

díxitos

NOVAS DO CENTRO DE SUPERCOMPUTACIÓN DE GALICIA - MAIO 2002

CESGA

ANUARIO DE ACTIVIDADE CESGA 2001

Os Equipos HPCN do CESGA
foron co-financiados polo FEDER
Fondo Europeo de Desenvolvemento Rexional



Mensaxe do Presidente

Miguel Angel Ríos



O CESGA ven incrementando e mellorando a súa actividade de forma continua e constante, como consecuencia do aumento do nivel de esixencia das comunidades investigadoras ás que presta os seus servizos e da súa vontade de contribuír máis eficazmente ó desenvolvemento de Galicia.

Complementariamente á importante labor realizada no contorno de cálculo intensivo e comunicacións, ó longo do ano 2001 o CESGA participou activamente na organización e posta en marcha de novos servizos e actividades entre os que cabe destacar: O Consorcio de Bibliotecas Universitarias de Galicia (BUGalicia), o Centro de Competencias en Comercio electrónico de Galicia e a Rede de Bioinformática de Galicia.

Tamén cabe salienta a actividade desenvolvida polo

Transcurrió el año 2001 manteniéndose la tónica constante de incremento de disponibilidad de nuevos recursos técnicos del CESGA y el uso de los mismos, tal y como se refleja pormenorizadamente en la presente memoria.

A modo de resumen debe mencionarse la evolución experimentada en la actividad del cálculo y las comunicaciones. Durante el pasado año los usuarios del CESGA incrementaron el número de horas de cálculo en un 57%, aumentaron el intercambio de información mediante la "Rede de

CESGA na área de promoción de proxectos relacionados coa supercomputación, as comunicacións e a implantación innovadora das novas tecnoloxías, área na que se acadou a adjudicación de proxectos tan significativos como o CrossGrid ou o e-Minder correspondentes ó V Programa Marco da UE, entre outros moitos.

Cabe esperar que todo este conxunto de actividades sexan unha importante contribución á mellora da investigación e o desenvolvemento da nosa sociedade. É vontade do Consello de Administración, que me cabe a honra de presidir, continuar facendo os esforzos precisos para que o CESGA siga mellorando día a día o nivel das súas prestacións co obxectivo de favorecer o avance científico e de impulsar un maior progreso de Galicia.

Ciencia e Tecnoloxía de Galicia" en un 60% y almacenaron más de 3 Terabytes de datos en el subsistema de almacenamiento del CESGA, lo que supuso un 600% de incremento respecto al año anterior.

Con objeto de seguir dando respuesta al incremento de necesidades que las comunidades investigadoras usuarias del CESGA precisan, durante el año 2001 se realizaron las acciones técnico-económicas necesarias para la puesta en marcha de nuevo equipamiento de cálculo y comunicaciones a lo largo del año 2002.

díxitos

S.A. Xestión Centro de Supercomputación de Galicia

Sociedade participada pola Xunta de Galicia e o Consejo Superior de Investigaciones Científicas.

DIRECCIÓN Javier García Tobío
COORDINACIÓN Fernando Bouzas Sierra
REDACCIÓN Dr. Ignacio López Cabido

Dr. Andrés Gómez Tato
José Carlos Pérez Gómez
Dr. Carlos Fernández Sánchez
Francisco Landeira Vega
Juan Villasuso Barreiro
Pablo Rey Mayo
María José Rodríguez Malmierca

ILUSTRACIONES E PORTADA Juan Sanmiguel
DISEÑO E MAQUETACIÓN Oscar Otero
FOTOMECÁNICA Macrom

Depósito legal: C-1604-1998
ISSN: 1139-563X

EDITA
CESGA

Avenida de Vigo, s/n (Campus Sur)
15705 SANTIAGO DE COMPOSTELA
A Coruña (España)
Teléfono 981 569810 Fax 981 594616
Correo electrónico: dixitos@cesga.es
Enderezo Web: www.cesga.es/dixitos

INDICE

Consello de administración	03
Temas salientables 2001	03
Usuarios de computación de altas prestacións	04
Computación de altas prestacións	05
VPP 300E	06
AP 3000	07
HPC 4500	08
SVG	09
Aplicacións de cálculo	10
Rede de ciencia e tecnoloxía de Galicia	12
Sistema de almacenamento de datos	14
Proxectos	15
Teleformación	16
Accións de Formación	16
Laboratorio de visualización científica	17
GIS	18
Difusión	19

Mensaxe do Vicepresidente

Emilio Lora Tamayo



Xunta de Galicia	Presidente Secretaria	Miguel Angel Ríos Fernández Carmen Bouso Montero	Secretario Xeral de Investigación e Desenvolvemento. Xefa do Gabinete Lexislativo, Asesoría Xurídica Xeral, Consellería da Presidencia e Administración Pública. Director Xeral de Universidades, Consellería de Educación e Ordenación Universitaria. Director Xeral do CIXTEC, Consellería de Economía e Facenda. Director Xeral de Organización e Sistemas Informáticos, Consellería de Presidencia e Administración Pública. Director Xeral do Centro de Información e Tecnoloxía Ambiental, Consellería de Medio Ambiente. Subdirector Xeral de Investigación, Secretaria Xeral de Investigación e Desenvolvemento.
	Vocal	José Eduardo López Pereira	
	Vocal	Gonzalo Gómez Montaña	
	Vocal	Javier Franco Tubío	
	Vocal	Juan José Casares Long	
	Vocal	Antonio Blanco Seoane	
Consejo Superior de Investigaciones Científicas (CSIC)	Vicepresidente	Emilio Lora-Tamayo d'Ocon	Vicepresidente de Investigación Científica y Técnica. Coordinadora Institucional do CSIC en Galicia Subdirector General de Actuación Económica. Colaborador Científico, Instituto de Investigaciones Marinas.
	Vocal	M ^a Cruz Mato Iglesias	
	Vocal	José Carlos Rubio García	
	Vocal	Julio Rodríguez Banga	
Universidade de Santiago	Vocal	José Ramón Leis Fidalgo	Vicerrector de Investigación
Universidade da Coruña	Vocal	Esther Fernández Fernández	Vicerrectora de Investigación
Universidade de Vigo	Vocal	Pedro Merino Gómez	Vicerrector de Desenvolvemento e Transferencia

TEMAS SALIENTABLES 2001

Os eventos máis relevantes que aconteceron ó longo do 2001, sintetízanse nos seguintes:

Cálculo intensivo:

- Incorporouse o SVG (Superordenador Virtual Galego) ó proceso productivo do CESGA.
- Incrementouse o número de horas de cálculo nun 57% respecto ó ano 2000.

Almacenamento de datos:

- O subsistema de almacenamento de datos soportado por 1 TB en disco e 8 TB en librería de cintas, incrementou os datos de usuario nun 60% que xunto cos datos de backup sumaron máis de 5 TB de ocupación.

Rede de Ciencia e Tecnoloxía de Galicia (RECETGA):

- Substituíronse os equipos que soportan os servizos da rede por outros de maior capacidade.
- Incrementouse o tráfico intercambiado en RECETGA nun 60%
- A saída a internet mediante RedIRIS ampliou o ancho de banda nun 230% respecto o ano anterior.

Cursos e Seminarios

- Durante o ano 2001, o CESGA organizou 16 cursos e seminarios relativos a cálculo, comunicacións e Sociedade da Información.

Programas de cálculo

- Ó longo do pasado ano houbo unha forte demanda de novos programas por

parte dos usuarios. Instaláronse 15 novas aplicacións, sendo a meirande parte delas relativas a Bioinformática.

Proxectos

- Cabe salientar a participación do CESGA en 5 proxectos europeos do VPM a través dos cales intensifícanse as relacións de carácter internacional.



USUARIOS DE COMPUTACIÓN DE ALTAS PRESTACIONES

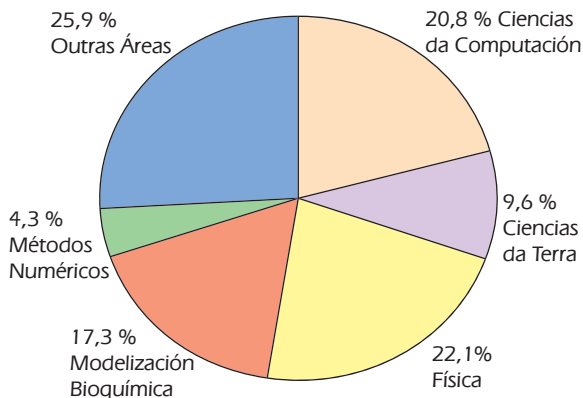
Ó longo do pasado ano, o Centro continuou prestando servizos de acceso a recursos de computación de altas prestacións ós seus usuarios. Os principais beneficiarios destes servizos no ano 2001 foron, como ven sendo habitual, investigadores das tres universidades galegas, dos centros de investigación da Xunta de Galicia e máis dos centros do Consejo Superior de Investigaciones Científicas do Estado. Como excepcións a estes colectivos de usuarios no 2001, cabe destacar, por unha banda, o uso por parte de usuarios do Centro de Informática Científica de Andalucía, que accederon a estes servizos en virtude do acordo de colaboración vixente entre os centros de cálculo científico de Andalucía

(CICA), Catalunya (CESCA) e Galicia (CESGA); e por outra, o acceso por parte de usuarios de empresas privadas como Dygra Films (o consumo destes usuarios aparece reflexado nas estatísticas como "CESGA").

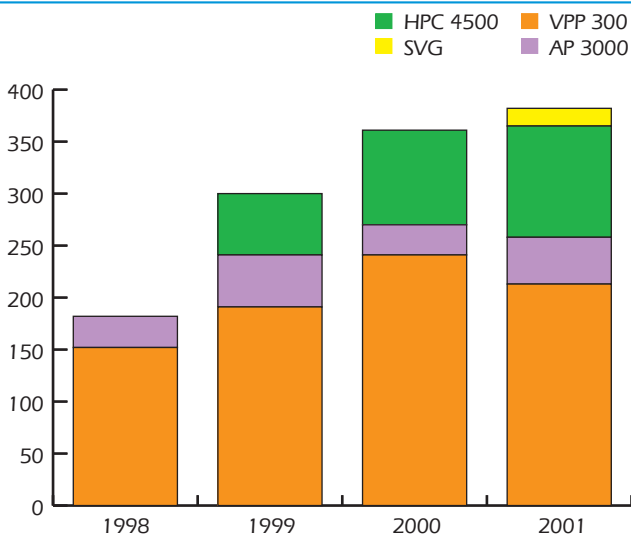
O número de contas de usuario activas experimentou un lixeiro ascenso ó longo do ano 2001. Este incremento debeuse fundamentalmente ós incrementos de contas rexistradas nos sistemas Sun HPC 4500, Fujitsu AP3000 e os usu-

rios do Superordenador Virtual Galego (SVG). Ademais da evolución das contas de usuario activas por sistema e ano, nas seguintes gráficas amósanse: o reparto de horas consumidas no centro por áreas de traballo, a distribución do uso dos sistemas por institucións, así como unha táboa cos usuarios máis activos no Cesga.

