

Métodos de Evaluación del Rendimiento

- Enfoque sistemático de evaluación del rendimiento.
- Metodología de planificación de capacidad.

Errores comunes en la Evaluación del Rendimiento

- Ausencia de objetivos
 - Elección de métricas, carga de trabajo y metodología dirigida por los objetivos
- Metas sesgadas (*Biased Goals*)
 - Elección de métricas y cargas de trabajo que dan resultados sesgados
- Enfoque no sistemático
- Análisis sin Comprensión del problema
 - "Un problema bien definido ya está mitad resuelto".
- Métricas de rendimiento incorrectas
- Carga de trabajo no representativa
- Elección incorrecta de la técnica de evaluación

Errores comunes en la Evaluación del Rendimiento (cont.)

- Pasar por alto parámetros importantes.
 - Establecer una lista completa de las características del sistema y de la carga de trabajo que afectan al rendimiento del sistema.
- Ignorar factores significativos.
 - Factor: parámetros que varían en el estudio
- Diseño de experimentos inadecuados.
- Nivel de detalle inadecuado.
- Ausencia de análisis .
- Análisis erróneos.
 - Ej: utilización solamente de la media en tiempo de respuesta.
- Análisis poco sensibles.
 - Análisis de sensibilidad: cómo afecta la carga o los parámetros a los resultados.
- Análisis demasiado complejos.

Errores comunes en la Evaluación del Rendimiento (cont.)

- Ignorar los errores de entrada.
 - En la utilización de variables para estimar parámetros hay que ajustar el nivel de confianza en la estimación.
 - Distribución de los errores de entrada.
- Tratamiento inadecuado de los valores extremos (*outliers*).
- Presunción de estabilidad y ausencia de cambios en el futuro
- Ignorar la variabilidad
- Presentación inadecuada de los resultados.

- Ignorar aspectos sociales.
- Omisión de las suposiciones y limitaciones asumidas-

Método de evaluación de rendimiento

1. Establecer los objetivos y definir el sistema (sus límites).
2. Lista de los servicios y consecuencias.
3. Selección de las métricas.
4. Lista de parámetros del sistema y de la carga de trabajo.
5. Selección de los factores (y valores/niveles) a estudiar.
6. Selección de la técnica de evaluación.
7. Selección de la carga de trabajo.
8. Diseño del experimento.
9. Analizar e interpretar los resultados.
10. Presentar los resultados. Empezar otra vez si es necesario.

[Jain 91, pp 26]

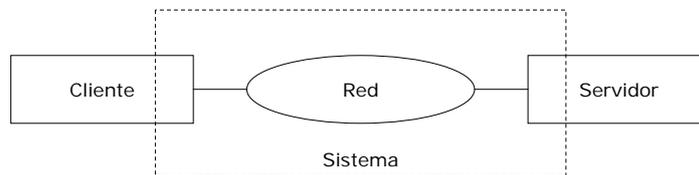
Método de Evaluación de Rendimiento Caso de estudio

- El problema consiste en la comparación, en un sistema distribuido con una arquitectura cliente-servidor, de dos mecanismos de invocación remota: remote pipes y rpc (remote procedure call)
 - RPC
 - El programa que invoca se bloquea
 - El control pasa al procedimiento invocado junto con los parámetros
 - Cuando el procedimiento se completa, los resultados y el control se devuelve al programa que invoca
 - Remote pipes (objetos de tipo procedimiento)
 - Cuando se invocan el invocador no se bloquea
 - La ejecución del pipe sucede de forma concurrente a la ejecución del invocador
 - Si se producen resultados se devuelven de forma asincrónica.

