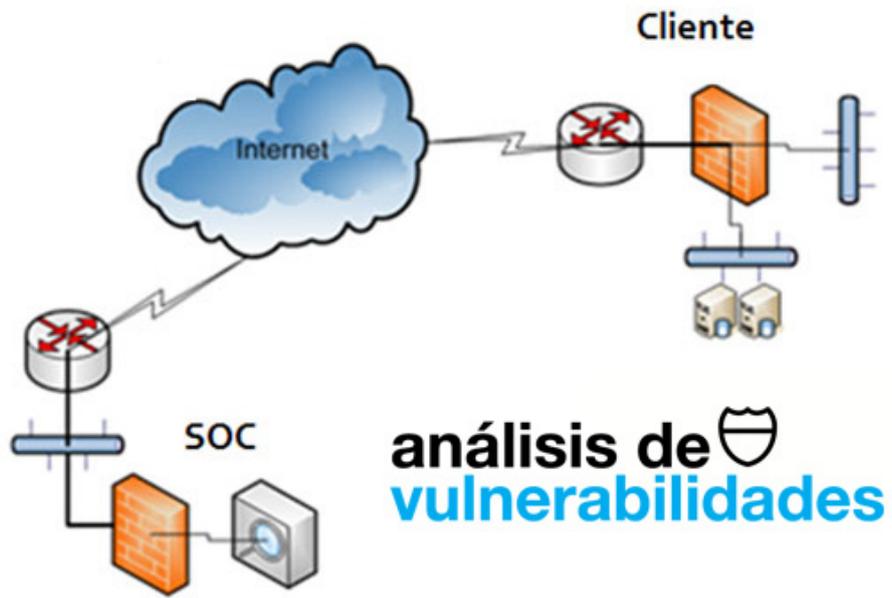
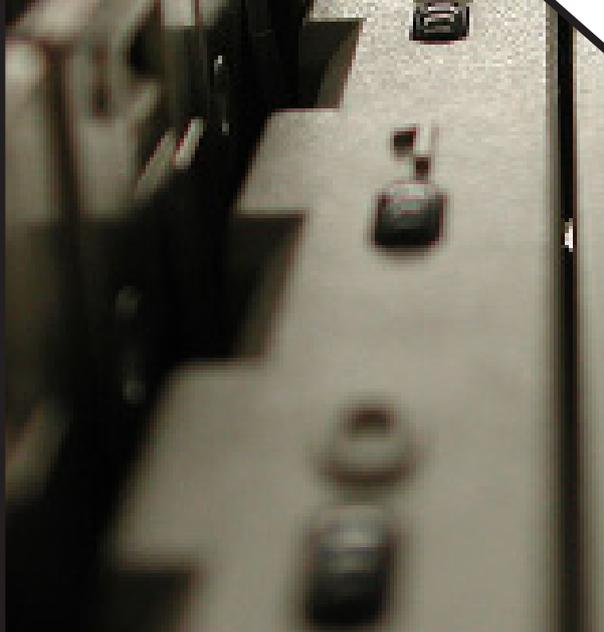
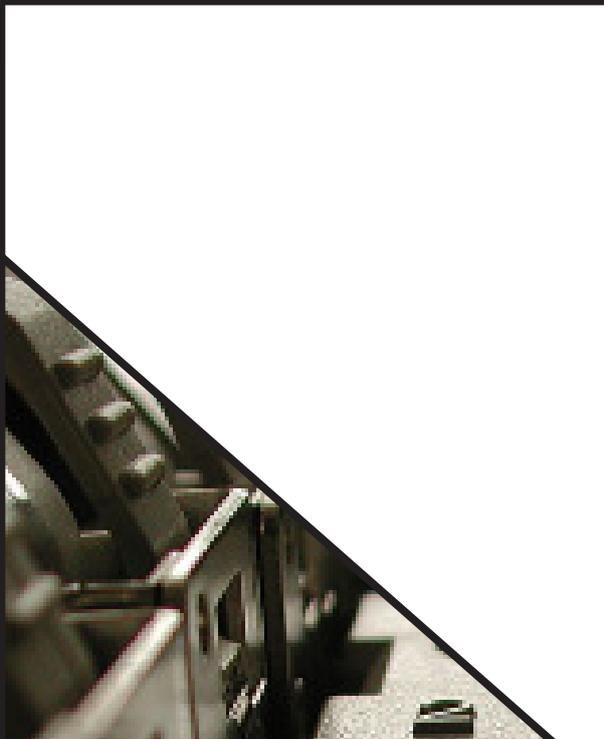


Figura 42: Esquema de análisis e vulnerabilidades



Sistemas / HPC

El mercado de los servidores x86 va sorprendiéndonos con importantes mejoras cada pocos meses. La familia de servidores multiCPU multicore acerca la familia x86 a los grandes mainframes en capacidad de cómputo. Eso sí, a unos precios muy asequibles y altamente competitivos.

El equipo de HPC del CICA ha podido probar estos super-servidores x86 con varias arquitecturas de este tipo de hardware gracias a INTEL y a fabricantes como HP y DELL que nos han prestado durante varios meses equipos con varias CPUs de 6 cores y de 8 cores. Las pruebas han sido exhaustivas tanto por parte del personal del CICA como por los investigadores de las universidades andaluzas necesitados de adquirir este tipo de hardware. El resumen de las pruebas es que para tareas de cálculo intensivo con un alto nivel de paralelismo, un exceso de cores y de CPUs en un único servidor provoca una importante caída de prestaciones en algunas arquitecturas, siendo mejores en coste/prestaciones los servidores algo menos masivos. Pero sin embargo, estos equipos más masivos son perfectamente aptos para el Cloud Computing, en el cual este año 2010 el CICA se ha metido de lleno.

Por otra parte, ha seguido el fuerte uso de los sistemas de cálculo intensivo del CICA por parte de los investigadores de las universidades andaluzas y de los centros de investigación ubicados en Andalucía.

Respecto a los servicios generales, se ha aumentado la tolerancia a fallos y el balanceo de cargas de todos los servicios de forma que en 2010 se han conseguido tener todos los servicios con sistemas tolerante a fallos. Y ya en 2011 se ha dado un paso más con la virtualización de dichos servidores, basada en KVM sobre CentOS, lo que hace que, usando las capacidades de replicación, la capacidad de servicio se adapte fácilmente a la demanda.

Por último, en 2011 ya se ha arrancado la nueva versión del Sistema de Información Científica de Andalucía (SICA), la aplicación de administración electrónica básica para la gestión de la investigación pública en Andalucía.



ISTEMA DE INFORMACIÓN CIENTÍFICA DE ANDALUCÍA

Este sistema informático conocido por sus siglas S.I.C.A. es la principal aplicación de administración electrónica ubicada en el CICA, en concreto es el repositorio de información de la producción científica de los investigadores andaluces, y a su vez es la aplicación que produce mayor impacto en el público de todos los servicios que se ofrecen a los investigadores andaluces.

1. ACTUACIONES EN LA INFRAESTRUCTURA DEDICADA A SICA.

El principal objetivo durante el año 2010 en el esquema de la infraestructura para SICA (Sistema de Información Científica de Andalucía) ha sido mantener la calidad del servicio ofrecida hasta la fecha (y avalada en los informes de Oracle) para la siguiente convocatoria, teniendo en cuenta ahorros de energía y de presupuesto.

Para ello se han llevado a cabo las siguientes acciones:

- Creación de una nueva vlan de SICA (26-64 IPs).
- Realización de cambios y pruebas de los mismos durante el verano, cuando menos puede afectar al servicio.

La infraestructura que soportaba el servicio en 2010 ha sido:

- Servidor Sun V890 para producción con base de datos SICA.
- Servidor Fujitsu PrimePower 650 para standby de la base de datos, la cuál es la más parecida a la anterior.
- Servidor Fujitsu PrimePower 450 para desarrollo de la base de datos.
- Servidor TTL 7200 con 16 cores y 128GB para convocatorias. Principal máquina para atender peticiones.
- Nodos TTL 300 como resto de servidores que atienden las peticiones. Es posible aumentar la

memoria RAM hasta 8GB y clonar con cobbler.

- Nodo TTL 300 con web service.
- Nodo TTL 300 para preproducción.
- Servidores HP BL25p como balanceadores.
- Máquinas de backups físicos de la base de datos.
- Servidor HP BL35 para hacer los exports solicitados por las universidades.
- Servidor BL25p para backups de respaldo a instalaciones remotas (proyecto FENIX de la Junta de Andalucía).
- Máquina virtual que contiene un portal en desarrollo de SICA2, la nueva versión del aplicativo.
- Máquina virtual dónde se almacena las estadísticas y las gráficas de los servidores de SICA.

A continuación se describe otras mejoras que se han realizado este año sobre la infraestructura de SICA.

1.1. Base de datos Standby.

A nivel de base de datos se ha dedicado la máquina ya mencionada Fujitsu PrimePower 650 para standby, es decir, se encuentra por un lado la base de datos Oracle de producción de la que se alimenta la aplicación y por otro lado existe otra base de datos en modo de espera (stand by) que deberá ser activada de forma manual en caso de fallo de la base de datos primaria. La base de datos Standby no es más que una base de datos idéntica a la que se encuentra en funcionamiento y que está disponible en caso de avería de los servidores que alojan a la primaria.

Una base de datos en Standby es una configuración Oracle de alta disponibilidad y contingencia relativa, que respalda el nodo de producción primario de la instalación Oracle, en este caso dedicada a SICA. En general, una base de datos Standby tiene como propósitos principales los siguientes:

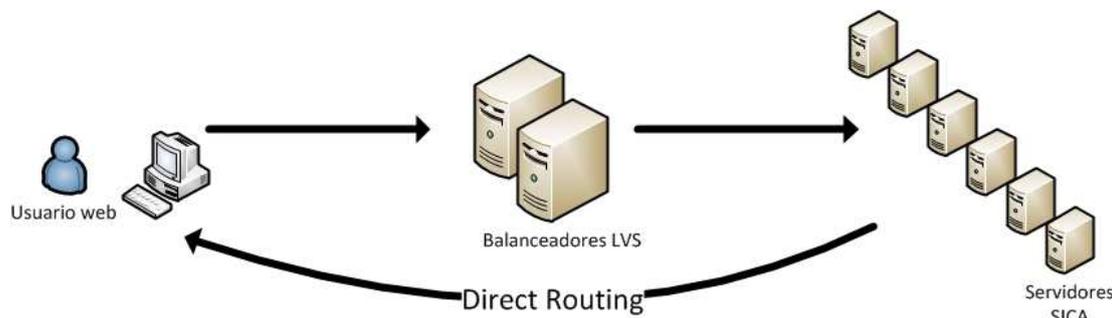


Figura 43: Esquema de comunicación aplicado en SICA usando Direct Routing

- Protección ante desastres.
- Protección en caso de corrupción de los datos.
- Eventualmente como una base de datos de consulta.

Con una base de datos adicional en modo standby en caso de fallo el tiempo de restauración es reducido y en cuestión de minutos se hace el cambio permitiendo continuidad en el servicio.

1.2. Direct routing.

Según se detalló en el anterior anuario de 2009, en el cluster dedicado a la aplicación de SICA, todas las peticiones de los clientes son filtradas por los firewalls del CICA. Las peticiones validadas por dichos firewalls, pasan a los balanceadores que enrutan y se encargan de equilibrarlas y dirigir las a los servidores reales en función de la carga y otros algoritmos.

En este año se ha optimizado la respuesta a las peticiones de los clientes mediante la técnica Direct Routing (ver figura 43). Como se puede observar en la imagen, mediante esta técnica el balanceador sólo gestionará las peticiones del cliente hacia el servidor con lo cual es una solución altamente escalable. La dirección virtual (VIP) es compartida por el balanceador y los servidores. De esta manera el balanceador recibe las peticiones y las envía a los servidores que procesan las peticiones y dan servicio directamente a los clientes.

1.3. Implantación del nuevo Sistema de Información Científica de Andalucía (SICA2).

Ya en 2011, se ha estado trabajando, en colaboración con diferentes empresas y organismos, en el proyecto de implantación del nuevo sistema SICA.

Es responsabilidad del centro mantener plenamente operativa la parte técnica de los servidores que contienen la información: implantación, gestión, securización, mantenimiento de la alta disponibilidad, recuperación ante desastres, etc. Ello incluye asegurar el correcto y continuo funcionamiento del entorno de SICA2.

Por ser SICA2, una aplicación completamente distinta a la antigua SICA, y para afectar lo menos posible a los usuarios de la misma, se ha trabajado duramente en la integración del nuevo sistema en los servidores que trabajaban para SICA; y por su mayor capacidad de almacenar y ofrecer información también en 2011 se ha renovado parte de la infraestructura con nuevos equipos, de mayores prestaciones y menor consumo energético.

El siguiente esquema de la figura 44 muestra la estructura que se ha instalado. La versión actual es parecida en el diseño pero no tanto el funcionamiento, lo cual ha determinado que haya habido que reorientar la infraestructura por las necesidades del nuevo aplicativo.

El cambio principal que puede verse en la figura mencionada se ha producido en los servidores webs (PWEBs en el esquema), disponiendo actualmente de 5 servidores. Uno de ellos cuenta con 4 procesadores Intel Xeon 7330 y 128 GB de RAM. El resto son servidores de apoyo,

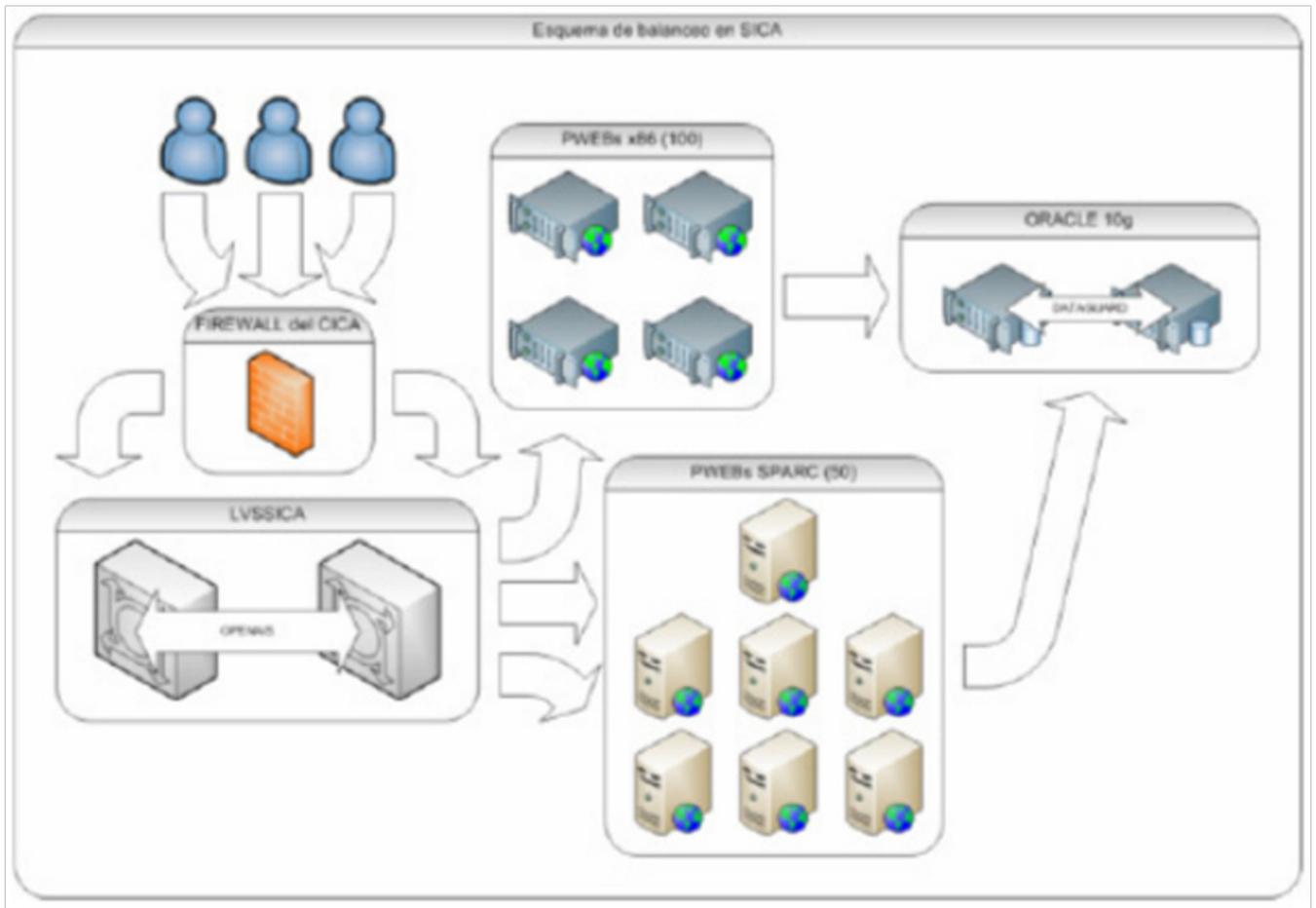


Figura 44: Esquema de SICA2

cada uno con procesadores Intel Xeon 2 Duo y 8 GB de RAM.

Así, los pasos que sigue la infraestructura en producción son los siguientes:

1. Todas las conexiones de los usuarios, entran inicialmente por el firewall del CICA. Éste se encarga de filtrar las peticiones hasta el balanceador de carga (lvssica)
2. El balanceador reparte todas las peticiones asignándolas a los servidores que estén más descargados en cada momento, distribuyendo las conexiones por los servidores (denominados pwebs).
3. Estas peticiones llegan a los programas que sirven las páginas web (software apache) de los servidores pwebs con distinto peso (el servidor más grande es capaz de soportar 5 veces más la carga de uno de los pequeños).
4. A partir de aquí, el programa apache balancea la carga por aplicación, enviando las peticiones a las

distintas herramientas y módulos que componen el aplicativo.

5. Por último, las consultas que requieran obtener datos de la base de datos, realizan las conexiones hasta el gestor de bases de datos (Oracle), que está instalado en un servidor con ocho procesadores SUN Ultra Sparc IV dual core con 32 GB de RAM. Además, contamos con otro servidor de similares características del fabricante FUJITSU, también con CPUs SPARC del modelo Sparc64, con el software Data-Guard de Oracle, trabajando en pasivo, pero listo para funcionar en caso de fallo del servidor principal.

También contamos con un “sistema de desarrollo”, en el cual, las empresas encargadas de crear y mantener el nuevo aplicativo SICA2, hacen sus pruebas y desarrollan el software necesario.

Una vez testeado, la nueva versión de la aplicación empaquetada es desplegarla en producción por el personal técnico del CICA.

En total, hay más de 15 servidores dedicados al proyecto

Secretaría General de Universidades, Investigación y Tecnología
Consejería de Economía, Innovación y Ciencia

SICA

Entrada

Nombre de usuario
.....

Recordarme [¿Has olvidado tu contraseña?](#)

Español English

¿Qué es SICA2?

Investig-AN

Conect-AN
Export-AN
Agrup-AN

Gestor curricular de la producción y actividad científica de los agentes del Sistema Andaluz del Conocimiento estructurado según el estándar CVN.

Noticias

Actualización en la generación de CV

Solventadas las incidencias en la obtención del CV generado por la FECYT. También se encuentra ya disponible el CV generado por SICA2 en formatos .odt, .pdf y .doc.

Restablecido el servicio de generación de CVN de la FECYT

Se ha restablecido el servicio de generación de CVN de la FECYT (21/12/2011 12:40), por lo que vuelve a estar disponible el formato de exportación PDF-FECYT.

Servicio de generación CVN de la FECYT no disponible

El servicio de generación de CVN de la FECYT está fuera de servicio a fecha 21/12/2011, por lo que SICA2 no puede ofrecer el formato PDF-FECYT. Por favor, seleccione otro de los formatos disponibles si desea generar su CVN.

[Abiertos a sugerencias](#)

Dimensión del sistema

18.878	Entidades
1.115.289	Producción científica
436.636	Actividades científicas
49.201	Usuarios

Mercado de ideas

Espacio abierto para oferentes y demandantes de tecnologías innovadoras

Entidades colaboradoras

10	Universidades
2.891	grupos PAIDI
142	Entidades privadas

Consejería de Economía, Innovación y Ciencia. | C.E.I.C. | L.O.P.D. | Manual | Contacto | faq

Figura 45: Esquema de SICA2

SICA2, cuya portada web de entrada a la aplicación se muestra en la figura 45.

2. NOVEDADES EN EL ÁREA DE SISTEMAS.

2.1. Cluster de virtualización

En el año 2010 se ha mejorado la infraestructura y organización del cluster de virtualización. Anteriormente, un único servidor NFS gestionaba el almacenamiento en un único sistema de almacenamiento de discos en red, y cualquier problema en dicho servidor o en la red propiciaba numerosos errores en las máquinas virtuales. En estos casos, se hacía necesario restaurar el servidor NFS y las conexiones de red antes de que el servicio de almacenamiento estuviese de nuevo disponible con el correspon-

diente retardo en la reentrada en servicio de los sistemas.

Actualmente, la parte de disco y memoria se encuentra en diferentes máquinas. Para la parte de memoria se ha usado el software libre hugepage. Esto supone acceder a una mayor cantidad de memoria con las mismas entradas en la tabla de paginación reduciendo así el consumo de CPU y por tanto, acelerar las aplicaciones de base de datos y virtualización y optimizar el rendimiento de los servidores virtuales. Estas máquinas tienen tanto una interfaz privada (conectada la VLAN privada) como una interfaz pública. Por su parte, la parte de disco físico se encuentra de manera centralizada en un servidor de discos en red de altas prestaciones de la marca Netapp y sirve a las máquinas virtuales a través de una VLAN privada.

Esta nueva estructura (ver figura 46) ha permitido mantener aproximadamente 30 huéspedes simultáneos para diferentes propósitos: hostings, web del CICA, reloj de picadas, dhcp, eduroam, mensajería instantánea de comu-

nicaciones, etc., de forma que cada uno de ellos tiene su parte de disco y memoria de forma separada y un error en alguna de ellas no compromete al resto.

Por último, destacar que, aparte de los seis servidores de producción (de virtual01 a virtual06), se han instalado dos nuevos servidores (virtual07 y virtual08) similares para pruebas o desarrollos previos a producción.

2.2. Infraestructura de backup.

En 2010 se ha modificado la estructura hardware del sistema de copias de seguridad del CICA que sigue usando el software de fuentes abiertas Bacula y servidores con sistema operativo LINUX para mejorar las prestaciones del sistema global aumentando su tolerancia a fallos (ver figura 47).

2.2.1. Servidores.

El antiguo sistema de backup dependía de un subsistema master y otro esclavo. La principal desventaja de esta in-

fraestructura era que el subsistema esclavo no podía realizar directamente copias en cintas debido a que dependía del máster. Este año la infraestructura de backup ha sido renovada para realizar estas tareas con una máquina llamada Sirhius (servidor con sistema operativo LINUX), que lleva instalado un gestor de base de datos MySQL y va conectada a un cargador de hasta 8 cintas. Esta máquina conecta de forma directa con los sistemas Bacula de los clientes a través de unos puertos específicos facilitando de esta forma la automatización de las copias de seguridad. Las copias completas (full backup) en cintas se realizan una vez por mes.

Para las copias a disco, se usa un sistema de almacenamiento en red basado en el protocolo AoE (ATA over Ethernet) de alta capacidad, prestaciones bajas (no se necesita más para realizar backups) y muy económico. Sobre este sistema de almacenamiento de la marca CORAID se realizan copias de seguridad incrementales todos los días además de una diferencial y otra total cada mes. La realización de estas dos últimas copias se realiza de forma separada a lo largo de cada mes para permitir, en caso de error, que la restauración de los datos sea lo más rápida posible.

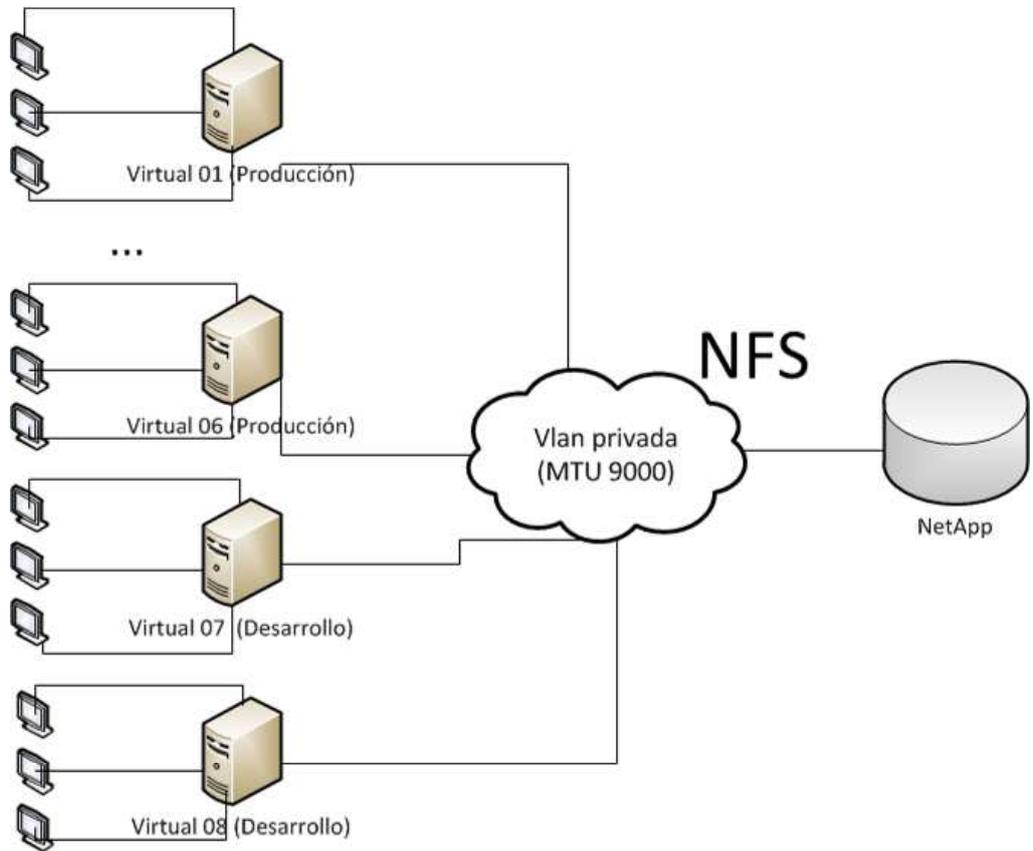


Figura 46: Estructura del cluster

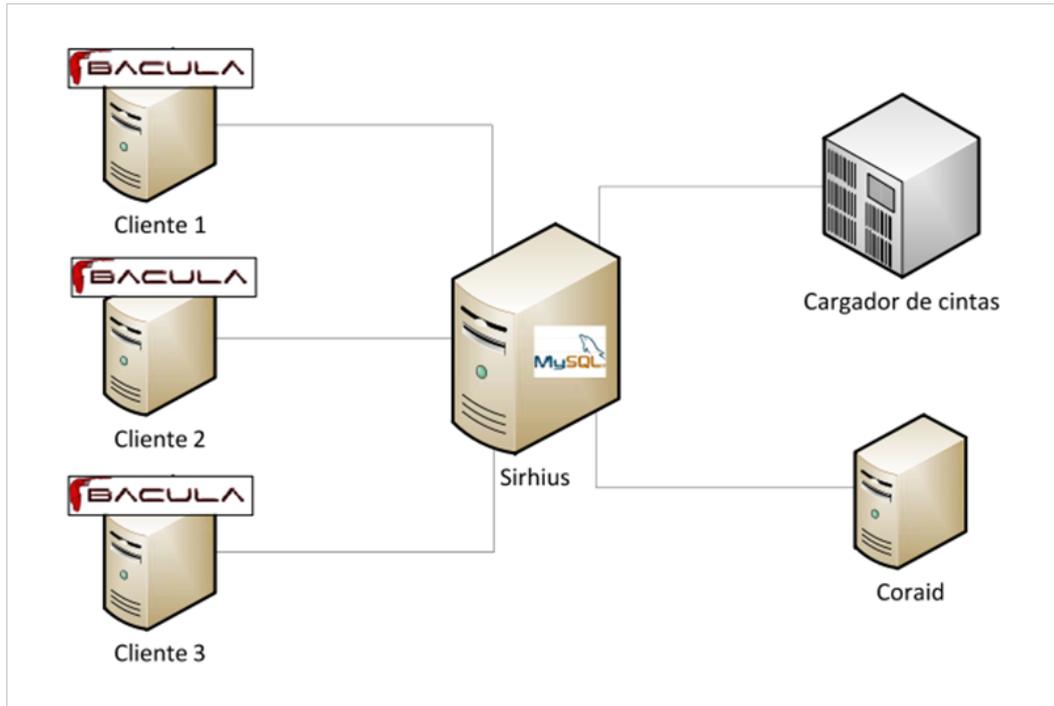


Figura 47: Diagrama de la infraestructura de backup

2.2.2. PCs de escritorio.

Para las copias de seguridad de los puestos de trabajo, en 2010 se ha implementado también un nuevo sistema para poder mantener apagado los PCs de escritorio cuando el trabajador no está presente y a la vez hacer los backup por la noche para evitar molestias al trabajador.