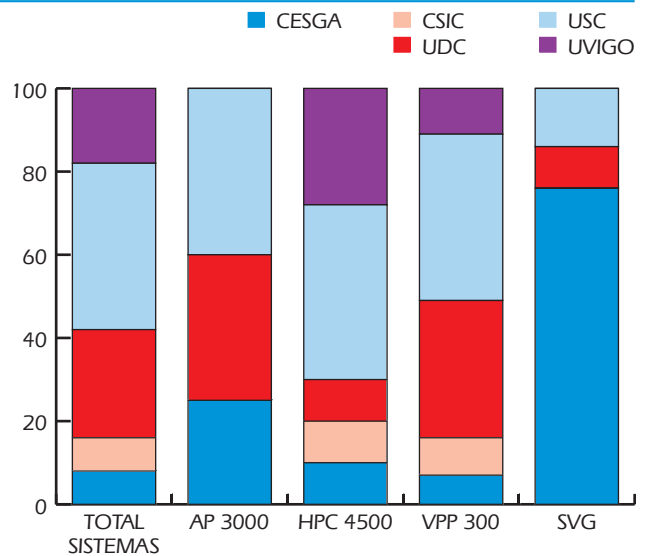
Áreas de traballo máis comúns segundo as horas de CPU consumidas no centro



Evolución das contas de usuario activas por sistema e ano



Distribución de uso dos sistemas por institucións



USUARIOS MÁIS ACTIVOS NO 2001

Usuario	Institución	Horas
Pereiro Lopez, Manuel	USC	25938
Fuentes Lopez, Marcelino	UDC	9115
Perez Muñuzuri, Vicente	USC	8392
Blanco Perez, Vicente	USC	7501
Becerra Permuy, Jose Antonio	UDC	6437
Unidade Pred. Meteorolóxica	USC/Consell.M.A.	3967
Perdines Lence, Inmaculada	USC	3910
Avila Blanco, Gustavo	CSIC	3323
Fernandez Perez, Maria Isabel	UDC	2978
Rodriguez Gutierrez, Santiago	UDC	2968
Sanz Anchelergues, Adolfo	USC	2877
Lorenzo, Maria de las Nieves	USC	2872
Perez Muñuzuri, Alberto	USC	2633
Gonzalez Noya, Eva	USC	2570
Canle Lopez, Moises	UDC	2536
Santaballa, Arturo	UDC	2438
Rodriguez Vazquez, Luis	UDC	2288
Perez Villar, Vicente	USC	2254
Rodriguez Nogueiras, Maria	USC	2226
Rodriguez Gelpi, Ivan Rafael	USC	2140
Rodriguez Ramos, Daniel	UDC	2080
Dieguez Lopez, Oswaldo	USC	2048
Sanchez Gil, Jose Antonio	CSIC	1964

COMPUTACIÓN DE ALTAS PRESTACIONES

O Centro dispón de servidores de cálculo de diferentes arquitecturas para permitir ó investigador elixir sempre o sistema que mellor se adecúa ás súas necesidades de cálculo. De acordo cos requisitos impostos polo tipo de investigación na que estean traballando e segundo o tipo de algoritmos empregados, os usuarios do CESGA poden seleccionar entre as seguintes arquitecturas: vectorial, vectorial paralela, escalar, escalar paralela con memoria compartida e escalar paralela con memoria distribuída.

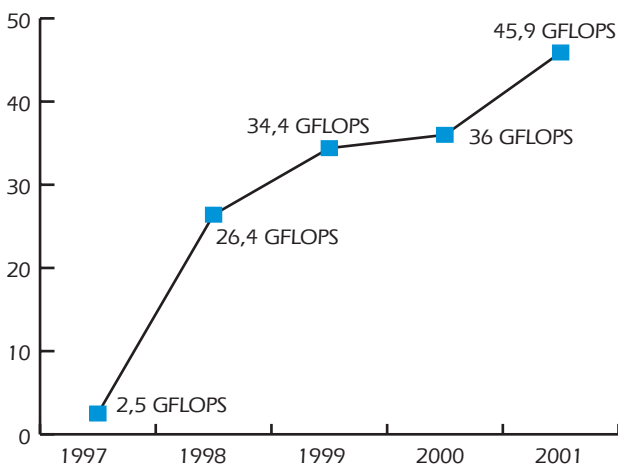
A este repertorio existente, o Centro engade un novo sistema que entra en produción en xaneiro de 2001 para mellor atender a demanda dun número crecente de grupos de investigación que nos seus traballos empregan algoritmos evolucionistas. Trátase do SVG. O SVG é un self made PC cluster desenvolvido ó longo do ano 2000 como resultado do proxecto de investigación "Superordenador Virtual Galego" levado a cabo polo Grupo de Sistemas Autónomos da Escola Politéc-

nica Superior da UDC en Ferrol xunto co CESGA.

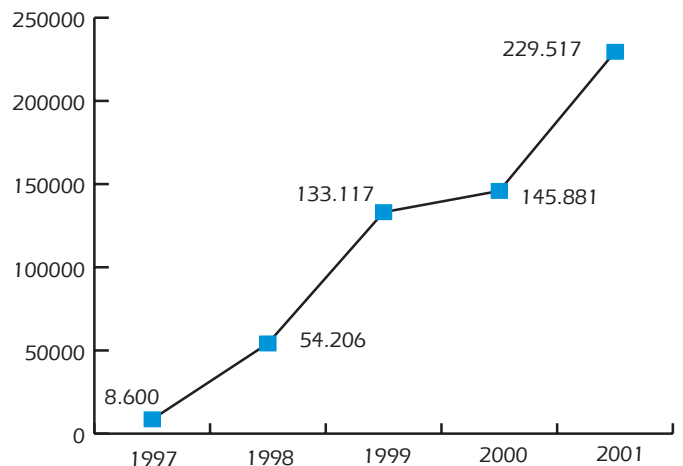
Ó longo do ano 2001 a capacidade de cálculo total instalada no Centro viuse incrementada nun 27.5% pasando de 36 a 45.9 GFLOPS. Como se pode apreciar na gráfica, o total de horas de CPU consumidas polos usuarios experimentou un crecemento dun 57% con respecto ó anterior ano pasando de 145.881 h. en 2000 a 229.517 horas en 2001.

Evolución da Capacidade de Cálculo Instalada no CESGA

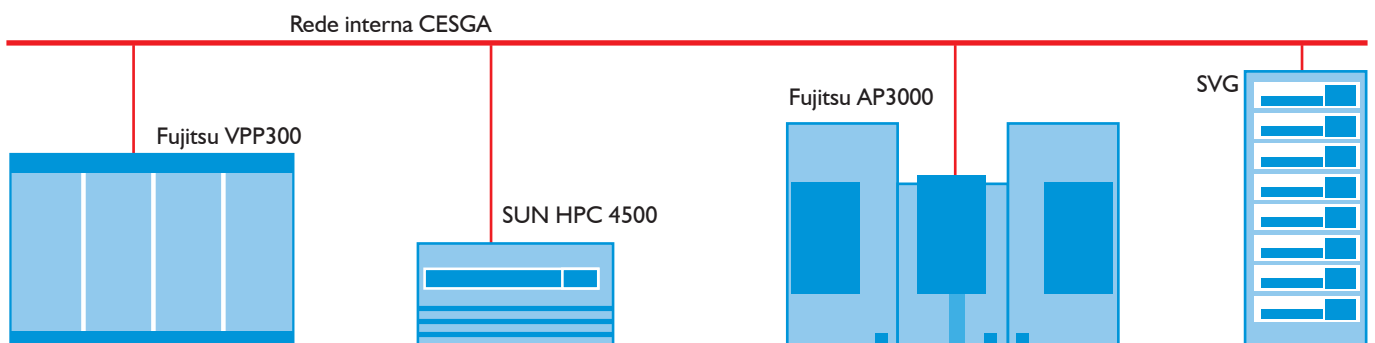
1 GFLOP= 1.000.000.000 operacións de punto flotante por segundo

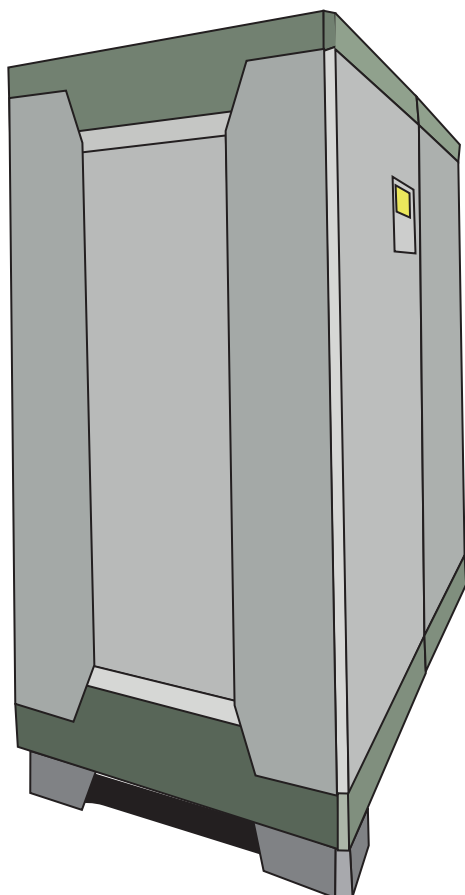


Evolución das Horas de Cálculo Empregadas polos Usuarios



Sistemas de computación de altas prestaciones





ESTADÍSTICAS DE USO FUJITSU VPP300E

O superordenador VPP300E tivo un funcionamento regular ó longo do ano 2001, cunha dispoñibilidade do 99'3% e unha ocupación media do 99'1%.

Os índices de evolución de uso deste superordenador durante o ano 2001, comparados cos do ano 2000 son como segue:

- Aumento de uso nun 6'7%.
- Dispoñibilidade similar á do ano anterior.
- Incremento da ocupación media nun 6'2%.

