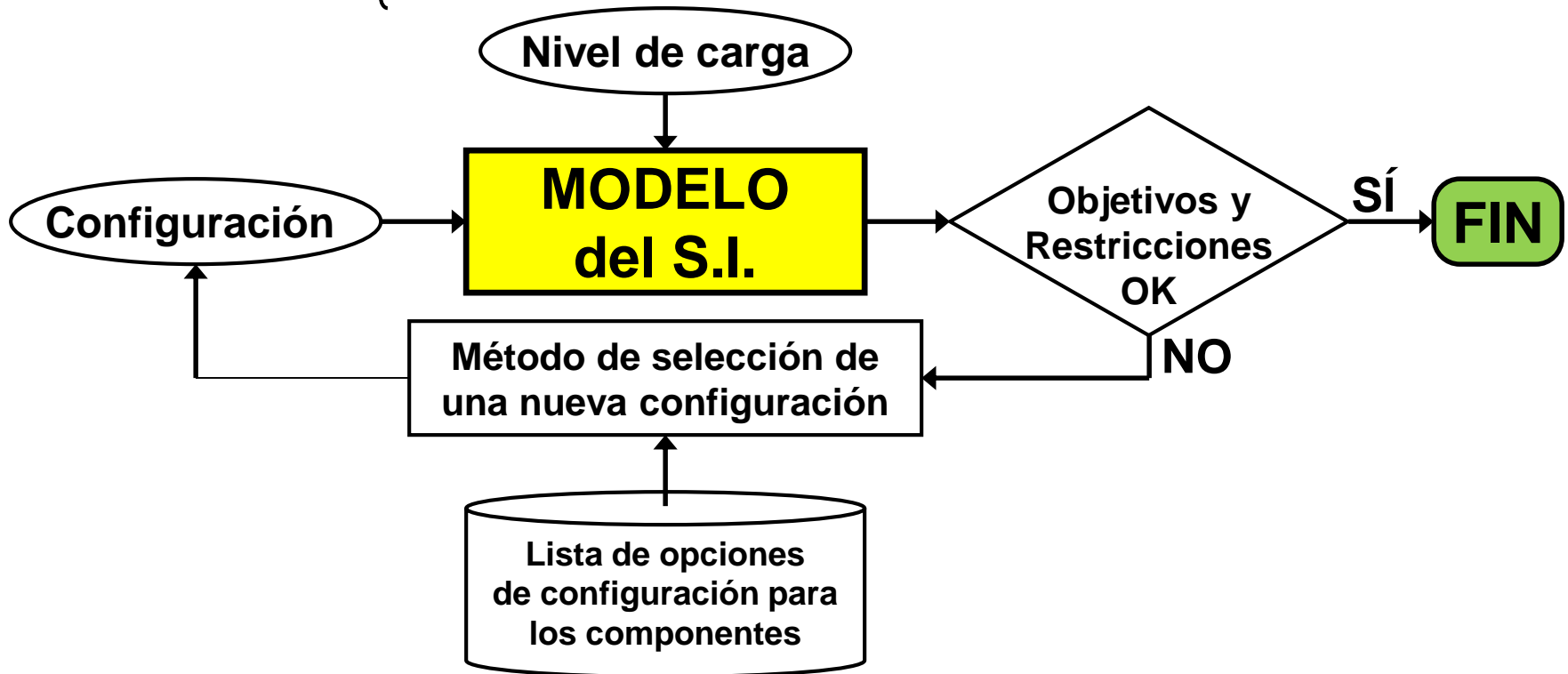


INTRODUCCIÓN A LA CONFIGURACIÓN DE SI

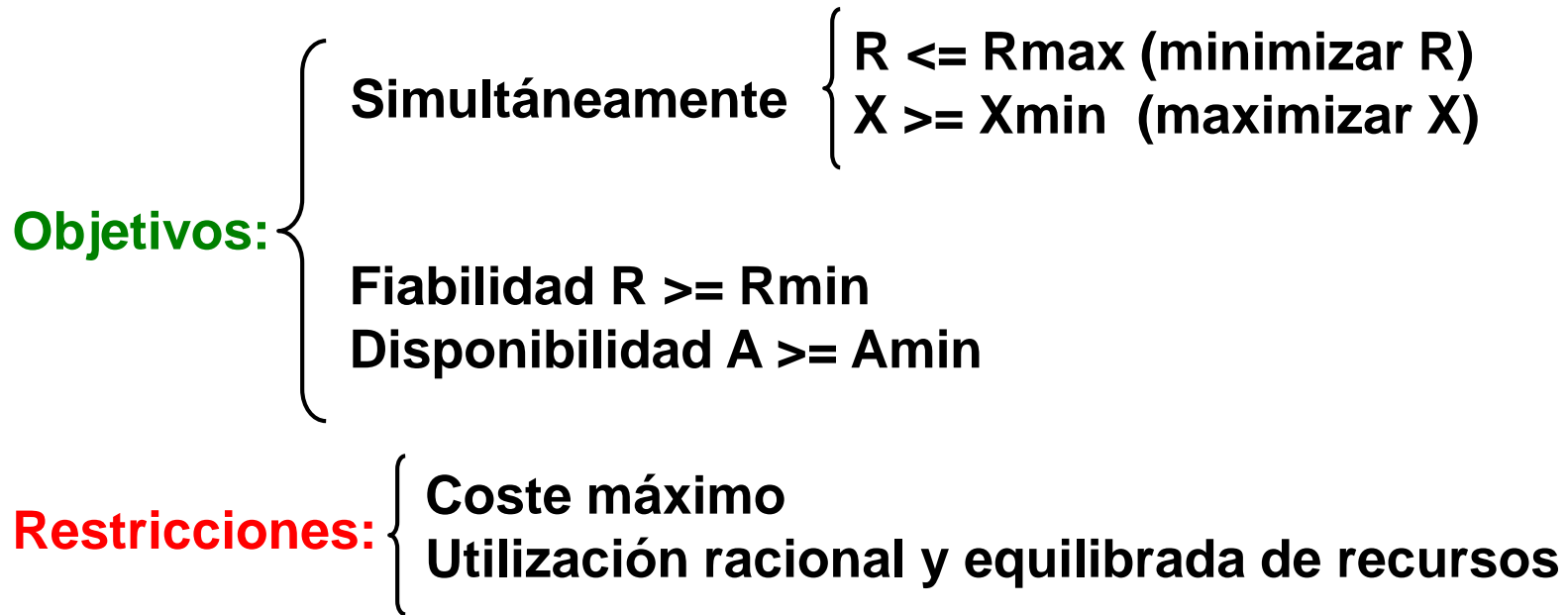
La configuración (diseño) de un SI es un prob. de optimización
Es necesario “buscar” la configuración óptima

Óptimo → { Máximas prestaciones, fiabilidad, disponibilidad, ...
Mínimo coste económico



