

# **ANALISIS DE SISTEMAS I**

**AÑO 2017**

## **ANALISIS ESTRUCTURADO DE SISTEMAS**

### **Carpeta de trabajos prácticos**

#### **Guía de orden de la carpeta a presentar**

- ⇒ Descripción de la Empresa
- ⇒ Organigrama
  - Real
  - Funcional
- ⇒ Relevamiento
  - entrevistas
  - cuestionario
  - observación personal
- ⇒ Objetivos
- ⇒ Límites y Alcances
- ⇒ Cronograma
- ⇒ Informe Inicial
- ⇒ D.F.D.
  - Actual
  - Propuesto
  - Explosiones
- ⇒ Cursograma
- ⇒ Diccionario de Datos
  - Elementos datos
  - Flujos de datos
  - Procesos
  - Almacenamiento de datos
- ⇒ Arboles y Tablas de decisión
- ⇒ D.I.A.D.
- ⇒ Factibilidades
- ⇒ Informe final
- ⇒ Anexos

# **ANALISIS DE SISTEMAS**

## **Teoría General de los Sistemas**

### **INTRODUCCION**

La teoría de la organización y la práctica administrativa han experimentado cambios sustanciales en años recientes. La información proporcionada por las ciencias de la administración y la conducta ha enriquecido a la teoría tradicional. Estos esfuerzos de investigación y de conceptualización a veces han llevado a descubrimientos divergentes. Sin embargo, surgió un enfoque que puede servir como base para lograrla convergencia, el enfoque de sistemas, que facilita la unificación de muchos campos del conocimiento. Dicho enfoque ha sido usado por las ciencias físicas, biológicas y sociales, como marco de referencia para la integración de la teoría organizacional moderna.

El primer expositor de la Teoría General de los Sistemas fue Ludwing von Bertalanffy, en el intento de lograr una metodología integradora para el tratamiento de problemas científicos.

La meta de la Teoría General de los Sistemas no es buscar analogías entre las ciencias, sino tratar de evitar la superficialidad científica que ha estancado a las ciencias. Para ello emplea como instrumento, modelos utilizables y transferibles entre varios continentes científicos, toda vez que dicha extrapolación sea posible e integrable a las respectivas disciplinas.

La Teoría General de los Sistemas se basa en dos pilares básicos: aportes semánticos y aportes metodológicos.

### **APORTES SEMANTICOS**

Las sucesivas especializaciones de las ciencias obligan a la creación de nuevas palabras, llegando a formar casi un verdadero lenguaje que sólo es manejado por los especialistas.

De esta forma surgen problemas al tratarse de proyectos interdisciplinarios, ya que los participantes del proyecto son especialistas de diferentes ramas de la ciencia y cada uno de ellos maneja una semántica diferente a los demás.

La Teoría de los Sistemas, para solucionar estos inconvenientes, pretende introducir una semántica científica de utilización universal.

### **Teoría General de los sistemas**

#### **Enfoque sistémico**

#### **Definición de Sistema**

Un conjunto de objetos reunidos, con relaciones entre dichos objetos y entre sus atributos, conectados o relacionados entre si y con su ambiente de tal modo que forman una suma total o totalidad. Von Bertalanffy. 1940.

Cabe aclarar que los objetos o partes que componen al sistema, no se refieren al campo físico (objetos), sino más bien al funcional. De este modo las cosas o partes pasan a ser funciones básicas realizadas por el sistema. Podemos enumerarlas en: entradas, procesos y salidas.

## **Sistema de Información (S I)**

Un conjunto formal de procesos que, operando sobre una colección de datos estructurada según las necesidades de la empresa, recopilan, elaboran y distribuyen la información (o parte de ella) necesaria para las operaciones de dicha empresa y para las actividades de dirección y control correspondientes (decisiones) para desempeñar su actividad de acuerdo a su estrategia de negocio. R. Andreu. 1991.

### **Entradas:**

Las entradas son los ingresos del sistema que pueden ser recursos materiales, recursos humanos o información.

Las entradas constituyen la fuerza de arranque que suministra al sistema sus necesidades operativas.