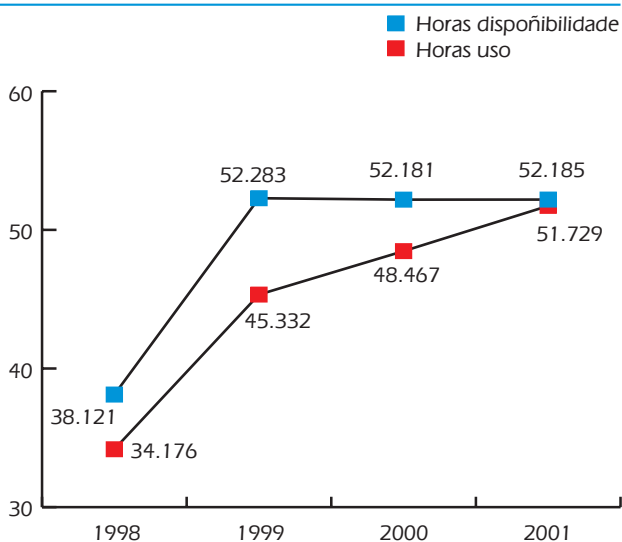
O Fujitsu VPP300E é un multiprocesador vectorial deseñado especificamente para realizar de forma eficiente operacións nas que se ven involucrados elementos de matrices, denominados vectores. Os ordenadores vectoriais resultan especialmente útiles para ser empregados en cálculos nos que as operacións con vectores e matrices son amplamente utilizadas.

As táboas e gráficas amosan as especificacións técnicas deste servidor, a distribución por institucións das horas de CPU nel consumidas e os índices de evolución do uso e dispoñibilidade durante o ano 2001 comparados cos do 2000.

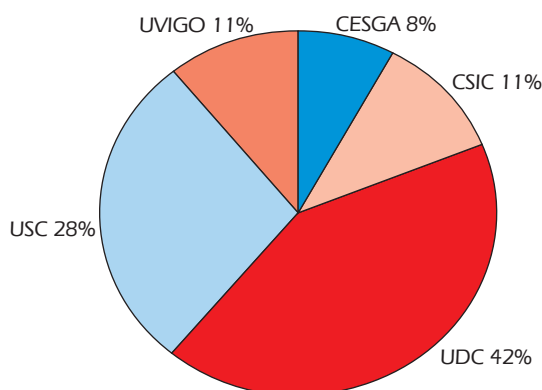
VPP300E	Ano 2000	Ano 2001	% evolución
Horas de uso	48.467	51.729	+6'7%
Horas de dispoñibilidade	52.181	52.185	+ 0%
% dispoñibilidade	99	99'3	+0'3%
% ocupación	92'9	99'1	+6'2%

Especificacións técnicas	
SERVER NAME	Fujitsu VPP300 E
ARCHITECTURE	Parallel Vector Multiprocessor
NUMBER OF PROCESSORS	6
TYPE OF PROCESSOR	Vector (Fujitsu Proprietary)
PEAK PERFORMANCE	14.4 GFLOPS
INTERCONNECT	615 MB/s X 2 Crossbar
MEMORY	12 GB SDRAM
DISC	560 GB
OS	UXP/V

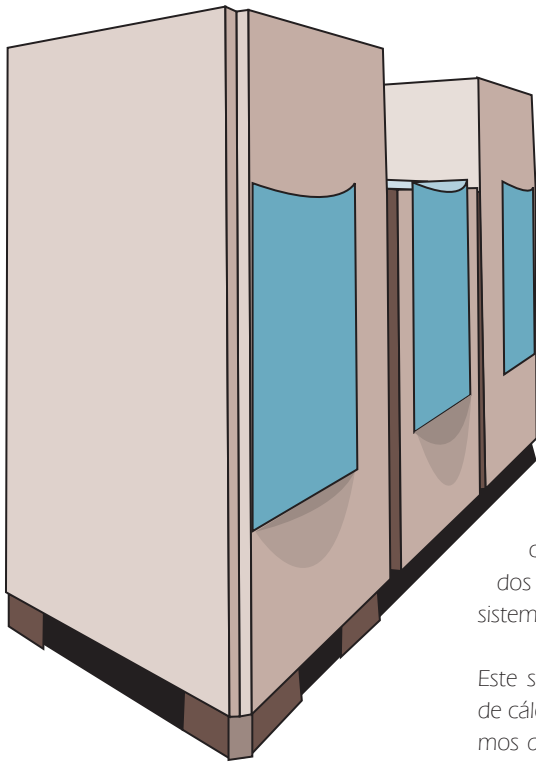
Evolución das Horas de Dispoñibilidade e de Uso no VPP 300E



Distribución por institucións das horas de CPU consumidas no VPP300E durante o 2001



ESTADÍSTICAS DE USO FUJITSU AP3000



O superordenador AP3000 incorpora unha arquitectura paralela con memoria distribuída, mediante a cal a comunidade científica está levando a cabo experiencias relativas ó deseño e uso de algoritmos de paralelización.

A obsolescencia tecnolóxica do AP3000 fai que as súas prestacións se vexan facilmente superadas por granxas de PC's de baixo custo localizados nos Departamentos dos usuarios. Por esta razón, o uso deste sistema é reducido.

Este sistema non se orienta á produción de cálculo senón á investigación de algoritmos de paralelización, polo que o nivel de ocupación media das CPU's foi dun 20'4% cunha dispoñibilidade do 99'4%.

O incremento de uso deste sistema viu determinado polas seguintes razóns:

- Novos proxectos de paralelización.
- Dispoñibilidade para uso en procesos non paralelizables.

Os índices de evolución de uso do superordenador durante o ano 2001 comparados co ano 2000, son como segue:

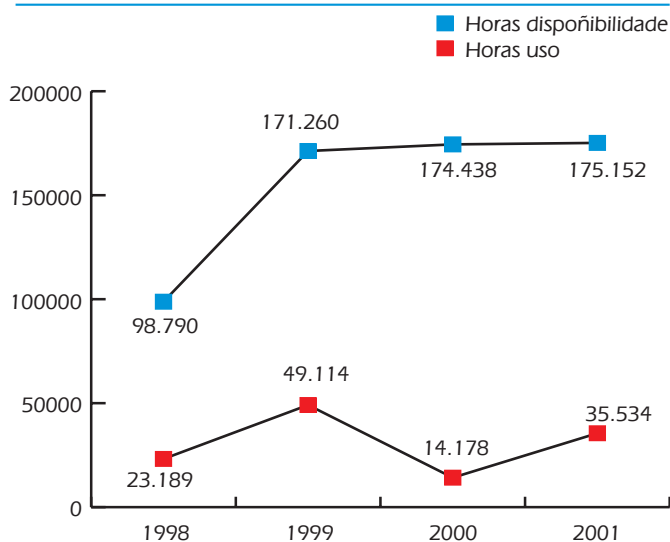
- Aumento en horas de uso nun 150'6%
- Dispoñibilidade similar á do ano anterior.
- Aumento do nivel de ocupación nun 12'3%

As táboas e gráficas amosan as especificacións técnicas deste servidor, a distribución por institucións das horas de CPU nel consumidas e os índices de evolución do uso e dispoñibilidade durante o ano 2001 comparados cos do 2000.

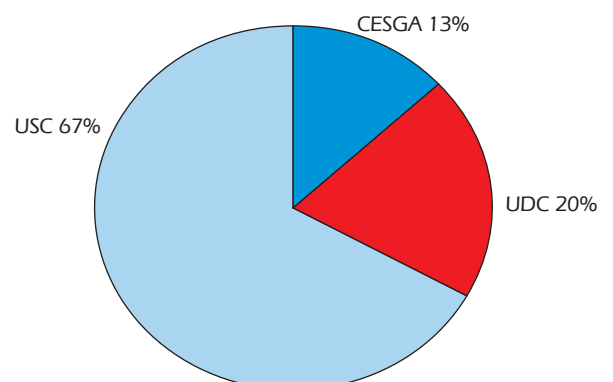
AP3000	Ano 2000	Ano 2001	% evolución
Horas de uso	14.178	35.534	150'6%
Horas de dispoñibilidade	174.438	175.152	+0'4%
% dispoñibilidade	99'3	99'9	+0'6%
% ocupación	8'1	20'4	12'3%

Especificacións técnicas	
SERVER NAME	Fujitsu AP 3000
ARCHITECTURE	Parallel Distributed Memory Multiprocessor
NUMBER OF PROCESSORS	20
TYPE OF PROCESSOR	UltraSPARC-2 300 MHz
PEAK PERFORMANCE	12 GFLOPS
INTERCONNECT	Torus 2D 2 X 200MB/s
MEMORY	2.5 GB SDRAM
DISC	67.2 GB
OS	SUN SOLARIS 2.6

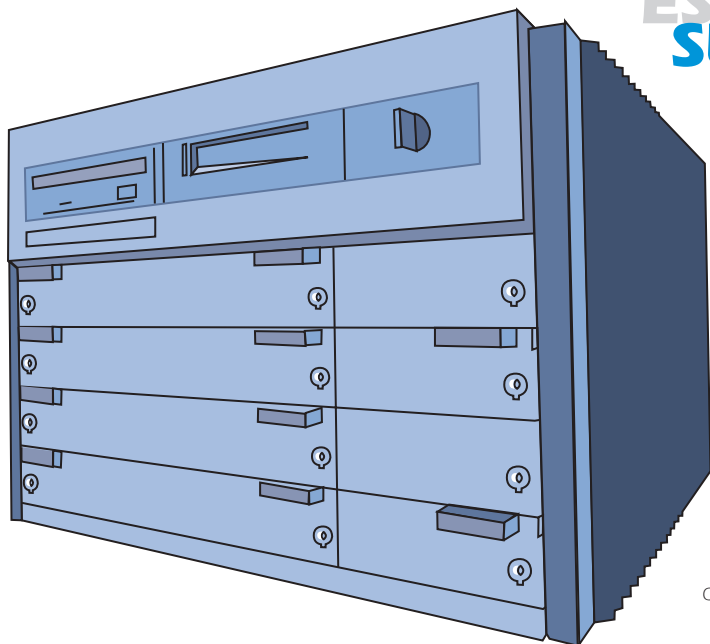
Evolución das Horas de Dispoñibilidade e de Uso no AP3000



Distribución por institucións das horas de CPU consumidas no AP3000 durante o 2001



ESTADÍSTICAS DE USO SUN HPC 4500



O servidor HPC4500 incorpora unha arquitectura SMP (multiprocesador con memoria compartida) utilizada en aplicacións de cálculo escalar. Este equipo dedícase fundamentalmente a explotar códigos de cálculo, sendo reducido o seu uso en deseños e probas de algoritmos de paralelización.