Caso de estudio

Establecer los objetivos y definir el sistema

- Objetivo: Comparar las prestaciones de aplicaciones que utilizan *rpipes* (*remote pipes*) con aquellos parecidos que utilizan *rpc* (*remote procedure calls*)
- Componente clave: Canal (procedimiento o *pipe*)
- Sistema: Dos ordenadores conectados vía red
- Peticiones se envían a través del canal desde el cliente al servidor
- El estudio se desarrollará con la menor interferencia posible de los elementos que no entran en el estudio



Caso de estudio

Lista de los servicios y consecuencias

- Dos tipos de llamada de canal: *rpc*, *rpipe*
- Los recursos utilizados por las llamadas de canal dependen de los parámetros que se pasan y la acción requerida sobre esos parámetros
- Elegimos la transferencia de datos como aplicación y las llamadas se clasifican en pequeñas y grandes dependiendo de la cantidad de datos a transferir a la máquina remota.
- Servicios:
 - Transferencia de pocos datos
 - Transferencia de muchos datos

Caso de estudio

Selección de las métricas

- El estudio se limita a operaciones correctas. No se estudian las limitaciones de recursos, los errores y los fallos.
- Se compararán
 - Velocidad a la que se realiza el servicio
 - Tiempo consumido en el servicio
 - Recursos consumidos: cliente, servidor, enlace
- Métricas
 - Tiempo transcurrido por invocación: *Elapsed time for call*
 - Tasa máxima de invocaciones por unidad de tiempo (aprox. el tiempo necesario para completar un bloque de n invocaciones sucesivas).
 - Tiempo de CPU local por invocación
 - Tiempo de CPU remota por invocación
 - Número de bytes enviados en el enlace por llamada

Caso de estudio

Lista de parámetros

- Sistema : parámetros que afectan al rendimiento de una determinada aplicación y tamaño de los datos. Parámetros software y hardware
 - Velocidad de la CPU local
 - Velocidad de la CPU remota
 - Velocidad de la red
 - Overhead del sistema operativo para la conexión con los canales
 - Overhead del sistema operativo para la conexión con la red
 - Número de retransmisiones necesarias. (confiabilidad de la red)
- Carga de trabajo: característicos de las peticiones de los usuarios.
 - Tiempo entre llamadas sucesivas
 - Número y tamaño de los parámetros de llamada
 - Número y tamaño de los resultados
 - Tipo de canal
 - Otras cargas en las CPU's local y remota
 - Cargas adicionales en la red

Caso de estudio

Selección de los factores a estudiar

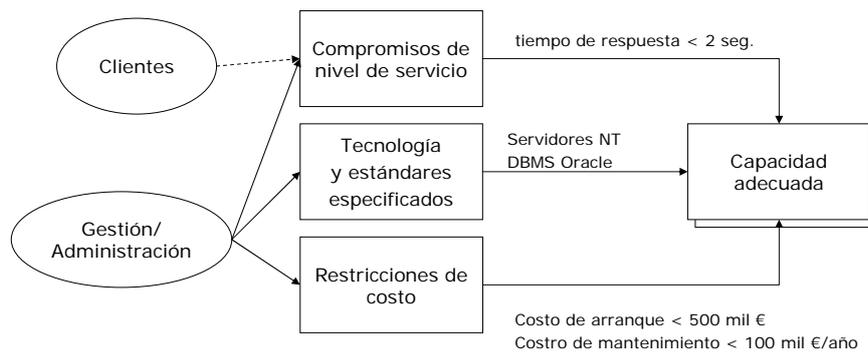
- Los parámetros que pueden variar en la evaluación se denominan *factores* y sus valores se denominan *niveles (levels)*. El propósito es comprender cómo afecta el cambio de los niveles en el rendimiento del servicio.
 - Tipo de canal (2):
 - RPC
 - rpipes
 - Velocidad de la red (2):
 - Distancia corta: Campus
 - Distancia larga: Ciudad
 - Tamaño de los parámetros de llamada a transferir (2):
 - Pequeño
 - Grande
 - Número de llamadas consecutivas (11):
 - 1, 2, 4, 8, 16, 32, 64, 128, 256, 512 y 1024
- Otras consideraciones
 - Parámetros fijos: Tipo de CPU, Sistema Operativo
 - Errores de transmisión ignorados
 - Experimentos realizados con poca carga en la red y en los hosts.