Para ello se usa un servidor que se dedica a “despertar” de forma automática a los PCs por la noche para realizar la copia de seguridad y al finalizar, vuelve a apagarlo. Para dicho despertar automático se usa estándar Wake on LAN (WOL). Una vez que los escritorios están encendidos, el sistema Bacula se encarga de realizar las copias de seguridad pertinentes y a través del comando shutdown, vuelve a apagar los equipos una vez terminado el proceso de backup.

2.3. Implantación del gestor de proyectos REDMINE

Hasta 2010, para el desarrollo de proyectos y la gestión adecuada de los mismos en el CICA se usaba Flyspray, un sistema de gestión de errores y tareas basado en PHP y MySQL. Esta herramienta (ver figura 48) permitía realizar un seguimiento de las tareas asignadas a cada uno de los departamentos del CICA y era especialmente usada en

los equipos de desarrollo y sistemas.

A partir de este año, se ha decidido usar la herramienta Redmine, un completo sistema de gestión de proyectos, desarrollado en este caso sobre la plataforma Ruby on Rails. Permite gestión de usuarios, con distintos roles, y que pueden tener acceso a distintos proyectos simultáneamente. Hoy en día, Redmine es uno de los gestores de proyectos más ágiles y mejor valorados, caracterizándose por su facilidad de uso y la calidad de la documentación. Las principales razones para pasar de Flyspray a Redmine han sido las siguientes:

- Vínculos fáciles hacia sistemas de gestión de versiones (de código) como Git, Subversion, Mercurial, etc.
- Gestión integrada del tiempo por tarea, permitiendo una evaluación del tiempo pasado por proyecto.
- Soporte integrado de diagramas de Gantt, permitiendo un mejor planeamiento de recursos humanos sobre nuestros proyectos.
- Gestión del proyecto más flexible (incluyendo sub-proyectos)
- Posibilidad de ejecutar comandos por e-mail (lo que permite, indirectamente, trabajar desconectado, usando una cola de correos)

#	Tipo	Estado	Prioridad	Tema	Actualizado	Categoría
5894	Feature	New	Normal	Allow bulk actions for issues in one project, especially status changes	2010-07-16 12:30	
5893	Feature	New	Normal	Filter Options: Allow to filter for description and notes	2010-07-16 12:26	
5892	Defect	New	Normal	Changing status from contextual menu opens the ticket instead	2010-07-16 11:22	Issues
5891	Patch	New	Normal	New IssuesController hook...	2010-07-16 00:33	Issues
5889	Feature	New	Normal	Progress bar for large file uploads	2010-07-15 22:59	
5884	Defect	New	Normal	PDF export doesn't work for Hebrew	2010-07-15 14:00	
5883	Defect	New	Normal	correct label_project_latest Chinese translation	2010-07-15 07:34	Translations
5882	Feature	New	Normal	Email alerts on new files	2010-07-15 00:30	
5881	Feature	New	Normal	Subtasks do not always appear below the parent task in Gantt	2010-07-14 23:18	Time tracking

Figura 48: Panel principal de Redmine

- Una apariencia visual más agradable que Flyspray.
- Paneles globales más visibles y completos.
- Una comunidad más amigable.
- Los almacenes de información pueden publicarse por HTTP, FTP, rsync o mediante un protocolo nativo, ya sea a través de una conexión TCP/IP simple o a través de cifrado SSH.

2.4. Implantación del gestor de versiones GIT.

Una de las principales razones de la implantación de Redmine para la gestión de tareas es su integración con un sistema de control de versiones llamado Git, un software que permite el mantenimiento de versiones de aplicaciones cuando estas tienen un gran número de archivos de código fuente. Entre las características más relevantes se encuentran:

- Git incluye herramientas específicas para navegar y visualizar un historial de desarrollo no-lineal.
- Gestión distribuida. Git le da a cada programador una copia local del historial del desarrollo entero, y los cambios se propagan entre los repositorios locales.

- Gestión eficiente de proyectos grandes, dada la rapidez de gestión de diferencias entre archivos, entre otras mejoras de optimización de velocidad de ejecución.

- Todas las versiones previas a un cambio determinado, implican la notificación de un cambio posterior en cualquiera de ellas a ese cambio (denominado autenticación criptográfica de historial).

- Realmacenamiento periódico en paquetes (ficheros).

2.5. Sistema de autoencendido de ordenadores.

Apagar el PC cuando no se usa es un consejo básico para el ahorro de energía. Corregir esta situación permite reducir costes y, al mismo tiempo, beneficiar al medio am-

biente. Con esto no solo se ahorra energía y se limita el consumo, sino que también el PC deja de ser vulnerable a las posibles subidas de voltaje.

El CICA y su personal, desde años atrás y especialmente durante este pasado año 2010, ha tomado conciencia de la necesidad del ahorro energético y el uso sostenible de la energía eléctrica. Por ello, cuando el personal no se encuentra presente en el Centro, se mantienen apagados los equipos de sobremesa en cada uno de los puestos de trabajo.

Sin embargo, se hace imprescindible un método que permita a cada trabajador acceder a su puesto de trabajo desde el exterior en todo momento para resolver problemas o incidencias sin necesidad de tener que desplazarse al centro fuera del horario laboral o cuando se está de viaje. Y para ello se dedicó un servidor que se usa para “despertar” a los PCs de escritorio con objeto de que el trabajador pueda luego conectarse a él y así resolver los incidentes que se hayan podido producir.

Este servidor usa la tecnología Wake on LAN (WOL), estándar que permite encender remotamente escritorios apagados cuando los usuarios no se encuentran presentes. Para el funcionamiento de esta tecnología es necesario que los equipos tengan instaladas tarjetas que cuenten con esta opción y que la placa base soporte el estándar. Cuando el sistema está apagado, la tarjeta de red usa una fuente de alimentación alternativa para monitorizar la red y esperar el envío de un paquete desde el servidor, esta energía es proporcionada por la placa base a la tarjeta a través del propio cable. La tensión proporcionada por la fuente de alimentación es de 5V y de bajo amperaje cuando se sitúa en modo standby. Una vez que recibe este paquete, la tarjeta de red alerta al sistema para que se en-

cienda.

De esta forma, el personal puede acceder a cualquier hora al contenido de su equipo y los servicios de CICA desde el exterior a través de Internet, ya sea para consultar algún documento o para resolver alguna incidencia. Sólo se necesita acceder a la VPN del CICA mediante el certificado correspondiente, y una vez dentro a través de la máquina despertador.cica.es y las credenciales LDAP, elegir el escritorio asociado al usuario para su encendido y arranque.

2.6. Configuración de IPv6 en servidores.

Durante el año 2010, el departamento de comunicaciones implantó IPv6 en los conmutadores y VLANs (redes virtuales) del CICA. Por su parte, el departamento de sistemas ha trabajado para configurar IPv6 en todos los servidores. Así, todos los servicios del CICA son ofrecidos por IPV6: webs, backups, sistema de conexión remota, etc. En la actualidad, debido a las características de los sistemas instalados, no es posible balancear los servicios centrales mediante IPv6 (estos son el correo, ftp, web del CICA, etc.). Para solucionar el contratiempo uno de los servidores reales de cada servicio se ha configurado con Dual Stack, de forma que recoge las peticiones IPv6 y a su vez balanceará las de IPv4.

A continuación se muestra una figura con un ejemplo en el que se simula el acceso a un mismo servicio FTP mediante IPv4 e IPv6. En el esquema aparecen dos servidores, uno es ftp1, que atiende las peticiones con IPv4, y el otro es ftp2, que atiende las peticiones con IPv4 e IPv6. En caso de peticiones bajo el protocolo IPv4, la petición, una vez autorizada por el firewall, llega a los balanceadores

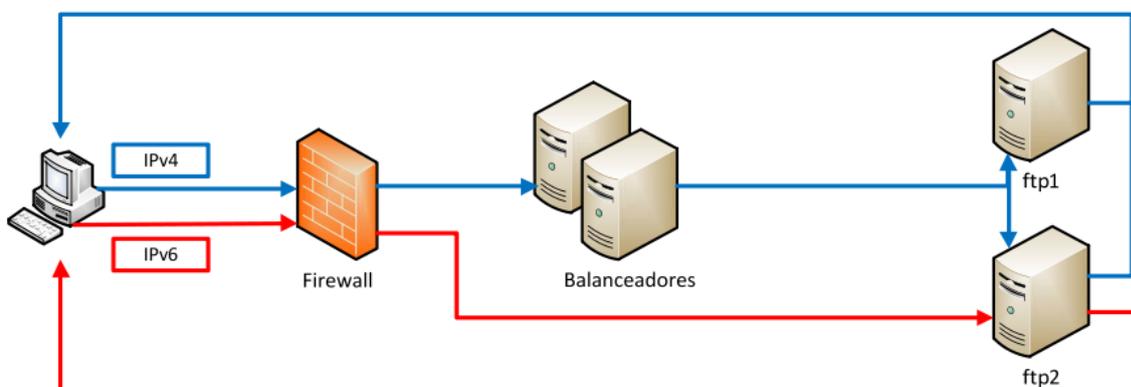


Figura 49: Ejemplo de acceso con IPv4 e IPv6

de carga que deciden en qué servidor se atenderá la petición, ya que ambos poseen IPv4. Para peticiones bajo IPv6, al ser sólo ftp2 el servidor que atiende las peticiones mediante dicho protocolo, no es necesario acceder a los balanceadores y se pasa directamente del firewall a la máquina ftp2 para atender a la petición.

3. CICA Y EUROPA.

Recientemente CICA ha comenzado a invertir muchos de sus esfuerzos en su proyección hacia Europa. Haciendo uso de sus servicios, CICA es un centro preparado para dar el salto internacional participando en proyectos junto a empresas españolas y del exterior.

Aunque es evidente que el centro está estrechamente relacionado con las Tecnologías de la Información y las Comunicaciones, CICA puede participar en proyectos y adherirse a consorcios de diversa índole. Esto es debido a que su principal objetivo es dar cobertura a todos los investigadores independientemente de su área de conocimiento (Informática, Medicina, Aeronáutica, etc.), siempre y cuando necesiten soporte relacionado con los servicios que CICA ofrece, tales como el almacenamiento y cálculo masivo, el cloud-computing, la supercomputación o las redes de comunicaciones de banda ancha.

Actualmente, CICA y parte de su personal se encuentra involucrado en la búsqueda de socios para la preparación y desarrollo de nuevas propuestas de cara a la participación en programas europeos, principalmente dentro del VII Programa Marco (FP7), y más concretamente, en los subprogramas relacionados con las TIC y las Infraestructuras de Investigación. Para ello, el personal dedicado estudiará con detenimiento todas aquellas propuestas de colaboración en proyectos que deseen la participación de la entidad y pondrá a disposición todos sus servicios y personal que sean necesarios.

3.1. Proyecto EDGI.

Como primer resultado de la búsqueda de colaboraciones en proyectos europeos CICA ya está participando en uno de ellos.

En Octubre de 2010 el proyecto EDGI (European Desktop Grid Initiative) lanza una llamada para la selección de dos nuevos socios como proveedores de infraestructura. El proyecto EDGI es un proyecto dentro del VII

Programa Marco, ya aprobado y comenzado, y cuyo consorcio está formado por centros tan relevantes como el MTA SZTAKI Computer and Automation Research Institute (Hungria), Cardiff University (Reino Unido), Centre National de la Recherche Scientifique (Francia) o la Universidad de Zaragoza. De forma resumida, este proyecto trata sobre el desarrollo de un middleware que sirva de unión de los resultados obtenidos en el proyecto EDGIs referidos a la extensión de servicios grid con grids de escritorio haciendo más eficiente el soporte a diferentes comunidades de usuarios que requieren una gran capacidad de computación en términos de CPUs y cores. En la figura 50 se observa el esquema de EDGI.

CICA compitió con otros centros para la oferta de proveedor de infraestructura y resultó seleccionado. Por tanto, CICA y más concretamente su equipo de HPC comenzó en Marzo de 2011 su participación en este proyecto ofreciendo parte de su infraestructura para la conexión al grid ya existente y en la cual se lanzó un conjunto de aplicaciones proporcionadas por el consorcio. En concreto las funciones de CICA dentro del proyecto están siendo las siguientes:

- Planteamiento: Descripción de la infraestructura aportada por CICA y diseño de integración en EDGI.
- Implementación e Integración: Implementación de la solución de integración, lanzamiento del software proporciona y pruebas.
- Operación: Mantenimiento, monitorización y supervisión de la actividades llevadas a cabo sobre la infra

Esta primera experiencia debe servir a CICA para comenzar su expansión en Europa, poniendo sus recursos, personal e infraestructura al servicio de proyectos de gran calado a nivel europeo donde se encuentren participantes destacados provenientes de varios países.

4. PROYECTO DE ADECUACIÓN PARA LA SOSTENIBILIDAD Y EFICIENCIA ENERGÉTICA EN EL CICA.

La sociedad actual se enfrenta al problema del importante gasto energético que necesita para su desarrollo y bienestar, y consiguientemente a su impacto sobre el medio ambiente. Se hace por tanto necesario buscar otros modelos de desarrollo o bien optimizar los existentes. Es por tanto prioritario, en pro de un desarrollo sostenible, el fomento de iniciativas que generen ahorro energético e

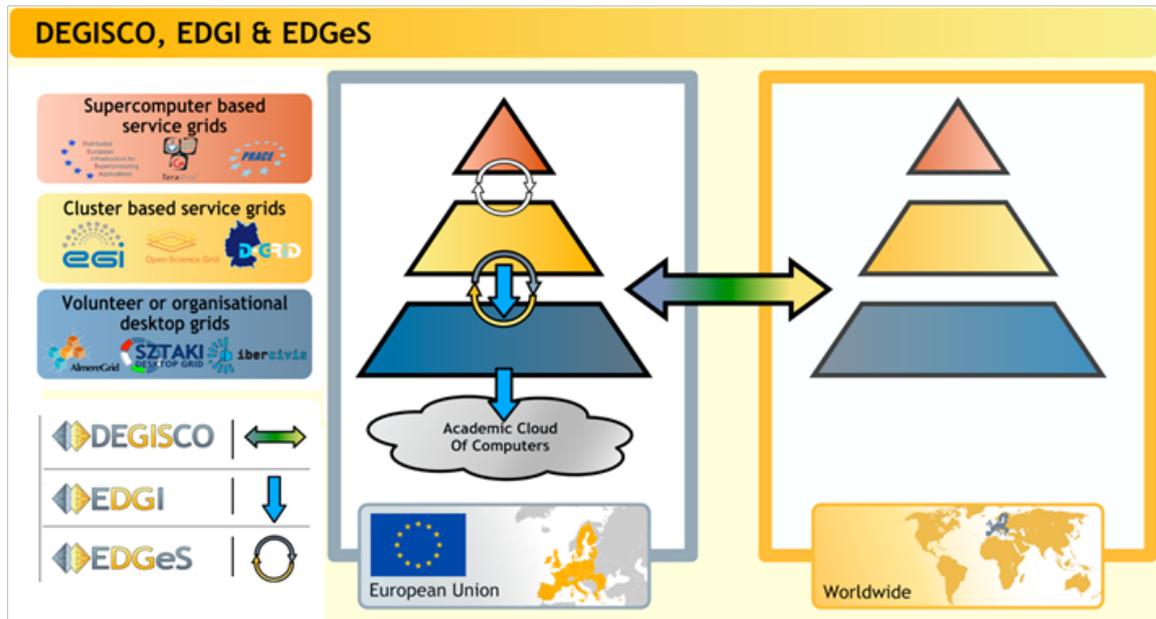


Figura 50: Arquitectura de EDGI

incentiven la eficiencia energética. Los centros de procesamiento de información como el CICA, son centros de recursos con alta demanda energética y por ello no son ajenos a esta problemática. En el momento actual, el consumo energético es uno de los costes más importantes que debe contemplarse en el presupuesto de un centro. Es por ello que surge la necesidad de reducir la huella energética que el centro produce, fomentando el uso de energías limpias y a la vez reduciendo el coste económico.

El proyecto "Adecuación para la Sostenibilidad y Eficiencia Energética en el CICA" pretende realizar la mejora y actualización de la infraestructura del centro de procesamiento de información del Centro Informático Científico de Andalucía (CICA), para permitir el ahorro energético, analizando en primer término la situación actual y proponiendo alternativas de acondicionamiento y consumo que perfectamente serán extrapolables al resto del edificio. En el informe de auditoría realizada en el centro en 2008 y la auditoría energética de las instalaciones del CICA realizada para el año 2009 por la Agencia Andaluza de la Energía describen en detalle las zonas de actuación donde se necesita mejorar o actualizar la infraestructura, así como las mejores medidas que permitan mejorar la eficiencia energética y la adaptación a las normativas y reglamentos actuales.

En este contexto, el proyecto propone innovar en el uso de las nuevas tecnologías para lograr la reducción directa

del consumo eléctrico, adquiriendo para ello las infraestructuras necesarias que permitan adaptar y optimizar los recursos disponibles en el CPD de manera eficiente, sin dejar de ofrecer todos los servicios que el CICA ofrece desde el año 1989 a los Agentes del Sistema Andaluz del Conocimiento.

De manera resumida, el proyecto propone los siguientes objetivos:

- Determinar mediante estudios y auditorías internas y externas (algunas de ellas ya realizadas) las necesidades del CPD del CICA para lograr una reducción del consumo energético notable.
- Definir un plan energético sostenible para el CPD, adquiriendo para ello la infraestructura necesaria para la gestión eficiente de la energía en el hardware, racks e instalación eléctrica, optimizando así los recursos y aplicando las buenas prácticas que permitan un ahorro energético, y por consiguiente, la reducción de la cantidad de CO2 emitida a la atmósfera y el ahorro económico que supone.
- Integración y expansión de las medidas tomadas en el CPD a todo el entorno del edificio, promoviendo así un gasto sostenible no sólo en el centro de datos, sino en todo el CICA. Este cambio debe ser realizado de manera gradual sin interferir en el funcionamiento normal del centro y en la prestación de sus servicios.

- Establecer un plan de difusión especial para el correcto uso de las nuevas infraestructuras, la concienciación del ahorro energético tanto al personal de CICA como a los usuarios de sus servicios y la publicación de las soluciones adoptadas y resultados obtenidos a la ciudadanía y empresas.

Durante el año 2010, las acciones a llevar a cabo para la realización de este proyecto se han recogido en un documento llevado a cabo por una empresa externa. Desde entonces se dispone de un proyecto de adecuación del centro que será la hoja de ruta para las actuaciones pendientes.

Entre finales de 2010 e inicios de 2011 ya ha comenzado la ejecución de una parte del proyecto, que se comentan en los siguientes apartados, y en el próximo año 2012 se espera llevar a concurso la ejecución de la obra.

4.1. Cambio de suelo técnico.

En el año 2010 se ha realizado el cambio del suelo técnico (ver imagen 51) bajo el cual se ubican las canalizaciones para el paso de cableado de datos, la red eléctrica y el sistema de detección y extinción de incendios. Las baldosas del suelo técnico retirado estaban en mal estado debido a su antigüedad de más de 20 años. Las nuevas baldosas del falso suelo son de material ignífugo, a 60cm de distancia respecto al suelo real.

La estructura sobre la que se apoya el suelo se ha mantenido gracias al buen estado de conservación de la misma.



Figura 51: Nuevo suelo del CPD

4.2. Adquisición de máquinas de aire acondicionado.

Asimismo, se han instalado tres nuevas máquinas de aire acondicionado específicas para CPDs, marca STULZ modelo Compact DX CSD 852 (ver imagen 52), que mejoran la refrigeración de la sala y el consumo energético de la misma. Estos equipos son fáciles de instalar y permiten el acceso a ellos por la parte delantera mejorando su mantenimiento. Apoyan a las tres antiguas máquinas AirDATA con más de 20 años de antigüedad y que usan gas R22 ya retirado del mercado por su efecto nocivo sobre la capa de ozono. Las máquinas antiguas solo se arrancan en periodo estival a la espera de completar la segunda fase que consiste en la sustitución de las mismas por otros equipos similares a los recientemente instalados.

Estos son sólo los primeros pasos para mejorar el CPD del CICA, haciéndolo más seguro, sostenible, gestionable y respetuoso con el medio ambiente.



Figura 52: Una de las nuevas máquinas de aire acondicionado

COMUNICACIONES

SEGURIDAD

SISTEMAS / HPC

APLICACIONES

FORMACIÓN Y EVENTOS



COMPUTACIÓN CIENTÍFICA DE ALTAS PRESTACIONES Y E-CIENCIA

1. USUARIOS Y PROYECTOS DE INVESTIGACIÓN

1.1. Usuarios y áreas de conocimiento

Investigadores de toda Andalucía han hecho uso de los servidores de cálculo para el apoyo de sus investigaciones, y así lo refleja el constante aumento de usuarios durante estos últimos años.

Durante el año 2010 se ha producido un aumento significativo de nuevos usuarios y de grupos de investigación registrados, abarcando varias áreas de

conocimiento. De tener en años anteriores grupos de investigación relacionados con Química, Biología, Minería de Datos, etc.. en este año se han incorporado investigadores pertenecientes al área de la Física Nuclear y Arquitectura, en este caso para el estudio de estructuras.

Como podemos observar en la gráfica de la figura 49, hubo un total de 75 investigadores activos en 2010. Esta cifra ha estado en constante aumento desde 2008. Sin embargo, los datos de 2011 reflejan una reducción de nuevas altas y usuarios significativo.

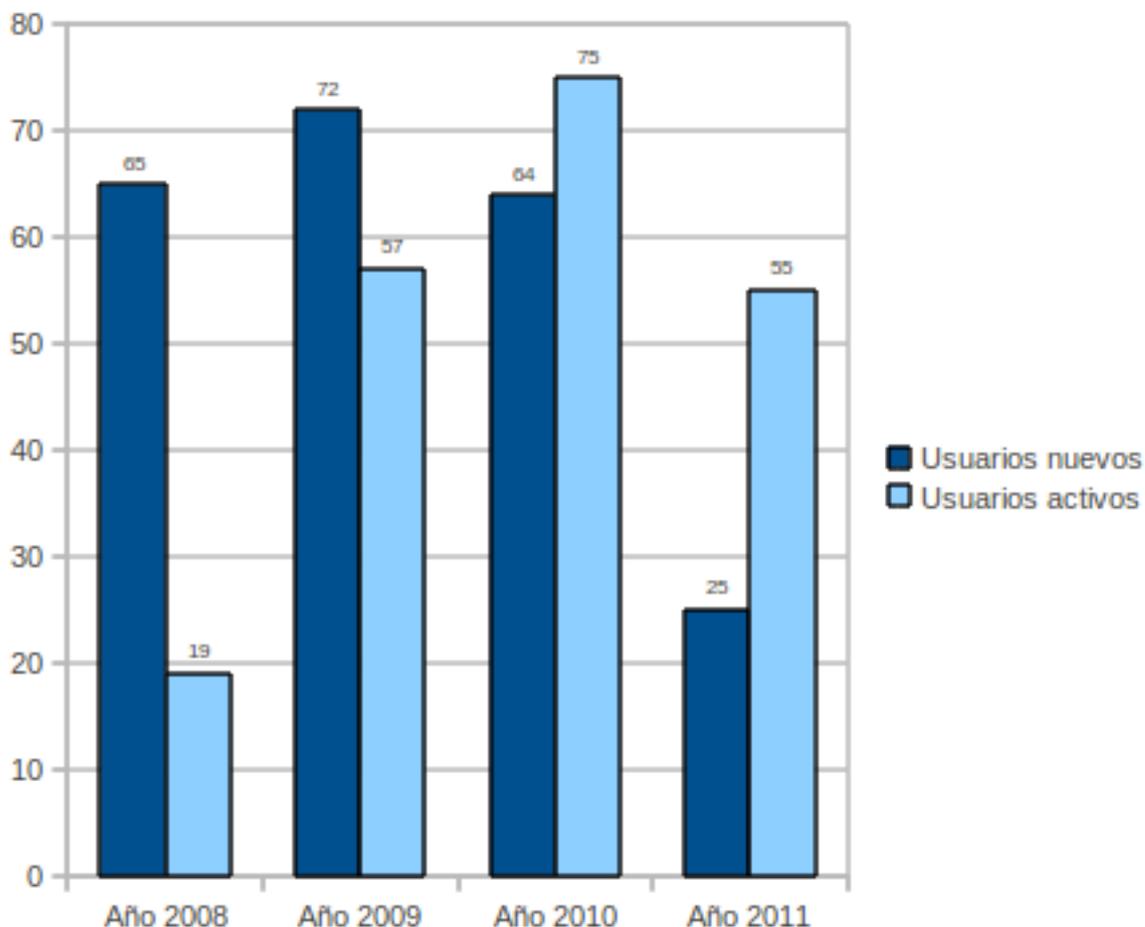


Figura 53: Usuarios totales

La reducción de nuevas altas puede ser debido a que los usuarios ya atendidos cubrían ya una buena parte de los potenciales usuarios. Respecto al menor número de usuarios activos esperamos que sea un descenso coyuntural, aunque puede deberse a la reducción de servidores de cálculo a disposición de los investigadores debido a que la entrada del CICA como centro de recursos del Proyecto Europeo EDGI, del 7º Proyecto Marco, ha obligado a reservar un considerable número de servidores para atender los compromisos contraídos.

De todas formas en 2011 se han seguido llevando a cabo cálculos correspondientes a investigaciones relacionadas con Química, Biología, Minería de Datos, Física Nuclear, Arquitectura o Biología. Como podemos observar en la gráfica 53 ya mencionada, hay un total de 55 investigadores activos en 2011.