O servidor HPC4500 tivo un funcionamento regular durante o ano 2001, cunha dispoñibilidade dun 99'3% e unha

ocupación media do 75'8%. Os índices de evolución do uso deste servidor durante o ano 2001, comparados cos do ano 2000 son como segue:

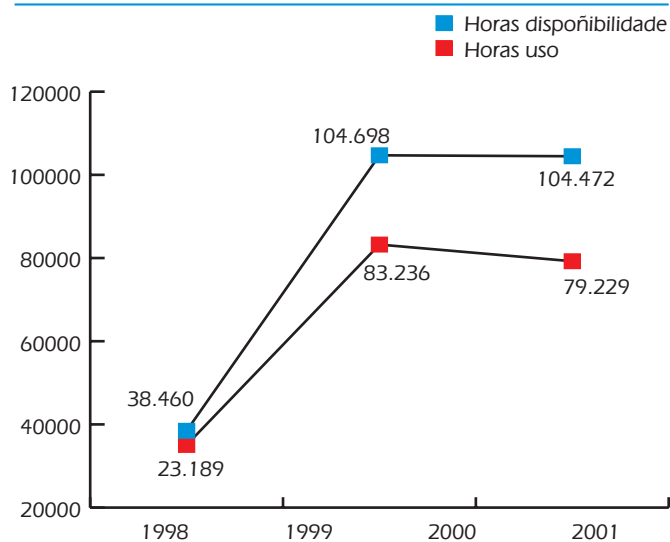
- Diminución de uso nun 4'8% debido fundamentalmente ó uso do AP3000 en aplicacións non paralelizables.
- Dispoñibilidade similar á do ano anterior.
- Redución da porcentaxe de ocupación debido ó menor número de horas consumidas.

As táboas e gráficas amosan as especificacións técnicas deste servidor, a distribución por institucións das horas de CPU nel consumidas e os índices de evolución do uso e dispoñibilidade durante o ano 2001 comparados cos do 2000.

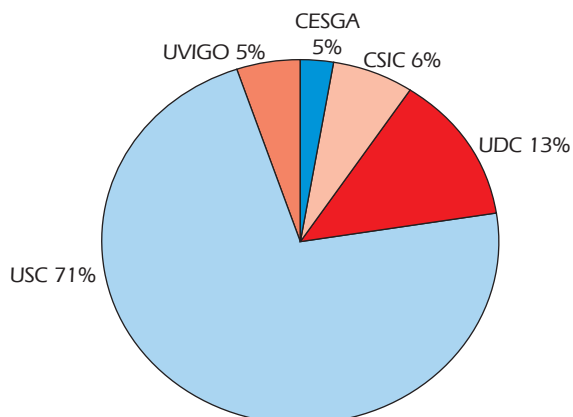
HPC4500	Ano 2000	Ano 2001	% evolución
Horas de uso	83.236	79.229	-4'8%
Horas de dispoñibilidade	104.698	104.472	+0'4%
% dispoñibilidade	99'3	99'3%	0%
% ocupación	79'5	75'8%	-3'7%

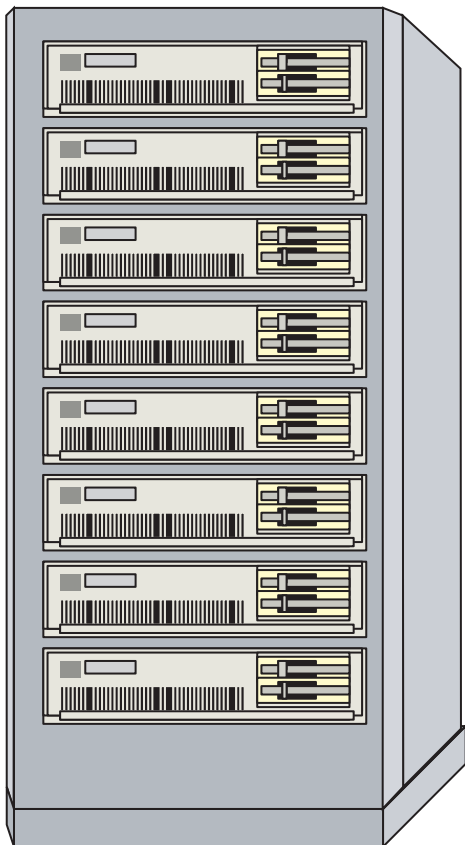
Especificacións técnicas	
SERVER NAME	Sun Microsystems HPC 4500
ARCHITECTURE	Shared Memory Processor
NUMBER OF PROCESSORS	12
TYPE OF PROCESSOR	UltraSPARC-2 400MHz
PEAK PERFORMANCE	9.6 GFLOPS
INTERCONNECT	Torus 2D 2 X 200MB/s
MEMORY	4 GB SDRAM
DISC	36 GB
OS	SUN SOLARIS 2.6

Evolución das Horas de Dispoñibilidade e de Uso no HPC4500



Distribución por institucións das horas de CPU consumidas no HPC4500 durante o 2001





ESTADÍSTICAS DE USO SVG

En xaneiro de 2001 entra en produción como un servidor de cálculo máis ó dispor dos usuarios o cluster que conforma o Nodo CESGA do Superordenador Virtual Galego. O Superordenador Virtual Galego nace dun proxecto de investigación realizado conxuntamente polo Grupo de Sistemas Autónomos da UDC e o Cesga no 2000. Este proxecto contou coa financiación da Secretaría Xeral de Investigación e Desenvolvemento.

Este sistema componse de tres nodos interconectados coa rede de banda ancha RECETGA.

O SVG emprégase basicamente en procesos de renderización, algoritmia xenética ou evolucionista e paralelización.

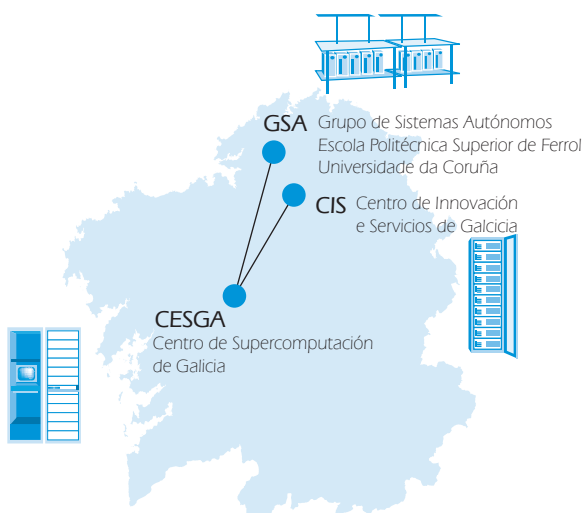
Nas táboas amósanse: unha descrición das especificacións técnicas do SVG, a configuración dos clusters nas distintas localizacións e os índices de uso do Nodo localizado no Cesga.

O SVG é un equipo "self made" moi versátil e de múltiples aplicacións. Durante o ano 2001 realizáronse probas con resultados satisfactorios empregando aplicacións para as áreas de bioinformática e física de partículas. O SVG foi tamén amplamente empregado na realización de experiencias GRID.

SVG (Nodo CESGA)	Ano 2001
Horas de uso	63.025
Horas de dispoñibilidade	156.420
% dispoñibilidade	99'2
% ocupación	40'8

Especificacións técnicas	
SERVER NAME	SVG
ARCHITECTURE	PC Farm
NUMBER OF PROCESSORS	Over 50
TYPE OF PROCESSOR	Intel Pentium III 550MHz up to 1GHz
INTERCONNECT	Ethernet 100
MEMORY	512 MB - 1GB per node
DISC	9 up to 36 GB per node
OS	Linux Red Hat 6.2

Principais nodos do SVG



Configuración SVG 2001	
NODO CESGA:	18 procesadores Pentium III a 550 MHz + 2 K7 a 1GHz, 7 GB memoria e 106 GB de disco e sistema de almacenamento masivo, conexión á Rede de Ciencia e Tecnoloxía de Galicia (RECETGA).
NODO GSA - UDC:	11 procesadores K7 a 700 MHz, 4 GB de memoria e 140 GB de disco, conexión á RECETGA.
NODO CIS Galicia:	18 procesadores Pentium III a 800 MHz, 4.5 GB memoria, 120 GB disco, conexión á RECETGA.

Aplicacións instaladas por máquina

Temas	Programas/aplicacións	Instalados en
Análise Científico	PAW, PAW++, ROOT	HPC4500, AP3000
Bases de Datos Científicas	CSD, CRYSTMET ZENTRALBLATT MATH	db.cesga.es zmath.cesga.es
Bioinformática	GeneHunter MUMmer Clustalw GlimmerM Combiner BLAST Phylip	HPC4500, AP3000 SVG HPC4500, AP3000 SVG SVG HPC4500, AP3000 HPC4500
Cálculo Estructural	ANSYS Multiphysics	HPC4500, AP3000
Flúidos e Magnetismo	GENESIS	HPC4500
Simulación	MATLAB GEANT 3.21 EGSnrc GEANT 4.3.2	HPC4500, VPP300 HPC4500, AP3000 HPC4500, AP3000 HPC4500, AP3000, SVG
Elementos Finitos	MODULEF	VPP300
Información Xeográfica	ARC-INFO / ARC-VIEW	PCGIS,PCGIS2,FGIS
Librerías Matemáticas	CERNLIB BLAS/VP DASSL DSS/2 IMSL LAPACK/VP SSLII/VP-VPP	HPC4500, AP3000 VPP300 VPP300 VPP300
Química Molecular	AMBER GAMESS GAUSSIAN MOPAC93 GAMESS NWChem MOLDEN	VPP300 VPP300 HPC4500 HPC4500, AP3000
Visualización Científica e animación	AVS-Express MAYA NCAR Graphics	HPC4500 ani.cesga.es HPC4500, AP3000
Xestión de software	CMZ, CVS	HPC4500, AP3000

Actualizacións do 2001

Aplicación	Descrición
Gaussian	Migración á última versión no VPP: Gaussian 98
GameSS	Migración da aplicación ó HPC4500
ANSYS	Engadiuse unha nova licencia e adquiriuse a actualización de versión
MATLAB	Instalouse a última versión e engadíronse novos módulos
ZentralBlatt	Actualizouse á última versión de software, separando nunha WEB independente. Ademais realizáronse diversas actualizacións das bases de datos
CSD	Instalouse unha nova versión, ademais de actualizar as bases de datos.
ARC-INFO	Actualización á última versión