FASE 1 DEL PROCESO DE CONFIGURACIÓN

Establecer Objetivos y Restricciones



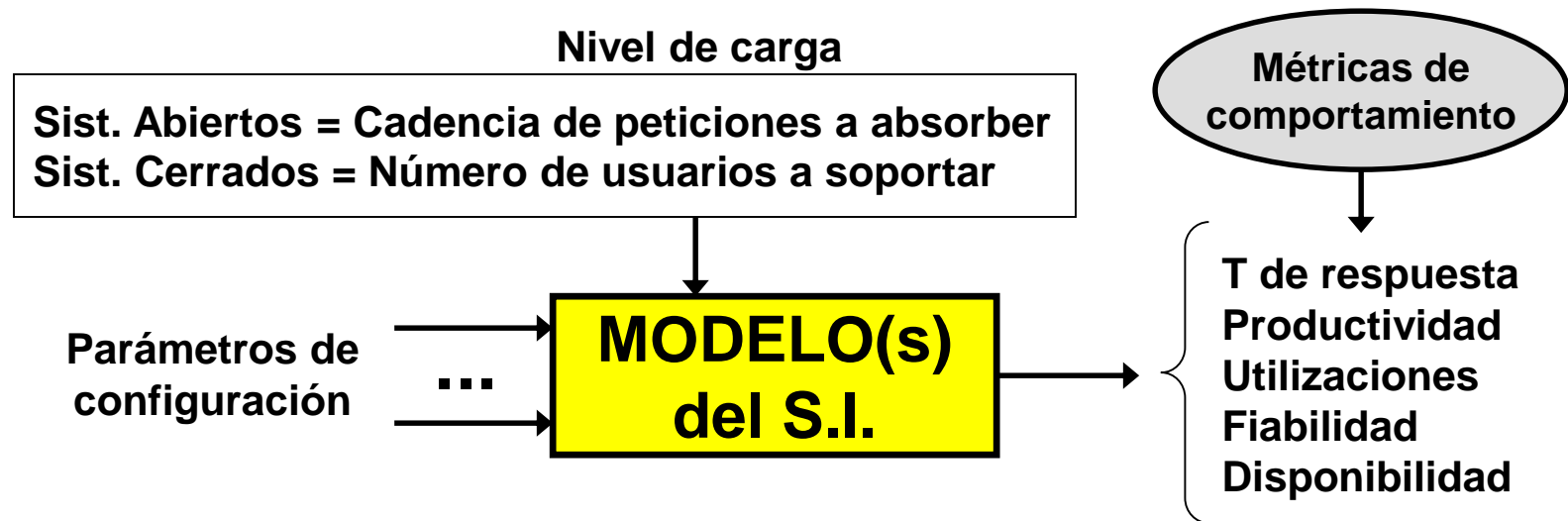
SIN un presupuesto máximo se pueden cambiar los objetivos

Maximizar la relación	Prestaciones / Coste
Minimizar la relación	Coste / Prestaciones

FASE 2 DEL PROCESO DE CONFIGURACIÓN

Predicción del comportamiento del SI

Es necesario disponer de **modelos que permitan predecir** (aunque de forma aproximada) el impacto de las cambios de componentes y/o la arquitectura del sistema en las métricas de comportamiento



Hay que disponer de un “**modelo de costes**” (incluye el mantenimiento)
En su versión más simple es una lista de precios aproximados

FASE 3 DEL PROCESO DE CONFIGURACIÓN

Método de selección de una nueva configuración

Hay que usar un método para ir seleccionando nuevos valores de los parámetros de configuración hasta cumplir los objetivos

Un proceso de configuración con múltiples objetivos es COMPLEJO
Conviene dividir el proceso en 2 fases separadas ...

PASO 1 de configuración

Dimensionar la **capacidad** que debe tener el sistema
PARA soportar a los usuarios requeridos con la calidad de servicio requerida

PASO 2 de configuración

Añadir la **redundancia** precisa
PARA dar el servicio con la fiabilidad y disponibilidad requeridas

Generalmente, al añadir redundancia al SI se modifica la capacidad de servicio del sistema y por tanto sus métricas de comportamiento
PERO en una configuración aproximada pueden ignorarse los cambios

CONFIGURACIÓN: DIMENSIONADO

Dimensionado de la capacidad de servicio de SI

Hay que usar un método para ir seleccionando nuevos valores de los parámetros de configuración hasta cumplir los objetivos

Realizar un experimento de medición + modelado con una **configuración BASE** del sistema = un PC

Reconfiguración NECESARIA SI

$$\left\{ \begin{array}{ll} R > R_{\max} & \longrightarrow (R / R_{\max}) \times 100 \\ \text{y/o} & \\ X < X_{\min} & \longrightarrow (X_{\min} / X) \times 100 \end{array} \right. \begin{array}{l} \swarrow \\ \searrow \end{array} \begin{array}{l} \text{Indican la magnitud de} \\ \text{la Reconfiguración del} \\ \text{sistema base} \end{array}$$

Para realizar la reconfiguración hay que responder a estas preguntas:

- 1) ¿Sobre qué componentes hay que actuar?
- 2) ¿Cómo hay que actuar?
- 3) ¿Con qué magnitud hay que actuar?