Las diferentes áreas de conocimiento de los investigadores se representan en las figuras 54 y 55 para 2010 y 2011 respectivamente. En dichas gráficas podemos observar la



Figura 54: Porcentaje de investigadores por área 2010

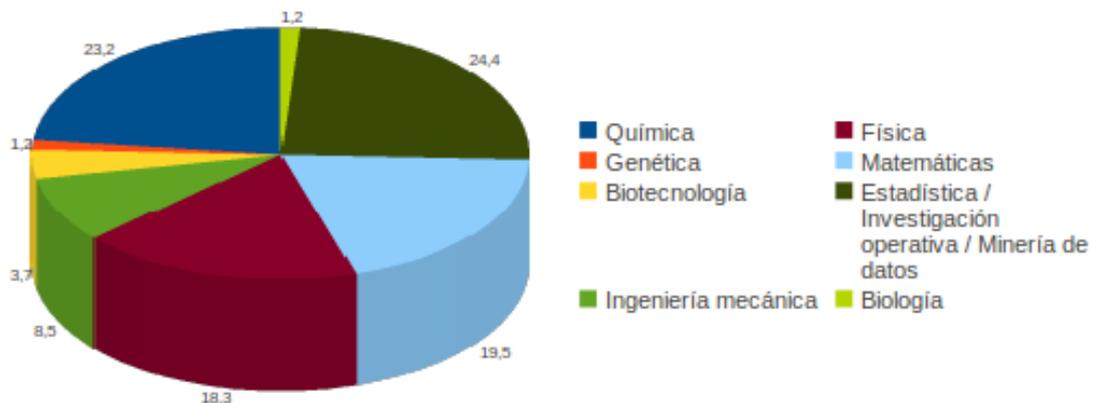


Figura 55: Porcentaje de investigadores por área 2011

variabilidad de ámbitos en los que se trabaja en CICA, cada uno con diferentes necesidades de computación (memoria compartida / distribuida con cálculo paralelo, gran capacidad de cálculo, etc.).

Siguiendo la tendencia de años anteriores, el mayor porcentaje de cálculos sigue siendo de las disciplinas de química y estadística-matemáticas, aunque con un mayor auge en cuanto a minería de datos e investigación operativa.

En estos años ha ido en aumento, con respecto a años anteriores, el ámbito de Investigación Operativa que junto a Química es uno de los campos en los que más se ha trabajado en el cluster.

1.2. Grupos de investigación registrados.

Los grupos de Investigación que han realizado algún cálculo o colaborado con nosotros, son los siguientes:

- Departamento de Supercomputación y Algoritmos de Biotecnología - Código: TIC-146

Descripción: Reconstrucción 3D de imágenes tomográficas usando entornos Orientados a Objetos concurrentes. Se buscan medios de aprovechar la heterogeneidad de las plataformas de cómputo y de explotar al máximo las arquitecturas multicore con modelos de programación hebrados.

- TeLoYDisRen - Código: FQM-270

Descripción: Investigación relacionada con Minería de datos, Diseño de encuestas, Imputación, Estadística Medioambiental, Teoría de Localización y Optimización.

- Departamento de Informática y Análisis Numérico - Código: SEJ-110

Descripción: Ecuaciones diferenciales. Simulación numérica y desarrollo de software. Aprendizaje y Redes Neuronales artificiales.

- Departamento de Genética de la Universidad de Córdoba - Código: FQM-124, AGR-214

Descripción: Análisis estadísticos datos moleculares.

- Instituto de Astrofísica de Granada - Código: TIC-114

- Departamento de Arquitectura y Tecnología de los Computadores - Código: TIC-117, TIN-2007-60587

- Grupo de Modelización y Diseño Molecular - Código: FQM-174

- Departamento de Física de la Atmósfera - Código: CGL2007-61151/CLI

- Departamento Radiation Physics and Environment - Código: PAI-1995

Descripción: Implantar un modelo de predicción meteorológica acoplado a un modelo de dispersión que permita simular la evolución de diferentes contaminantes en Andalucía a una resolución espacial de 3 Km. Se trabajará en modo análisis (estudio de episodios pasados) y en modo predicción.

- Departamento de Economía Agraria - Código: SEJ-110

- Departamento de Tecnología de la Información - Código: TIC-198

- Servicio de Informática y Comunicaciones de la Universidad de Huelva

- Grupo de Investigación de Control y Robótica - Código: TEP-192

- Grupo de Geoquímica Ambiental - Código: RNM-347

- Departamento de Sistemas Inteligentes y Minería de Datos - Código: TIC-207

Descripción: Realización de experimentos relacionados con computación evolutiva y algoritmos co-evolutivos desarrollados en tanto en C como en JAVA.

- Departamento de Estadística Teórica y Aplicada e Investigación Operativa - Código: FQM-245

- Departamento de Estructura y Dinámica de Sistemas Químicos - Código: FQM-337

- Grupo de Modelización de la Atmósfera y Radiación Solar - Código: TEP-220

- Departamento de Química Física Teórica y Experimental - Código: FQM-173

- Grupo de Oceanografía Física de Málaga - Código: RNM-137

- Grupo de Técnicas Computacionales en la Ingeniería - Código FIS2009-12864
- Departamento de Lenguajes y Sistemas informáticos
- Grupo ITÁLICA - Código: TIC-134
Descripción: Procesamiento del lenguaje Natural. Compilación de modelos declarativos de software.
- Departamento de Matemática Discreta: Teoría de Grafos y Geometría Computacional - Código: FQM-164
Descripción: Algorítmica en Grafos Infinitos. Problemas de Inmersión de Grafos. Grafos de Visibilidad. Grafos Extremales. Geometría Computacional en Superficies no Planas.
- Departamento de Física Atómica y Molecular - Código: P06-FQM-01869
- Departamento de Química Teórica - Código: FQM-132
- Departamento de Matemática Aplicada de la Facultad de Arquitectura - Código: MTM2006-01275
- Departamento de Química Física - Código: FQM-282
- Grupo de Estudio de Microorganismos Halofilos - Código: BIO-213
- Departamento de Metalurgia e Ingeniería de los Materiales - Código: TEP-123
- Grupo de Modelado Matemático y Simulación de Sistemas Medioambientales - Código: FQM-120
- Departamento de Estudio, Ingeniería y Experimentación de Estructuras - Código: TEP-245
- Grupo de Lógica, Lenguaje e Información - Código: HUM-609
- Departamento de Organometálicos y Catálisis Homogénea - Código: FQM-223
- Grupo de Álgebra Computacional en Anillos no Conmutativos y Aplicaciones - Código: FQM-333
- Grupo de Química Organometálica y Catálisis Homogénea - Código: FQM-3151, FQM119
- BioInformatics Research Group - Código: TIN-68084-C02-00

- Departamento de Mecánica de Medios Continuos. Teoría de Estructuras e Ingeniería del Terreno - Código: TEP-114
- CITIUS - Código: RNM-138
- Física Atómica, Molecular y Nuclear - Código: FQM-160
- Departamento de Ciencias de la Computación e Inteligencia Artificial - Código: TIC-193

1.3. Publicaciones de resultados

Durante estos años 2010 y 2011 el CICA ha sido mencionado en los agradecimientos de los artículos publicados por el Grupo de investigación del departamento de Química Física y Analítica de la Facultad de CC. Experimentales de la Universidad de Jaén, entre otros.

Los usuarios Andrés Garzón, Jose M. Granadino-Roldan, Tomás Peña Ruiz y Manuel Fernández Gómez junto a otros compañeros han publicado en las siguientes revistas científicas de gran impacto mundial:

- The Journal of Chemical Physics nº 132 , de Enero de 2010, el artículo “Density functional theory study of the optical and electronic properties of oligomers based on phenyl-ethynyl units linked to triazole, thiadiazole and oxadiazole rings to be used in molecular electronics”.
- Australian journal of Chemistry nº 63, de Junio de 2010, el artículo “The Role of linear Alkyl and Alkoxy Side chains in the modulation of the structure and electrical properties of Bithiophene: a Theoretical study”
- The Journal of Physical Chemistry C 2010 nº114, de Junio de 2010, el artículo “Theoretical study of Bis(phenylethynyl)thienoacenes as precursors of molecular wires for molecular electronics”.

2. ASISTENCIA A EVENTOS

2.1. II Jornadas Ibéricas de Supercomputación.

El Centro de Supercomputación de Galicia (CESGA) organizó las II Jornadas Ibéricas de Supercomputación, que tuvieron lugar en Braga el 27 de Mayo de 2010. Este evento brindó a investigadores y a centros de supercom-

putación de España y Portugal la oportunidad de exponer sus necesidades y recursos de supercomputación.

El programa contó con la participación de Francisco Herrada Martín (MICINN, España) y Pedro Alberto (UC-LCA, Portugal) que expusieron el estado de la Supercomputación en la Península Ibérica.

Los centros CESGA, RES-BSC, CESCA, FCSCS, CémitS, FPCMUR, CETA-CIEMAT, UC-LCA y CICA contaron las novedades del último año y sus políticas de acceso.

El fabricante Bull presentó la ponencia “Camino del Exaflop. Persiguiendo la mayor eficiencia energética y de espacio”. Esta ponencia corrió a cargo de Jean Marc Denis, responsable del área de Supercomputación en Europa de Bull.

Los proyectos emblemáticos presentados por cada centro fueron los siguientes:

- CémitS, presentó “Estudio de tropopausa extratropical mediante el modelo WACCM”, a cargo de Agustín García García.
- CESGA, presentó “Retos computacionales en el Finis Terrae”, a cargo de Aurelio Rodríguez López.
- RES-BSC, “Grandes proyectos de supercomputación en la RES”, a cargo de David Vicente.
- CESCA, “Modelling solar energetic particle (SEP) events by using 3D magnetohydrodynamic simulations”, a cargo de David Tur.

- FCSCS, “Política Científico-Tecnológica de la FCSC, áreas estratégicas”, a cargo de Carlos Redondo.
- FPCMUR, “Caso de éxito del Centro de Supercomputación de la FPCMUR: interacciones entre el cambio climático y la contaminación atmosférica”, a cargo de José Guillén Mercader.
- UC-LCA, “Lattice QCD@Milipeia”, a cargo de Orlando Oliveira.
- CICA, “TiO2 functionalization by N-ion implantation. A theoretical DFT study”, a cargo de Antonio M. Márquez Cruz

Las jornadas contaron con el patrocinio del MCTES (Ministerio de Ciencia, Tecnología e Ensino Superior de Portugal), MICINN (Ministerio de Ciencia e Innovación), Universidad de Porto, Universidad de Murcia, CESGA, FCT (Fundación para la ciencia y la tecnología), LIP, UT Austin Portugal, Bull, HP, Oracle, Microsoft e IBM.

2.2. III y IV Jornadas de Usuarios de la Red Española de Supercomputación (RES).

En el marco de la Red Española de Supercomputación (RES), el 1 de Julio de 2010 se celebró en Zaragoza la III Jornada de Usuarios de la RES.

El objetivo de las jornadas fue la difusión de los resultados obtenidos por los diferentes grupos de investigación de cada área contemplada (Biología y Ciencias de la



Figura 56: Aspecto de la sala

Vida; Astronomía y Ciencias de la Tierra; Física e Ingeniería, y Química y Ciencias de los materiales) gracias a las horas de Supercomputación concedidas por el comité de acceso en el entorno de la RES que se coordina a través del BSC.

En la reunión IV reunión de usuarios de la RES celebrada el día 15 de diciembre de 2010 en Madrid, se presentó el comité de Acceso y su funcionamiento. Se dedicó especial atención al recientemente formado comité de Usuarios que expuso su funcionamiento, permitiendo la participación de los asistentes en una mesa redonda que ayudó a detectar futuras propuestas de mejora.

Los seminarios incluyeron presentaciones de prestigiosos investigadores que desarrollaron sus actividades en la RES, entre otros contó con la presencia del investigador de la Universidad de Sevilla Javier Fernández Sanz, usuario de los recursos de CICA, que presentó la ponencia “Ab initio Modeling of Chemical Reactivity at Solid Surfaces”, también contó con una mesa redonda en la que participaron miembros del comité de acceso.

2.3. IV Reunión Plenaria de la Red Española de e-Ciencia.

El 2 de Diciembre 2010 se celebró la 4ª Reunión Plenaria de la Red Española de e-Ciencia, se celebró en el Campus de Excelencia Internacional de Universitat Autònoma de Barcelona, en Bellaterra y fué organizado por el Port d'Informació Científica.

En esta reunión se trató el mapa actual de la Red Española de e-Ciencia, de la mano del coordinador de la Red Vicente Hernández, así como el estado de las distintas áreas de e-Ciencia, Infraestructura Grid por Isabel Campos, Supercomputación por Sergi Girona y Aplicaciones por Ignacio Blanquer.

En esta reunión se celebraron sesiones paralelas donde los usuarios de e-Ciencia expusieron sus mejoras y experiencias dentro de las áreas de Infraestructura y Aplicaciones. Ambas sesiones fueron moderadas por Isabel Campos e Ignacio Blanquer, respectivamente.

En la segunda parte del día, tubo lugar la sesión de Cloud computing moderado por Rubén S. Montero. Se expusieron los distintos desarrollos para computación científica relacionados con tecnología Cloud.

3. HARDWARE Y SOFTWARE.

3.1. Introducción a la tecnología hardware.

Dentro del campo del cálculo científico existen distintos modelos computación. Por un lado nos encontramos con los conceptos de HPC, del inglés High Performance Computing, computación de alto rendimiento, y HTC, High Throughput Computing, computación de alta productividad.

En HPC se busca conseguir el más alto rendimiento para la ejecución de una única aplicación. Se va buscando el mayor nivel de paralelismo que admita la aplicación. A mayor nivel de paralelización, menor tiempo real de ejecución (aunque aumente el tiempo de CPU que se necesita ya que hay que sumar todos los tiempos de CPUs de cada uno de los cores que trabajan en paralelo y de forma simultánea para realizar los cálculos).

En el modelo HTC trata de la computación distribuida, de larga o corta duración, definido por servidores separados físicamente y conectados por una red de comunicaciones. Un ejemplo de este modelo sería disponer de servidores de 8 cores cada uno con 16 GB de RAM interconectados por red Gigabit, de forma que si el usuario lanza 32 ejecuciones, que requieren un core y 1 GB de RAM, secuenciales e independientes, usaría 4 servidores de estas características. En el tiempo de una ejecución se hacen 32.

3.2. Hardware para cálculo científico.

En el CICA actualmente se dispone de ambos modelos de cálculo:

- HTC formado por más de 100 servidores de 1 CPU de 2 cores, 3 Gbytes de RAM por core y unidos por una red Gigabit Ethernet, otras decenas de servidores de 4 cores e iguales características de los anteriores.
- HPC, formado por decenas de servidores de 8 cores unidos por Infiniband y servidores de mayor número de cores: 16, y 32 cores constituidos por varias CPU de 4, 6 y 8 cores cada una, según el modelo.

Los recursos disponibles en el CICA, por lo tanto, ofrecen al investigador andaluz distintas opciones dentro de los modelos de HTC y HPC, en pequeña escala. CICA ofrece estos recursos controlados por un gestor de colas, que distribuye de forma equitativa y por necesidades, indicadas por el investigador, las tareas de cómputo. El

gestor de colas empleado en el CICA actualmente, es Sun Grid Engine (SGE), que es gratuito y de libre distribución.

Los recursos hardware están agrupados en varios clusters. Todos son de arquitectura x86, tanto INTEL como AMD, de distintos fabricantes. Mediante el gestor de colas, que se encarga de distribuir las tareas de los investigadores, éstos acceden a los servidores de cálculo, seleccionando la cola de trabajos en función de sus necesidades de cores y memoria por servidor. Más adelante se describen las colas disponibles. Además, esta misma información más detallada está disponible en el Portal de e-Ciencia de Andalucía <https://eciencia.cica.es/cluster/recursos-e-ciencia-cica.html>

Los cluster disponibles en 2010 y 2011 han sido:

1. **Cluster de memoria distribuida:** se trata de un cluster de 100 servidores independientes biprocesadores (200 cores), con 2 GB de RAM por core, arquitectura de 64 bits y unidos por una red Gigabit Ethernet.
2. **Cluster de memoria compartida:** está formado por los siguientes servidores:
 - Servidor Sun X4600, compuesto por 16 cores y 64 GB de RAM.
 - Servidor Sun X4600 M2, compuesto por 24 cores y 64 GB de RAM.
 - Servidores Dell Power Edge, compuestos por dos servidores de 16 cores y 16 GB de RAM.
 - Servidores Nehalem formados por cuatro intel x5550 de 8 cores y 24 GB de RAM.
 - Servidor Nehalem 2, que es un intel x5570 de 8 cores y 12 GB de RAM.
 - Servidor Nehalem 3, que es un intel x5670 de 12 cores y 6 GB de RAM.
3. **Cluster de baja latencia Nova:** está formado por 32 servidores dotados de 2 CPUs Intel Xeon de 4 cores y 4 GB de RAM por core interconectados por una red de baja latencia Infiniband.

Aparte de todo este hardware descrito y englobado en los cluster de cálculo, se dispone de un servidor IBM pSeries 550, compuesto por procesador Power6 dual core modelo P550 a 4,2 GHz y 16 GB RAM.

Este servidor es el único con una arquitectura distinta de la x86 común para todos los demás. Es usado por investigadores que, o bien necesitan un servidor de alta velocidad de CPU para programas que no paralelizan o el

nivel de paralelización es bajo o bien para investigadores que quieren adaptar sus programas para que se puedan ejecutar en la infraestructura de cálculo de la RES (Red Española de Supercomputación). Las librerías que usa este servidor son distintas de las del resto de servidores pero común a la usada en la RES. Esto supone que el investigador que quiera usar sus programas dentro de la RES pueden usar este servidor para retocar sus aplicaciones (necesitarán una recompilación con las librerías de subprogramas específicas de IBM) y así poder mejorar su calificación en la petición de tiempo de cómputo en las convocatorias de asignación de la RES.

Por último, respecto al almacenamiento, cada servidor dispone de discos locales donde se recomienda que los investigadores coloquen los ficheros temporales (o de "scratch") y todos acceden simultáneamente a un sistema de almacenamiento distribuido que usa el software de libre distribución LUSTRE.

Este sistema está soportado por varias cabinas de disco exportadas a través de fibra al servidor destinado a LUSTRE, siendo éste el sistema de ficheros usado en el servicio de supercomputación de CICA. El espacio total exportado es de 12 TB, compartido por todos los nodos y usuarios.

3.3. Software.

Todos los servidores usan el mismo sistema operativo Linux. Tanto en 2010 como en 2011 la versión disponible ha sido la CentOS 5.5.

Respecto a las aplicaciones, compiladores y librerías, sobre las librerías y programas ya instalados en 2009 se ha ido añadiendo todo aquello que los investigadores han solicitado durante 2010 y 2011. A finales de 2011 estaban disponibles:

- * Compiladores: INTEL Fortran, INTEL C/C++ y GNU/gcc.
- * Librerías (bibliotecas de subprogramas):
 - OpenMPI:
 - Openmpi-1.4.4 compiladas para Intel
 - Openmpi-1.2.3 compiladas para gcc
 - Lam-mpi 7.1.4
 - Mpich 1.0.8-p1
 - Mpich 2-1.0.4p1

- MKL 10.1, MKL 10.2 e Intel MKL 9.0
 - Lapack, Scalapack
 - Blas, GotoBlas
 - FFTW, Libtool, PVM, GTK+, Atlas, AMD
 - Hips/Hypre/Jasper, Metis/Parmetis
 - Netcdf, scilab, umfpack
- * Aplicaciones: Perl, Lenguaje R, Octave, Gamess, Geant-4, GnuPlot, Gerris, Lammmps, WRF, Gaussian (bajo Licencia del usuario), WPS, VASP, NCARG, NWCHEM, Python, ant, SIESTA, NAND, Macaulay, Maxima, Sage, 4ti2, Blast, MrBayes.

Actualmente las aplicaciones de pago son adquiridas directamente por el investigador que las va a usar. El equipo de HPC de CICA se limita a instalarlas y ayudar al investigador a su correcto uso.

Las aplicaciones gratuitas las puede solicitar cualquier investigador y el equipo de HPC se encarga de localizarlas, instalarlas y ayudar al investigador para su uso. Como se observa en la lista prácticamente todas las aplicaciones son gratuitas.

3.4. Pruebas de rendimiento.

3.4.1. Introducción.

Durante 2010 y gracias a la colaboración de INTEL y de HP se pudieron realizar pruebas sobre servidores con hardware x86 de última generación, antes de su aparición en el mercado.

Las pruebas de rendimiento se realizaron con aplicaciones de HPC (High Performance Computing), en concreto para tareas fuertemente paralelas. Los servidores evaluados estaban equipados con distintos modelos de procesadores INTEL de las familias Nehalem (4 cores por CPU), Nehalem EX (8 cores por CPU) y Westmere (6 cores por CPU).

Para determinar las prestaciones ofrecidas por los diferentes servidores se ha ejecutado la suite HPCC con diferentes configuraciones, adaptadas en cada caso a la cantidad de procesadores, cores y memoria.

3.4.2. Descripción de los servidores.

En las pruebas se usaron tanto los servidores prestados para la ocasión como hardware propio de referencia, para

hacer los estudios comparativos.

En este apartado se describen las características físicas de las diferentes máquinas sobre las que se realizaron las pruebas, así como el software instalado en cada una de ellas. En la lista siguiente se agrupan las máquinas clasificadas en tres bloques: Servidores con 2 procesadores, servidores con 4 o más procesadores y un último grupo donde se describe una infraestructura virtualizada también testada.

Grupo 1. Servidores con 2 procesadores:

Nehalem12:

- Hardware: 2 procesadores Intel Xeon 5670 2,93GHz (total 12 cores) y 6 GB DDR3 1067 Mhz (6 placas de 1GB).
- Software:
 - S.O. Centos 5.4.
 - Compiladores Intel 10.1.
 - HPCC 1.3.1
 - OpenMPI 1.3.2
 - Librerías MKL 10.1.

Nehalem8:

- Hardware: 2 procesadores Intel Xeon 5570 2,93GHz. (total 8 cores) y 12 GB DDR3 1333 Mhz (6 placas de 2GB).
- Software:
 - S.O. Centos 5.4.
 - Compiladores intel 10.1.
 - HPCC 1.3.1
 - OpenMPI 1.3.2
 - Librerías MKL 10.1.

Chuck1, chuck2, chuck3 y chuck4:

- Hardware: 2 procesadores Intel Xeon 5550 2,63GHz (total 8 cores) y 24 GB DDR3 1333 Mhz (6 placas de 4GB).
- Software:
 - S.O. Centos 5.5.

- Compiladores intel 10.1.
- HPCC 1.4.1
- OpenMPI 1.4.2
- Librerías MKL 10.2.

Grupo 2. Servidores con más 4 o más procesadores:

Nehalem32:

- Hardware: 4 procesadores Intel Xeon 7550 2,00 Ghz (total 32 cores) y 64 GB DDR3 1067 Mhz (16 placas de 4GB).
- Software:
 - S.O. Centos 5.4.
 - Compiladores intel 10.1.
 - HPCC 1.3.1
 - OpenMPI 1.3.2
 - Librerías MKL 10.1.

Grupo 3. Virtualización.

Virtualización con KVM:

- Hardware: 8 procesadores AMD Opteron 8220 2,8GHz (total 16 cores) y 64 GB DDR1 400 Mhz (32 placas de 2GB).
- Software:
 - S.O. Centos 5.4.
 - qemu-kvm 0.10.6
 - libvirt 0.6.2
 - Compiladores gcc 4.1.
 - HPCC 1.3.2
 - OpenMPI 1.3.2
 - Librerías acml.

Virtualización con XEN:

- Hardware: 1 procesador Intel 6400 2,13GHz (total 2 cores) y 4 GB DDR2 533 Mhz (4 placas de 1GB).
- Software:
 - S.O. Debian 5.0.4.
 - libvirt 0.4.6
 - xen 3.2
 - Compiladores intel 10.1.
 - HPCC 1.3.1
 - OpenMPI 1.3.2
 - Librerías MKL 10.1.

Los resultados de las pruebas de cada batería de servidores se encuentran en un anexo final adjunto al documento.

3.5. Nuevos servidores

Para el servicio de supercomputación en 2010 se han adquirido 5 servidores INTEL de 2 fabricantes distintos, equipados con procesadores de la serie x5500 Nehalem con las siguientes características:

- 1 servidor con 2 procesadores x5570 (2,93GHz) , 12 GB de memoria RAM DDR-3 1333 MHz , 2 discos duros de 320GB, controladora de discos LSI Logic / Symbios Logic M1064E MegaRAID SAS y tarjeta InfiniBand Mellanox Technologies MT25204.
- 4 servidores cada uno equipado con 2 procesadores x5550 (2,63GHz), 24GB de memoria RAM DDR-3 1333 MHz, disco duro 160GB y tarjeta InfiniBand Mellanox Technologies MT25204.

Como podemos comprobar en los test, estos procesadores son muy idóneos para tareas de HPC, pues tienen una buena frecuencia de reloj, con 4 cores por procesador, lo que permite alcanzar una muy buena capacidad de cálculo. Así mismo, el ancho de banda de acceso a memoria es suficiente para que las tareas con más demanda no sufran en exceso una pérdida de prestaciones.

La inclusión de tarjetas InfiniBand permiten que tareas con necesidades de paralelización superiores a los 8 cores (los disponibles en un solo servidor) se vean considerablemente menos penalizadas que si se hiciera uso de la

interconexión por ethernet, la cual tiene una latencia mucho mayor.

Otra opción a tener en cuenta para la adquisición fueron los procesadores de la serie x7500, pero pese a tener más capacidad de cálculo el procesador, cada uno de los cores trabajan a frecuencias algo más bajas y una menor capacidad de cómputo teórica, además teniendo en cuenta las características generales de los trabajos de los usuarios, se sacará más provecho de los procesadores adquiridos.

En 2011 se incorporaron a la infraestructura de cálculo intensivo dos servidores HP ProLiant DL585 G7 con 4 procesadores AMD Opteron Processor 6174 de 12 núcleos cada uno, 132 GB de RAM y discos de 500 GB, aportando a la capacidad de computación dos excelente servidores de memoria compartida de 48 cores cada uno.

4. ESTADÍSTICAS ANUALES

4.1. Consumo de CPU y memoria RAM.

Durante el año 2010 se llevó a cabo una iniciativa de ahorro energético. En dicha iniciativa, tanto los diferentes grupos de investigación como los propios centros de

la Junta de Andalucía intentan tomar medidas que minimicen costes y sean beneficiosas para el medio ambiente.

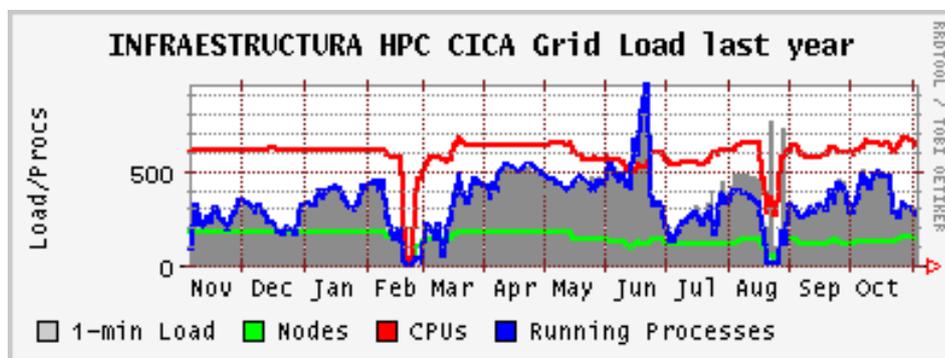
En el CICA, se han implantado diversos métodos para apoyar esta iniciativa. Uno de ellos es el de crear un script que monitorice el estado de todos los servidores, viendo si se ha estado haciendo uso de él durante los últimos 15 minutos. En el caso que no haya reportado actividad, se deja en suspensión, estando a la espera de que haya más actividad en el gestor de colas.

Una vez que haya un número determinado de tareas esperando, se activan nuevamente los servidores a través de Wake On Lan, de forma que siempre se intenta tener ocupados al 100% los servidores encendidos.