Aplicacións de nova instalación no 2001

Nome	Clase	Descrición
NWChem	Química Molecular	NWChem é un paquete de química computacional deseñado para a súa execución sobre superordenadores paralelos de altas prestacións ó igual que sobre clusters de estacións de traballo convencionais. Pretende ser escalable tanto na súa capacidade para tratar grandes problemas de un xeito eficiente como no seu uso dos recursos de computación paralelos dispoñibles.
ROOT	Análise científica	ROOT proporciona unha serie de paquetes Orientados a Obxectos con todas as funcionalidades necesarias para tratar e analizar grandes cantidades de datos de forma eficiente.
CERNLIB	Librerías científicas	Librerías de cálculo científico do CERN de uso en varias das aplicacións instaladas.
GEANT	Simulación	Describe o paso das partículas elementais a través da materia. Orixinalmente foi deseñado para experimentos de Física de Altas Enerxías, pero hoxe en día ten aplicacións tamén en áreas de Medicina, Bioloxía, protección radiolóxica e astronáutica.
EGSnrc	Simulación	Este programa é un Monte Carlo para a simulación do transporte de electróns y fotóns. A utilidade do mesmo é variada aínda que se está utilizando para o cálculo de dosimetrías en tratamentos médicos.
Phylip	Bioinformática	Programa para a realización de árbores filoxenéticas.
ClustalW	Bioinformática	Programa de aliñamento de secuencias de xens ou proteínas e a realización de árbores filoxenéticas.
GeneHunter	Bioinformática	Esta aplicación da solución a problemas de optimización que utilizan as metodoloxías máis avanzadas de algoritmia xenética (multipoint analysis of pedigree data including: non-parametric linkage analysis, LOD-score computation, information-content mapping, haplotype reconstruction)
BLAST	Bioinformática	BLAST® (Basic Local Alignment Search Tool) é un conxunto de programas deseñados para explorar as bases de datos de secuencias dispoñibles independentemente de se é un fragmento de DNA ou de proteína.
PAW,PAW++	Análise científica	PAW é un instrumento de axuda ós físicos na análise e representación dos seus datos. Proporciona representación gráfica interactiva e análise estatística ou matemática. Forma parte dun paquete máis xeral: as librerías de cálculo científico do CERN, tamén instaladas.
MUMmer	Bioinformática	É un sistema de aliñamento de secuencias que permite aliñar secuencias de milleiros de nucleótidos.
Glimmer/GlimmerM	Bioinformática	É un buscador de xens. Está baseado nos algoritmos de programación dinámica que consideran tódalas combinacións de exóns para incluílas no modelo de xen.
Combiner	Bioinformática	Permite a combinación das saídas de varios programas de busca de xens.
CMZ	Xestión de Software	CMZ é un xestor de código fonte, interactivo, rápido, con función de autodocumentación, personalizable e independente tanto da plataforma como do xestor de documentos do sistema.
CVS	Xestión de Software	Concurrent Versions System é o sistema de control de versións dominante na actualidade no desenvolvemento de software non comercial. Permite tanto a xestión do software de pequenos proxectos individuais como grandes proxectos distribuídos entre moitas localizacións, gracias á súa arquitectura cliente-servidor. Permite o acceso ó código dende calquera parte a través de Internet, evitando conflitos no desenvolvemento de software. CVS é empregado por coñecidos proxectos de código aberto como Mozilla, GIMP, XEmacs, KDE, e GNOME

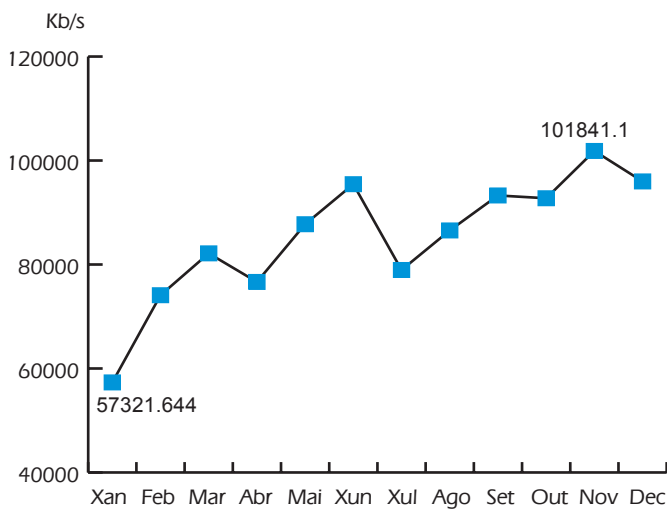
REDE DE CIENCIA E TECNOLOXÍA DE GALICIA

O Cesga continuou, ó longo do ano 2001, a xestionar a Rede de Ciencia e Tecnoloxía de Galicia. A RECETGA é unha infraestrutura de comunicacións propiamente galega mantida conxuntamente con RETEGAL, institución dependente da Consellería de Cultura, Comunicación Social e Turismo. O CESGA é o nodo de RedIRIS en Galicia, a rede de comu-

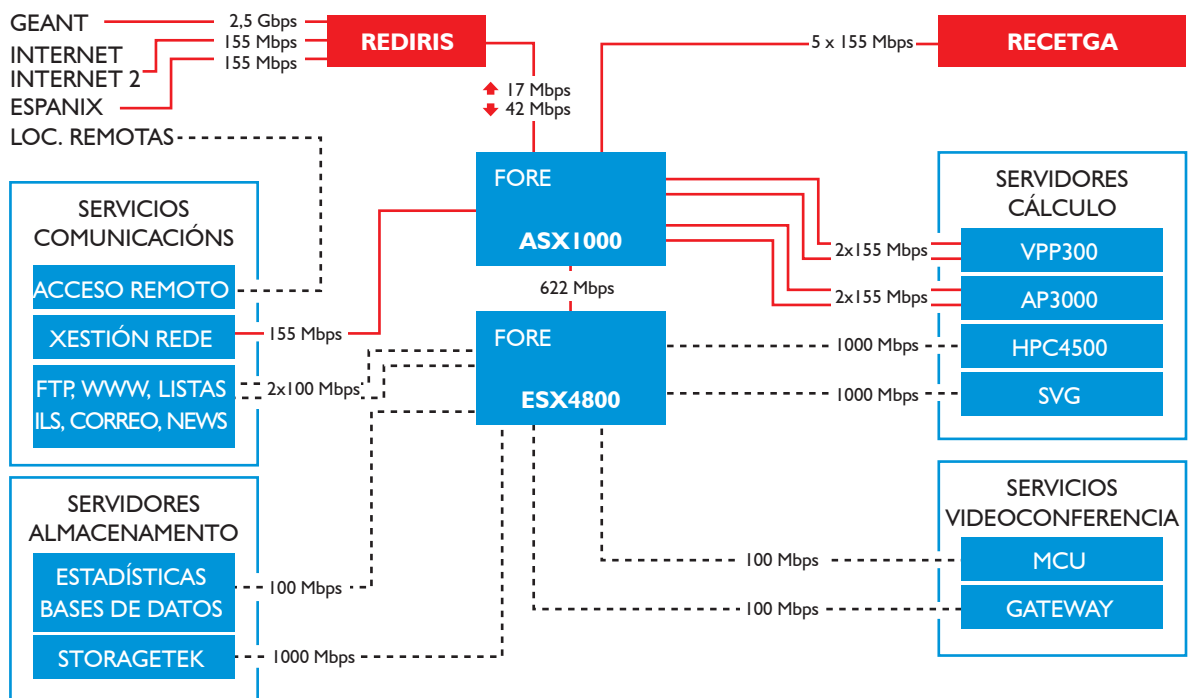
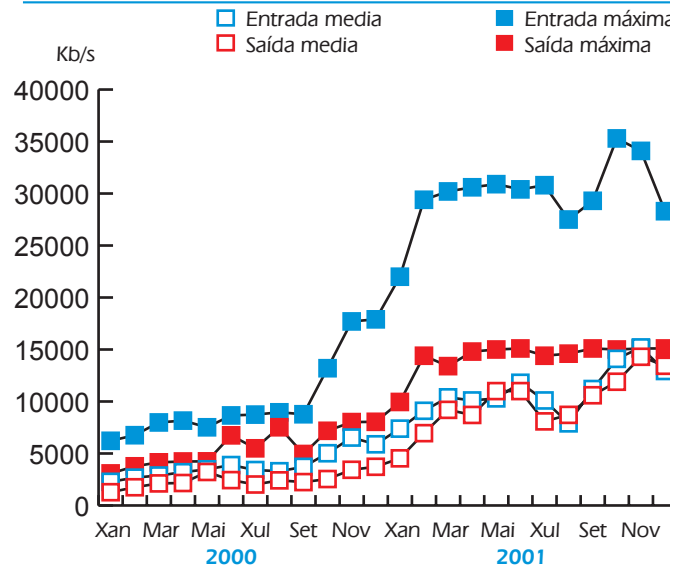
nicacións da comunidade científica española. A actividade de RECETGA incrementouse notablemente durante o ano 2001 debido, entre outras, ás seguintes razóns: ampliación do ancho de banda de RedIRIS, aumento do número de usuarios e maior intercambio de información entre centros.



Intercambio de Tráfico en 2000 no RECETGA



Intercambio de Tráfico RECETGA-RedIRIS 2000-2001



Como feitos significativos cabe destacar: (1) un incremento nun 60% no tráfico intercambiado en RECETGA, (2) un incremento dun 230% no ancho de banda de acceso a RedIRIS que pasou de 18 a 42 Mbps (Se ben estes incrementos son significativos vironse freados no 2001 pola limitada dispoñibilidade de ancho de banda na conexión de RECETGA con outras redes) e (3) a instalación de novos servidores de servizos xenericos de rede para mellor prover servizos á comunidade académica de de investigación en Galicia.

ESPECIFICACIÓNS TÉCNICAS RECETGA

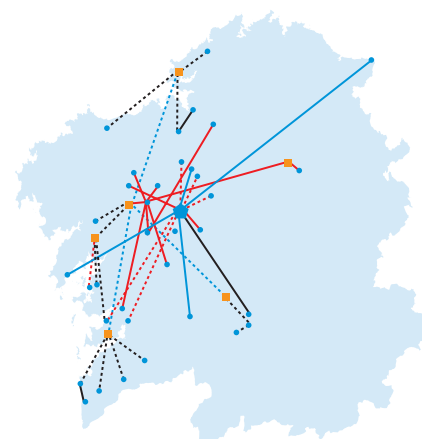
Rede Troncal	Baseada en radio enlaces SDH de 155 Mbps. Conmutadores ATM de FORE ASX-200 e ASX-1000. Conmutadores ATM de CISCO LS1010
Rede de Acceso	Fibra Óptica, Radio enlaces SDH de 155 Mbps., Radio enlaces 4x2 Mbps., Enlaces RDSI. Conmutadores PH7000, PH6000 e LAX-20 de FORE Routers CISCO
Xestión de Rede	Baseada en SPECTRUM de Aprisma e en desenvolvementos propios
Conexión a RedIRIS	42 Mbps entrada, 17 Mbps saída
Rede interna CESGA	Rede ATM - Fast Ethernet - Gigabit Ethernet Conmutadores ATM de FORE ASX-200 ASX-1000 Conmutadores Fast Ethernet 3COM e 2810 Conmutadores FORE ESX-4800 Gigabit ATM ata o posto nos servizos que o requiren.