CONFIGURACIÓN: DIMENSIONADO

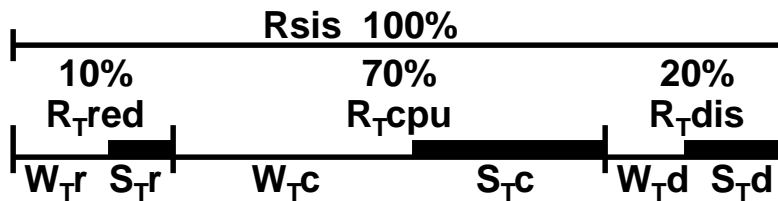
Selección del componente en que hay que actuar

Observar cómo contribuye cada componente del sistema (o modelo) a los valores obtenidos de las métricas de prestaciones del sistema (al tiempo de respuesta y la productividad observadas)

Rsis	Xsis
$R_{Tred} = W_{Tred} + S_{Tred}$	Ured
$R_{Tcpu} = W_{Tcpu} + S_{Tcpu}$	Ucpu
$R_{Tdis} = W_{Tdis} + S_{Tdis}$	Udis

$$Xsis = (1/Di) \times Ui$$

Ui es la proporción de la máxima cadencia de peticiones que puede proporcionar el componente que es usada para obtener una productividad del sistema Xsis



- PARA reducir Rsis actuar sobre el componente(s) que es responsable de un % elevado del tiempo de respuesta
- PARA aumentar Xsis actuar sobre el componente(s) que presenta una U = 1 pues está limitando la productividad

CONFIGURACIÓN: DIMENSIONADO

Decisión de la forma de actuar

Técnicamente hay 2 opciones → { Mejorar el componente
Replicar el componente

Para reducir R

Depende de la descomposición del tiempo agregado de respuesta del componente

Si el % del tiempo de espera W es elevado, indica que el componente está saturado
Ambas opciones son validas: mejorar el componente o replicarlo

Si el % del tiempo de espera W es reducido, indica que el componente NO está saturado
Tan solo un componente más potente puede reducir el T de respuesta ($\approx T$ de servicio)

Para aumentar X

Ambas opciones son válidas, pues cualquiera de ellas aumenta la productividad

Restricciones tecnológicas

Réplica CPU: Sencillo si el software servidor es multihilo y el S.O. dispone de SMP

Réplica Disco: Muy complicado sin técnicas de stripping o RAID

Réplica Red: Se dispone de tarjetas con 1-4 canales ethernet en paralelo

Cuando ya se están usando los mejores componentes la única opción es la replicación

CONFIGURACIÓN: DIMENSIONADO

Estimación de la magnitud de la actuación

Los componentes presentan unas **opciones** de configuración **discretas**

Ej CPU: PIII 750Mhz 128Kbc / PIII 750Mhz 256Kbc / ... / PIV 1.1Ghz 204Kbc / ...

Ej Disco: IDE 7500 rpm ... / SCSI 9000rpm / ... / SCSI 15000rpm / ...