Caso de estudio

- Selección de la técnica de evaluación.
 - Se construyeron prototipos y se utilizaron las medidas para evaluar. Se utilizó un modelo analítico para justificar la consistencia de los valores medidos.
- Selección de la carga de trabajo.
 - Programas sintéticos que generaban los tipos de peticiones de canal en análisis. Este programa también monitorizaba los recursos consumidos y guardaba los resultados.
- Diseño del experimento.
 - Diseño factorial completo. $2 \times 2 \times 2 \times 11 = 88$ experimentos
- Analizar e interpretar los resultados.
 - Análisis de la varianza para cuantificar los efectos de los tres primeros factores.
 - Regresión para cuantificar los efectos de n llamadas sucesivas.
- Presentar los resultados.
 - Gráfico de función (variable: tamaño de bloque n)

Metodología de planificación de capacidad [Menascé & Almeida 2002]

- Decimos que un servicio Web tiene una capacidad adecuada si los compromisos de nivel de servicio se satisfacen de forma continua para una tecnología y estándares especificados, y si los servicios son proporcionados dentro de unas determinadas restricciones de costes

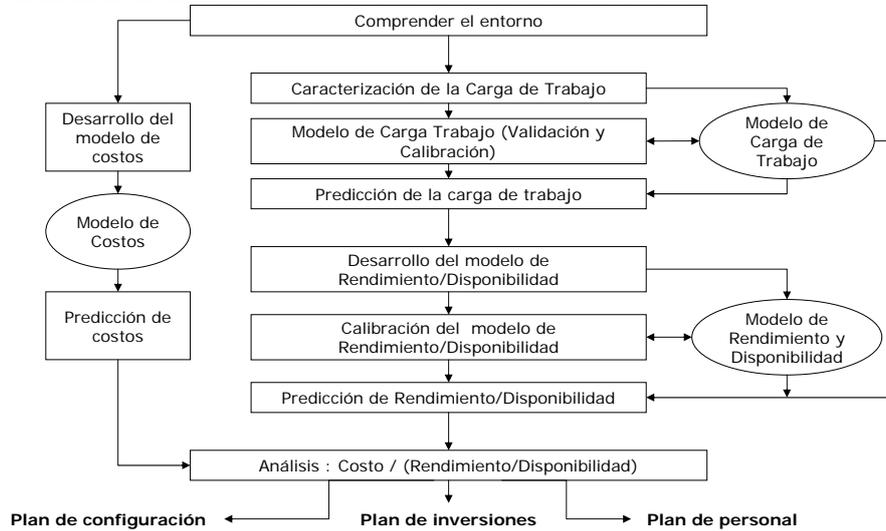


[Menascé & Almeida 2002, pp 177]

Metodología de planificación de capacidad

- Compromisos/acuerdos de nivel de servicio (SLA)
 - Los valores deseables o aceptables para las métricas de rendimiento (tiempo de respuesta, throughput, disponibilidad, etc.)
 - Centrado en métricas que el usuario pueda comprender
 - Conjunto de objetivos de medida fácil
 - Los valores del SLA son específicos de cada organización y determinados por el gestor.
 - Los SLA tienen que definir los tiempos de operación del sitio, de la red y los tiempos de respuesta.
 - Tº de respuesta del servidor para peticiones de búsqueda < 2 seg.
 - Tº de respuesta de descarga página para modem 56kbps < 8 seg
 - Sitio web disponible el 99,99%

Metodología de planificación de capacidad



[Menascé & Almeida 2002, pp 179]