CENTROS CONECTADOS A RECETGA

Universidade de A Coruña Campus de Coruña Campus de Ferrol	155 Mbps 155 Mbps
Universidade de Santiago de Compostela Campus de Santiago Campus de Lugo	2 x 155 Mbps 155 Mbps
Universidade de Vigo Campus de Vigo Campus de Ourense Campus de Pontevedra	155 Mbps 34 Mbps 155 Mbps
Centros Tecnolóxicos e de Investigación Centro de Investigación Lingüística e Literaria Ramón Piñeiro Centro de Experimentación en Acuicultura Centro de Cultivos Mariños Centro de Control de Calidade do Medio Mariño Centro de Investigacións Mariñas Centro de Información e Tecnoloxía Ambiental Centro de Investigacións Ambientais e Forestais. Lourizán Centro de Innovación e Servizos Centro de Investigacións Agrarias de Mabegondo Estación de Viticultura e Enoloxía de Leiro Fundación Empresa-Universidade de Galicia Aula de Produtos Lácteos Centro Superior Bibliográfico de Galicia Consortio de Bibliotecas Universitarias de Galicia CESGA Centro de Supercomputación de Galicia	155 Mbps 64 Kbps 64 Kbps 155 Mbps 2 Mbps 2 Mbps 155 Mbps 155 Mbps 2 Mbps 64 Kbps 155 Mbps 2 Mbps 155 Mbps 155 Mbps 622 Mbps
Hospitais Complexo Hospitalario "Cristal-Piñor" Complexo Hospitalario "Xeral-Cies" Hospital Clínico Universitario de Santiago de Compostela Unidade de Investigación del Hospital do Meixoeiro Complexo Hospitalario Universitario Juan Canalejo	64 Kbps 155 Mbps 155 Mbps 155 Mbps 155 Mbps
IEO Instituto Español de Oceanografía (A Coruña) Instituto Español de Oceanografía (Vigo)	64 Kbps 128 Kbps
CSIC Instituto de Estudos Gallegos 'Padre Sarmiento' Instituto de Investigacións Mariñas Misión Biolóxica de Galicia Instituto de Investigacións Agrobiolóxicas de Galicia Delegación de Galicia	2 Mbps 155 Mbps 2 Mbps 1 Gbps 64 Kbps

SERVICIOS Ó DISPOR DOS CENTROS CONECTADOS

DNS
ALOXAMENTO WEBS
CORREO ELECTRÓNICO
LISTAS DE CORREO
WEBMAIL
ESTADÍSTICAS DE USO
FTP
NEWS
MIRRORS (contidos de interese)
SERVICIOS ACCESO REMOTO
SERVICIOS SEGURIDADE
MULTICAST
VIDEOCONFERENCIA
MCU/GATEWAY
ILS
PROXY CACHÉ



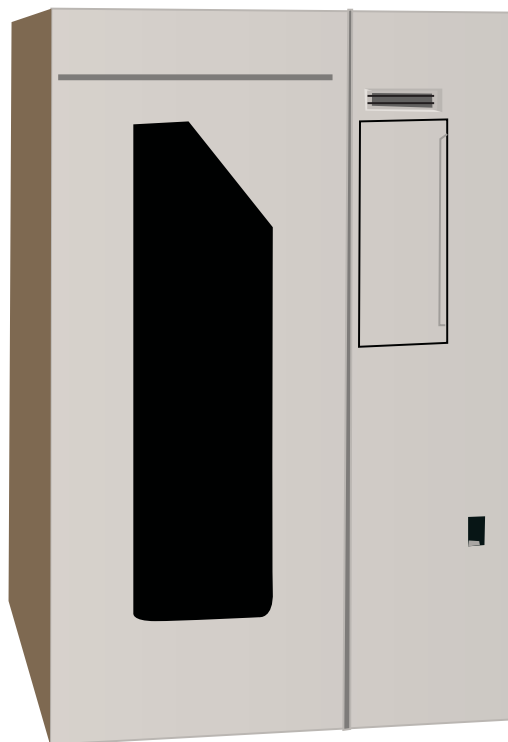
SISTEMA DE ALMACENAMENTO DE DATOS

Durante o 2001, o sistema de almacenamento de datos instalado no centro realizou con normalidade as funcións de: (1) soporte de arquivos de traballo dos sistemas de cálculo, (2) almacenamento de arquivos de usuario e (3) backup automático dos sistemas do Cesga.

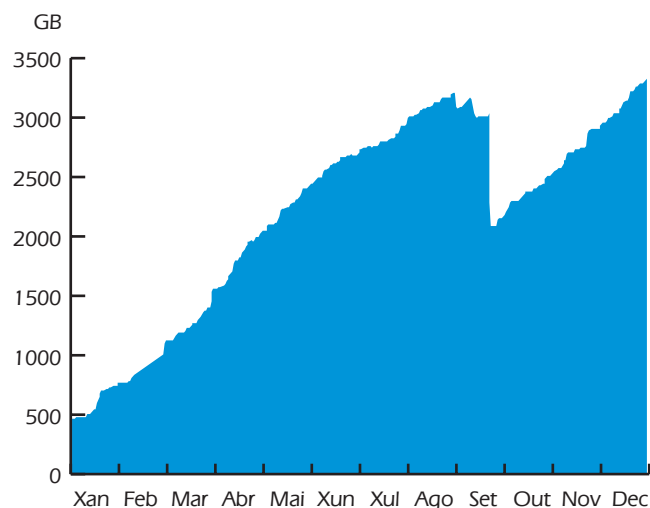
O sistema dispón dunha capacidade en cinta de ata 8TB (sen compresión). O total de capacidade ocupada no 2001 foi de 6 TB, incluíndo 1'6 TB adicados a Backup de sistemas e arquivos de usuario.

O sistema de almacenamento esta constituído por un servidor, dous arrays de discos e un robot de cintas magnéticas, todo isto gobernado por un paquete de software de almacenamento xerárquico da información. O servidor é un Sun Ultra Enterprise 3500, dotado con 4 procesadores UltraSPARC a 400 MHz e 768 MB de memoria. Este servidor dispón de dous discos de 18 GB para albergalo sistema operativo e o software do sistema de almacenamento nunha configuración en espeillo, de maneira que unha avaría nun dos discos non alteraría a dispoñibilidade do sistema cara ós usuarios. Este servidor leva conectados mediante dous bucles de

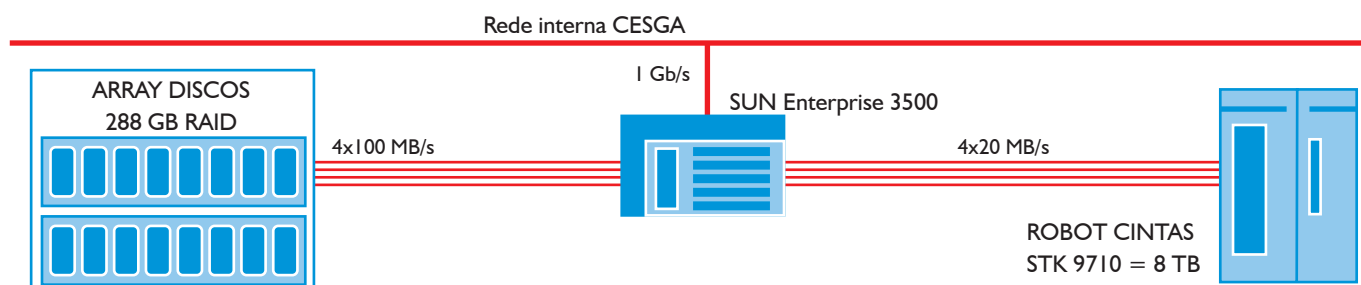
fibra dous arrays de discos Sun A5100 de tecnoloxía fiber-channel, contribuíndo cada un con 8 discos de 18 GB, tendo un total de 288 GB de capacidade de almacenamento en disco. O robot de cintas é un Storagetek 9710, dotado de 4 unidades de lectura/escritura Storagetek 9840 e un total de 420 slots para cintas. As unidades de cinta STK 9840 proporcionan unha capacidade de 20 GB por cinta sen compresión e posúen un índice de compresión típico de 4:1. Polo tanto, o robot dispón dunha capacidade total en liña de 8 Terabytes (sen compresión).



Ocupación do sistema de almacenamento



Sistemas de almacenamento



PROXECTOS

Ó longo do ano 2001 realizáronse tarefas de investigación e desenvolvemento a través dos seguintes proxectos e contratos de servizos:

BUGALICIA, Consorcio de Bibliotecas Universitarias de Galicia

Financiado pola Consellería de Cultura, Comunicación Social e Turismo, pola Secretaría Xeral de Investigación e Desenvolvemento e pola Consellería de Educación e Ordenación Universitaria (Dirección Xeral de Universidades).

Coa participación de: Universidade da Coruña, Universidade de Santiago de Compostela, Universidade de Vigo e Centro de Supercomputación de Galicia.



RIGA, Desenvolvemento, Posta en Producción e Mantemento da Base de Datos

Distribuída para a xestión do Rexistro de Investigadores de Galicia

Financiado pola Secretaría Xeral de Investigación e Desenvolvemento.

Proxecto A PONTE, An aPproach to the depLOyment of New Technologies in Education.

Financiado pola Unión Europea a través do Programa ESPRIT.

Coa participación de SEMA Group (hoxe Schlumbergersema), Asociación de Universidades da Região Norte de Portugal, Context European Educational Ltd. E South Bristol Learning Network.



Proxecto METEO, Modelo de Predicción Meteorolóxica.

Financiado polo Centro de Información e Tecnoloxía Ambiental da Consellería de Medio Ambiente.

Coa colaboración do Grupo de Física non Lineal, Departamento de Física da Materia Condensada, Facultade de Física, Universidade de Santiago de Compostela.

Proxecto RSU, Optimización de Rutas de Transporte de Residuos Sólidos Urbanos.

Financiado pola Sociedade Galega de Medio Ambiente (SOGAMA).

Proxecto e-Galego, Desenvolvemento dun Curso Multimedia Interactivo de Aprendizaxe de Lingua Galega para Adultos.

Financiado polo Fundación Caixa Galicia e a Consellería de Educación e Ordenación Universitaria.

Coa participación do Instituto da Lingua Galega da Universidade de Santiago de Compostela, da Dirección Xeral de Política Lingüística da Consellería de Educación e Ordenación Universitaria e da empresa Eido Media, S.L.

Proxecto SIGREMAR (GIS - CIMA), Aplicación dos Sistemas de Información Xeográfica ó Estudio de Dinámica e Xestión de Poboacións de Invertebrados Mariños con Interese Económico.

Financiado pola Consellería de Pesca, Marisqueo e Acuicultura.

Dirixido polo Centro de Investigacións Mariñas (CIMA) de Corón.

Proxecto e-CANNED, Comercio Electrónico na Industria Conserveira.

Financiado pola Unión Europea, V Programa Marco a través da liña IST (Information Society Technologies).

Coa participación de SEMA Group (hoxe Schlumbergersema, España), ANFACO (Asociación Nacional de Fabricantes de Conservas de Pescados e Mariscos), Escurís, S.A., Justo López Valcárcel, S.A., Conservas Antonio Alonso, S.A., Modesto Carrodegas, S.L. e a Compagnie Bretonne de la Coquille Saint-Jacques (COBRECO).