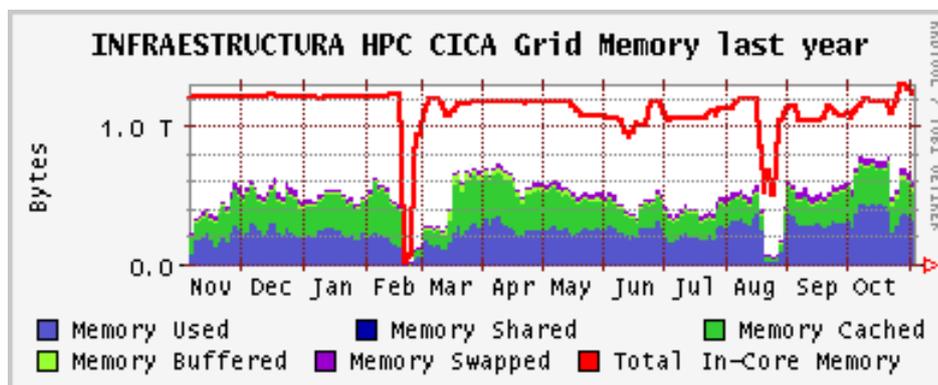
Las estadísticas de uso de la infraestructura de cálculo se pueden ver en las gráficas 57 y 58.

Estas dos últimas gráficas reflejan la actividad de los servidores y la memoria utilizada de las 3 infraestructuras que dispone el cluster CICA (cluster e-Ciencia, cluster BULL e infraestructura para el GRID Nacional, enmarcado dentro del proyecto europeo NGI/EGI).

El descenso brusco que se observa en los meses de Febrero y Agosto se corresponde con las paradas téc-



Gráfica 57: Consumo de recursos del cluster CICA en 2010



Gráfica 58: Consumo de memoria física del cluster CICA en 2010

nicas que se realizaron para labores de mantenimiento.

En la gráfica referente a la carga, podemos ver el uso del script de ahorro energético explicado anteriormente y que se puso en funcionamiento en Mayo de 2010, en el cual los servidores se dejan en suspensión cuando no haya una actividad en alguna de las colas de ejecución.

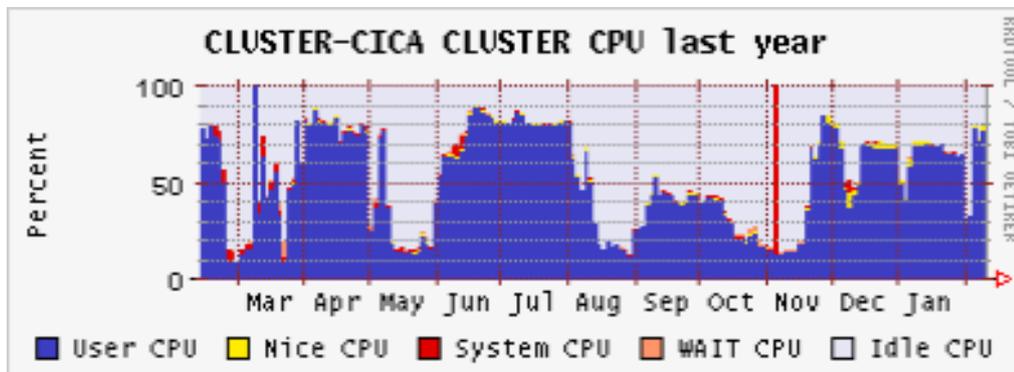
Durante los años 2010 y 2011 se ha seguido con la línea de años anteriores. Se ha apostado por el uso de software de ejecución paralela y multinúcleo, aprovechando casi en un 100% el hardware de la máquina, unido a últimas incorporaciones con máquinas que tienen conectividad Infiniband.

La conexión Infiniband es una tecnología que hemos introducido en el cluster CICA. Es una red cuya latencia teórica es de 160 ns, 10 veces inferior a la de la red de comunicaciones estándar (Ethernet). Gracias al protocolo de transmisión utilizado por Infiniband (RDMA), se consigue liberar al procesador de gestionar el paso de información de un servidor a otro, accediendo directamente a la memoria. Todo esto, unido a la capacidad de utilizar simultáneamente el canal para enviar/recibir in-

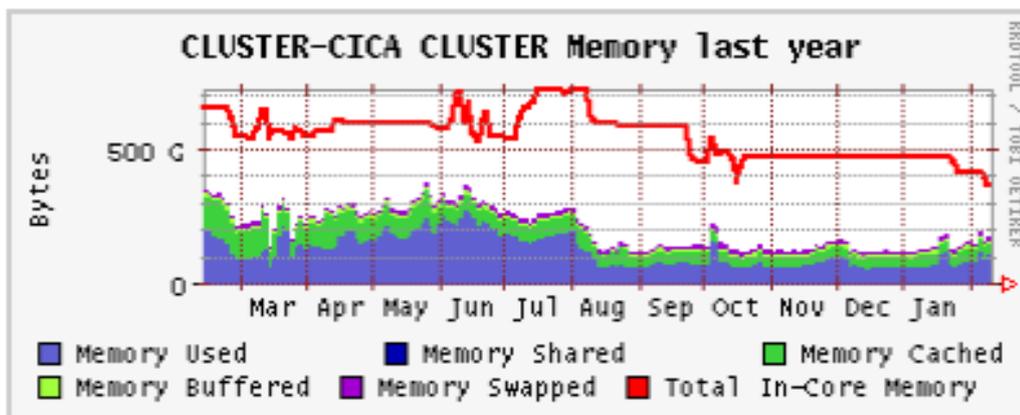
formación por varios dispositivos, hace que sea idóneo su uso en la paralelización de aplicaciones e imprescindible en la computación basada en servidores independientes (memoria distribuida) pero que funcionan conjuntamente como un superordenador de memoria compartida.

Se han abordado trabajos de naturaleza muy diferente, desde multihilo con Java para minería de datos, uso de Gaussian para investigaciones químicas, al modelado meteorológico con WRF, por citar solo algunos ejemplos. Precisamente este tipo de aplicaciones de Química-Física son los que más aprovechan la paralelización, siendo las de porcentaje mayor dentro de las tareas totales que se han ejecutado en el grupo de tareas con una duración de más de 24 horas.

Durante 2011 se ha seguido en la línea marcada en el año anterior, siendo las gráficas correspondientes a dicho periodo las 59 y 60. En las mismas se puede observar que se han producido ciertos periodos valle debido a que dependiendo de los trabajos que se estén realizando se producen picos y valles en el uso del CPU.



Gráfica 59: Consumo de recursos del cluster CICA en 2011



Gráfica 60: Consumo de memoria del cluster CICA en 2011

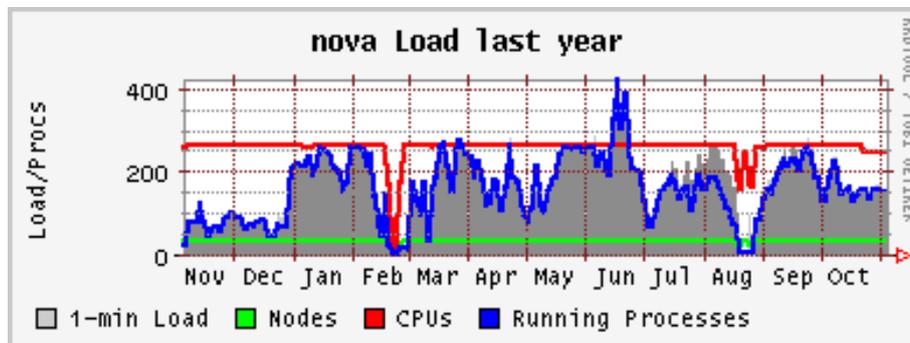
En cuanto al uso de memoria del cluster CICA en 2011 observamos que aunque se ha usado bastante memoria, el cluster esta preparado para cargas incluso mayores de lo consumido lo que permite afirmar que los investigadores han dispuesto en 2011 de todos los recursos de memoria que han necesitado.

4.2. Estadísticas del cluster NOVA BULL

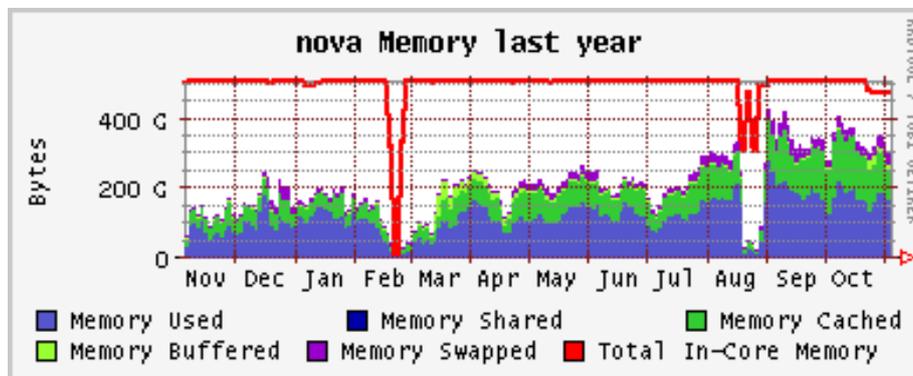
Durante el año 2010 este cluster se ha consagrado consagrado como uno de los que tiene más demanda de uso, que se ha mantenido en 2011. En 2010 su utilización media ha sido de aproximadamente 70-80%,

con épocas en la que la carga ha sido muy superior a la esperada por la cantidad de trabajos ejecutándose. Por el tipo de procesadores que disponen sus servidores, así como la conectividad y la memoria disponible en cada uno de ellos, hacen que sea idóneo para ejecuciones de cálculo muy intensivo y una duración superior a 7 días, tanto paralelo como en serie.

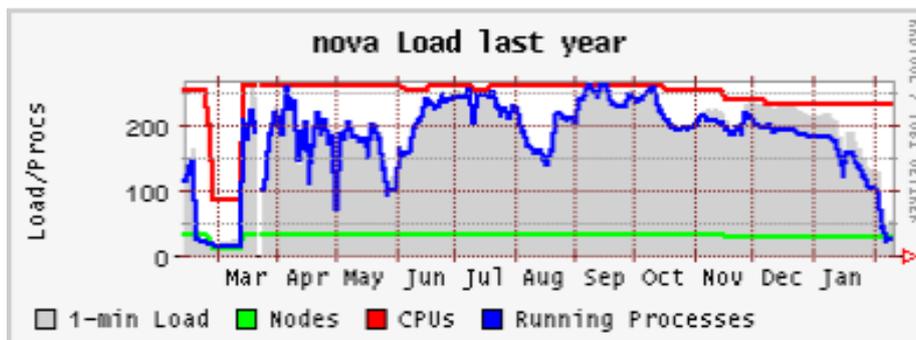
En las gráficas 61 y 62 se puede ver el uso de este cluster. En él ha habido períodos de muy alta demanda, lo que ha hecho que el tiempo medio de espera haya sido ligeramente superior a lo esperado. Por ello, le ofrecemos el acceso a este cluster principalmente a los usuarios que consideramos que pueden aprovechar al máximo las características de él.



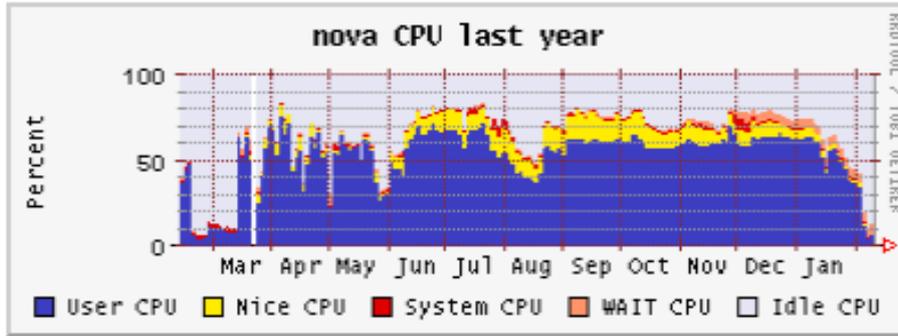
Gráfica 61: Carga del cluster NOVA en 2010



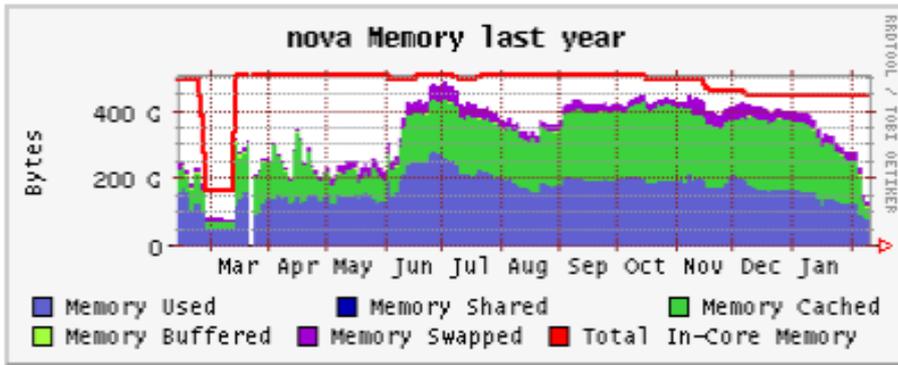
Gráfica 62: Uso de memoria del cluster NOVA en 2010



Gráfica 63: Carga del cluster NOVA en 2011



Gráfica 64: Consumo de CPU del cluster NOVA en 2011



Gráfica 65: Consumo de memoria del cluster NOVA en 2011

El cluster nova ha sido muy demandado el último año, teniendo gran carga de procesos durante todo el año 2011 como se puede observar en las gráficas 63 y 64.

Además, los procesos cargados en el cluster NOVA BULL han tenido un gran uso de CPU, estando su uso medio en torno al 60% con picos cercanos al 80%.

El uso de memoria en 2011 también ha sido alto (ver gráfica 65) y nuestros investigadores han usado mas de 300GB de media con picos cercanos a los 500GB.

4.3. Ejecución de tareas, tiempo medio de espera.

Al analizar los resultados, durante este año se ha producido un incremento de un 92% en el número de tareas ejecutadas en los clusters CICA, considerando todas las colas de ejecución. En las siguientes gráficas (de la 66 a la 69) se representan el del porcentaje de tareas ejecutadas para programas que han tardado menos de 24 en ejecutarse o un tiempo superior, tanto por área de conocimiento como por universidad.

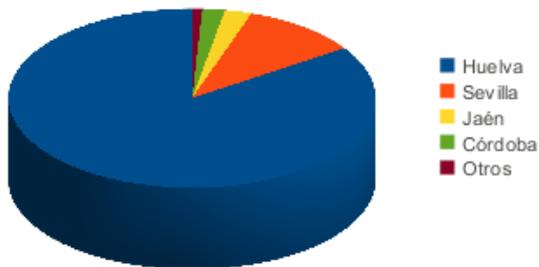


Gráfica 66: Año 2010



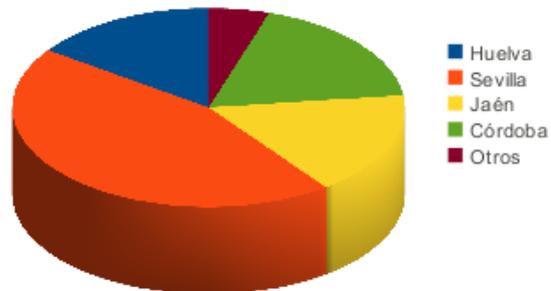
Gráfica 67: Año 2010

% Tareas por Universidad
Tareas con duración < 24 horas



Gráfica 68: Año 2010

% Tareas por Universidad
Tareas con duración > 24 horas



Gráfica 69: Año 2010

Respecto al año 2011, se han ejecutado 327851 trabajos, considerando todas las colas de ejecución. Por universidad, los resultados se muestran en las gráficas 70 y 71.

Se observa un evidente incremento de los trabajos atendidos para la Universidad de Sevilla, aunque en el cluster no existe ningún tipo de limitación a la procedencia de las tareas.

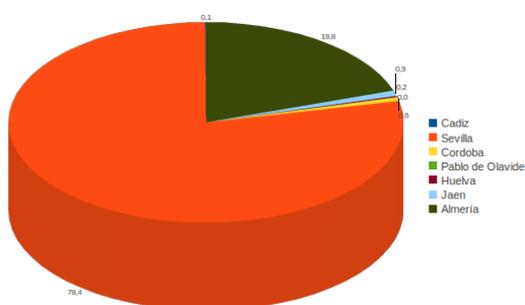
Por último es interesante observar el reparto de tareas por tiempo de ejecución tomando como referencia las 24 horas comentadas, cuyos datos en 2011 se pueden ver en la gráfica 72.

En la gráfica 73 correspondiente a 2010, se observa que en ciertos meses donde, los trabajos se agolpan en las colas de batch, los tiempos de espera se alargan consi-

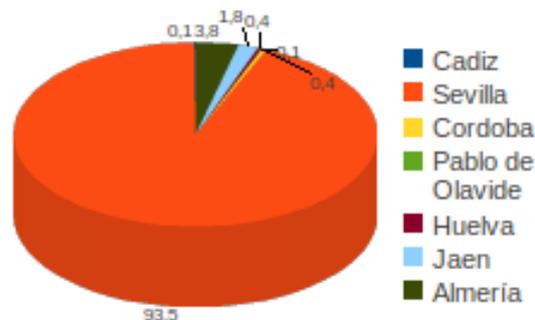
derablemente. Y esto ocurre en todas las colas disponibles para los investigadores. Sin embargo esos tiempos de espera ha sido asumibles durante 2010.

En 2011, los tiempos de espera se han reducido a finales del año en todas las colas de trabajos, como se puede ver en la gráfica 74. El fuerte incremento en dicho tiempo de espera en las colas en marzo y abril de 2011 se debe a la paulatina retirada de equipos para incorporarlos al proyecto europeo EDGI del que el CICA es uno de los dos centros de recursos existentes.

Por otro lado, en la gráfica 75 correspondiente a 2010, se observa el porcentaje de consumo de tiempo de CPU por parte de las tareas que se han ejecutado, agrupadas por Universidades, y sobre un total de 2294467168 segundos (total del wallclock time registrado por Sun



Gráfica 70: Año 2010

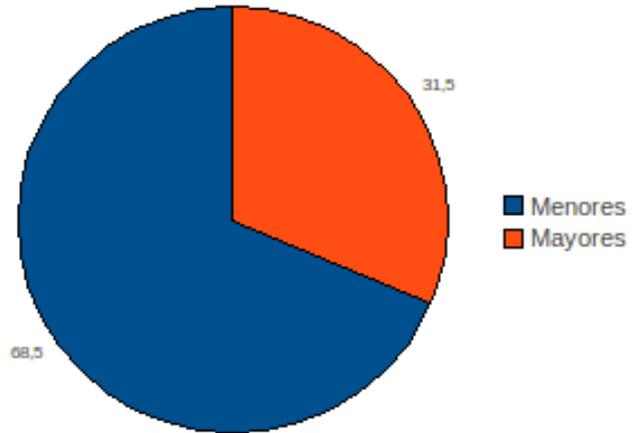


Gráfica 71: Año 2010

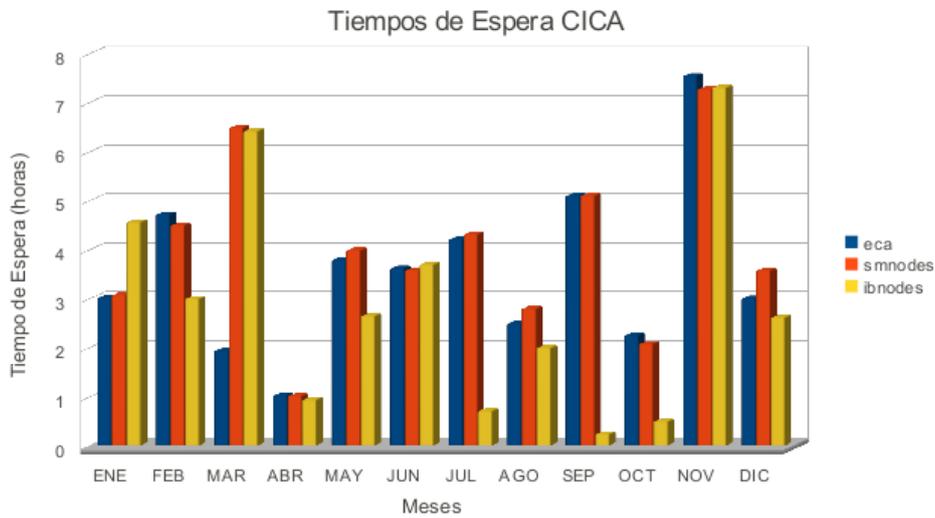
Grid Engine - SGE). Dicho tiempo es el obtenido por la diferencia entre el instante de finalización y el instante de inicio de la tarea.

Durante este año 2010 las Universidades de Sevilla, Jaén, Huelva y Córdoba han sido los usuarios más activos en el cluster de computación, copando casi el 100 % de uso del cluster.

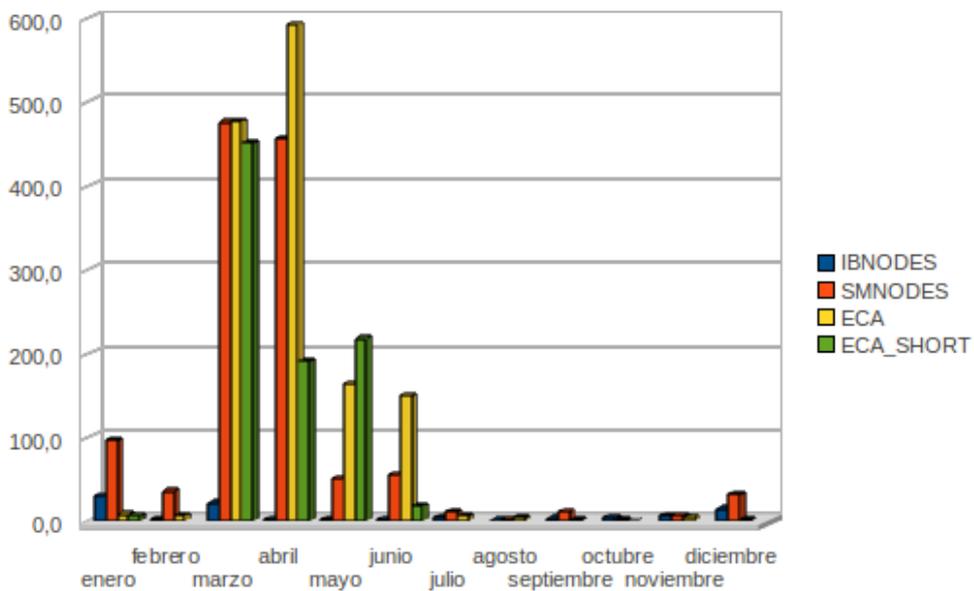
En 2011 (ver la gráfica 76) el reparto ha sido muy distinto, pero es debido a la procedencia de los investigadores que usan estas infraes-



Gráfica 72: Número de tareas según su duración, en 2011

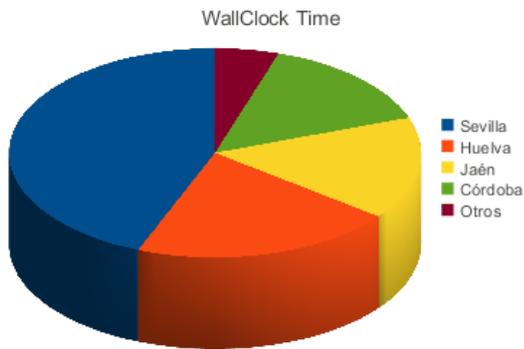


Gráfica 73: Tiempo de espera de los trabajos en las colas, en 2010

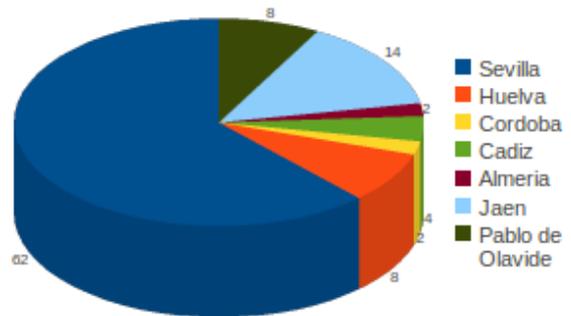


Gráfica 74: Tiempo de espera de los trabajos en las colas, en 2011

Consumo de CPU por Universidades



Gráfica 75: Datos de 2010



Gráfica 76: Datos de 2011

estructuras, no a limitaciones internas.

Por otro lado, para 2011, como se puede observar en la tabla de la figura 77, la mayoría de las operaciones realizadas en dicho año son de un solo procesador. Estos resultados no significan que no se estén ejecutando procesos con paralelismo de grano fino (memoria compartida) sino que refleja, de forma evidente, que la forma de resolución de algoritmos para hacer cálculos y simulaciones basados en la gridificación de aplicaciones es importante.

Esta forma de trabajo se explica fácilmente: en vez de tener un algoritmo que paralelice la ejecución de un modelo, se trabaja con una gran cantidad de parámetros (decenas de miles y hasta millones) y se ejecuta un programa muy liviano para cada grupo de parámetros, pro-

Número de cores	Porcentaje de uso
1	98,1
2	0,3
3	0,6
4	0,4
más de 8	0,5

Figura 77: Porcentaje de tareas por core

gramas que tardan segundos o minutos en ejecutarse. Esta forma de paralelización, que no es tal, se denomina paralelismo de grano grueso. De esta manera en vez de tener un programa que tarde semanas en ejecutarse y que resuelva el modelo en una única ejecución, se eje-

cuta un programa muy pequeño decenas o centenares de miles de veces, cada vez con unos valores de entrada diferentes. Todas estas ejecuciones breves se hacen simultáneamente en cada uno de los distintos cores de cada una de las CPU de cada uno de los servidores disponibles.

Si hay que hacer cien mil ejecuciones pero se dispone de mil cores, y cada ejecución son 3 minutos, se tardarán trescientos minutos en ejecutarlas todas. Pero para la gráfica saldrán cien mil tareas de 1 core frente a 1 tarea para, por ejemplo, 32 cores, del modelo de paralelismo puro.

4.4. Colas de ejecución. Descripción y uso.

Debido a que las necesidades de cálculo de los investigadores son muy diversas, constantemente vamos adaptando los sistemas de cálculo a las necesidades de los mismos.

Por ello, se han preparado varias colas de ejecución, atendiendo a las diferentes necesidades de cálculo de los investigadores. En la nueva organización de las colas, se ha habilitado una orientada a las tareas que no requieren más de 24 horas para su ejecución, que no existía en 2009.

En esta cola, con límite máximo de 24 horas de tiempo de CPU, se ejecutan aproximadamente el 85% de las tareas totales. Éstas están principalmente orientadas a la estadística/minería de datos e investigación operativa, ya que no requieren de mucha capacidad de cálculo en cuanto a velocidad de procesador ni ejecución multicore. Las tareas de este tipo que se han lanzado en 2010

han sido de poca duración: aunque han sido miles de tareas su tiempo de ejecución medio ha sido de 11 minutos y, al usar varios servidores simultáneamente para ejecutar todas esas tareas en 1 día se han podido, en general atender a todas.

Al habilitar dicha cola y no interferir en la eficiencia del/ de los sistemas de cálculo intensivo, se ha incrementado notablemente en la entrada de nuevas tareas para el resto de ejecuciones por parte de otros investigadores. De esta forma, tenemos en el cluster CICA las siguientes colas de ejecución:

- COLA eca: Para investigadores pertenecientes a la e-Ciencia Andaluza y el proyecto NGI/EGI. Son nodos en su mayoría dual-core con 4 GB de RAM cada uno, interconectados a través de una red Gigabit Ethernet. Dicha cola de ejecución se divide a su vez en 2, una con el límite impuesto de 24 horas de ejecución y otra para tareas que requieran de varios días de cálculo.
- COLA smnodes: En dicha cola se incluyen servidores de memoria compartida, con un mínimo de 8 cores y 16 GB por servidor. Este tipo de nodos están dedicados a tareas cuyo cálculo requiere de gran cantidad de memoria RAM.
- COLA ibnodes: En esta cola hemos incluido los servidores del cluster de BULL, cada uno con 8 cores y 16 GB de RAM e interconectados con red de baja latencia Infiniband.