M.A.V.S. oct-10

Dpto. Informática – ETSII – U. Valladolid

15

Comprensión del entorno

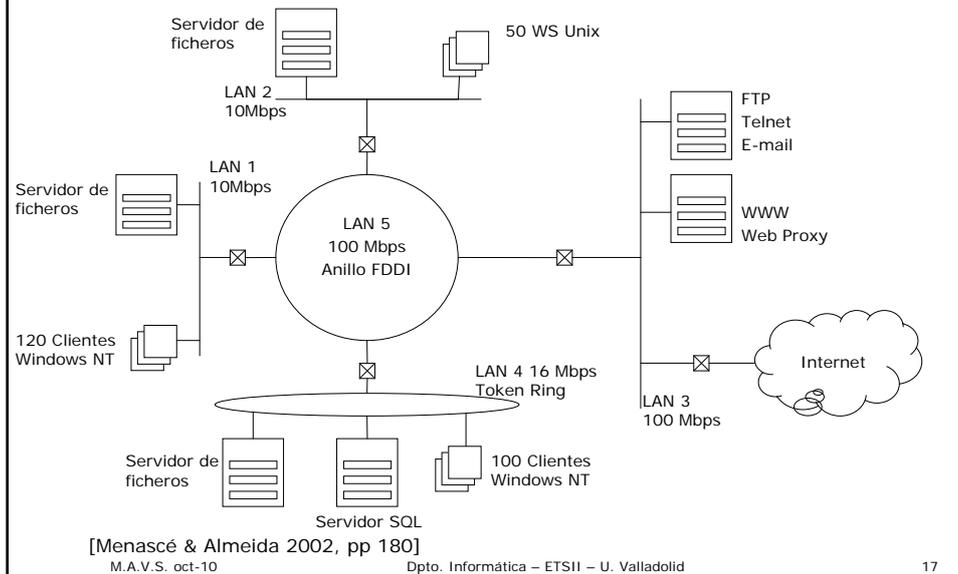
- Conocer el tipo de hardware (clientes y servidores)
- Conocer el tipo de software (sistemas operativos, *middleware*, aplicaciones)
- Conectividad de redes de comunicaciones
- Protocolos de comunicación
- Identificación de los periodos de picos de utilización
- Estructuras de gestión
- SLA's
- Técnica a utilizar: Recogida de la información.

M.A.V.S. oct-10

Dpto. Informática – ETSII – U. Valladolid

16

Ejemplo



Ejemplo (cont)

- Clientes y servidores (gráfico anterior)
- Aplicaciones
 - LAN2 para I+D; simulación experimentos de biología
 - Almacenan los datos en el servidor de ficheros de LAN2
 - Los servicios de ftp, web, email, telnet del LAN3
 - Clientes de LAN1 y LAN4: ofimática y aplicaciones de gestión C/S transaccionales contra el servidor SQL de Oracle.
- Patrones de utilización.
 - Las aplicaciones C/S pico entre las 10:00 – 12:00 y 14:00 – 16:00
 - Las aplicaciones de formación sobre web 16:00 – 17:00 horas
 - Servicios web distribuidos sin picos entre las 9:00 y las 21:00 horas

Ejemplo (cont)

- SLA
 - Tiempo medio de respuesta de las transacciones C/S contra el servidor SQL
 - Sencillas < 2 s
 - Complejas < 5 s para el 95 % de las transacciones
 - El servidor de correo ha de procesar al menos 60000 mensajes en las 8 horas de trabajo y cada mensaje ha de ser procesado en menos de 30 s.
 - Considerar en la carga de correo la carga adicional causada por SPAM*
 - Web. Acceso a la página principal, páginas de descripción de productos y las búsquedas de información de productos
 - Usuarios externos, modem 56.6 kbps < 4s.
 - Usuarios internos, intranet < 1.5 s
 - Disponibilidad de servicios web 99.9%
- Soporte de administración y soporte
 - 4 administradores del sistema
- Políticas de adquisición
 - Comité de usuarios

Ejemplo Estadísticas Spam 2006 (Don Evett)

- Fuentes: Google, Brightmail, Jupiter Research, eMarketer, Gartner, MailShell, Harris Interactive, and Ferris Research.