Proxecto METAFOR, Desenvolvemento e Validación dunha Plataforma de Teleformación vía Satélite en Empresas no Medio Rural.

Financiado pola Unión Europea, V Programa Marco a través da liña IST (Programa

TEN-TELECOM)

Coa participación de Alcatel Space (Francia), Aula de Productos Lácteos da Universidade de Santiago de Compostela en Lugo, Vivial (Francia), Renater (Francia), ENESAD (Francia), SLU (Suecia), Món Blau Verd (Catalunya).

Este proxecto conta co apoio adicional de empresas do sector lácteo en Galicia como son: Queserías Prado (Monforte de Lemos), Feiraco S.C.L. (Negreira), Arqueixal (Palas de Rei), Lactalis (Vilalba), Lácteos Lorán (Abadín), CAPSA (A Robra), CAPSA (Vilagarcía), e LAGASA (Pontedeume).

Proxecto Herdanza, Aplicación de Sistemas de Información Xeográfica ó Estudio, Ordenación, Catalogación, Xestión e Manexo de Xacementos Arqueolóxicos.

Financiado pola Consellería de Cultura, Comunicación Social e Turismo.

Dirixido polo equipo técnico da Dirección Xeral de Patrimonio Cultural.

Proxecto e-MINDER, Desenvolvemento dunha Rede Europea de Centros de Competencias en Comercio Electrónico.

Financiado pola Unión Europea, V Programa Marco a través da liña IST (Information Society Technologies)



Coa participación de SchlumbergerSema (España), Presidencia da Pomerania (Polonia), Computer Center Regional Data Bank (Polonia), Virtual IT (Chipre), Cyprus Telecommunications Authority (Chipre), University of Cyprus (Chipre), Universidade de Santiago de Compostela, Universidade da Coruña.



TELEFORMACIÓN



No ano 2001, a rede de aulas de teleformación desplegada nos sete campus universitarios galegos operou con normalidade. Universidades galegas empregaron estas aulas para a realización de cursos de formación remota inter-campus para persoal e alumnos propios. O CESGA participou do uso destas aulas a través da realización dos cursos do programa de formación para usuarios de servidores de cálculo de altas prestacións. Entre outros impartíronse cursos de Introducción á Programación dos Superordenadores e de Introducción á Programación Paralela.

Adicionalmente, o Centro de Innovación e Servizos (CIS Galicia) no Ferrol promoveu a realización dun curso sobre a aplicación práctica do paquete Rhinoceros de modelado 3D para deseños industriais empregando esta rede de aulas. Así mesmo, a Universidade de Vigo promoveu a realización dun curso de extensión universitaria a través destas aulas sobre Tecnoloxías de Desenvolvemento Competitivo de Software: Patróns de Deseño e Java, que contou tamén coa participación de profesores da Universidade da Coruña.

Os técnicos e investigadores da área de Teleformación do Cesga traballaron ó longo do 2001 no desenvolvemento e realización de tres proxectos de innovación tecnolóxica aplicada a teleformación:

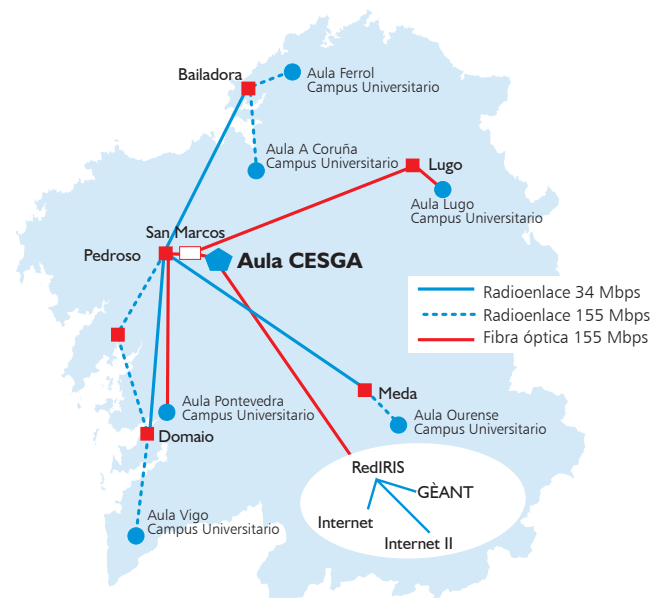
- Proxecto A PONTE, An aPproach to the deplOyment of New Technologies in Edu-cation. Este proxecto chegou o seu remate no primeiro trimestre de 2001 e estivo financiado pola Unión Europea a través do Programa ESPRIT.

- Proxecto METAFOR, Desenvolvemento e Validación dunha Plataforma de Teleformación vía Satélite en Empresas no Medio Rural. Este proxecto comezou en Febreiro de 2001. Está financiado pola Unión Europea, V Programa Marco a través da liña IST (Programa TEN-TELECOM).

- Proxecto e-Galego, Desenvolvemento dun Curso Multimedia Interactivo de Aprendizaxe de Lingua Galega para Adultos. Este proxecto contou coa financiación da Fundación Caixa Galicia e da Consellería de Educación e Ordenación Universitaria.



Rede de aulas de tele-ensino



ACCIÓN DE FORMACIÓN

Curso Modelado 3 D, utilizando o programa Rhinoceros 1.1.

22 ó 26 de Xaneiro

Promovido e impartido polo Centro de Innovación e Servizos CIS Galicia no Ferrol. Celebrado simultaneamente en tres Aulas de Teleensino: no CESGA, no Campus de Lugo e no Campus de A Zapateira (Coruña)

Curso GIS Arc View

26 febreiro ó 2 de Marzo

Centro de Supercomputación de Galicia (CESGA)

26 ó 30 de Marzo

Centro de Supercomputación de Galicia. (CESGA) dirixido a alumnos de Xeografía e Historia da USC.

Formación Universitaria a través de videoconferencia.

Experiencia piloto de impartición de sesións formativas a alumnos da Universidade de Vigo ó longo do curso académico 2000-2001 na asignatura "Introducción ós servizos de Internet".

Tecnoloxías de desenvolvemento competitivo de software: Patróns de deseño e Programación Java

4 ó 13 de setembro de 2001

Vicerrectorado de Extensión Universitaria, Universidade de Vigo. Aulas teleformación de Vigo, Coruña, Lugo e Santiago. Dirixido a membros da Comunidade Universitaria, persoas tituladas e profesionais.

Programa de Formación para Usuarios dos Servidores de Cálculo do CESGA 2001.

Utilizando as Aulas de Tele-ensinanza instaladas no CESGA e no Edificio de departamentos da E.T.S. de Arquitectura (Universidade da Coruña), realizáronse os cursos que a continuación se detallan e que estaban orientados ó uso óptimo dos sistemas de supercomputación instalados no CESGA no referente a programas desenvolvidos polos propios usuarios.

Introducción á Programación dos Supercomputadores (06/Nov/2001)

Introducción á Programación Paralela (26-30/Nov/2001)

LABORATORIO DE VISUALIZACIÓN CIENTÍFICA

O laboratorio de visualización continuou dando apoio ós usuarios no emprego dos seus recursos. Estes servicios consistiron fundamentalmente na xeración de vídeos, diapositivas e material impreso para presentación de resultados.

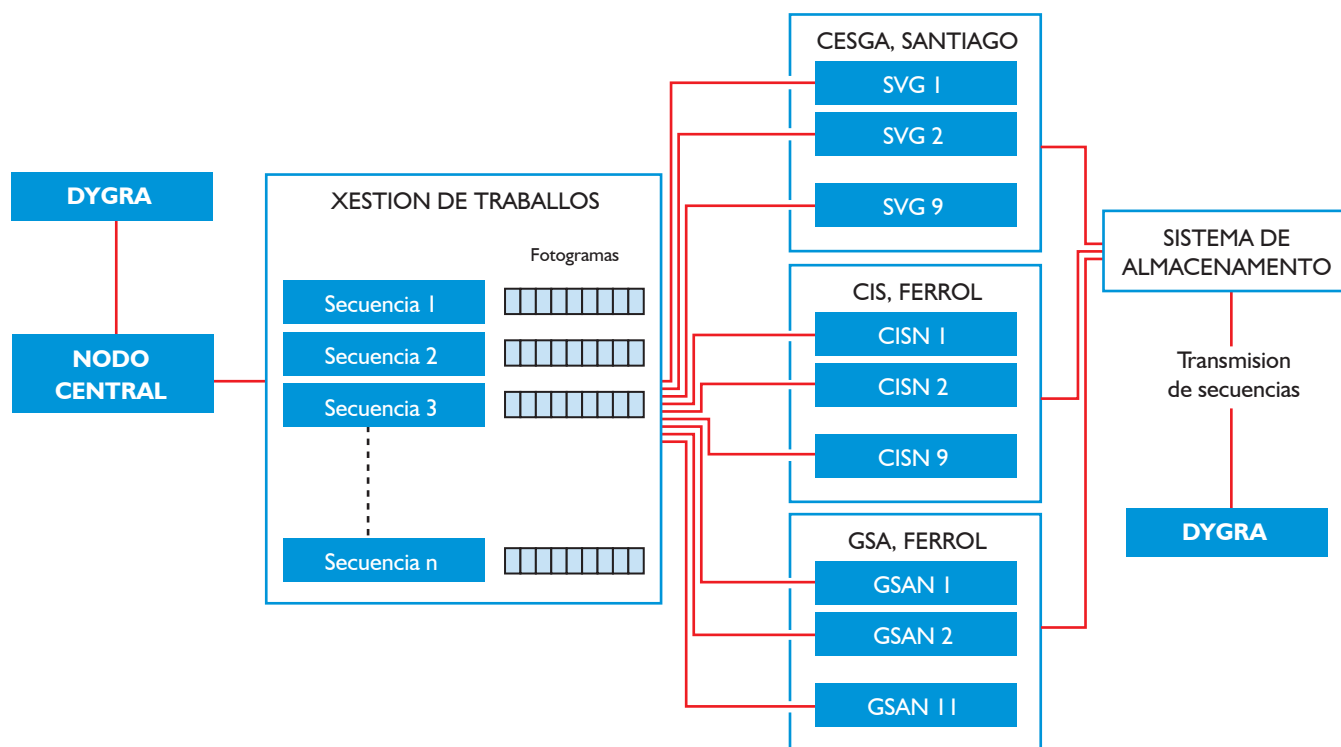
Adicionalmente, no 2001, o Cesga participou, dando soporte á renderización da primeira longametraxe de animación 3D

producida en Europa por Dygra Films, El Bosque Animado, que foi galardoada con dous premios Goya. Para isto, a plataforma do Superordenador Virtual Galego (desenvolvido polo Grupo de Sistemas Autónomos da UDC xunto co Cesga) tomouse nun contorno de renderización de alta produtividade empregando o software para render de MAYA sobre Linux. Neste traballo participaron todos os nodos do Superordenador

Virtual Galego instalado na Escola Politécnica Superior de Ferrol, no CIS Galicia e no Cesga. A realización deste traballo de grande custo computacional supuxo unha proba definitiva para testar a bondade desta plataforma.