4.5. Evolución en el período 2008-2010

4.5.1. Consumo de recursos

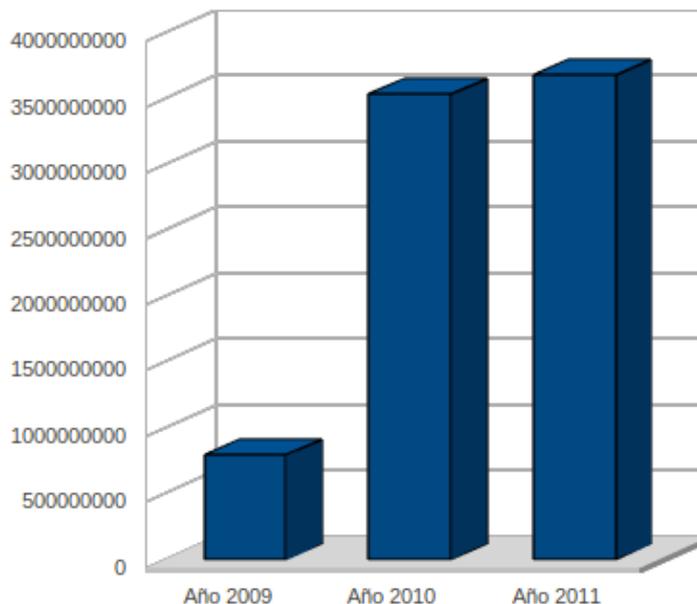
Durante estos años se observa un claro crecimiento del tiempo de CPU consumido por todos los trabajos, como se puede ver en la gráfica 78.

En dicha gráfica se representa el WallClock Time de todas las tareas de cada año, siendo una medida que refleja el tiempo de uso desde que se inicia la ejecución hasta que finaliza. Este tiempo se ha obtenido a través de las estadísticas de uso que tiene registrado el software de cluster Sun Grid Engine (SGE).

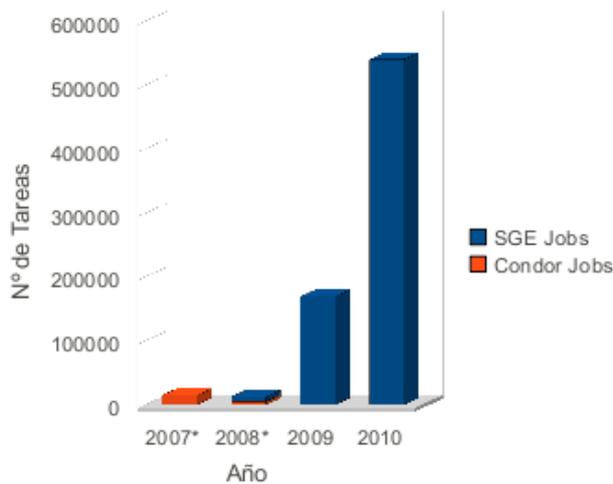
Se observa un leve incremento del consumo de recursos en 2011, que parece estabilizarse pero que debe esperarse a años posteriores para ver la evolución.

Otra gráfica de interés (gráfica 79) es la que refleja el número de tareas ejecutadas en el cluster desde el año 2008 al año 2010, usando el gestor de colas de SGE.

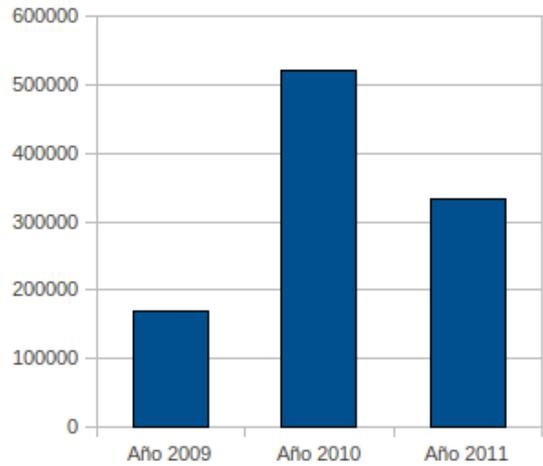
En la misma gráfica 79 se representa la evolución del número de trabajos ejecutados entre los años 2007 y 2010. En 2007 solo existía el gestor de colas CÓNDROR. En 2008 coexistieron los gestores de colas CÓNDROR y SGE y a partir de 2009 se retiró el primero ya que tras las pruebas experimentales realizadas en 2008 sobre SGE se observó que éste era más adecuado por versatilidad y facilidad de uso.



Gráfica 78: Consumo de recursos 2009-2011



Gráfica 79: Número de tareas en 2007-2010



Gráfica 80: Número de tareas en 2009-2011

Los datos del número de tareas ejecutadas que incluyen el año 2011 están en la gráfica 80, ya solo con el gestor de colas SGE.

4.5.2. Tiempos de espera en las colas

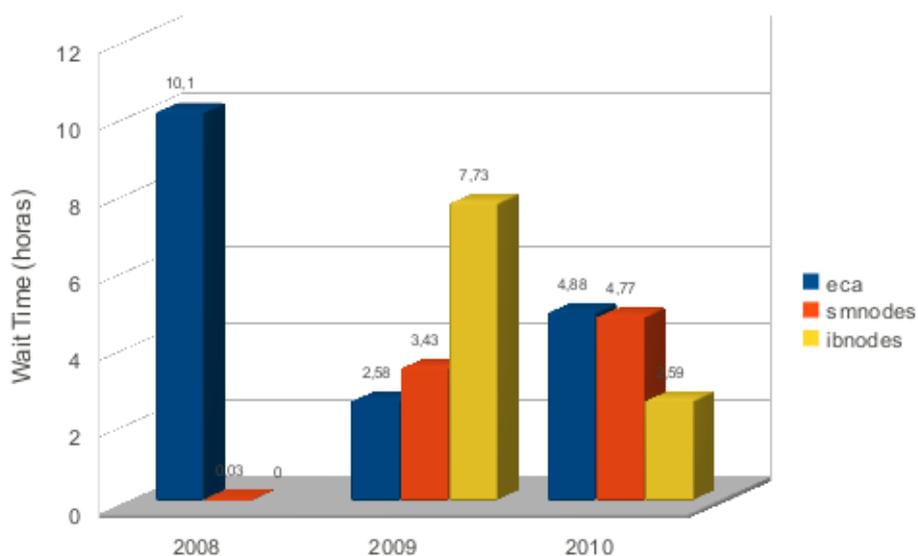
En la gráfica 81 se representa la evolución desde 2008 al 2010, del tiempo de espera en las colas, es decir, el tiempo transcurrido desde que el investigador manda el trabajo a una cola hasta que el trabajo comienza a ejecutarse al quedar cores de cálculo libres.

En dicha gráfica podemos observar como a medida que se ha incrementado el número de máquinas, el tiempo

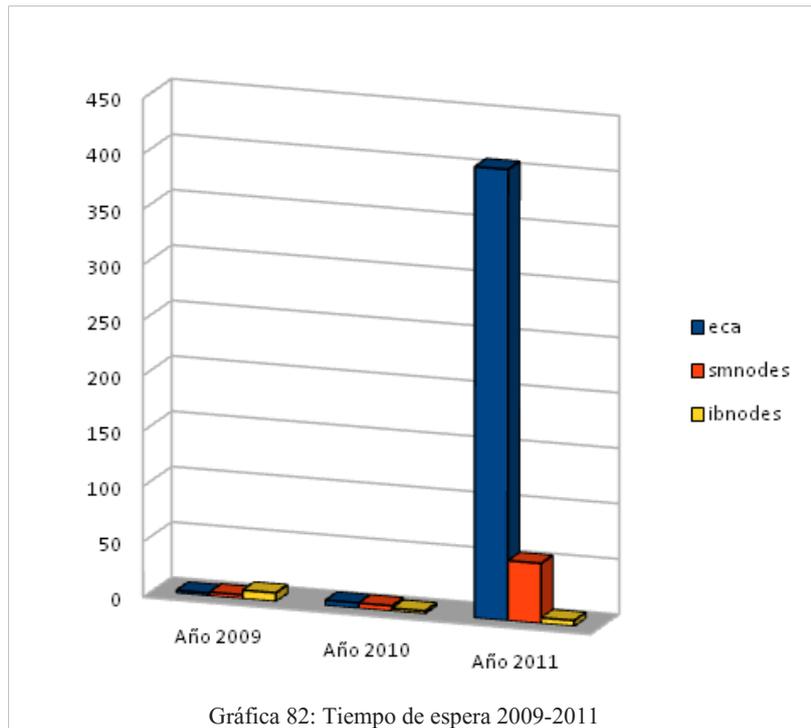
de espera medio se ha ido estabilizando. Durante el año 2008 tan sólo había una cola principal donde ejecutar las tareas (cola eca), por lo que el tiempo de espera era muy alto (aproximadamente 10 horas).

Durante los años posteriores, el hecho de incorporar nuevas máquinas de memoria compartida hizo que se redujeran estos tiempos, pasando a 2 horas (año 2009) y 4 horas (año 2010) en cada una de las colas de ejecución principal (eca y smnodes).

Analizando la cola ibnodes que corresponde al cluster de servidores BULL, cabe destacar que durante el año 2009 tan sólo tuvo 8 servidores durante gran parte del



Gráfica 81: Tiempo de espera en las colas de trabajo



Gráfica 82: Tiempo de espera 2009-2011

año, lo que hizo que el tiempo de espera fuese muy superior al resto de colas, problema que se solventó a partir de finales del 2009 cuando se amplió a 32 servidores. Una vez hecha esta ampliación, se observó que se redujeron drásticamente los tiempos de espera de los trabajos enviados a dicha cola, siendo inferior al del resto de las colas.

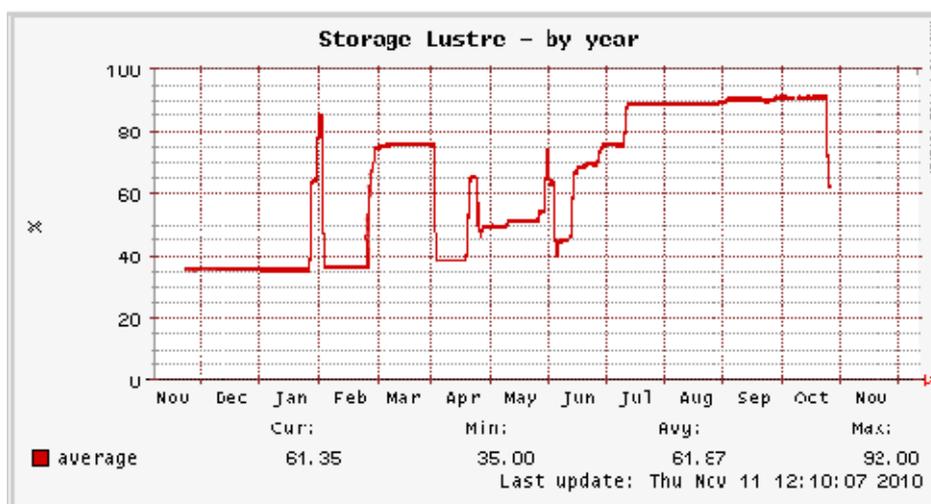
En la gráfica 82 se representa el periodo 2009 a 2011, habiendo sufrido un gran incremento en el tiempo medio de espera debido a la parada que hubo de realizarse para preparar la infraestructura de la cola ECA para adaptarla al proyecto europeo EDGI en el que CICA participa como centro de recursos.

4.6. Almacenamiento en disco compartido

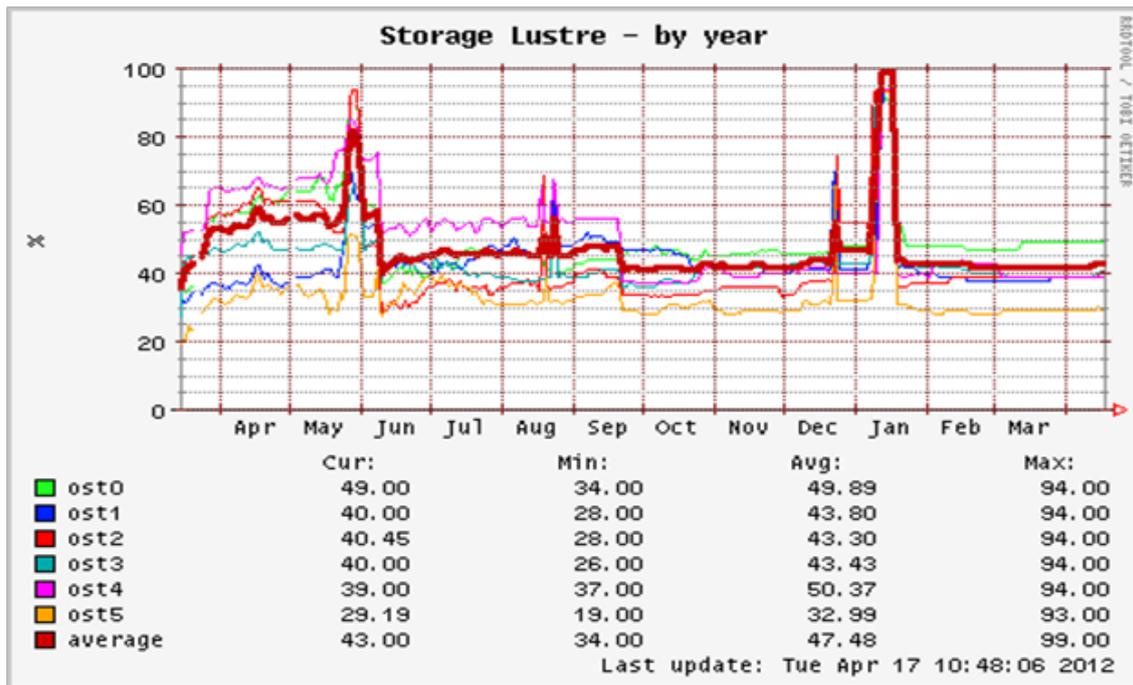
Durante el año 2010 se llevó a cabo una ampliación del sistema de ficheros que comparten todos los servidores de cálculo usando el software de fuentes abiertas LUSTRE. En concreto se añadieron 1.8 TB, para atender la constante demanda que se hacía de él.

Como puede verse en la gráfica 83, durante ese año ha habido épocas con una alta ocupación, dependiendo ésta del número de tareas que se estaban ejecutando en cada momento.

En Agosto de 2010 se realizó una parada de manteni-



Gráfica 83: Consumo de almacenamiento en disco compartido, en 2010



Gráfica 84: Consumo de almacenamiento en disco compartido, en 2010

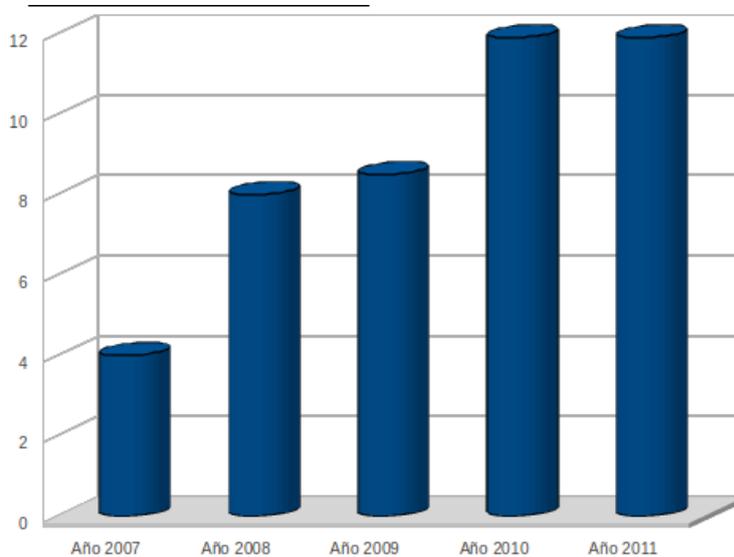
miento del sistema de ficheros, viendo la alta demanda que estaba teniendo durante los 3 meses anteriores (llegando a un límite de 92% de ocupación). En este punto fue necesaria la ampliación mencionada para evitar posibles riesgos de corrupción en los nuevos ficheros, ya que varios discos del sistema de almacenamiento se habían sobrecargado y podían dar lugar a fallos.

En el CICA se incorpora la monitorización los OST. Un OST (Object Storage Target) es básicamente cada uno de los diferentes discos lógicos (formados por particiones de discos físicos) que componen el sistema de alma-

cenamiento distribuido LUSTRE. Dichos OSTs forman un conjunto visible por todos los clientes, y administrado a su vez por el MGS (Management Server) y el MDT (Metadata Target).

Todo este conjunto hace de LUSTRE un sistema robusto y perfectamente escalable para que sea usado de forma compartida y simultánea por muchos servidores.

En la gráfica 84 se representa la monitorización del estado de cada OST en 2011, tal y como hemos comentado, que se comenzó a recolectar en Noviembre de 2010.



Gráfica 85

Evolución de la capacidad de almacenamiento en disco

En la gráfica 85 se representa la evolución de la capacidad máxima del sistema de ficheros Lustre de 2007 a 2010. Este sistema de ficheros distribuido ha estado funcionando tanto con el gestor de colas CÓNDROR, como con SGE e incluso con ambos coexistiendo, lo que nos ha permitido hacer cambios y modificaciones en estos años sin necesidad de paradas importantes de reestructuración de los sistemas de archivos.

Al comienzo de la implantación del cluster de cálculo, se planteó un sistema con aproximadamente 4 TB de almacenamiento, junto con el gestor de metadatos de todo el sistema que ocupaba otros 200 MB adicionales.

En sucesivas operaciones de mantenimiento y ampliación del sistema de ficheros, se han ido añadiendo discos de 2 TB cada uno, junto a un gestor de metadatos de 400 MB, necesario para tener un control de todos los ficheros almacenados.

5. ACCESO A LOS RECURSOS

5.1. ¿Cómo hacerse usuario del CICA?

La solicitud de alta a los recursos de computación y e-Ciencia de CICA se realiza a través del portal de eCiencia de Andalucía (<https://eciencia.cica.es>).

Este servicio podrá ser solicitado por cualquier ciudadano investigador andaluz o que trabaje en Andalucía y que cumpla uno de los siguientes requisitos:

- Ser investigador o docente en una de las 10 universidades públicas de Andalucía
- Ser un Investigador perteneciente a un grupo de investigación oficial de la Junta de Andalucía, pudiendo ser docente en un Instituto de Enseñanza Secundaria o de FP.
- Ser un investigador con proyectos dotados de fondos públicos que trabaja en un centro ubicado en Andalucía.
- Ser investigador de uno de los centros del CSIC ubicados en Andalucía,
- Ser investigador de uno de los centros del IFA-PA
- Colaboradores externos (españoles o extranjeros) de investigaciones dirigidas por un investigador de los grupos anteriores.

El proceso de alta es el siguiente:

1. Registro como usuario.

En este primer paso se solicitarán unos datos previos: nombre y dirección de correo. Con el fin de agilizar el proceso, se solicitará una dirección de correo con dominio de institución pública.

2. Solicitud de datos de usuario.

El usuario una vez aprobada su alta recibirá un correo solicitando los datos necesarios.

3. Envío de instrucciones

Una vez recibidos los datos solicitados al usuario, se le envía un email facilitando las instrucciones de acceso al servicio.

5.2. Descripción del portal de e-Ciencia

El Portal de e-Ciencia del CICA pretende ser el lugar de encuentro de los investigadores que usan las TIC para su investigación, cálculos y difusión de los mismos.

Cada año se añaden nuevas funcionalidades y es el lugar de entrada a los sistemas de cálculo intensivo del CICA

A finales de 2010 el portal (ver figura 86) incorpora varios módulos, cada uno con su funcionalidad. Cada una de éstas funcionalidades a través de sus correspondientes menús se describen más a continuación, indicando las principales características de cada una de ellas:

a. Menú lateral

En este menú (figura 87) se encuentran los enlaces para acceder a los distintos apartados de la web, siendo algunos de ellos los siguientes:

- *Noticias:* En esta sección se pueden encontrar las noticias de diversa índole re-lacionadas con la supercomputación, tanto a nivel interno del CICA, incidencias en los servicios, anuncios de conferencias, cursos organizados, actualizaciones y compras de nuevo equipamiento, o a nivel nacional e internacional, congresos o cursos que pueden ser de interés para los investigadores que hacen uso de nuestros servicios, etcétera.
- *Recursos CICA:* Se muestran los recursos hardware y software disponibles para supercomputación. No obstante, aunque se encuentra disponible un amplio abanico de posibilidades y herramien-

Portal de e-Ciencia de Andalucía

Portada | Noticias | Contactar | Mapa Web | Encuesta | Información de Acceso | JIS 2009 | Colaboraciones | Farbwahl + - R



Información General

- Inicio
- Noticias
- Recursos CICA
- Alta Grupos Inv.
- Enlaces Grupos Inv.
- Cluster Looking Glass
- Monitorización
- Manuales y Descargas
- Publicaciones
- Información de Interés
- Notas de prensa y Eventos

Acceso

Usuario

Clave

Recordarme

[¿Recuperar clave? Regístrate aquí](#)

Noticias RSS

Centro Informático Científico de Andalucía

- Restaurado servicio webmail IMAP y POP
- Corte en servicio de webmail IMAP y POP
- Corte en las comunicaciones de RICA
- Acceso a WoK a través de la FECyT
- Parada del portal wokv3.cica.es

Sindicación

RSS	0.91
RSS	1.0
RSS	2.0
ATOM	0.3
OPML	SHARE IT!

Publicación artículo e-Ciencia

miércoles, 27 de abril de 2011

El Departamento de Química Física y Analítica de la Universidad de Jaén ha publicado recientemente un nuevo artículo titulado "Optoelectronic and Charge Transport Properties of Oligomers Based on Phenylethynylene Units Linked to Thieno-acenes: A DFT Study", en el que una parte ha sido apoyada en CICA a través de sus recursos de Supercomputación. Pueden acceder a su contenido a través del enlace a Artículos y Presentaciones en e-Ciencia:

Optoelectronic and Charge Transport Properties of Oligomers Based on Phenylethynylene Units Linked to Thieno-acenes: A DFT Study

Más...

- EDGI demonstrates how to submit 10,000 jobs to a Volunteer Desktop Grid from within the EGI
- Joint European DCI Summer School. Budapest, Hungary, 11-16 July 2011.
- ICTS-CESGA: Call for access proposals 2011 - 2nd period
- [Parada Técnica No Programada] Servidores de cluster CICA
- [Parada Técnica] Finalización Labores de Mantenimiento

<< Inicio < Anterior 1 2 3 4 5 6 7 8 9 10 Siguiente > Final >>

Resultados 1 - 6 de 70

Últimas noticias

- Publicación artículo e-Ciencia
- EDGI demonstrates how to submit 10,000 jobs to a Volunteer Desktop Grid from within the EGI
- Joint European DCI Summer School. Budapest, Hungary, 11-16 July 2011.
- [Parada Técnica No Programada] Fin de la Intervención

Más visitados

- 1ª Conferencia IBERGRID 2007
- Red Andaluza de Supercomputación Científica - e-Ciencia Andaluza
- Computación Voluntaria
- 1º Congreso Internacional FLOSS



CONSEJERÍA DE ECONOMÍA, INNOVACIÓN Y CIENCIA

Información de Acceso

- Esquema General
- Conexión SSH
- Sistema de colas SGE
- Almacenamiento SGE a través de la web

Aplicaciones

Manuales

Powered by JoomlaGadgets

CIENCIA DIRECTA ES UNA WEB DE DESBREE

Fundación Andaluza para la Divulgación de la Innovación y el Conocimiento

Andalucía Investiga

www.andaluciainvestiga.com



versión beta

Red Andaluza de Supercomputación Científica



PKIRIS

powered by projects



BOINC

Figura 86: Portal web de e-Ciencia del CICA en 2011

tas a disposición de los investigadores, aquellos usuarios que lo requieran pueden consultar por la disponibilidad de nuevo software no instalado u otros recursos hardware que satisfagan sus necesidades.

- *Grupos de investigación*: Una selección de enla-

ces a grupos de Investigación Andaluces, de Supercomputación y de e-Ciencia.

- *Cluster Looking Glass*: Enlace para acceder al servicio web desarrollado en CICA con el fin de facilitar el acceso de los investigadores a los recursos de supercomputación ofrecidos en el centro.



Figura 87: Menú

- **Monitorización:** En esta sección se pueden visualizar el estado de los nodos de cálculo, siendo esta una útil herramienta para los usuarios, pues pueden comprobar de una forma fácil y cómoda los recursos que están usando las tareas que tienen en ejecución o el estado de ocupación de los diferentes sistemas que componen las colas a las que envían sus trabajos.
- **Manuales y descargas:** Podrá encontrar guías de uso del cluster, cómo enviar sus trabajos haciendo uso de las diferentes facilidades que se ofrecen, así como otros manuales y guías que se estimen interesantes para que los usuarios puedan hacer un correcto uso de los recursos ofrecidos.
- **Publicaciones:** Artículos o referencias a artículos donde se citan los recursos de computación ofrecidos por CICA, o que han sido escritos por el departamento de supercomputación de CICA.

b. Noticias RSS

En este módulo (figura 88) aparecen las noticias publicadas en el portal de CICA, con ello se pretende que las noticias publicadas en dicho portal tengan



Figura 88: Menú RSS

la mayor difusión posible dentro de los usuarios de nuestros servicios.

c. Módulo de información directa

La finalidad de este módulo (figura 89) es hacer más cómodo el acceso a recursos de información útil para los usuarios, como son guías de usuario de los diferentes servicios o descripciones de las colas de ejecución disponibles.

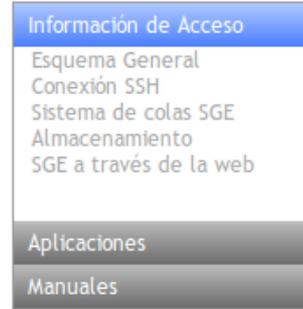


Figura 89: Menú de información directa

d. Sección de banners y enlaces

Esta sección (ver figura 90) está dedicada a albergar banners y enlaces a diferentes webs de proyectos con los que colaboramos.