Email considered Spam	40% of all email
Daily Spam emails sent	12.4 billion
Daily Spam received per person	6
Annual Spam received per person	2,200
Spam cost to all non-corp Internet users	\$255 million
Spam cost to all U.S. Corporations in 2002	\$8.9 billion
States with Anti-Spam Laws	26
Email address changes due to Spam	16%
Estimated Spam increase by 2007	63%
Annual Spam in 1,000 employee company	2.1 million
Users who reply to Spam email	28%
Users who purchased from Spam email	8%
Corporate email that is considered Spam	15-20%
Wasted corporate time per Spam email	4-5 seconds

- <http://spam-filter-review.toptenreviews.com/spam-statistics.html>
- US-Billon: 10⁹

Ejemplo Estadísticas Spam 2009 (Cisco)

- Origen del spam en 2009
 1. Brazil: 7.7;
 2. USA: 6.6;
 3. India: 3.6;
 4. South Korea: 3.1;
 5. Turkey: 2.6;
 6. Vietnam: 2.5;
 7. China: 2.4;
 8. Poland: 2.4;
 9. Russia: 2.3;
 10. Argentina: 1.5.
- Costo para los negocios a nivel mundial \$130 billion
- Costo para los negocios en U.S.A. \$42 billion
- US-Billon: 10⁹

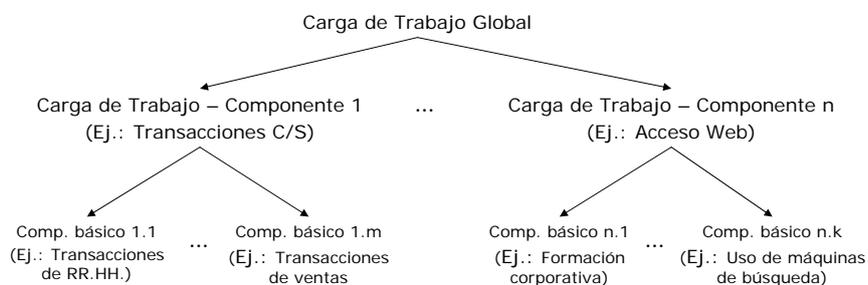
M.A.V.S. oct-10

Dpto. Informática – ETSII – U. Valladolid

21

Caracterización de la carga

- Es el proceso de describir de forma precisa la carga de trabajo global del sistema en términos de sus componentes principales.
- Los componentes básicos son caracterizados por los parámetros de intensidad y demanda de servicio en cada recurso del sistema.



[Menascé & Almeida 2002, pp 184]

M.A.V.S. oct-10

Dpto. Informática – ETSII – U. Valladolid

22

Ejemplo de Parámetros de componentes básicos y tipo

- Transacciones de ventas
 - Núm. de transacciones enviadas por cliente WI
 - Núm. de clientes WI
 - **Núm. total** de E/S a la BD de ventas SD
 - Utilización CPU del servidor de BD SD
 - Media de mensajes enviados/recibidos por el servidor de BD. SD
- Formación basada en web
 - Núm. medio de sesiones/día de entrenamiento WI
 - Tamaño medio de ficheros de vídeo por sesión SD
 - Tamaño medio de documentos HTML descargados SD
 - Tamaño medio de ficheros de imagen descargados/sesión SD
 - Tamaño medio de documentos descargados/sesión SD
 - Utilización media del CPU del servidor Web SD
- Procesamiento de correo electrónico
 - Núm. de mensajes recibidos por cliente/día WI
 - Núm. de mensajes enviados por cliente/día WI
 - Núm. de clientes WI
 - Tamaño medio de mensajes SD
 - Utilización CPU del demonio de correo electrónico SD

WI: Workload intensity (Intensidad de la Carga de Trabajo)
SD: Service Demand (Demanda de servicio)