Flujo de traballos para o render de "El Bosque Animado" no SVG



GIS

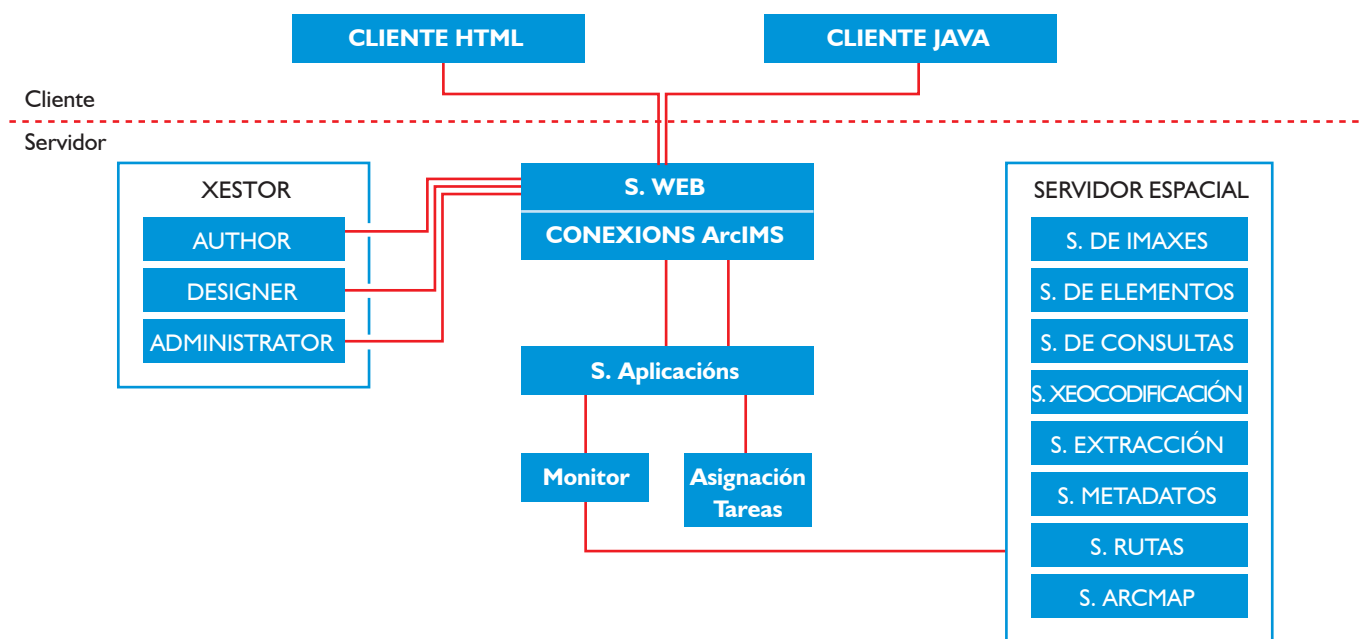
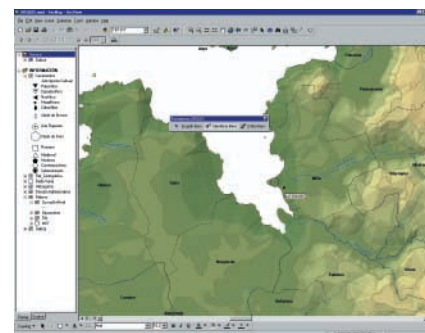
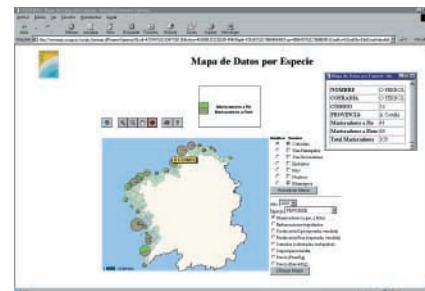
SISTEMAS DE INFORMACIÓN XEOGRÁFICA

Ó longo do ano 2001, o Cesga continuou dando soporte en servizos GIS. Destacables nese ano son a participación e soporte ós seguintes proxectos:

- Proxecto RSU, Optimización de Rutas de Transporte de Residuos Sólidos Urbanos. Financiado pola Sociedade Galega de Medio Ambiente (SOGAMA).
- Proxecto SIGREMAR (GIS - CIMA), Aplicación dos Sistemas de Información Xeográfica ó Estudo de Dinámica e Xestión de Poboacións de Invertebrados Mariños con Interese Económico. Financiado pola Consellería de Pesca, Marisqueo e Acuicultura.

Dirixido polo Centro de Investigacións Mariñas (CIMA) de Corón.

- Proxecto Herdanza, Aplicación de Sistemas de Información Xeográfica ó Estudo, Ordenación, Catalogación, Xestión e Manexo de Xacementos Arqueolóxicos. Financiado pola Consellería de Cultura, Comunicación Social e Turismo. Dirixido polo equipo técnico da Dirección Xeral de Patrimonio Cultural.
- Soporte os Proxectos desenvolvidos polos investigadores do Dpto. de Bibliometría do Centro de Información e Documentación Científica do CSIC (Proxectos GISSTI e Proxecto HI+DROLOGIA)



ACCESIBILIDADE GIS

No ano 2001 incorporouse o novo servidor cartográfico Arc IMS3. Para labores de consulta, edición de información, administración ou xestión, os usuarios poden interactuar co servidor cartográfico Arc IMS3 instalado no Cesga facendo uso de distintas plataformas de acceso remoto. Para consultar a cartografía no servidor, os usuarios poden facer uso dos clientes Java ou html soportados polos seus PCs ou Palm Tops. No seguinte esquema descríbese a funcionalidade do servidor Arc IMS3.



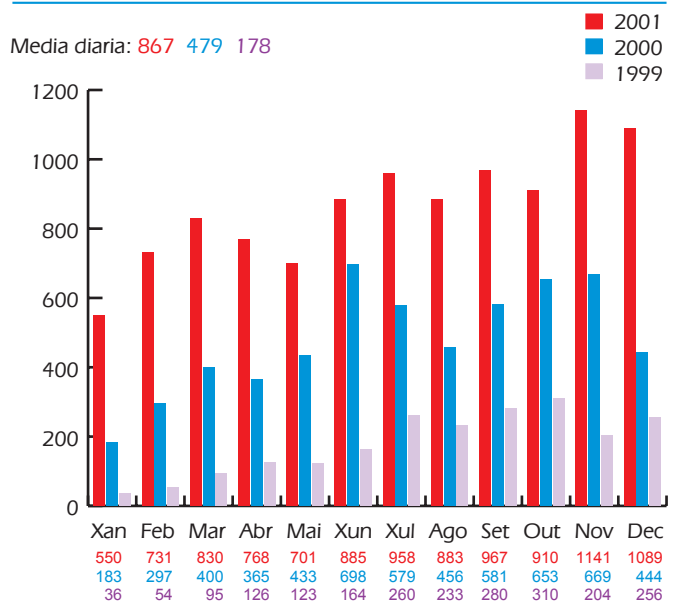
DIFUSIÓN

O CESGA continuou os seus esforzos no 2001 por dar difusión á actividade investigadora da comunidade de usuarios e á propia actividade do Centro. Así, catro novas tiradas de "díxitos" veron a luz nese ano. Unha destas foi un número especial co anuario de actividade do Centro. O sitio web do Cesga continuou incorporando contidos útiles para os investigadores e rexistrou un significativo incremento con respecto ós anteriores anos no tocante á media diaria de sesións de usuarios tal como se pode apreciar na gráfica. A Workshop HPCN 2001, que contou co patrocinio de GreenData, serviu unha vez máis para achegar un foro no que usuarios e administradores de recursos poideran compartir as súas experiencias, necesidades e inqedanzas. Tamén no 2001, o Cesga acolleu bo número de visitas de grupos de

investigación e de alumnos de centros de secundaria.

Salientable, entre as accións de difusión realizadas no 2001, foi a participación do Cesga na organización de dúas xornadas sobre Bioinformática, celebradas baixo o paraugas do Foro Científico-Tecnolóxico da Secretaría Xeral de Investigación e Desenvolvemento que culminarían coa creación da Rede Galega de Bioinformática.

Promedio diario de sesións de usuario en www.cesga.es



PERSONAL CESGA:
PRINCIPAL ACTIVO DO CENTRO NO 2001

WORKSHOP

HPC 2002 High Performance Computing

Salón de Actos
Facultade de Informática, Universidade da Coruña

Xoves 30 de Maio

Salón de Actos



PROGRAMA E
INSCRIPCIÓNS: <http://www.cesga.es/Cursos/>

El Único con 100% de capacidad útil



V960

Disco Virtual Aprovecha toda su capacidad

- Es la base para un modelo de almacenamiento heterogéneo escalable y 100% disponible, con la última tecnología avanzada en entornos críticos, ahora ya disponible para sistemas abiertos.
- Gestión inteligente y automática de las necesidades inesperadas del día a día, con grandes ahorros en costes, tiempo y recursos.



Disco Tradicional

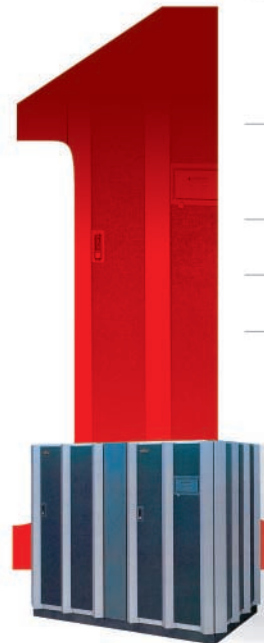
Disco Virtual



STORAGETEK

Soluciones Globales de Almacenamiento
www.storagetek.es

PRIMEPOWER Líder en rendimiento



Por Fiabilidad

Procesadores SPARC con protección ECC en todos los niveles de caché para minimizar la posibilidad de fallo.

Por Escalabilidad

Arquitectura cross-bar que asegura máxima escalabilidad y utilización de recursos del sistema, hasta 128 procesadores.

Por Flexibilidad

El sistema puede ser dividido para permitir configurar aplicaciones diferentes.

Así, los Servidores Solaris™ PRIMEPOWER™ de Fujitsu han obtenido la máxima calificación como número uno en los benchmarks SAP, SPEC, Oracle System y TPC.

FUJITSU

THE POSSIBILITIES ARE INFINITE

FUJITSU ICL ESPAÑA, S.A. www.fujitsu.es • 901 100 900