IRIX

- Standards
- User Shells
- FileSystems: XFS
- Memory Mangement & SMP
- Performance I/O
- Scheduling
- Real Time
- Cellular IRIX
- HA (High Availability)
- Integración con UNICOS
- Checkpoint – Restart

Standards

3

- *X/OPEN XPG4 Base95 (API para portabilidad)*
- *SVID3 (SystemV Release 4 (SVR4)*
  Verificado con el gABI de UNISOF (7.800 tests)
- *MIPS ABI 1.2.2*
- *POSIX 1003.1, 1003.2, 1003.1b (1003.4 Draft14) for RT,*
  1003.1c (pthreads), 1003.2

  FIPS 151–2 certification
- *Single Unix Specification (SPEC 1170)*
- *X11R6, OSF Motif 1.2.4 (32 y 64 bit)*
  CID keyed fonts, X double buffering extension

User Shells

4

- ***Shells***
  - ***Bourne Shell (sh)***
  - ***Korn Shell (ksh)***
  - ***C Shell (csh)***
  - ***Tcshell (tcsh)***

Otros free–ware (secure shell, zsh, etc...)

Filesystems

5

- ***EFS (32 bit: 8GB Filesystems, 2GB files)***
- ***XFS (64 bit)***
- ***NFS (v3) + CFS***
- ***ISO (High Sierra & 9660 (CDFS)***
  con/sin Rock Ridge Extensions
- ***DOS + HFS/Mac***
  single/double density floppy
- ***swap & /proc***

Filesystems: XFS

6

- ***Journaled Filesystem***
- ***67 million files/dir***
- ***Block size: 512 Bytes –> 1 GB***
- ***Guaranteed I/O (GRIO)***
- ***XLV (striping, concatenation, mirroring)***
- ***DMAPI for HSM (EMASS & Hitachi uses it)***

Filesystems: XFS

	Filesystem	File
32 bit systems (Indy – 02)	1 TB	1 TB
64 bit systems	9 mill TB's	9 Mill TB's

- ***compatible tanto aplicaciones de 32 o de 64 bits***

Memory mngmt & SMP

- ***SVR4 Based (mmap() & mprotect())***
- ***Physical Memory –> 256 GB***
- ***1 TB virtual memory (VM) address space***
- ***Multithreated Kernel***
- ***Synchronization primitives in MIPS inst. set***
- ***sproc() Lightweight processes***

Performance I/O

9

I/O Type	Description	Purpose
UNIX read/write	Normal I/O UNIX methods	This is the standard way that processes obtain data from the filesystem
Memory mapped files (SVR4)	Map contents of file directly into process address space memory	Bypass read() operations and access disk files as though it was regular
Asynchronous (Posix 1003.4)	Allow process to continue without waiting for operations to complete	Keep I/O operations from delaying process execution
Direct I/O	Copy data directly from process space to disk via DMA	Avoid delay of copying data to kernel and having the buffer flushed to disk at a later time

La libreria de FFIO de Cray està ya portada sobre 6.4

Scheduling

10

sched. class	app. types	OS queue
AFFINITY	interactive jobs batch jobs	TIME–SHARING BATCH
GANG	parallel app's	GANG GANG–BATCH
ABSOLUTE PRIORITY	kernel real–time app's	KERNEL REAL–TIME

Scheduling

11

- *Process to memory affinity*
- *Mem. page locking (no page fault)*
- *Currency Based Scheduler*
- *Dynamic Shared Objects*
- *Kernel Threads*
- *POSIX threads syncronization*

Real Time

12

- *specific schedulings (rt & fifo)*
- *Mem. page locking (no page fault)*
- *Guaranteed interrupt latency*
- *Processor isolation, restriction*
- *Guaranteed I/O from XFS (GRIO)*
- *Graphic tool (WindView)*

Cellular IRIX

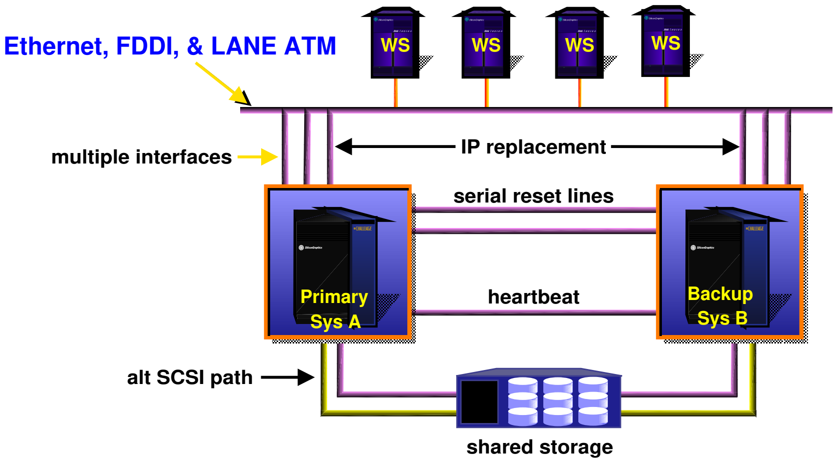
13

- ***Actual soporte de hasta 64 CPU’s en SMP***
- ***Hardware graph***
- ***Soporte de hasta 128 CPU’s en SMP en IRIX 6.5 (beta)***
- ***Células parciales (golden – slaves)***
- ***TCP/IP por Craylink***
- ***Kernel memory migration/replication***
- ***Large page support** (16KB, 64KB, 256KB, 1MB, 4MB, 16MB)*
- ***Actual soporte de hasta 64 CPU’s en SMP***

