

Mejorando el rendimiento de tu Linux

Jesús Espino García

Grupo de Usuarios de Linux
Universidad Carlos III de Madrid.



13 de Marzo de 2009

Índice

1 Introducción

- ¿Qué es?
- ¿Por qué?
- ¿Cuándo?
- ¿Dónde?
- ¿Cómo?

2 Entendiendo Linux

3 Tunning

4 Para terminar.

¿Qué es?

- Adaptar el equipo a nuestras necesidades.
- Equilibrar la capacidad y la demanda.
- Dimensionar correctamente nuestro hardware.

Índice

1 Introducción

- ¿Qué es?
- ¿Por qué?
- ¿Cuándo?
- ¿Dónde?
- ¿Cómo?

2 Entendiendo Linux

3 Tunning

4 Para terminar.

¿Por qué?

- La configuración genérica no cubre las necesidades.
- El rendimiento es una pieza clave del sistema.
- No todos los recursos son igual de importantes.

Índice

1 Introducción

- ¿Qué es?
- ¿Por qué?
- **¿Cuándo?**
- ¿Dónde?
- ¿Cómo?

2 Entendiendo Linux

3 Tunning

4 Para terminar.

¿Cuándo?

- Desde la planificación del proyecto.
- En cada una de las etapas del proyecto.
- Después de un backup!!!

Índice

1 Introducción

- ¿Qué es?
- ¿Por qué?
- ¿Cuándo?
- **¿Dónde?**
- ¿Cómo?

2 Entendiendo Linux

3 Tunning

4 Para terminar.

¿Dónde?

- En el hardware.
- En el software.
- En el entorno.

Índice

1 Introducción

- ¿Qué es?
- ¿Por qué?
- ¿Cuándo?
- ¿Dónde?
- ¿Cómo?

2 Entendiendo Linux

3 Tunning

4 Para terminar.

¿Cómo?

- Fuertes conocimientos sobre el sistema.
- Monitorización.
- Benchmarks.
- Prueba y Error.
- Mas arte que ciencia.
- Cambiando:
 - Parámetros del kernel.
 - Parámetros del sistema de ficheros.
 - Parámetros de los drivers.
 - Ampliando hardware.
 - Mejorando el entorno.
 - Modificando el software (kernel incluido).

Índice

1 Introducción

2 Entendiendo Linux

- Gestión de procesos
- Gestión de memoria
- Gestión de IO
- Métricas de rendimiento

3 Tunning

4 Para terminar.

¿Qué es un proceso?

- Estructura de datos en el SO.
- Proceso completo o hilo.
- Instancia de tarea a realizar por la CPU.
- Utiliza, reserva y comparte recursos.

Prioridades

- Los procesos tienen 2 prioridades.
- Estática (Asignada por el tipo de tarea y el nice)
- Dinámica (Calculada por el scheduler)
- El nivel nice determina el tamaño del quantum del proceso.

Context Switch

- Para simular multitarea se va alternando la ejecución de procesos.
- Cada cambio de contexto tiene unos costes de tiempo.

Interrupciones

- El hardware y el software produce interrupciones.
- Las interrupciones, normalmente producen cambios de contexto.
- Las interrupciones hardware, por lo general, se envían a todos los procesadores.

Scheduler

- Es el encargado de seleccionar un proceso u otro.
- Antes del 2.6.23 el scheduler era el $O(1)$.
- En el 2.6.23 y posteriores el scheduler es CFS.

Índice

1 Introducción

2 Entendiendo Linux

- Gestión de procesos
- **Gestión de memoria**
- Gestión de IO
- Métricas de rendimiento

3 Tunning

4 Para terminar.

Memoria principal

- Segmentos: Linux apenas los utiliza.
- Paginas: Lo que utiliza normalmente Linux.
- En sistemas de 32bits se debe dividir la memoria en 2 partes:
 - Una accesible directamente (hasta los 896M).
 - Y el resto mapeada a través de la zona accesible directamente (hasta 64G usando PAE).

Sistema de caches

- Los sistemas modernos tienen caches de memoria.
- Las caches son memorias mas rápidas que evitan el acceso a la memoria principal
- Usan ciertas políticas y algoritmos para determinar de que se guarda cache.

Swap

- Memoria basada en dispositivos de disco.
- Almacena paginas de memoria a las que no se acceden en mucho tiempo.
- En caso de falta de falta de memoria principal se utiliza la swap como tal.

Asignación de Memoria

- Al solicitar memoria, si hay paginas libres las asigna.
- Si no hay paginas libres, utiliza métodos de reclamado de pagina.
- Para esto utiliza el Buddy System, que se encarga de:
 - Mantener una lista de las paginas libres.
 - Solicita la liberación paginas a kswapd.
 - Evitar la fragmentación de la memoria.
- kswapd libera paginas basándose en ciertas políticas.

Índice

1 Introducción

2 Entendiendo Linux

- Gestión de procesos
- Gestión de memoria
- Gestión de IO
- Métricas de rendimiento

3 Tunning

4 Para terminar.

Sistemas de ficheros

- VFS: Capa de abstracción para el acceso a sistemas de ficheros.
- Journaling: Log de las operaciones realizadas.
- Sistema de Ficheros: Estructura de datos para almacenar ficheros (normalmente en disco).

Acceso a disco (IO)

- Los sistemas de ficheros acceden a disco a través del sistema IO.
- El sistema IO organiza los acceso a discos con el IO Scheduler.
- Y este finalmente escribe los datos a través del driver de disco.