El proceso de dimensionado parte de la configuración base medida y modelada

Para terminar en una configuración “óptima” usando varias configuraciones intermedias

Resumir el proceso en forma tabular

Config.	Sistema básico	CPU	Disco	Red	Prestaciones / Costes
BASE	Placa MonoP B €	1 P-III 500 128c C €	1 IDE ATA-100 D €	1 Eth100 R €	Rb Xb Ucb Udb Urb T €
Conf 1	...				
C Opt					

Hay 2 posibilidades para estimar la potencia del nuevo componente(s)

Seleccionar sucesivamente el siguiente en potencia al último utilizado hasta cumplir requisitos
 Estimar el incremento de potencia necesaria en el componente para cumplir los requisitos

CONFIGURACIÓN: DIMENSIONADO

Integración de un nuevo componente en el modelo

Para integrar en el modelo un componente nuevo en vez del componente base hay que estimar su tiempo de servicio a partir del tiempo del componente base

Generalmente, la potencia o capacidad de un componente se expresa mediante un índice de productividad máxima IP

Ej: CPU en MIPS (SpecCPUint) o MFLOPS (SpecCPUfp) o una combinación de ambos

Ej: Discos en Mbytes/seg

Ej: Red en Mbps (ancho de banda)

Se precisan los valores del mismo índice o benchmark de productividad (IP) para el componente base (IPb) y para el componente nuevo (IPn)

Relación entre las potencias de los componentes $\longrightarrow \left\{ \frac{IPn}{IPb} = \frac{\mu n}{\mu b} = \frac{1 / S_n}{1 / S_b} \right.$

Tiempo de servicio del
componente nuevo $\longrightarrow S_n = \frac{IPb}{IPn} S_b$

Ejemplo

Elección de un nuevo procesador

El tiempo de servicio determinado para el procesamiento de una tarea en el procesador base es de 0,2 seg.

El nuevo procesador ha de tener una capacidad de procesamiento de al menos 14,2 tareas/seg.

¿Qué procesador es necesario?

¿Cómo se refleja en el modelo?

Si se tiene en cuenta el coste y la posibilidad de utilizar sistemas biprocesadores, ¿Cambia la situación?

$$\frac{IP_n}{IP_b} = \frac{\mu_n}{\mu_b} = \frac{1 / S_n}{1 / S_b}$$

Modelo	Índice	Coste
Base	4	100 €
A1	6	120 €
B2	8	140 €
C3	10	300 €
D4	12	600 €
E5	14	1500 €

CONFIGURACIÓN: DIMENSIONADO

Integración de un nuevo componente en el modelo

Particularización para servidores múltiples (CPU)

Actualmente los procesadores están integrados por varios cores, son intrínsecamente servidores múltiples

Por otra parte, su índice de prestaciones viene expresado de forma global, es decir, para todo el procesador, NO por núcleo

En este caso las fórmulas anteriores cambian:

Relación entre las potencias de los componentes $\longrightarrow \left\{ \frac{IP_n}{IP_b} = \frac{\mu_n}{\mu_b} = \frac{N_n / S_n}{N_b / S_b} \right.$

N_b = número de núcleos o cores del procesador base

N_n = número de núcleos o cores del procesador nuevo a incorporar

Tiempo de servicio del componente nuevo $\longrightarrow S_n = \frac{N_n \times IP_b}{N_b \times IP_n} S_b$

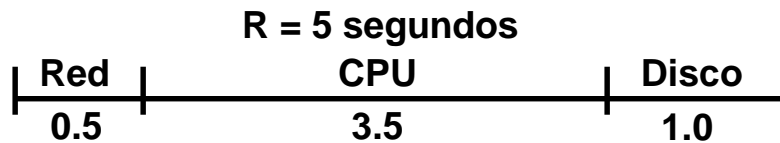
CONFIGURACIÓN: DIMENSIONADO

Estimación de potencia necesaria en un componente

PARA REDUCIR EL TIEMPO DE RESPUESTA

Sistemas Abiertos

1) Repartir la reducción total de tiempo necesaria entre los componentes



Objetivo Rmax = 2 segundos Decisiones ...