Figura 90: Banners y links de colaboraciones

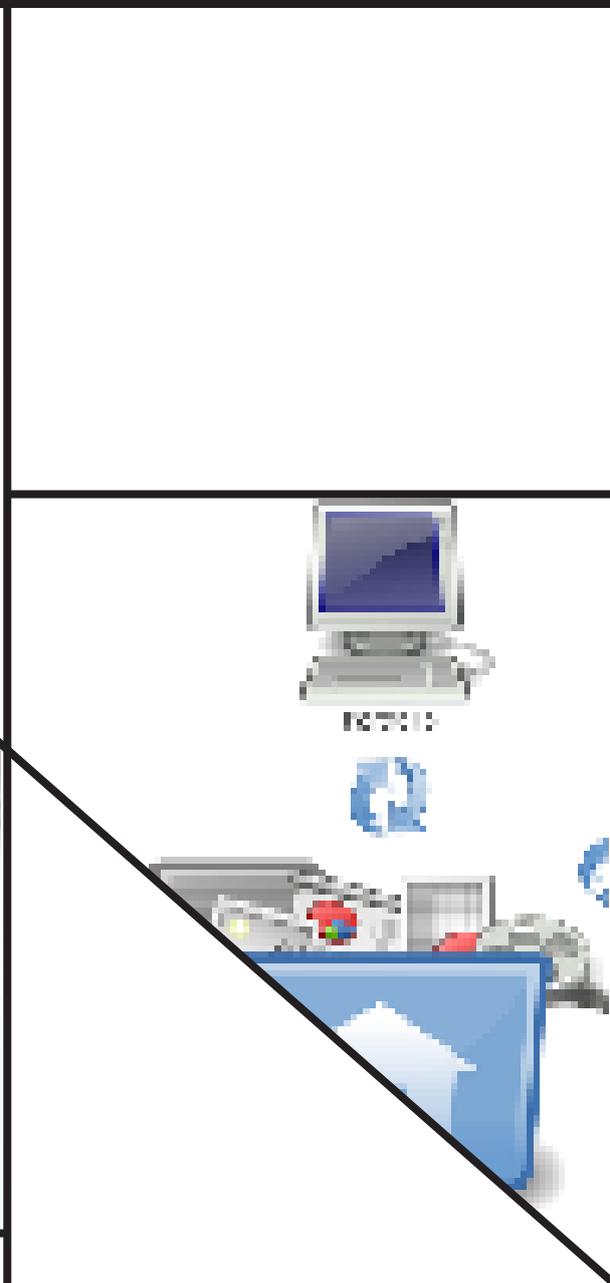
COMUNICACIONES

SEGURIDAD

SISTEMAS / HPC

APLICACIONES

FORMACIÓN Y EVENTOS



Aplicaciones

Como sucedió en 2009, tanto en 2010 como en 2011 el Sistema de Información Científica de Andalucía (SICA) se ha llevado buena parte del esfuerzo

del equipo de técnicos de CICA. De esta forma, se han seguido realizando mejoras sobre el sistema, todas ellas encaminadas a seguir ofreciendo un óptimo servicio a los usuarios.

Asimismo, en 2010 se introdujeron nuevos sistemas y hardware, se han corregido los problemas de memoria y se han introducido mejoras en la configuración de los servidores web, haciendo que durante dicho año y buena parte de 2011 el funcionamiento de SICA fuese satisfactorio para los investigadores andaluces.

Ya en 2011, CICA ha puesto a disposición de la empresa que está desarrollando la nueva versión de SICA los equipos de pruebas y desarrollo, de forma que a finales de 2011 ya se puso en explotación sobre el mismo hardware que se usaba en la versión anterior de SICA el acceso a la nueva versión.

Por otro lado, durante 2010 y 2011 CICA ha seguido prestando el servicio de soporte de infraestructura informática al Catálogo Colectivo del Consorcio de Bibliotecas Universitarias Andaluzas así como a la Forja de Software Libre de RedIRIS-CICA y a portales web institucionales de Difusión de la Ciencia.





ATÁLOGO COLECTIVO DEL CONSORCIO DE BIBLIOTECAS UNIVERSITARIAS ANDALUZAS

Este catálogo aglutina los libros, revistas y documentos disponibles en las bibliotecas universidades públicas andaluzas, por lo que cualquier alumno o profesor, y a través de una única interfaz web, puede consultar todos los libros existentes en todas las bibliotecas universitarias y hacer su reserva. Y en 2011 se incorporó el servicio de préstamo interbibliotecario automático con el módulo de circulación (ver figuras 91 y 92).

La aportación del CICA al proyecto incluye los servidores, almacenamiento en red SAN (Storage Array Network) de alta disponibilidad, copias de

seguridad, seguridad y el personal para administración de sistemas y soporte a los bibliotecarios de sistema encargados del proyecto.

Por otro lado, en 2010 se sigue prestando el servicio de acceso a las digitalizaciones de publicaciones antiguas (libros y documentos) de las bibliotecas de las universidades del Reino Unido, producto conocido como EEBO (Early English Books Online). Este proyecto dispone de más de 125.000 títulos, entre los que se incluyen los textos completos y las descripciones catalográficas de los libros publicados en Gran Bretaña entre los años 1475 y 1799.

La información es accesible tanto en el servidor del editor de la colección, como desde los servidores ubicados en CICA. En concreto, en éste último es-

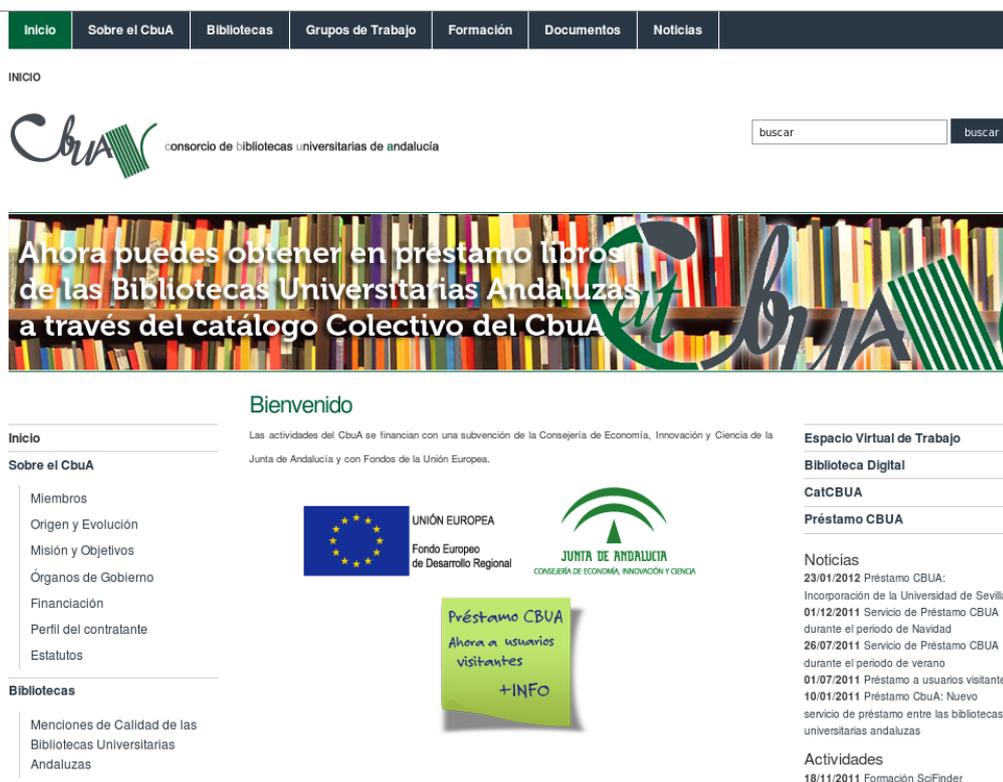
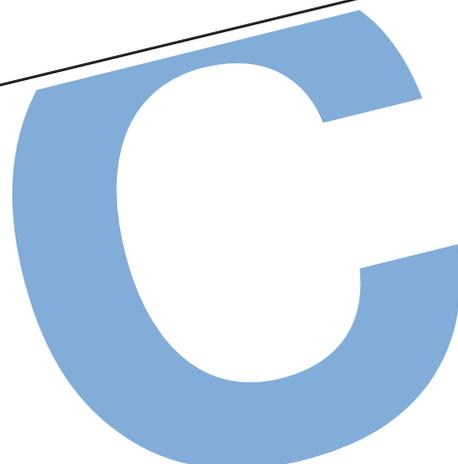


Figura 91: Entrada a la WEB del CBUA



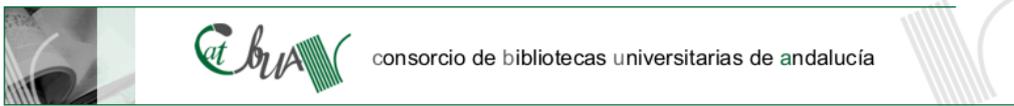


Préstamo CBUA
 Ahora a usuarios
 visitantes
 +INFO

Consultar Catálogo Consorcio Consulte los fondos de la biblioteca para encontrar libros, libros electrónicos, recursos digitales, imágenes, revistas y mucho más.

PALABRAS CLAVE Búsqueda Avanzada

- servicios**
 - ▶ Préstamo CBUA
 - ▶ Otras bibliotecas vía z39.50
 - ▶ Guías del catálogo
- bibliotecas**
 - ▶ Bibliotecas integrantes



WebPAC PRO © Innovative Interfaces, Inc.

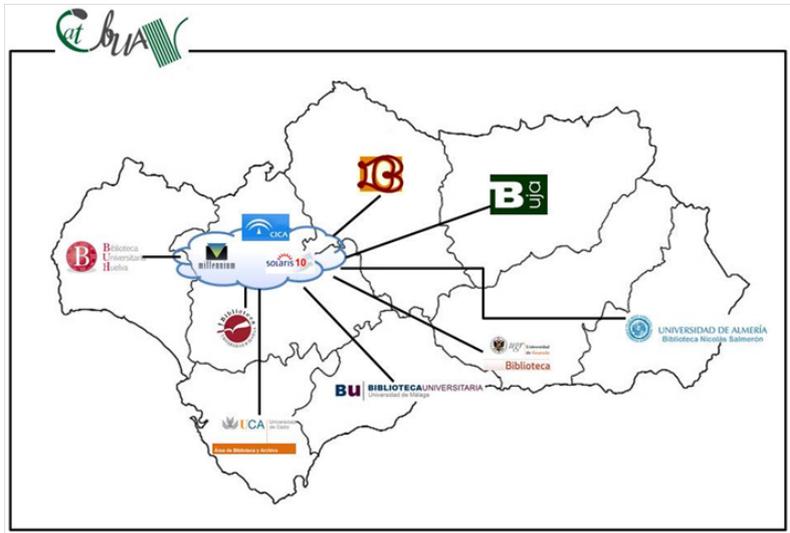
Figura 92: Página de entrada al catálogo Colectivo del CBUA

tán accesibles los libros completos en formato PDF.

Para este servicio CICA aporta el almacenamiento y distribución vía ftp de dichos fondos digitalizados a través de los catálogos de las bibliotecas universitarias de Andalucía, incluido el Catálogo Colectivo del CBUA.

Estos fondos, en inglés antiguo, fueron adquiridos por el mismo CBUA para enriquecer la oferta local de las universidades andaluzas.

El acceso a esta bibliografía se ha mostrado más eficiente tras las mejoras en el hardware aportados por CICA, al



utilizar un almacenamiento en red de alto rendimiento y con tolerancia a fallos basado en una red SAN (Storage Area Network).

Por último, indicar que durante 2010 y 2011 se ha trabajado en el proceso de cambio de la infraestructura hardware/software del sistema informático que atiende este Catálogo Colectivo, habiéndose decidido migrar a una nueva versión basada en LINUX Red Hat que permitirá prescindir de hardware y sistema operativo propietario (SUN SPARC y Solaris) Utilizado hasta ahora.

2. PROYECTO CONFIA

2.1. Introducción

Confía es la federación de identidades de las universidades andaluzas. Su principal objetivo es permitir a las

personas que pertenecen a esta comunidad utilizar servicios distribuidos manteniendo el nombre de usuario y la contraseña de su universidad de origen.

La idea parte de las propias universidades y está avalada por la Asociación de Universidades Públicas de Andalucía (AUPA), a través de su Comisión Sectorial de Tecnologías de la Información y Comunicación (AUPA-TIC).

Una vez creada la plataforma de Identidad Federada se decide por parte de la AUPA-TIC que el servidor que hace de coordinador esté en el CICA, para lo que este centro aporta al proyecto los servidores y se ayuda a la instalación. En la imagen 93 se puede ver el portal del sitio web de CONFÍA (<https://confia.aupa.info>).

La primera función del servicio es apoyar al Campus Andaluz Virtual de forma que no hayan que intercambiarse los datos de los alumnos de una universidad a otra sino que usen el sistema CONFIA. Con éste cada alumno que



Figura 93: Portada del sitio web de CONFIA



acceda a los cursos de una universidad distinta a la suya usa los sistemas de identificación de su propia universidad.

Ello implica que una universidad dará acceso a sus asignaturas (las que ofrece al Campus Andaluz Virtual) a los alumnos que entren desde otra de las universidades andaluzas "confiando" en el visto bueno que éstas le den para ese alumno que pretende acceder.

Aparte de este servicio, durante 2010 y 2011 se han arrancado otros nuevos que se ofrecen entre las universidades públicas andaluzas usando este sistema de acceso federado.

2.2. Infraestructura hardware y servicios aportados.

El CICA aporta para este proyecto:

- Equipamiento: 4 servidores físicos y 2 servidores virtuales.
- Alojamiento de una base de datos en uno de los servidores multipropósito del CICA.
- Servicio de soporte:
 - Ayuda del equipo de administración de sistemas cuando lo han demandado desde la empresa que gestiona el proyecto o por parte de los técnicos informáticos que en cada universidad gestionan el funcionamiento de la parte local del proyecto.
 - Instalación de sistema operativo en servidores.
 - Definición de las conexiones a la red académica andaluza (RICA).
 - Instalación del software de los balanceadores de carga del servicio.
 - Backup diario de la instalación de coordinación.
 - Sistemas de alimentación eléctrica redundante

2.3. Descripción del estado actual de los elementos comunes.

En CONFIA existen actualmente 3 entornos comunes, desplegadas en máquinas del CICA, que se utilizan para distintos fines:

- Entorno de producción. Es el que se conecta con los IdPs (Identity Provider – proveedor de identidad) y SPs (Service Provider – proveedor de servicio) de producción de las diferentes universidades que a su vez van conectados con las fuentes de datos usuarios y LMS (Learning Management System) finales que participan en el Campus Andaluz Virtual (CAV).
- Entorno de test. Es el que sirve para validar un IdP o un SP antes de pasar al entorno de producción.
- Entorno de laboratorio. Es el que sirve para experimentar con los IdPs, SPs, conectores y diferentes desarrollos que se hagan bajo el amparo de CONFIA.

En cada uno de estos entornos se dispone de los siguientes elementos:

- IdP-Homeless. Proveedor de identidad para aquellas personas que no pertenecen a ningún IdP de ninguna universidad pero que participan en CONFIA.
- SP (test). Proveedor de servicio que se utiliza para hacer pruebas del IdP, mientras no se dispone de un SP propio.
- WAYF (Where Are You From). Nodo central que se utiliza como servicio de descubrimiento de los diferentes IdPs de la federación y como punto común en el que se valida la información y se aplican filtros comunes de seguridad.
- Gestor de metadatos (Janus). Registra los diferentes metadatos de las entidades que forman la federación CONFIA. Metadatos que luego exporta para que sean recogidos por el módulo metarefresh de cada IdP/SP para automatizar la gestión de los metadatos.

3. FORJA RedIRIS - CICA

El Software Libre ofrece una tecnología de última generación que permite el desarrollo de software potente y de calidad. Además puede ser un idóneo punto de partida en la formación de nuestros jóvenes universitarios.

Una solución Software Libre se basa principalmente en un desarrollo distribuido donde los programadores se dividen las tareas de forma que cada uno genera la suya localmente y la envía a un repositorio central donde reside la aplicación completa.

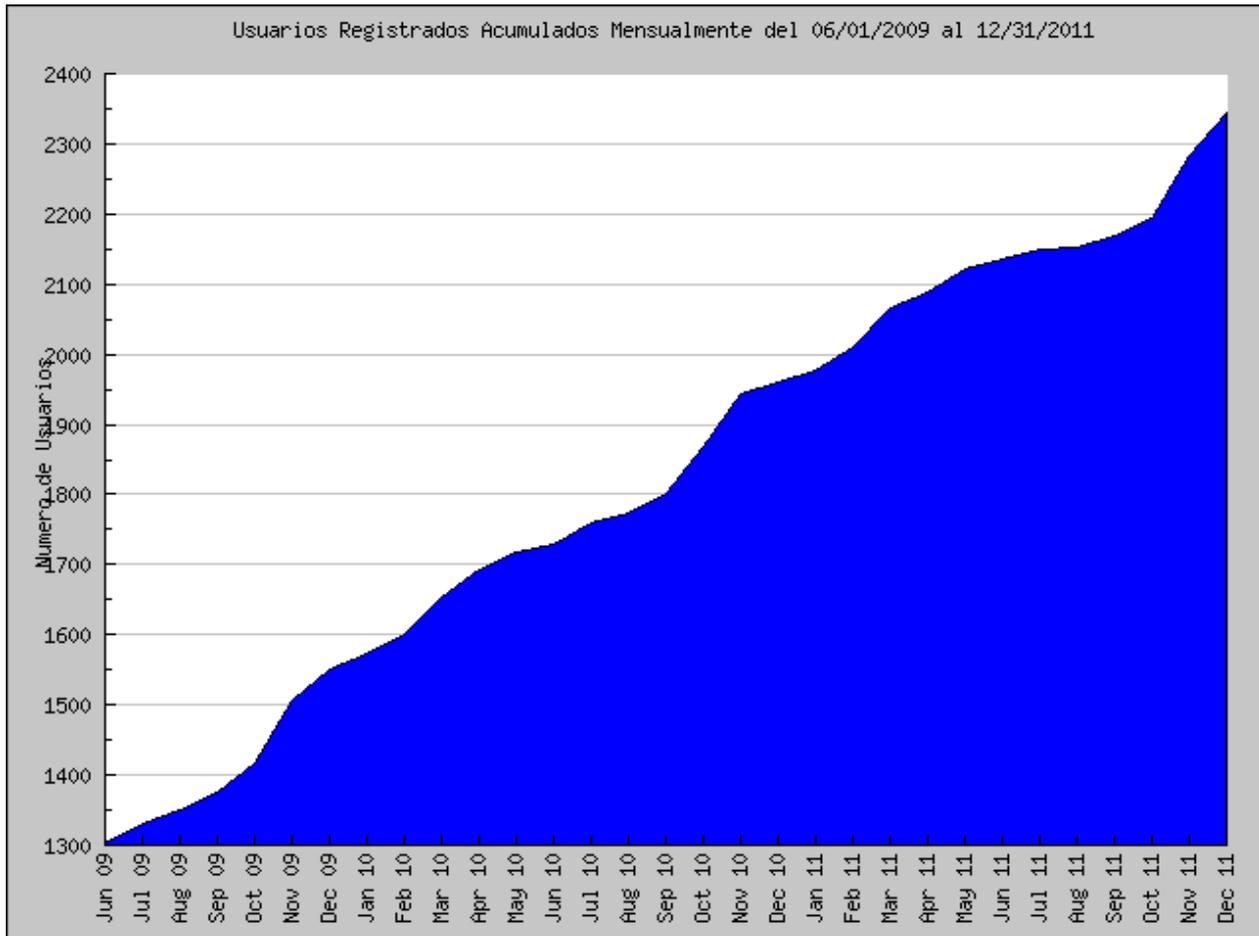


Figura 94: Total de nuevos usuarios registrados 2009-2011

Una herramienta de administración de un repositorio de software junto con herramientas de comunicación y organización del proyecto forman un modelo de desarrollo de software distribuido que puede aplicarse en proyectos en los que los desarrolladores están físicamente en diferentes lugares así como en una aplicación desarrollada por una empresa en la que sus trabajadores están en la misma oficina.

Un entorno de desarrollo que contiene todas estas herramientas y alberga diferentes proyectos es lo que se conoce como Forja de Software

Hace ya cuatro años surgió la idea de crear una forja en colaboración con las Universidades españolas y RedIRIS que albergara proyectos relacionados con la comunidad universitaria, investigadora y docente. Con este objetivo se ha montado la “Forja de Conocimiento Libre de la comunidad RedIRIS” ubicada físicamente en el CICA.

La Forja RedIRIS-CICA es un proyecto de ámbito nacional que se inició en 2007, avalado por RedIRIS y hospedado en el CICA. En la gráfica 94 puede verse la evolución de los usuarios registrados desde 2009 hasta el

presente. En ella pueden desarrollar proyectos alumnos universitarios, profesorado, personal de investigación o el PAS.

Como vemos, en cada edición del Concurso Universitario de Software Libre produce un incremento de altas de usuarios y proyectos en la forja de RedIRIS, alcanzando en los meses de Octubre y Noviembre de cada año sus cotas más altas. En relación a los proyectos alojados, extraemos los siguientes datos desde 2009 a finales de 2011 (ver gráfica 95).

Entre todos ellos podemos encontrar proyectos relacionados con la educación, con tareas de administración de sistemas o con Internet. Un breve paseo por su árbol de proyectos puede mostrarnos la calidad y la cantidad de proyectos que hay alojados en la forja de RedIRIS-CICA.

Como no podía ser de otra manera, la forja RedIRIS-CICA, está basada en software libre. El proyecto se encuentra alojado en un servidor Debian Etch (v.4.0) con seguridad reforzada con Grsecurity, incrementando por tanto la fiabilidad y la protección de todo el sistema.

Subversión es la herramienta elegida para manejar los repositorios que se ponen a disposición de cada proyecto creado. El acceso a estos repositorios y al resto de herramientas, como listas de correo, foros, gestor de tareas, publicación de ficheros, etc, se integran mediante el software Gforge 4.5.16, que ha sufrido varias correcciones y modificaciones para poder adaptarlo a las necesidades concretas de nuestro proyecto.

Este software proporciona un entorno web con interfaz PHP que permite el uso de cada una de las herramientas comentadas además de proporcionar diversas funcionalidades adicionales como el seguimiento de la actividad de los proyectos, estadísticas o noticias publicadas por los usuarios. Además, se realizan backups diarios y pruebas semanales de seguridad y actualizaciones para mantener la fiabilidad del servidor.

Podemos concluir diciendo que fomentar los desarrollos de software libre, así como servir de soporte a iniciativas de interés en el entorno académico-científico es el objetivo principal de este proyecto.

4. PORTAL DE DIFUSIÓN DE LA CIENCIA Ainnova

4.1. Introducción

Durante estos dos años el CICA ha desarrollado la nueva versión del portal de Andalucía Investiga o también llamado "Ainnova". Andalucía Investiga es el portal dedicado a la investigación y la promoción de la ciencia y la tecnología en Andalucía. Está gestionado por la Consejería de Economía, Innovación y Ciencia y mediante este portal se pretende potenciar la comunicación de la ciencia en el territorio andaluz, dinamizando estrategias que permitan un mejor conocimiento de la investigación realizada tanto en la iniciativa privada como en la pública para contribuir a una mayor proyección social de la ciencia en nuestra comunidad. Los principales objetivos de este portal son:

- Potenciar los contenidos científicos en los medios de comunicación y darlos a conocer de manera inteligible, ofreciendo a la ciudadanía instrumentos

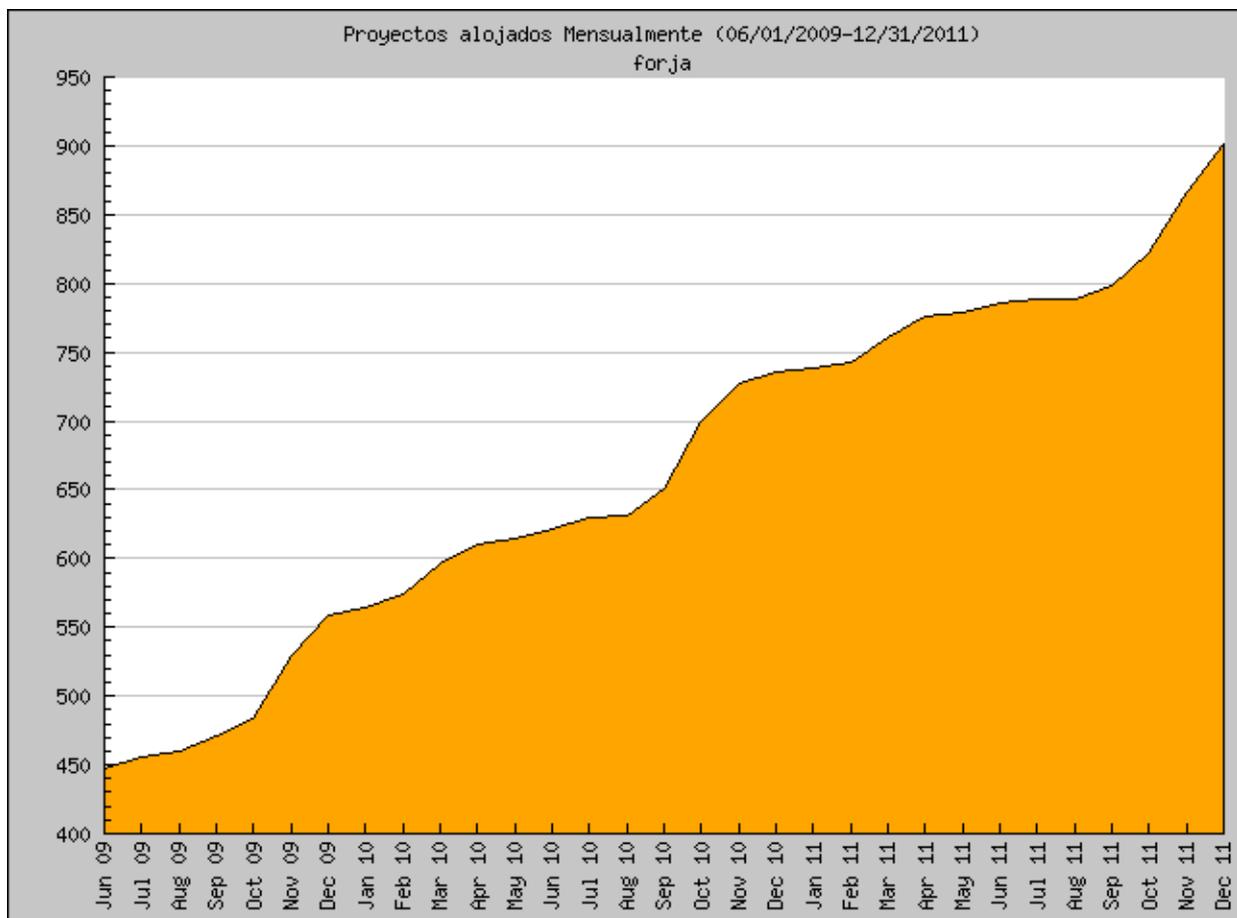


Figura 95: Proyectos registrados 2009-2011

para la comprensión de los mismos y su adaptación a la civilización tecnológica.

- Concienciar a los investigadores del sistema I+D de la necesidad de divulgar sus actividades y resultados para generalizar el conocimiento social acerca de los descubrimientos científicos.
- Favorecer el rol de divulgador científico o tecnológico entre los diferentes agentes del sistema I+D.
- Servir como puente entre los científicos, los periodistas y la sociedad, con el fin de promocionar la difusión nacional e internacional de los resultados de la investigación producida en Andalucía.
- Incrementar el interés por la ciencia y la tecnología y sensibilizar a la ciudadanía

4.2. Desarrollo inicial.

Para el desarrollo del nuevo portal se ha usado Joomla! 1.5. Joomla! es un gestor de contenidos web (CMS o Content Management System) que permite crear sitios web. Joomla! permitiría que el desarrollo del nuevo portal no empezara desde cero y que muchas de las funcionalidades requeridas fueran satisfechas a través de componentes, módulos o plugins desarrollados previamente (ver figura 96).

En una primera fase de diseño del portal, el personal de desarrollo de CICA tuvo varias reuniones con el equipo de periodistas encargados de la publicación de contenido en el portal para concretar los requisitos de la nueva web. Durante este tiempo, se tomó como punto de partida la anterior web de Andalucía Investiga con el fin de establecer las necesidades de funcionalidad del nuevo portal.

Respecto al aspecto de la web, el CICA se encargó de llevar a cabo la implementación de una plantilla previamente diseñada por una empresa externa especializada. Durante las siguientes reuniones esta plantilla fue evolucionando hasta obtener el resultado deseado, parecido al de otros periódicos digitales.