Recogida de información

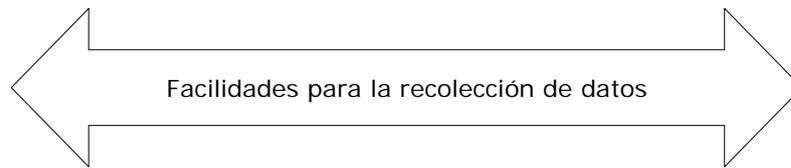
- Determinación de los valores de los parámetros para cada componente básico
 - Medidas directas: sistemas de seguimiento (*Accounting systems*), monitores de rendimiento
 - Ej: Monitorizar el envío de mensajes durante 1 hora.
 - Tamaño del mensaje
 - Tiempo de CPU del servidor de correo
 - Tiempo de E/S en el servidor para cada mensaje
 - Componente básico: formación soportada por web
 - Recogida de datos de varios días
 - 11.2 videos cada minuto con 3 minutos de duración media
 - Procesando estos datos se obtiene un consumo del 10% del ancho de banda de LAN3
 - Benchmarks
 - Demanda de servicio de la CPU con el SPEC CINT2000

Alternativas de recogida de datos para la caracterización de la carga

Utilización exclusiva de benchmarks, prácticas de industria, ROTs (Rules of thumb)

Utilización de benchmarks, prácticas de industria, ROTs y mediciones

Utilización exclusiva de mediciones



Ninguna

Alguna

Detallada

M.A.V.S. oct-10

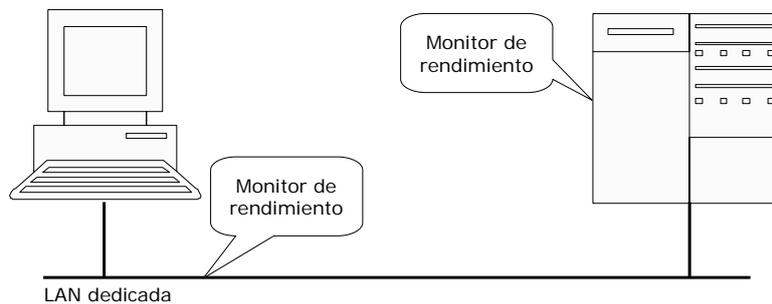
Dpto. Informática – ETSII – U. Valladolid

25

Entorno controlado para el benchmarking de componentes de carga de trabajo

Cliente ejecutando un script para la aplicación A

Servidor dedicado



M.A.V.S. oct-10

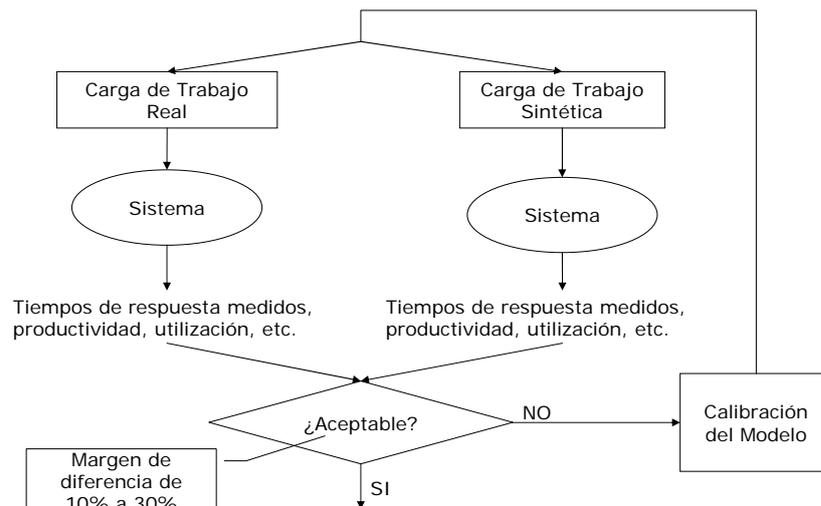
Dpto. Informática – ETSII – U. Valladolid