Reducción en Red NINGUNA

Reducción en CPU y Disco POR IGUAL

4.5 seg deben reducirse a 1.5 seg / Factor = 0.33 / CPU = 1.16 = 0.33x3.5 y DISCO = 0.33 = 0.33x1.0

2) Estimar el nuevo μ o s para los componentes

Se puede usar la formula de la cola M/M/1 para estimar el valor de $\mu_n = \lambda + 1/R$

Otros tipos de colas presentan una formulación demasiado compleja para ser práctica

Observar que la fórmula supone un flujo de tareas λ fijo

No es cierto es sistemas cerrados y puede ser aproximado en sistemas abiertos

3) Seleccionar el nuevo componente a usar

Seleccionar el componente “más barato” que supere el valor IPn calculado

Si el IPn obtenido es muy alto se pueden usar dos componentes en paralelo

$$IP_n = \frac{\mu_n}{\mu_b} IP_b \quad \text{Tras seleccionar el nuevo componente obtener su tiempo de servicio para integrarlo en el modelo} \quad S_n = \frac{IP_b}{IP_n} S_b$$

CONFIGURACIÓN: DIMENSIONADO

Estimación de potencia necesaria en un componente

PARA AUMENTAR LA PRODUCTIVIDAD en Sistemas ABIERTOS (NO interactivos)

1) Estimar el nuevo μ o s para los componentes

Sistemas Abiertos

Con la configuración base ...
Se obtiene una productividad X del sistema
con una utilización U_i de cada componente

$$X \cdot D_i = U_i$$

Con la configuración nueva ...
Se desea una productividad λ del sistema
con una utilización $U_i < 0.9$ en los componentes

$$\lambda \cdot D_i = U_i < 0.9$$

Entonces todos los componentes deben verificar la siguiente ecuación:

$$\mu_i \geq V_i \lambda / 0.9$$

2) Evaluar si el tiempo de respuesta es aceptable

Se puede usar la formula de la cola M/M/1 para estimar el valor de $R_i = \frac{1 / \mu_i}{1 - \lambda / \mu_i}$
Y la ley general del tiempo de respuesta para estimar el valor de R_{sis}

3) Seleccionar el nuevo componente a usar

Proceder como se ha visto anteriormente

CONFIGURACIÓN: DIMENSIONADO

Estimación de potencia necesaria en un componente

PARA AUMENTAR LA PRODUCTIVIDAD en Sistemas CERRADOS (interactivos)

1) Estimar el nuevo μ o s para los componentes

- En los sistemas interactivos NO hay un objetivo concreto de productividad
- El objetivo es soportar N usuarios con un $R \leq R_{max}$

Usar la ley del tiempo de respuesta interactivo para estimar la productividad a superar para soportar N usuarios con $R \leq R_{max}$

$$R = (N / X) - Z \leq R_{max}$$

Todos los componentes deben superar esta productividad

$$X \geq N / (R_{max} + Z)$$

A partir de aquí proceder como en los sistemas abiertos: $\mu_i \geq V_i X / 0.9$

2) Evaluar si el tiempo de respuesta es aceptable

En principio NO sería necesario por basarnos en la ley del tpo de respuesta interactivo

3) Seleccionar el nuevo componente a usar

Proceder como se ha visto anteriormente

CONFIGURACIÓN: COSTES

Integración de los costes en la configuración

Se pueden presentar dos situaciones:

HAY un presupuesto concreto que no se debe superar

NO HAY un presupuesto concreto

Situación: CON PRESUPUESTO

Hay 2 formas típicas de proceder

- 1) Realizar todo el proceso de diseño (la sucesión de configuraciones) y sólo AL FINAL comprobar si se supera el presupuesto
- 2) En cada paso de configuración analizar el coste de la mejora