HA (High Availability)

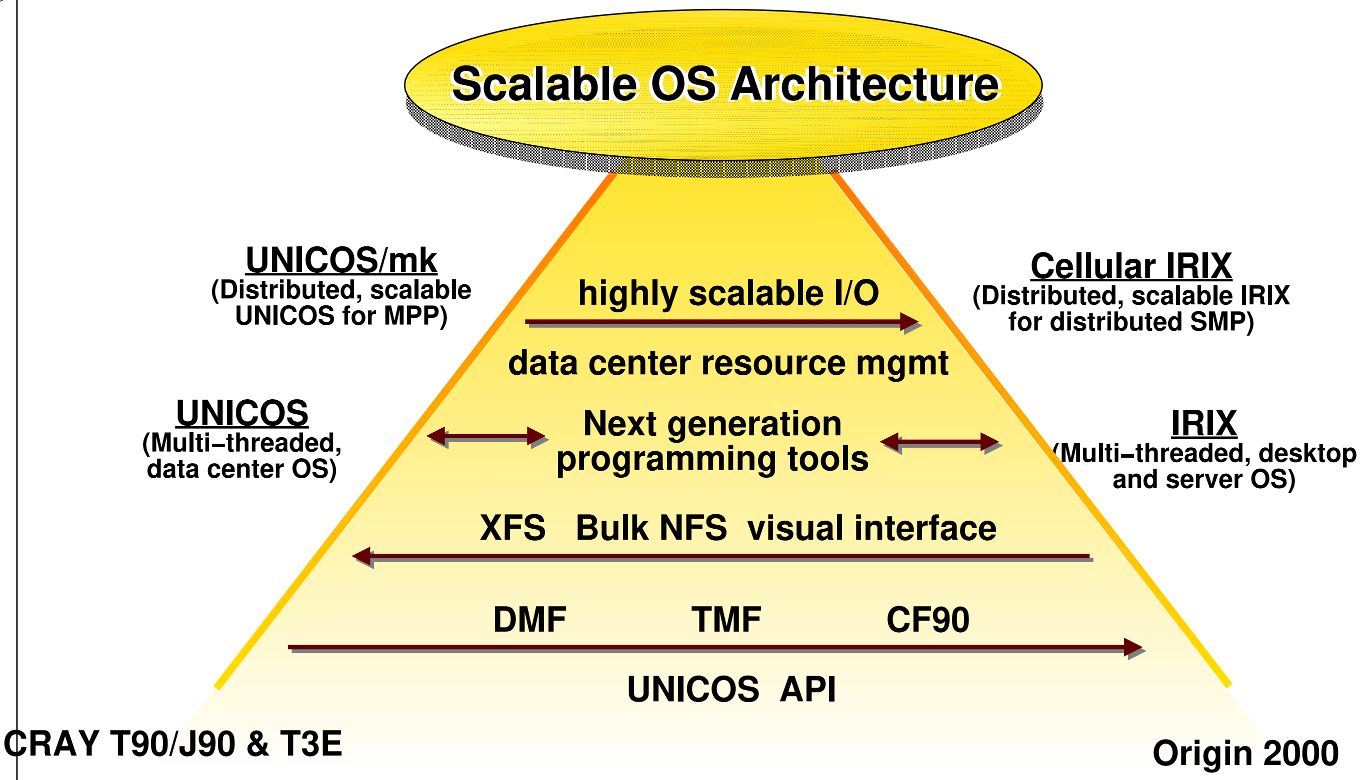
14

- ***Detectar una anomalía en una aplicación y arrancarla en otro server, manteniendo las direcciones IP y los discos.***
- ***A los clientes les parece como si el servidore hubiese rebotado rápidamente.***
- ***Pluginns:***
  - Web – Netscape® servers
  - NFS – WITH NFS file locks
  - Oracle® RDB
  - INFORMIX® RDB
  - Sybase® RDB
  - Gauntlet™ Firewall
  - WebFORCE® MediaBase



Integración con UNICOS

15



Checkpoint – Restart

16

– Parar un job running, guardarlo a disco para re–arrancarlo más tarde. También funciona para un grupo de procesos (prog paralelo, pthreads,...), para sesiones de array y para un arbol de procesos (p.e. todos los de un usuario).

– Importantísimo para HPC (jobs muy largos) para:

  * paradas de mantenimiento HW o SW

  * verificacion de la convergencia

– integrado con NQE (al parar las colas de guardan los procesos en marcha)

– man cpr

Checkpoint – Restart

17

- Command Line:
  - checkpointing processes

    e.g. cpr –c ~/Jan8_backup –p 1234:HID –k
  - querying checkpoint statefiles

    cpr –i ~/Jan8_backup ...
  - restarting processes

    cpr [–j] –r ~/Jan8_backup ...
  - deleting statefiles

    cpr –D ~/Jan8_backup ...

- API (POSIX + SGI Extensions)
  - *ckpt_setup(), ckpt_create(), ckpt_restart(),*

    ckpt_stat(), ckpt_remove()

- GUI
  - *cview [–display XwindowDisplay]*
  - cview checkpoint and restart panels

- CPR Event Handler
  - *atcheckpoint()*
  - *atrestart()*