4.3. Funcionalidad

De manera general, la nueva web de Andalucía Investiga soporta las siguientes funcionalidades:

- **Noticias:** parte fundamental del portal donde se gestionan todas las noticias que se actualizan a diario por el equipo de periodistas. Se distribuyen según temática.
- **Boletines:** de manera semanal se emiten boletines a nivel externo a suscriptores y a nivel interno al equipo redactor de la web.
- **Agenda:** en este módulo del proyecto se gestiona la



Figura 96: Portal, en su desarrollo inicial

publicación y descripción de eventos futuros.

- **Notas de prensa (*InnovaPress*):** igualmente, se puede enviar las notas de prensa a los medios suscritos o programarlas para su envío.
- **Publicaciones:** se hace cargo de la parte que lista las diferentes publicaciones en la web.
- **Contactos:** gestión de todos los contactos relacionados con Andalucía Investiga así como la inserción de palabras clave.

También cabe destacar la variedad de perfiles que deben actuar en la web a la hora de generar nuevo contenido. El acceso a las herramientas de la web de estos perfiles difiere según su responsabilidad.

4.4. Pruebas y mantenimiento

Una vez que el desarrollo fue terminado y la funcionalidad de la web había sido validada por el equipo de Andalucía Investiga, se procedió al volcado de los datos de la web anterior a la nueva y a la elaboración de manuales que facilitarían el manejo de la aplicación a sus usuarios. Así mismo, se crearon alojamientos restringidos de prueba donde los editores de Andalucía Investiga pudie-

ron comenzar a familiarizarse con la aplicación.

5. EL CICA EN LAS REDES SOCIALES

El CICA se ha hecho un hueco en las redes sociales para poder llegar a los investigadores por una vía alternativa y de rápido crecimiento en los usos habituales de Internet, dándose de alta y creando perfiles en los más utilizados a día de hoy y con una voluntad de futuro de irse añadiendo a los nuevos usos y modelos de intercambio de información y presencia en la red. En concreto, de momento en 2011 CICA tiene presencia en facebook, en twitter y en linkedin.

Facebook, la red social más importante de Internet. A través de su perfil, el usuario de los servicios del CICA puede estar permanentemente informado de las últimas noticias del centro.

Twitter, red social que permite mandar mensajes de texto plano de bajo tamaño con un máximo de 140 caracteres, llamados tweets. A través de esta página el usuario de CICA puede conocer de forma rápida las últimas noticias del centro.

Por último, el CICA a través de su equipo de profesionales se encuentra también presente en LinkedIn.



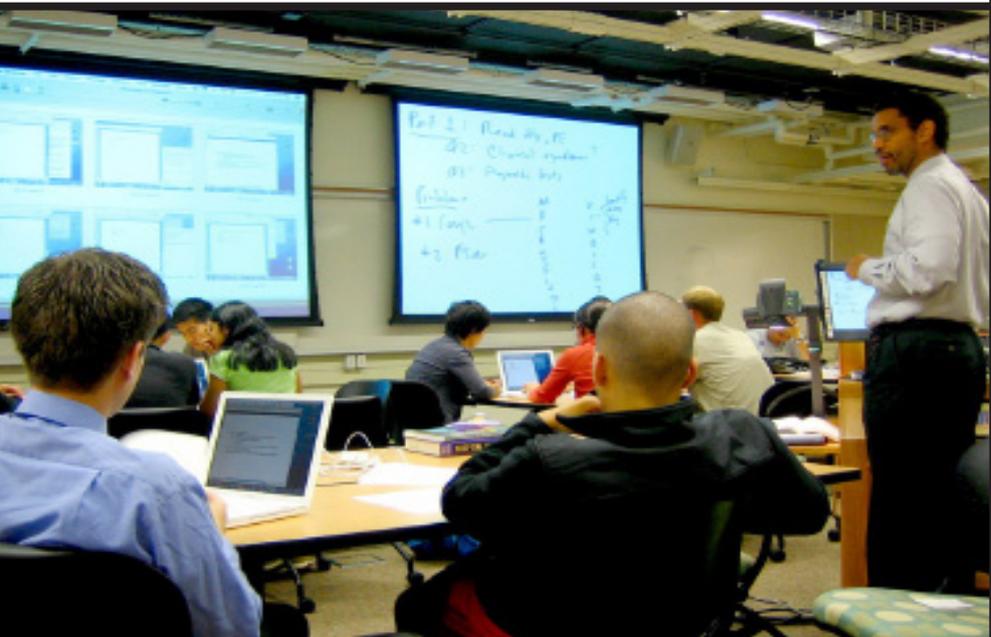
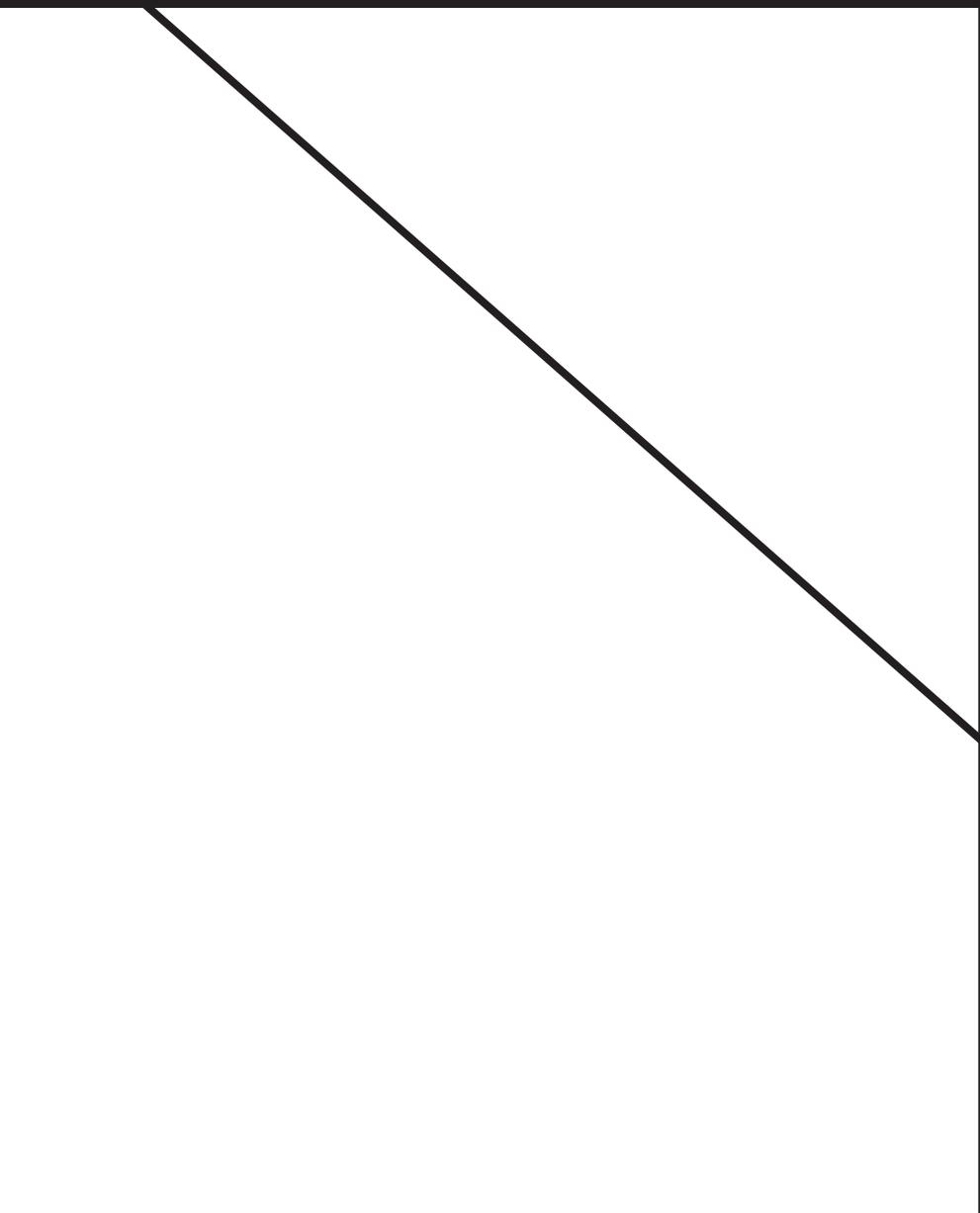
<http://www.facebook.com/CentroInformaticoCientificoAndalucia>



http://twitter.com/#!/cica_es



<http://www.linkedin.com/company/centro-inform-tico-cient-fico-de-andaluc-a-cica->



Formación y eventos

Durante 2010 y 2011 se ha continuado la labor de formación y participación en eventos, destacando la asistencia a las reuniones de coordinación del proyecto europeo EDGI donde CICA participa.

También destacan los múltiples cursos organizados por el Instituto Andaluz de Administración Pública (organismo dependiente de la Junta de Andalucía) en las aulas del CICA.

Asimismo continúa la afluencia de centros docentes que visitan las instalaciones del CICA como centro de referencia de servicios de comunicaciones y supercomputación de Andalucía.





Durante el periodo cubierto por este anuario, de la asistencia a eventos formativos se ha pasado a los colaborativos y a las reuniones de seguimiento de talleres de preparación para enfrentarse al reto de la participación en proyectos europeos en cualquiera de sus modalidades: colaborador principal, colaborador normal y como centro de recursos que atiende las necesidades de un proyecto ya en marcha.

Y como se decía en el preámbulo de este apartado, la elección del CICA como centro de recursos del proyecto europeo EDGI (European Desktop Grid Initiative) nos comprometía para la realización de una serie de reuniones de coordinación.

1. ASISTENCIA A EVENTOS Y TALLERES

1.1. Talleres para proyectos europeos.

Dentro del interés del Centro en la participación en proyectos europeos, durante el año 2010 personal del CICA asistió a varios talleres y conferencias relativas a la participación en diferentes convocatorias del VII Programa Marco.

Algunos de ellos han sido:

- Taller sobre Energía en Sevilla, el 17 de Junio de 2010.
- Taller sobre Tecnologías TIC en Málaga, el 16 de Septiembre de 2010.
- III Seminario de Preparación de Propuestas al VII Programa Marco en Antequera (Málaga), 17 y 18 de Noviembre de 2010.

Organizaciones como CITA-Andalucía, OTRIs de las universidades públicas de Andalucía o el CDTI ayudan con su labor a organizaciones tanto públi-

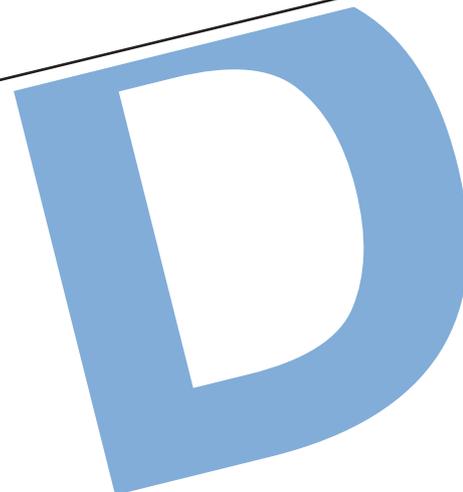
cas como privadas en su participación en programas europeos, asesorando en primer lugar con la búsqueda de socios, organización de eventos informativos y reuniones personales, y una vez dentro del proyecto, en el buen desarrollo de los trabajos y la comunicación con el consorcio. En este sentido, estas organizaciones, y especialmente CITA-Andalucía, han asesorado a CICA en todos estos temas.

1.2. Reuniones de coordinación de EDGI.

El CICA, como centro de recursos del proyecto europeo EDGI (European Desktop Grid Initiative) de 7º Proyecto Marco europeo (EDGI está dentro del "FP7 Capacities Programme" bajo el "grant agreement nr RI-261556") donde participa dentro del paraguas de la Universidad de Sevilla, se comprometió en su día a la asistencia a las reuniones de coordinación del proyecto, habiendo asistido de momento a las reuniones realizadas en Córcega en Abril de 2011 y en Sicilia en Octubre de 2011.

1.3. Otros eventos

En la siguiente tabla de la página siguiente se enumeran los otros eventos a los que ha asistido el personal del CICA.



Nombre del evento	Fecha	Lugar
Jornadas Técnicas RedIRIS 2010	Noviembre 2010	Córdoba
Presentación de la versión 10 del software ArcGIS de ESRI y del contrato ESRI-CICA	Junio 2010	Sevilla
Conferencia de supercomputación en las IV Jornadas Tecnológicas IES Padre Miravent	Marzo 2011	Huelva
4º Congreso HP-CAST Iberia	Octubre 2011	Málaga
Campus Party Milenio	Octubre 2011	Granada
IBM System X: Flexibilidad e integración para una muy alta densidad de computación	Octubre 2011	Sevilla

Durante 2010 y 2011 se han realizado varios cursos en las aulas destacando los organizados por la Consejería de Economía, Innovación y Ciencia, de la que depende el CICA y los del Instituto Andaluz de Administraciones Públicas y que se relacionan en las tablas 1 y 2 adjuntas en las páginas siguientes.

Abreviaturas:

- CICE: Consejería de Innovación, Ciencia y Empresa
- CICA: Centro Informático Científico de Andalucía
- CEIC: Consejería de Economía, Innovación y Ciencia
- CIBS: Consejería de Igualdad y Bienestar Social
- IAAP: Instituto Andaluz de Administración Pública

Por otro lado, el personal del CICA ha asistido a los siguientes cursos en dicho periodo:

Nombre del curso	Fecha	Lugar
Metodologías Ágiles y Técnicas de Productividad Personal	Abril 2011	Sevilla
Seminario Best Practices de Administración Oracle	Octubre 2011	Sevilla
Seminario de Alfresco	Octubre 2011	Sevilla
Seminario/Curso de Zentyal	Junio 2011	Sevilla
Fundamentos de funcionamiento del centro de Respaldo y Continuidad de la Junta de Andalucía	Junio 2011	Sevilla
Oracle Database 11g: Administration Workshop I	Noviembre 2011	Sevilla
Oracle Database 11g: SQL and PL/SQL New Features	Mayo 2011	Cádiz
Oracle Database: Program with PL/SQL	Mayo 2011	Cádiz

FECHA (2010)	TÍTULO DEL CURSO	ORGANIZADOR
ENERO	EL PROCEDIMIENTO ADMINISTRATIVO SANCIONADOR:TRAMITACIÓN Y SANCIONES	CICE
FEBRERO	TUTORIAL DE GRID: Curso EGEE gLite tutorial for users and Applications developers	RedIRIS y CICA
FEBRERO	CURSO MADEJA	CICE
FEBRERO	CURSO MADEJA	CICE
FEBRERO	CURSO FUNDAMENTOS EN GESTIÓN DE LA CONFIGURACIÓN MÓDULO CMDDB DE NAOS	CICE
MARZO	CURSO FORMACION A DESARROLLADORES DE NAOS:INTEGRACION CON TERCEROS (2ª Edición)	CICE
MARZO-ABRIL	FORMACIÓN SERVICIO DE CERTIFICACIÓN	CICA
ABRIL	CURSO SISTEMA INTEGRADO DE GESTIÓN DEL GASTO Y CONTABILIDAD PÚBLICA	CICE
ABRIL	INTRODUCCION A LA GESTIÓN DOCUMENTAL Y AL GESTOR DOCUMENTAL ALFRESCO	CICE
ABRIL	PROCEDIMIENTO DE LA ORDEN DE ECONOMÍA SOCIAL SOBRE PRESENTA-PCT	CICE
ABRIL	CURSO GESTIÓN DE COMUNICACIONES INTERIORES, PORTAFIRMA Y ECO	CICE
MAYO	Formador de Usuarios Finales de AlfrescoECM	CICE
JUNIO	Desarrollo Básico con ALFRESCO	CICE
JUNIO	ADMINISTRACIÓN BÁSICA CON ALFRESCO	CICE
JUNIO	CURSO LIBRAE	CICE
JUNIO	MADEJA: Curso Técnico de Introducción al Subsistema de Desarrollo	CICE
JULIO	PUES (Puesta en Funcionamiento)	CEIC
SEPTIEMBRE	PUES (Puesta en Funcionamiento)	CEIC
SEPTIEMBRE	ECO (COMUNICACIONES ELECTRONICA	CEIC
OCTUBRE	CURSO DE ALFRESCO: ADMINISTRACIÓN AVANZADA CON ALFRESCO ECM	CEIC
OCTUBRE	AUTORIZACIONES TIPO (TRAMITADOR AUTI	CICE
OCTUBRE	NORMATIVA ANDALUZA DE SUBVENCIONES: BASES REGULADORAS PROCEDIMIENTOS AMTVO. Y FORMULARIOS (I10F77PAE)	IAAP
OCTUBRE	OPEN OFFICE CALC	DELEGACION PROVINCIAL DE SEVILLA CIBS
OCTUBRE	NORMATIVA ANDALUZA DE SUBVENCIONES: BASES REGULADORAS PROCEDIMIENTOS AMTVO. Y FORMULARIOS (I10F74PAE)	IAAP
OCTUBRE	SISTEMA INTEGRADO DE GESTIÓN DEL GASTO Y CONTABILIDAD PÚBLICA JUPITER.	CEIC
OCTUBRE	PUESTA EN FUNCIONAMIENTO DE INSTALACIONES: TRAMITADOR PUES	CICE
NOVIEMBRE	DESARROLLO DE PORTALES CON ALFRESCO WebScripts y Surf	CEIC
NOVIEMBRE	NORMATIVA ANDALUZA DE SUBVENCIONES: BASES REGULADORAS PROCEDIMIENTOS AMTVO. Y FORMULARIOS (I10F75PAE)	IAAP

Tabla 1: Cursos realizados en el CICA en 2010

FECHA (2011)	TÍTULO DEL CURSO	ORGANIZADOR
ENERO	CURSO DE FORMACIÓN EN JAVA	JUAN ANTONIO ORTEGA
FEBRERO	CURSO DE FORMACIÓN EN JAVA	JUAN ANTONIO ORTEGA
FEBRERO	FORMACIÓN EN PLATAFORMA MERCADO DE IDEAS Y TECNOLOGIA	CEIC
FEBRERO	PORTAFIRMA Y Eco	CEIC
MARZO	CURSO DE FORMACIÓN EN JAVA	JUAN ANTONIO ORTEGA
MARZO	LibraE (presentación de la aplicación)	CEIC Y CULTURA
ABRIL	PORTAFIRMA Y eCO	CEIC
ABRIL	Sistemas de Gestión de Seguridad de las TI y Arquitecturas Seguridad	IAAP
ABRIL	CURSO METODOLOGIA AGILES Y PRODUCTIVIDAD	CICA
MAYO	Tecnología de Servicio Web	IAAP
MAYO	ORACLE	CICA
MAYO	MADEJA en el Proceso de Desarrollo de Software	IAAP
MAYO	CURSO REDES DE DISTRIBUCIÓN DE ENERGÍA ELÉCTRICA	IAAP
JUNIO	PORTAFIRMA Y eCO	CEIC
JUNIO	SEMINARIO NUEVO REPOSITORIO DE LIBRERÍA DE MADEJA	CEIC
JUNIO	CURSO GESTIÓN DE LA SEGURIDAD TIC	CEIC
JUNIO	BUENAS PRÁCTICAS EN LA GESTIÓN DE SERVICIOS TIC	CEIC
JUNIO	"Principios Básicos de Seguridad en la red: Buenas prácticas en el uso de Internet y protección ante los fraudes más comunes".	CEIC-IAAP
JUNIO	SEMINARIO ENS 1ª EDICIÓN	CEIC/IAAP
JUNIO	AGATA 1ª edic. Administradores	CEIC
JUNIO	AGATA 1ª edic. Sistemas	CEIC
JUNIO	AGATA 1ª edic. Usuarios	CEIC
JUNIO	AGATA 2ª edic. Administradores	CEIC
JUNIO	AGATA 2ª edic. Sistemas	CEIC
JUNIO	AGATA 2ª edic. Usuario	CEIC
JUNIO	SEMINARIO ENS 1ª EDICIÓN	CEIC/IAAP
JUNIO/JULIO	ORACLE DEVELOPING	CICA
JULIO	CONTRATACIÓN DE BIENES Y SERVICIOS TIC	CEIC
SEPTIEMBRE	BUSINESS INTELIGENCE	CEIC
SEPTIEMBRE	PORTAFIRMA Y eCO	CEIC
SEPTIEMBRE	SISTEMA INTEGRADO DE GESTIÓN DEL GASTO Y CONTABILIDAD PÚBLICA JUPITER	CEIC
SEPTIEMBRE	Seminario ENS 2ª Edición	CEIC
SEPTIEMBRE/ OCTUBRE	CURSO DE CREACIÓN DE MAPAS CON DATOS SECTORIALES	IAAP
OCTUBRE	TRAMITACIÓN TELEMÁTICA DE PROCEDIMIENTOS ADMITIVOS. EN MATERIA DE VEHÍCULOS DESTINADOS AL TRANSPORTE DE MERCANCÍAS ATP Y ADR	CEIC (Servicio de Personal)
OCTUBRE	SISTEMA DE GESTIÓN DE SEGURIDAD DE LAS TIC Y ARQUITECTURAS DE SEGURIDAD 2	IAAP
OCTUBRE	GESTIÓN DE OFICINAS DE CALIDAD	IAAP
NOVIEMBRE	GESTIÓN DE SEGURIDAD DE LAS TIC	CEIC
NOVIEMBRE	PORTAFIRMA Y eCO	CEIC
NOVIEMBRE	IMPLANTACIÓN DE TRAMITADORES TELEMÁTICO 3	CEIC (Servicio de Personal)
NOVIEMBRE	ORACLE	CICA
NOVIEMBRE	IMPLANTACIÓN DE TRAMITADORES TELEMÁTICO 2	CEIC (Servicio de Personal)
NOVIEMBRE	REDES, SISTEMAS Y SERVICIOS DE RED COOPERATIVA DE LA J. DE ANDALUCÍA	CEIC - IAAP
NOVIEMBRE/DICIEMBRE	GESTIÓN DE OFICINAS DE CALIDAD	CEIC

Tabla 2: Cursos realizados en el CICA en 2011

2. VISITAS AL CICA.

Al igual que años anteriores, durante 2010 y 2011 el Centro Informático Científico de Andalucía ha seguido organizando visitas a sus instalaciones. Esta actividad tiene como principal finalidad la de establecer un acercamiento del gran público y sobre todo de los alumnos de Institutos, a la organización y al trabajo de un centro de cálculo como es CICA, cuyo objetivo primordial es estar

al servicio de la sociedad, contribuyendo al desarrollo y avance tecnológico de la comunidad andaluza. Durante el pasado año, la presencia de los alumnos de secundaria fue muy presente, aspecto relevante ya que son el futuro de la ciencia andaluza y el CICA realiza una gran labor de fomento de las vocaciones científicas.

La lista de centros formativos que han acudido con sus alumnos a visitar el CICA se adjunta a continuación:

IES ANTONIO GALA de Palma del Río (05/03/10)

AYUNTAMIENTO DE CASTILLEJA DE LA CUESTA - CURSO FPE TECNICO SISTEMAS MICROINFORMATICOS de Castilleja de la Cuesta (25/06/10)

INSTITUTO INFORMÁTICO HISPANENSE de Sevilla (21/08/10)

CENTRO DE FORMACIÓN PROFESIONAL GORCA de Sevilla (19/10/10)

IES GONZALO NAZARENO de Dos Hermanas (25/10/10)

AULA ABIERTA PARA MAYORES DE 55 AÑOS DE CASTILLEJA DE LA CUESTA (UNIV. PABLO DE OLAVIDE) (01/12/10)

I.E.S. TRIANA de Sevilla (01/12/10)

IES LOS ALCORES de Mairena del Alcor (15/12/10)

I.E.S. ISIDRO ARCENEGUI Y CARMONA de Marchena (17/01/11)

IES CAVALERI de Mairena del Aljarafe (18/01/11)

IES. CIUDAD JARDIN de Sevilla (24/01/11)

IES ALARIFES RUIZ FLORINDO de Fuentes de Andalucía (03/02/11)

IES RAMÓN DEL VALLE INCLÁN de Sevilla (16/02/11)

INGENIUS de la Universidad de Sevilla (23/02/11)

IES TORRE DEL REY de Pilas (09/03/11)

IES RUIZ GIJÓN de Utrera (16/03/11)

IES LUIS VELEZ DE GUEVARA de Écija (24/03/11)

IES DELGADO HERNÁNDEZ de Bollullos del Condado (28/03/11)

I.E.S. JACARANDÁ de Brenes (06/04/11)

PRESCAL - CENTRO DE FORMACION PROFESIONAL PARA EL EMPLEO de Sevilla (19/04/11)

HUELVA ENSEÑANZA SOC. COOP. ANDALUZA de Ayamonte (03/05/11)

MASTER SIG (03/05/11)

ACADEMIA AFOBAN de Sevilla (15/06/11)

SODECSA (SOCIEDAD PARA EL DESARROLLO DE CAMAS) de Camas (30/06/11)

ACADEMIA AFOBÁN de Sevilla (18/09/11)

IES GONZALO NAZARENO de Dos Hermanas (12/10/11)

IES MAR DE CÁDIZ del Puerto de Santa María (14/10/11)

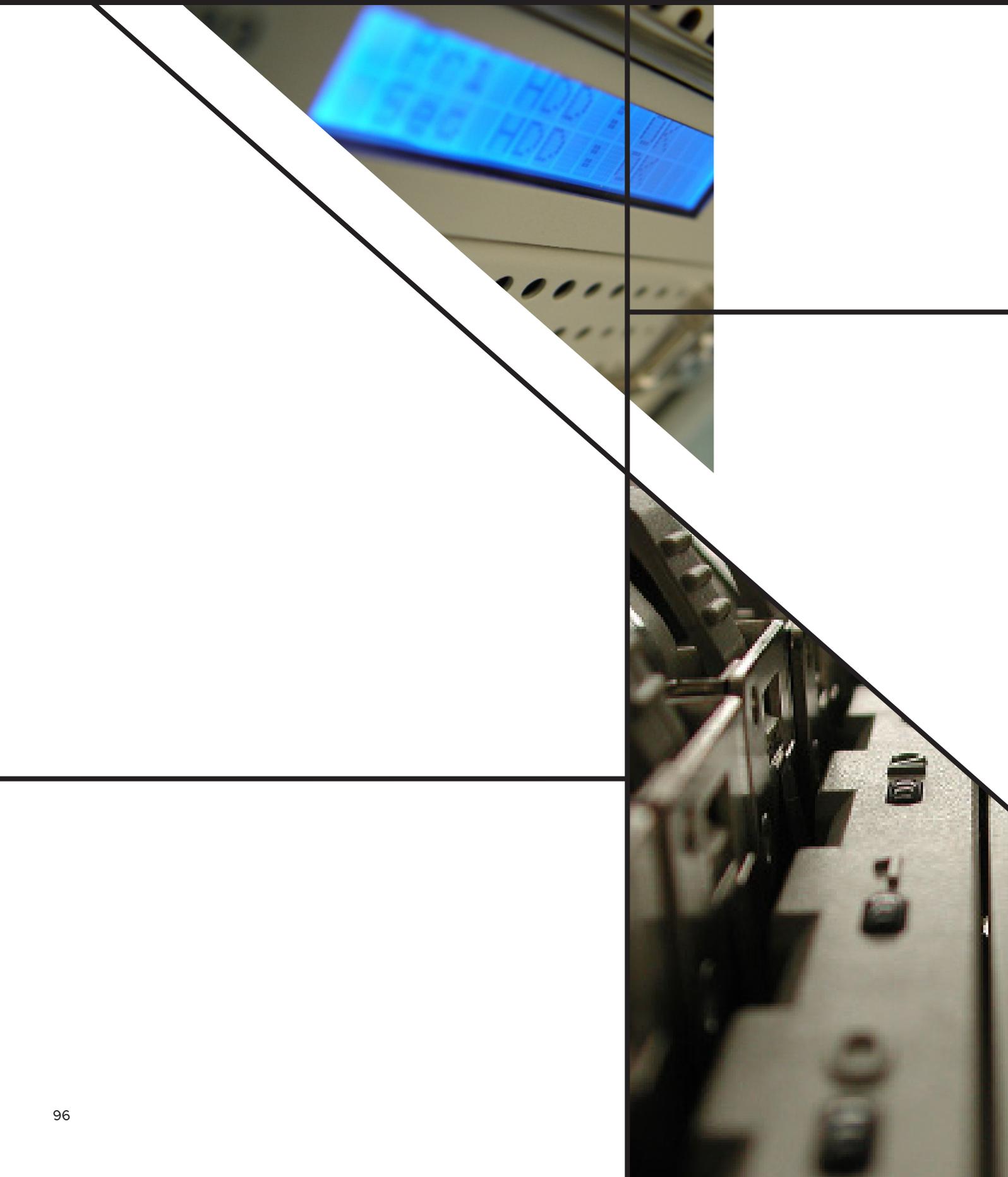
IES POLÍGONO SUR de Sevilla (14/10/11)

IES LOS ALCORES de Mairena del Alcor (21/10/11)

IES MARTÍNEZ MONTAÑÉS de Sevilla (14/12/11)

IES POLÍGONO SUR de Sevilla (14/12/11)

IES AXATI de Lora del Río (21/12/11)



Anexo: Pruebas

HPC



RESULTADOS E INTERPRETACIÓN DE LAS PRUEBAS DE CÁLCULO DE HTC Y HPC

1. DESCRIPCIÓN DE LAS PRUEBAS.

Para la realización de las pruebas se ha utilizado la suite HPC Challenge (HPCC) que consta de un conjunto de benchmarks que evalúan el rendimiento de diferentes capacidades de un servidor para cálculo intensivo desde el punto de vista de su adecuación para tareas de HPC fuertemente paralelas.