26

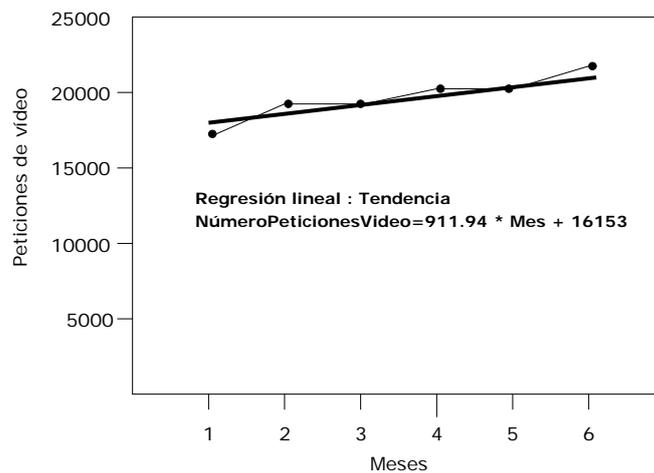
Ejemplo

- La demanda de servicio de CPU en el servidor para una aplicación es de 10ms.
 - Obtenida en un entorno controlado con un servidor con calificación SPEC CINT2000 de 431.
- ¿Cuál será la demanda si el servidor fuese más rápido, con una calificación SPEC CINT2000 de 518?
 - La demanda será $10 / (518/431) = 8.3$
- La selección del benchmark que debe utilizarse para escalar las medidas estará sujeta al tipo de la aplicación. CINT2000, CFP2000, etc.
- Formula general de la demanda de servicio.
 $DmdServ = DmdServMedida \times RelaciónProductividad$
 $DmdServ = 10 * (431/518) = 8.3$

Validación del modelo de carga



Proyección de la carga de trabajo



M.A.V.S. oct-10

Dpto. Informática – ETSII – U. Valladolid

29

Modelos y predicción del rendimiento

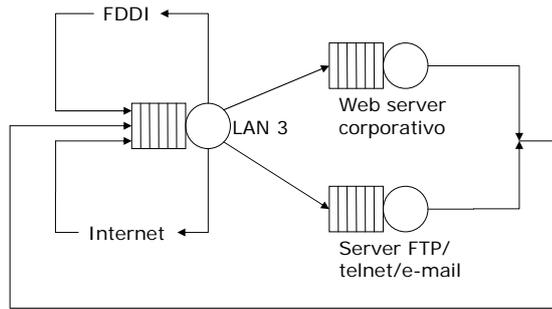
- Es el proceso de estimar las medidas de rendimiento de un sistema para un conjunto dado de parámetros.
- La predicción del rendimiento necesita la utilización de modelos.
 - Modelos de simulación
 - Modelos analíticos
- Ej: Modelo de colas
 - Entrada: Descripción del sistema y de la carga de Trabajo
 - Salida: Métricas de rendimiento:
 - Tiempo de respuesta
 - Throughput
 - Utilización, etc.