Inconveniente: Requiere más esfuerzo de evaluación

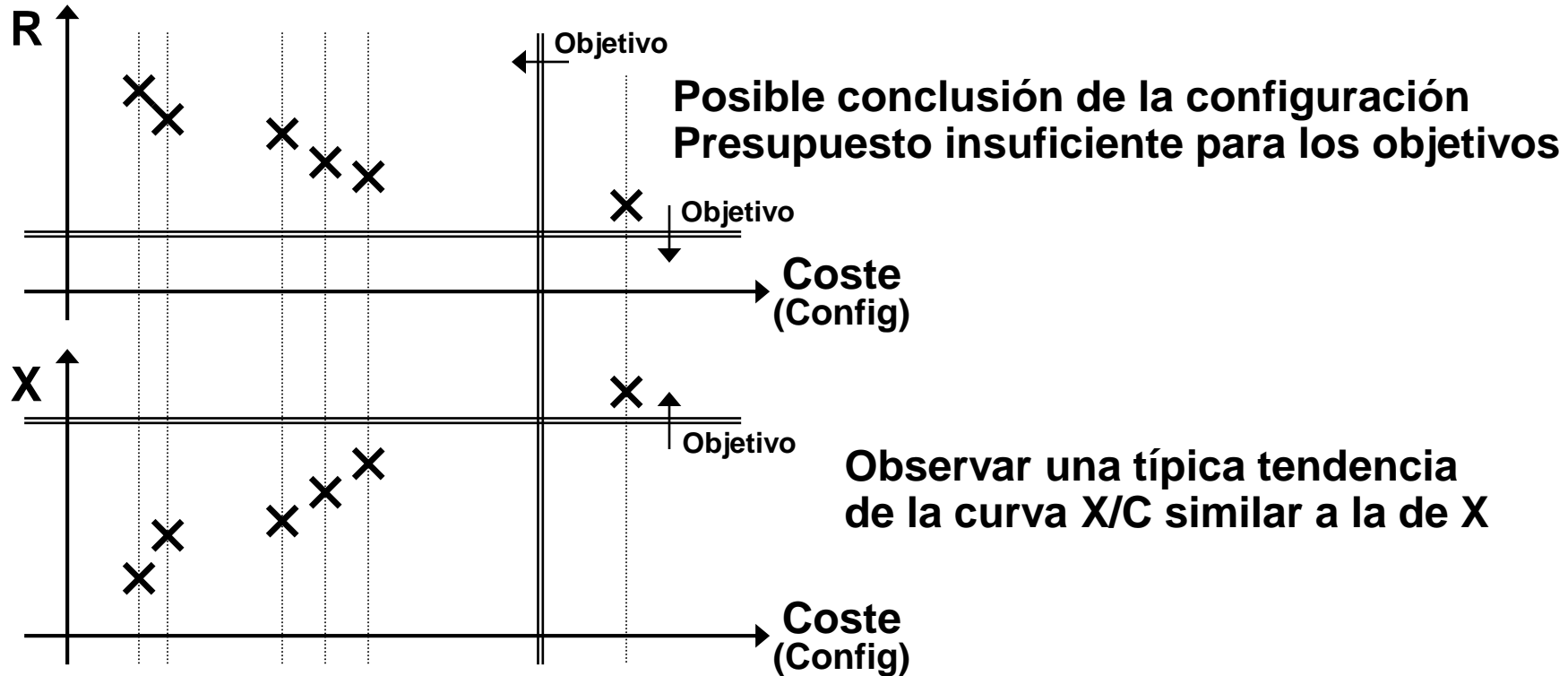
Ventajas: Evita realizar un proceso largo de configuración que puede ser inútil por falta de presupuesto

Permite ir observando (dibujando) el coste de las prestaciones que se van alcanzando durante el proceso de configuración

Importante ver los saltos y estabilizaciones de las curvas
(Al cambiar de tecnología o al escalar una misma tecnología)

CONFIGURACIÓN: COSTES

Integración de los costes en la configuración



Situación: SIN PRESUPUESTO

Ir incrementando progresivamente la capacidad del sistema
MIENTRAS QUE el incremento mantenga la relación prestaciones/coste y
HASTA QUE se obtenga la X deseada o se soporten los N usuarios