A continuación se describen las operaciones correspondientes a cada batería de pruebas:

- **PTRANS:** Realiza la traspuesta de una matriz aplicando paralelismo de grano fino. Debido a que este benchmark requiere de la comunicación entre pares de procesos, es interesante para medir la capacidad de la interconexión de la red de interconexión de servidores. Se mide en Gflops.
- **Rpeak Teórico:** Indica el máximo teórico de instrucciones en punto flotante que se pueden ejecutar. Se mide en GFlops. Su cálculo se realiza con la fórmula:
$$R_{peak} = \text{Frecuencia(Ghz)} \cdot \text{Instrucciones/ciclo} \cdot \text{cores.}$$
- **HPL: High Performance Linpack.** Mide las operaciones en punto flotante capaz de realizar el sistema probado resolviendo un sistema lineal de ecuaciones. Este Benchmark es usado por Top500.org para realizar la clasificación. Se mide en Gflops.
- **Eficiencia: HPL/Rpeak teórico.** Con este parámetro podemos medir la eficiencia la cual se calcula dividiendo los Gflops reales entre el máximo teórico. Es importante notar que el tiempo de inicialización tendrá una reper-

ción directa en este parámetro, por lo que cuanto mayor sea el problema a tratar (mayor memoria RAM disponible), menos significativo será el tiempo de inicialización.

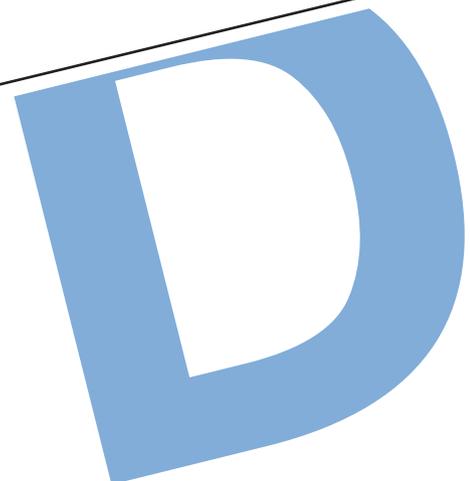
- **DGEMM:** Multiplicación de matrices de doble precisión. El algoritmo usado está diseñado para minimizar el número de accesos de memoria, logrando un alto grado de reutilización de los datos, siendo el valor obtenido muy utilizado como el valor máximo teórico de operaciones en aritmética de punto flotante por core. Tiene dos versiones, una multicore (StarDGEMM) y otra de un solo core (SingleDGEMM), pero debido a que la comunicación entre los procesos es mínima los resultados son muy similares. Se mide en Gflops.
- **STREAM:** Mide ancho de banda sostenido en el acceso a memoria mediante la realización de operaciones de punto flotante sobre vectores. Hay dos versiones de este test, SingleSTREAM, mide el ancho de banda máximo usado por un solo proceso haciendo uso exclusivo de memoria, y StarSTREAM que realiza la misma prueba pero trabajando en paralelo con todos los cores. Se trata de un benchmark sintético que realiza las siguientes cuatro operaciones sobre vectores:

Copy: $a(i) = b(i)$

Scale: $a(i) = q \cdot b(i)$

Add: $a(i) = b(i) + c(i)$

Triad: $a(i) = b(i) + q \cdot c(i)$



El valor System STREAM deriva de multiplicar el número de cores por el StarSTREAM para obtener el ancho de banda total del que es capaz el sistema.

MPIFFT y StarFFT: Ambas pruebas resuelven transformadas de Fourier con aritmética flotante en doble precisión. Se diferencian en el grado de paralelismo que emplean: MPIFFT emplea diferentes cores y MPI para distribuir la resolución de un solo problema, mientras que StarFFT resuelve diferentes problemas a la vez usando varios cores de forma paralela aunque no distribuida. El objetivo de este test es evaluar la eficiencia del procesador. Se mide en Gflops.

Ancho de banda/latencia MPI: Evalúa la latencia y el ancho de banda de la red de interconexión. La latencia se mide en microsegundos (el rendimiento es mejor a valores más bajos) y el ancho de banda en GB/s (el rendimiento es mejor a valores más altos).

2. REALIZACIÓN DE LAS PRUEBAS Y RESULTADOS.

Una vez detallada la batería de pruebas en general, a continuación se describen cómo se han realizado las mismas

sobre los diferentes grupos de servidores, junto con un breve análisis de los números más significativos obtenidos. Los diferentes grupos son los siguientes:

- 1. Servidores con dos procesadores:** análisis de las prestaciones de los servidores Nehalem12, Nehalem8 y Chuck (con Chuck hacemos referencia a cualquiera de los servidores Chuck1, Chuck2, Chuck3 o Chuck4, pues son iguales).
- 2. Interconexión:** análisis de las prestaciones al ejecutar el test haciendo uso de varias máquinas y haciendo uso de diferentes sistemas de interconexión (Ethernet e Infiniband), y varias máquinas Chuck.
- 3. Servidor con 4 procesadores:** realizaremos diferentes pruebas donde cambiaremos el número de cores del que se hace uso, con el fin de comprobar la existencia o no de cuellos de botella.
- 4. Hyperthreading:** comparativa habilitando y deshabilitando el hyperthreading en la máquina Nehalem12.
- 5. Virtualización:** análisis del comportamiento al introducir servidores virtualizados tanto con Xen como con KVM.

Servidor:		Nehalem12			
Nº Cores:	12				
	Intel Xeon 5670 2,93 Ghz				
PTRANS (GB/s)	2,14				
Rpeak Teórico (Gflops)	140,64				
HPL (Gflops)	115,29				
Eficiencia (%)	0,81975256				
StarDGEMM	10,66				
SingleDGEMM (Gflops)	11,02				
StarSTREAM (GB/s)	COPY	SCALE	ADD	TRIAD	
	1,86	1,86	2,1	2,1	
SingleSTREAM (GB/s)	COPY	SCALE	ADD	TRIAD	
	9,02	8,95	9,64	9,6	
System STREAM Triad (GB/s)	25,2				
MPIFFT (Gflops)	4,42				
StarFFT (Gflops)	1,09				
MPI LATENCY(ms)	min	avg	max		
	0,00030	0,00052	0,00080		
MPI BANDWIDTH (MB/s)	min	avg	max		
	4634,59	6477,37	8599,29		

Tabla A-01

Servidor:		Nehalem8			
Nº Cores:	8				
	Intel Xeon 5570 2,93 Ghz				
PTRANS (GB/s)	2,03				
Rpeak Teórico (Gflops)	93,76				
HPL (Gflops)	83,28				
Eficiencia (%)	0,888225256				
StarDGEMM (Gflops)	11,18				
SingleDGEMM (Gflops)	11,27				
StarSTREAM (GB/s)	COPY	SCALE	ADD	TRIAD	
	2,9	2,92	3,2	3,37	
SingleSTREAM (GB/s)	COPY	SCALE	ADD	TRIAD	
	8,25	8,15	8,63	8,42	
System STREAM Triad (GB/s)	26,96				
MPIFFT (Gflops)	4,71				
StarFFT (Gflops)	1,35				
MPI LATENCY(ms)	min	avg	max		
	0,00027	0,00053	0,00075		
MPI BANDWIDTH (MB/s)	min	avg	max		
	5064,05	7006,99	9409,54		

Tabla A-02

Servidor:		Chuck			
Nº Cores:	8				
	Intel Xeon 5550 2,67 Ghz				
PTRANS (GB/s)	2,82				
Rpeak Teórico (Gflops)	85,44				
HPL (Gflops)	81,91				
Eficiencia (%)	0,958684457				
StarDGEMM (Gflops)	10,86				
SingleDGEMM (Gflops)	11,21				
StarSTREAM (GB/s)	COPY	SCALE	ADD	TRIAD	
	3,33	3,35	3,72	3,71	
SingleSTREAM (GB/s)	COPY	SCALE	ADD	TRIAD	
	10,61	10,56	11,17	10,98	
System STREAM Triad (GB/s)	29,68				
MPIFFT (Gflops)	5,75				
StarFFT (Gflops)	1,44				
MPI LATENCY(ms)	min	avg	max		
	0,00027	0.00053	0.00074		
MPI BANDWIDTH (MB/s)	min	avg	max		
	5187,76	6956,64	9388,48		

Tabla A-03

2.1. Batería de pruebas 1: Servidores con 2 procesadores.

Los resultados están en las tablas A-01, A-02 y A-03. Respecto a esta última tabla, del servidor denominado Chuck, se ha de tener en cuenta que los resultados mostrados en la misma pueden verse incrementados en cierta medida por hacer uso de un software más moderno (como puede verse en la descripción del software usado), debido a que estas máquinas fueron adquiridas con posterioridad y por tanto los tests realizados tienen varios meses de diferencia. Debido a ello, solo se hará mención a éstos cuando las prestaciones se hayan visto afectadas por otras causas.

Se puede observar que el valor para DGEMM es similar en los 3 servidores mostrados, aunque algo superiores en el servidor Nehalem8, lo que nos indica que tratamos con procesadores con capacidades de cálculo similares.

En este aspecto de capacidad de cálculo, lo más interesante es el parámetro HPL y la eficiencia obtenida. Entre las dos primeras máquinas (Nehalem12 y Nehalem8) observamos que la eficiencia de la primera de ellas es notablemente inferior a la eficiencia obtenida para la segunda, lo cual es perfectamente achacable a las prestaciones de la memoria, pues al ser menor la RAM en Nehalem12 el problema a tratar ha de ser más pequeño, y junto a que posee más cores, hace que el tiempo de ini-

cialización será más significativo que en la máquina Nehalem8. También debemos de tener en cuenta que la memoria en Nehalem12 funciona a 1067Mhz, mientras que en Nehalem8 es a 1333Mhz.

Si ahora se compara la máquina Nehalem8 con las denominadas Chuck, se observa un importante incremento de la eficiencia, lo que indica que la potencia de cálculo entre ambos servidores es muy similar. Este incremento de la eficiencia viene motivado, además de que se usa una versión algo más actualizada del software como ya se ha comentado, por la mayor cantidad de RAM en las máquinas Chuck respecto a Nehalem8, pasando de 12 a 24GB o de 1.5GB/core a 3GB/core, lo que de nuevo supone tiempos de inicialización menos significativos.

Otros valores importantes a considerar son los correspondientes al test STREAM. Se observa que las prestaciones entre las máquinas Nehalem12 y Nehalem8 son parecidas, aunque se debe tener en cuenta que la frecuencia de la RAM de esta última es un 25% mayor, por lo que nos hace suponer que la primera máquina hace un acceso a memoria más eficiente que la segunda, necesario para poder soportar los 4 cores extras que tiene Nehalem12 frente a Nehalem8, y que no suponga el acceso a memoria un cuello de botella (como sí ocurría en los procesadores INTEL anteriores a las series Nehalem, Nehalem EX y Westmere).

En los resultados del test MPIFFT lo más interesante es que se observa una pequeña pérdida de prestaciones de Nehalem12 respecto a Nehalem8, siendo esto debido a la naturaleza del test, que paraleliza en un número de procesos en potencia de 2. Al disponer el primero de los procesadores de 12 cores, usa tan solo 8 de éstos, que teniendo en cuenta los valores observados para DGEMM el usar el mismo número de cores supone una pequeña pérdida de prestaciones.

Por último, si se observa la latencia y ancho de banda en la comunicación, MPI LATENCY y MPI BANDWIDTH, los resultados en la máquina Nehalem12 son sensiblemente peores a los de las otras máquinas, que son muy similares, debido a que al poseer más cores por procesador se incrementa la comunicación que se debe de realizar entre los dos procesadores. Como era de esperar, los parámetros que se ven más afectados por el tipo de interconexión entre los servidores son aquellos que miden la eficiencia del paralelismo. Un test que provoca una fuerte paralelización conlleva un elevado inter-

cambio de mensajes entre ambos servidores por la red y por tanto el tipo de interconexión por red influye mucho en los valores de dichos parámetros. Éstos son PTRANS, HPL, MPIFFT y los parámetros que miden MPI.

Observando los test MPI LATENCY y MPI BANDWIDTH vemos que la mejoría en prestaciones es prácticamente debida en exclusiva a la menor latencia que proporciona la red de Infiniband, que disminuye a una quinta parte, ya que el ancho de banda proporcionado, en media, es prácticamente el mismo.

2.2. Batería de pruebas 2: Análisis de la interconexión.

Las pruebas se realizan para servidores conectados en un caso por una red de baja latencia (en nuestro caso Infiniband) y en la otra prueba por una conexión de red normal (Ethernet). Los resultados se muestran en las tablas A-04 (Infiniband), y A-05 (Ethernet).

El efecto de la interconexión es claramente observable en esta batería de test. Por ejemplo, se puede apreciar una diferencia de prestaciones del 45% en el test PTRANS,

8% en HPL o 51% en MPIFFT.

Esa pérdida de prestaciones en el test HPL indica que la eficiencia baja considerablemente cuando usamos la interconexión Ethernet, observándose una eficiencia muy buena en el caso de la interconexión Infiniband, pues la eficiencia baja tan solo un 2%, donde no solo hemos de tener en cuenta la pérdida de prestaciones por tener que hacer uso de la red de interconexión, sino que hemos de tener también en cuenta la pérdida de prestaciones que supone el overhead debido al paralelismo.

2.3. Batería de pruebas 3: Servidor de 4 CPUs con uso simultáneo de número distinto de cores.

En este caso se trabaja con un único servidor de 4 CPUs de ocho cores cada una. En las tablas A-06 (uso simultáneo de 32 cores), A-07 (uso simultáneo de 16 cores) y A-08 (uso simultáneo de tan solo 8 cores de los 32 disponibles) se presentan los resultados obtenidos.

En una primera observación, vemos una eficiencia superior a 1, achacable al uso de “modo turbo” (incremento

Servidor:		Chuck1, Chuck2			
Nº Cores:	16				
	Intel Xeon 5550 2,67 Ghz				
PTRANS (GB/s)	1,47				
Rpeak Teórico (Gflops)	170,88				
HPL (Gflops)	160,97				
Eficiencia (%)	0,942006086				
StarDGEMM (Gflops)	10,96				
SingleDGEMM (Gflops)	11,13				
StarSTREAM (GB/s)	COPY	SCALE	ADD	TRIAD	
	3,35	3,36	3,72	3,72	
SingleSTREAM (GB/s)	COPY	SCALE	ADD	TRIAD	
	10,63	10,58	11,05	10,87	
System STREAM Triad (GB/s)	59,52				
MPIFFT (Gflops)	4,36				
StarFFT (Gflops)	1,5				
MPI LATENCY(ms)	min	avg	max		
	0,00041	0,00258	0,00437		
MPI BANDWIDTH (MB/s)	min	avg	max		
	772,36	3316,20	9320,67		

Tabla A-04: Conexión Infiniband

Servidor:		Chuck1, Chuck2			
Nº Cores:	16				
	Intel Xeon 5550 2,67 Ghz				
PTRANS (GB/s)	0,8				
Rpeak Teórico (Gflops)	170,88				
HPL (Gflops)	147,82				
Eficiencia (%)	0,865051498				
StarDGEMM (Gflops)	10,84				
SingleDGEMM (Gflops)	11,27				
StarSTREAM (GB/s)	COPY	SCALE	ADD	TRIAD	
	3,34	3,35	3,72	3,72	
SingleSTREAM (GB/s)	COPY	SCALE	ADD	TRIAD	
	10,43	10,4	10,94	10,74	
System STREAM Triad (GB/s)	59,52				
MPIFFT (Gflops)	2,14				
StarFFT (Gflops)	1,43				
MPI LATENCY(ms)	min	avg	max		
	0,00053	0,01128	0,02158		
MPI BANDWIDTH (MB/s)	min	avg	max		
	214,12	3351,56	9258,94		

Tabla A-05: Conexión Ethernet

de la frecuencia) u otros sistemas de mejora de prestaciones que pueden ser aprovechados por los procesadores Nehalem cuando el sistema no se encuentra sobrecargado.

Analizando el resto de los test, observamos que entre hacer uso de 8 y 16 cores las prestaciones prácticamente se duplican, a excepción del System STREAM Triad, que incrementa tan solo 1,3 Gflops, lo cual propicia que el StarSTREAM se reduzca a casi la mitad cuando hacemos uso de 16 cores. Esto nos indica que con 8 cores ya se ha alcanzado el máximo ancho de banda de acceso a memoria, también observable si nos fijamos en que el StarSTREAM es menor al SingleSTREAM.

Esto puede suponer un cuello de botella al incrementar el número de cores usados, pues al volver a duplicarlo el System STREAM Triad permanecerá invariable y el ancho de banda disponible por core será menor. Sin embargo, esto no parece ocurrir en exceso, pues las prestaciones, sin llegar a duplicarse al pasar a usar de 16 a 32 cores, experimentan unas mejoras cercanas al doble, teniendo en cuenta también que al estar siendo usados todos los cores no se posibilita al procesador el entrar en el “modo turbo”.

2.4. Batería de pruebas 4: Análisis del uso de Hyperthreading.

Como se puede observar en las dos tablas correspondientes (tabla A-09 sin usar hyperthreading y tabla A-10 usando esta característica), casi todos los test dan peores resultados a excepción de STREAM, en el que vemos una mejora del 150%. Observando los valores obtenidos para DGEMM se pone de manifiesto claramente que esta tecnología no aumenta la capacidad de cálculo, pues al habilitarla en el sistema hay el doble de cores pero de la mitad de capacidad de cálculo.

Pese a esta importante mejora en el test STREAM, esto no es suficiente para justificar el uso de la tecnología hyperthreading en HPC, pues los problemas a tratar, usualmente, tienen más similitudes con los test HPC o MPIFF, en los cuales se observa que se sufren importantes pérdidas de rendimiento. No obstante, es una

Servidor:		Nehalem32			
Nº Cores:	32				
	Intel Xeon 7550 2,00GHz				
PTRANS (GB/s)	3,84				
Rpeak Teórico (Gflops)	256				
HPL (Gflops)	221,37				
Eficiencia (%)	0,864726563				
StarDGEMM (Gflops)	7,49				
SingleDGEMM (Gflops)	7,63				
StarSTREAM (GB/s)	COPY	SCALE	ADD	TRIAD	
	0,92	0,92	1,04	1,05	
SingleSTREAM (GB/s)	COPY	SCALE	ADD	TRIAD	
	5,18	5,1	5,08	4,98	
System STREAM Triad (GB/s)	33,6				
MPIFFT (Gflops)	8,97				
StarFFT (Gflops)	0,52				
MPI LATENCY(ms)	min	avg	max		
	0,00053	0,00108	0,00137		
MPI BANDWIDTH (MB/s)	min	avg	max		
	3871,96	4330,57	5847,75		

Tabla A-06: Uso de los 32 cores

Servidor:		Nehalem32			
Nº Cores:	16				
	Intel Xeon 7550 2,00GHz				
PTRANS (GB/s)	3,24				
Rpeak Teórico (Gflops)	128				
HPL (Gflops)	132,41				
Eficiencia (%)	1,034453125				
StarDGEMM (Gflops)	8,83				
SingleDGEMM (Gflops)	8,9				
StarSTREAM (GB/s)	COPY	SCALE	ADD	TRIAD	
	1,84	1,84	2,04	2,05	
SingleSTREAM (GB/s)	COPY	SCALE	ADD	TRIAD	
	5,22	5,15	5,12	5,03	
System STREAM Triad (GB/s)	32,8				
MPIFFT (Gflops)	6,82				
StarFFT (Gflops)	0,83				
MPI LATENCY(ms)	min	avg	max		
	0,00045	0,00108	0,00141		
MPI BANDWIDTH (MB/s)	min	avg	max		
	3991,72	4576,12	6745,96		

Tabla A-07: Uso de 16 de los 32 cores

Servidor: Nehalem32

Nº Cores:	8			
	Intel Xeon 7550 2,00GHz			
PTRANS (GB/s)	2,68			
Rpeak Teórico (Gflops)	64			
HPL (Gflops)	65,76			
Eficiencia (%)	1,0275			
StarDGEMM (Gflops)	8,82			
SingleDGEMM (Gflops)	8,86			
StarSTREAM (GB/s)	COPY	SCALE	ADD	TRIAD
	3,56	3,54	3,83	3,82
SingleSTREAM (GB/s)	COPY	SCALE	ADD	TRIAD
	5,25	5,18	5,14	5,05
System STREAM Triad (GB/s)	30,56			
MPIFFT (Gflops)	4,36			
StarFFT (Gflops)	1,04			
MPI LATENCY(ms)	min	avg	max	
	0,00045	0,00108	0,00133	
MPI BANDWIDTH (MB/s)	min	avg	max	
	3763,39	4435,02	6789,64	

Tabla A-08: Uso de solo 8 de los 32 cores

Servidor: Nehalem12 No Hyperthreading

Nº Cores:	12			
	Intel Xeon 5670 2,93GHz			
PTRANS (GB/s)	2,14			
Rpeak teórico (Gflops)	140,64			
HPL (Gflops)	115,29			
Eficiencia (%)	0,81975256			
StarDGEMM (Gflops)	10,66			
SingleDGEMM (Gflops)	11,02			
StarSTREAM (GB/s)	COPY	SCALE	ADD	TRIAD
	1,86	1,86	2,1	2,1
SingleSTREAM (GB/s)	COPY	SCALE	ADD	TRIAD
	9,02	8,95	9,64	9,6
System STREAM Triad (GB/s)	25,2			
MPIFFT (Gflops)	4,42			
StarFFT (Gflops)	1,09			
MPI LATENCY(ms)	min	avg	max	
	0,00030	0,00052	0,00080	
MPI BANDWIDTH (MB/s)	min	avg	max	
	4634,59	6477,37	8599,29	

Tabla A-09

tecnología muy útil para otros fines como el Cloud Computing, donde no todos los cores están realizando la misma tarea simultáneamente.

2.5. Batería de pruebas 5: Virtualización.

La finalidad de esta batería de pruebas es medir la pérdida de prestaciones que se produce al virtualizar máquinas con los dos sistemas que usamos en el centro, por lo que se muestran tan solo aquellos test que tienen interés.

Los resultados se muestran en las tablas A-11 y A-12.

Podemos observar que por lo general la virtualización con XEN presenta unas pérdidas de prestaciones entorno al 10% en casi todos los test, una pérdida que asumíamos que pasaría pues el virtualizar supone cierto overhead.

Al virtualizar con KVM observamos que las pérdidas de prestaciones son mucho mayores, siendo en algunos casos superiores al 50%, como en el ancho de banda, aunque sorprendentemente presenta mejores resultados para el test PTRANS en sistemas virtualizados que en la máquina física.

Como observamos, a priori, en entornos de HPC preferiremos virtualizar con XEN. Pero la virtualización con XEN es más compleja tanto en instalación como en configuración, pues hay que aplicar parches al kernel, mientras que en KVM es todo más simple, además de que la mayoría de S.O. están apostando por la virtualización con KVM e incluyendo mejoras al soporte con éste y menos compatibilidad con XEN.

Servidor: Nehalem12 Hyperthreading

Nº Cores:	24			
	Intel Xeon 5670 2,93GHz			
PTRANS (GB/s)	1,71			
Rpeak teórico (Gflops)	140,64			
HPL (Gflops)	103,73			
Eficiencia (%)	0,737556883			
StarDGEMM (Gflops)	5,96			
SingleDGEMM (Gflops)	5,33			
StarSTREAM (GB/s)	COPY	SCALE	ADD	TRIAD
	2,42	2,46	2,58	2,66
SingleSTREAM (GB/s)	COPY	SCALE	ADD	TRIAD
	8,2	8,08	8,09	7,92
System STREAM Triad (GB/s)	63,84			
MPIFFT (Gflops)	4,42			
StarFFT (Gflops)	1,09			
MPI LATENCY(ms)	min	avg	max	
	0,00036	0,00072	0,00097	
MPI BANDWIDTH (MB/s)	min	avg	max	
	4291,94	5676,02	7678,36	

Tabla A-10

Servidor	Intel 6400				Intel 6400 XEN				Prestaciones
Nº Cores	2				2				
PTRANS (GB/s)	0,65				0,54				0,830769231
HPL (Gflops)	14,26				13,01				0,912342216
StarSTREAM (GB/s)	COPY	SCALE	ADD	TRIAD	COPY	SCALE	ADD	TRIAD	TRIAD
	1,56	1,57	1,9	1,93	1,55	1,59	2,17	2,03	1,051813472
MPIFFT (Gflops)	0,87				0,78				0,896551724
StarFFT (Gflops)	0,67				0,6				0,895522388
MPI LATENCY(ms)	0,00043				0,00053				1,23255814
MPI BANDWIDTH (MB/s)	1471,17				1464,73				0,995622532

Tabla A-11: Virtualización con XEN

Servidor	Matrix				KVM Matrix				Prestaciones
Nº Cores	16				16				
PTRANS (GB/s)	0,2				0,35				1,75
HPL (Gflops)	32,69				24,36				0,745182013
StarSTREAM (GB/s)	COPY	SCALE	ADD	TRIAD	COPY	SCALE	ADD	TRIAD	TRIAD
	0,18	0,18	0,2	0,2	0,17	0,18	0,21	0,22	1,1
MPIFFT (Gflops)	0,55				0,53				0,963636364
StarFFT (Gflops)	0,09				0,08				0,888888889
MPI LATENCY(ms)	0,00475				0,04079				8,587368421
MPI BANDWIDTH (MB/s)	1285,91				542,4				0,421802459

Tabla A-12: Vistualiación con KVM

COMUNICACIONES

SEGURIDAD

**SISTEMAS / HPC
ANEXO**

APLICACIONES

FORMACIÓN Y EVENTOS



JUNTA DE ANDALUCÍA