M.A.V.S. oct-10

Dpto. Informática – ETSII – U. Valladolid

30

Ejemplo: LAN 3 – Alto nivel

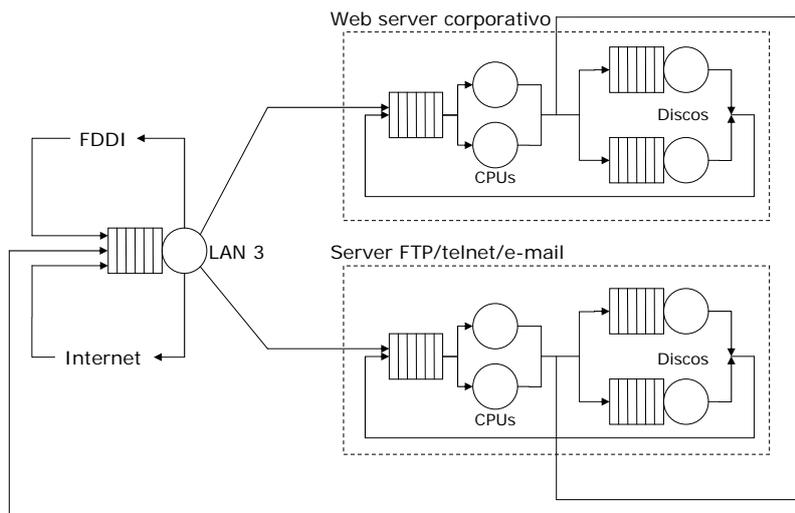


M.A.V.S. oct-10

Dpto. Informática – ETSII – U. Valladolid

31

Ejemplo: LAN 3 – Detallada



M.A.V.S. oct-10

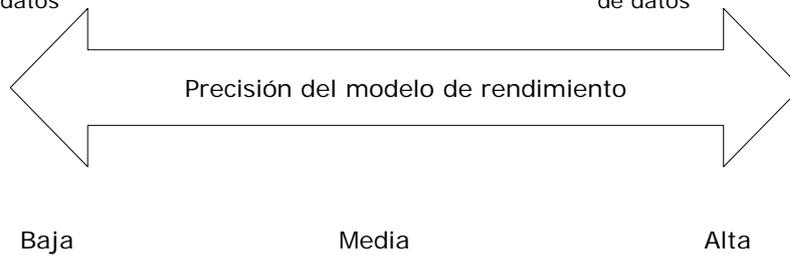
Dpto. Informática – ETSII – U. Valladolid

32

Precisión de los modelos de rendimiento

Modelos de granularidad gruesa:
Esfuerzo pequeño
para la recolección
de datos

Modelos de granularidad fina:
Esfuerzo grande
para la recolección
de datos

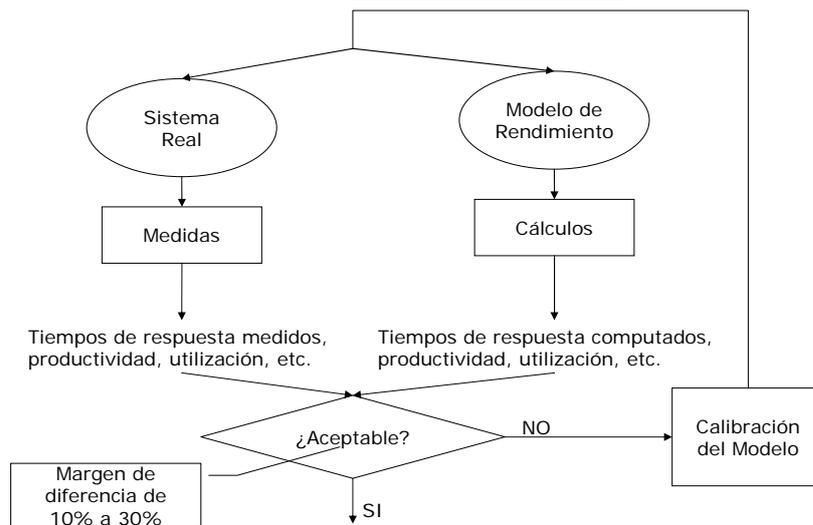


M.A.V.S. oct-10

Dpto. Informática – ETSII – U. Valladolid

33

Validación de los modelos de rendimiento



M.A.V.S. oct-10

Dpto. Informática – ETSII – U. Valladolid

34

Categorías del modelo de costes

- Costes Hardware:
 - Máquinas clientes y servidores, discos .
 - Routers, pasarelas, cableado.
 - Mantenimiento, etc.
- Costes Software:
 - Sistemas operativos, middleware.
 - DBMS, software de procesamiento de correo.
 - Automatización de oficinas, aplicaciones, etc.
- Costes Telecomunicaciones:
 - Servicios WAN, ISP, etc.
- Costes de Soporte:
 - Salarios de los administradores del sistema.
 - Soporte en línea.
 - Personal de red, etc.