IO Schedulers

- Se pueden escoger entre 4:
 - Anticipatory: Intenta guardar datos para escribir en tramas largas y continuas. Además usa deadline.
 - CFQ (Completely Fair Queuing): Implementa un reparto equitativo del IO entre procesos.
 - Deadline: Round robin con deadline.
 - NOOP: No hace ninguna operación de orden. Es un FIFO.

Índice

1 Introducción

2 Entendiendo Linux

- Gestión de procesos
- Gestión de memoria
- Gestión de IO
- Métricas de rendimiento

3 Tunning

4 Para terminar.

Métricas de procesador

- Uso de CPU.
- User time: Tiempo utilizado por los procesos de usuario (no kernel).
- System time: Tiempo utilizado por operaciones del kernel.
- Waiting: Tiempo esperando al subsistema IO.
- Idle time: Tiempo esperando tareas.
- Nice time: Tiempo utilizado para restablecer prioridades de los procesos.
- Load average: Carga del equipo (no es un porcentaje).
- Runnable processes: Numero de procesos listos para correr.
- Blocked: Procesos esperando IO.
- Context Switch: Cantidad de cambios de contexto.
- Interrupts: Cantidad de interrupciones

Métricas de memoria

- Memoria libre.
- Uso de Swap.
- Buffer y cache: Caches asignadas al sistema de ficheros o a dispositivos de bloque.
- Slabs: Uso de memoria del kernel (no swapeable).
- Memoria activa/inactiva: La memoria inactiva es la que es candidata de ser swapeada.

Métricas de red

- Paquetes recibidos y enviados.
- Bytes recibidos y enviados.
- Colisiones por segundo.
- Paquetes "tirados".
- Overruns: Veces que la tarjeta de red se ha quedado sin espacio en el buffer.
- Errors: Numero de tramas marcadas como fallidas.

Métricas de IO

- iowait: Tiempo de CPU gastado esperando al subsistema IO.
- Media de longitud de la cola: Media de peticiones IO en cola.
- Media de espera: Tiempo medio de respuesta de una petición IO.
- Transferencias por segundo: Numero de operaciones IO por segundo.
- Bloques por segundo r/w
- KB por segundo r/w

Índice

- 1 Introducción
- 2 Entendiendo Linux
- 3 Tunning**
 - Herramientas
 - En la CPU
 - En la memoria
 - En el disco
 - En la red
- 4 Para terminar.

Monitorizacion

- top
- vmstat
- uptime
- ps
- free
- iostat
- mpstat
- numastat
- pmap
- netstat
- wireshark
- strace
- /proc

Benchmarks

- lmbench
- iozone
- netperf

Índice

1 Introducción

2 Entendiendo Linux

3 Tunning

- Herramientas
- **En la CPU**
- En la memoria
- En el disco
- En la red

4 Para terminar.

En la CPU

- Herramientas: uptime, top, vmstat, lmbench...
- Determinar prioridades: Baja latencia, alta cantidad de proceso.
- Determinar las prioridades de los procesos.
- Buscar una disminución del context-switch por interrupción.
- Asignar interrupciones a una CPU en concreto (por ejemplo las de una tarjeta red).
- Podemos ampliar CPU.

Índice

1 Introducción

2 Entendiendo Linux

3 Tunning

- Herramientas
- En la CPU
- **En la memoria**
- En el disco
- En la red

4 Para terminar.

En la memoria

- Herramientas: top, vmstat, free, pmap, iostat, lmbench...
- Determinar prioridades: Aprovechamiento de la memoria, menor carga de proceso.
- Si hay mucho uso de swap, normalmente indica falta de memoria.
- Si hay muchas peticiones de paginas nos podemos plantear ampliar el tamaño de las mismas.
- Si tenemos poca memoria y muchos procesos muy pequeños, podemos disminuir el tamaño de pagina.
- Podemos analizar procesos concretos para ver en que consumen la memoria.
- Podemos asegurarnos de usar programas con bibliotecas dinámicas.
- Podemos ampliar memoria.

Índice

1 Introducción

2 Entendiendo Linux

3 Tunning

- Herramientas
- En la CPU
- En la memoria
- **En el disco**
- En la red

4 Para terminar.

En la disco

- Herramientas: vmstat, iostat, iotop, lmbench...
- Determinar prioridades: Aprovechamiento, menor carga de proceso, mas ancho de banda.
- Deberíamos determinar el sistema de fichero que nos interesa.
- Si usamos solo ficheros muy grandes, podríamos ampliar el tamaño de bloque del sistema de ficheros.
- Si usamos muchos ficheros pequeños, podríamos disminuir el tamaño de bloque del sistema de ficheros.
- Escoger el IO Scheduler que nos pueda interesar.
- Añadir mas RAM. Esto amplía la memoria de cache para disco.
- Establecer parámetros del sistema de ficheros (por ejemplo, noatime).
- Establecer las políticas de journaling correctas.
- Uso de sistemas RAID o volúmenes lógicos.

Índice

1 Introducción

2 Entendiendo Linux

3 Tunning

- Herramientas
- En la CPU
- En la memoria
- En el disco
- En la red

4 Para terminar.

En la red

- Herramientas: netperf, lmbench, netstat ...
- Determinar prioridades: Ancho de banda, baja latencia.
- Ampliar el tamaño del frame (jumbo frames, no soportados por todo el hardware!).
- Asignar interrupciones a un procesador en concreto.
- Usar mejores tarjetas de red (para mas ancho de banda o para mas baja latencia).
- Hacer bonding o balanceo de carga.
- El firewall ralentiza las comunicaciones.

Referencias

- <http://www.redbooks.ibm.com/redpapers/pdfs/redp4285.pdf>: RedPaper de IBM.

Dudas

...

Fin