Las entradas pueden ser:

- **en serie:** es el resultado o la salida de un sistema anterior con el cual el sistema en estudio está relacionado en forma directa.
- **aleatoria:** es decir, al azar, donde el término "azar" se utiliza en el sentido estadístico. Las entradas aleatorias representan entradas potenciales para un sistema.
- **retroacción:** es la reintroducción de una parte de las salidas del sistema en sí mismo.

### **Proceso:**

El proceso es lo que transforma una entrada en salida, como tal puede ser una máquina, un individuo, una computadora, un producto químico, una tarea realizada por un miembro de la organización, etc.

En la transformación de entradas en salidas debemos saber siempre como se efectúa esa transformación. Con frecuencia el procesador puede ser diseñado por el administrador. En tal caso, este proceso se denomina "caja blanca". No obstante, en la mayor parte de las situaciones no se conoce en sus detalles el proceso mediante el cual las entradas se transforman en salidas, porque esta transformación es demasiado compleja. Diferentes combinaciones de entradas o su combinación en diferentes órdenes de secuencia pueden originar diferentes situaciones de salida. En tal caso la función de proceso se denomina una "caja negra".

### **Caja Negra:**

La caja negra se utiliza para representar a los sistemas cuando no sabemos que elementos o cosas componen al sistema o proceso, pero sabemos que a determinadas entradas corresponden determinadas salidas y con ello poder inducir, presumiendo que a determinados estímulos, las variables funcionaran en cierto sentido.

### **Salidas:**

Las salidas de los sistemas son los resultados que se obtienen de procesar las entradas. Al igual que las entradas estas pueden adoptar la forma de productos, servicios e información. Las mismas son el resultado del funcionamiento del sistema o, alternativamente, el propósito para el cual existe el sistema.

Las salidas de un sistema se convierten en entrada de otro, que la procesará para convertirla en otra salida, repitiéndose este ciclo indefinidamente.

### **Relaciones:**

Las relaciones son los enlaces que vinculan entre sí a los objetos o subsistemas que componen a un sistema complejo.

Podemos clasificarlas en:

- **Simbiótica:** es aquella en que los sistemas conectados no pueden seguir funcionando solos. A su vez puede subdividirse en unipolar o parasitaria, que es cuando un sistema (parásito) no puede vivir sin el otro sistema (planta); y bipolar o mutual, que es cuando ambos sistemas dependen entre si.
- **Sinérgica:** es una relación que no es necesaria para el funcionamiento pero que resulta útil, ya que su desempeño mejora sustancialmente al desempeño del sistema. Sinergia significa "acción combinada". Sin embargo, para la teoría de los sistemas el término significa algo más que el esfuerzo cooperativo. En las relaciones sinérgicas la acción cooperativa de subsistemas semi-independientes, tomados en forma conjunta, origina un producto total mayor que la suma de sus productos tomados de una manera independiente.
- **Superflua:** Son las que repiten otras relaciones. La razón de las relaciones superfluas es la confiabilidad. Las relaciones superfluas aumentan la probabilidad de que un sistema funcione todo el tiempo y no una parte del mismo. Estas relaciones tienen un problema que es su costo, que se suma al costo del sistema que sin ellas puede funcionar.

### **Atributos:**

Los atributos de los sistemas, definen al sistema tal como lo conocemos u observamos. Los atributos pueden ser definidores o concomitantes: los atributos definidores son aquellos sin los cuales una entidad no sería designada o definida tal como se lo hace; los atributos concomitantes en cambio son aquellos que cuya presencia o ausencia no establece ninguna diferencia con respecto al uso del término que describe la unidad.

### **Contexto:**

Un sistema siempre estará relacionado con el contexto que lo rodea, o sea, el conjunto de objetos exteriores al sistema, pero que influyen decididamente a éste, y a su vez el sistema influye, aunque en una menor proporción, influye sobre el contexto; se trata de una relación mutua de contexto-sistema.

Tanto en la Teoría de los Sistemas como en el método científico, existe un concepto que es común a ambos: el foco de atención, el elemento que se aísla para estudiar.

El contexto a analizar depende fundamentalmente del foco de atención que se fije. Ese foco de atención, en términos de sistemas, se llama límite de interés.

Para determinar este límite se considerarían dos etapas por separado:

a) La determinación del contexto de interés.

b) La determinación del alcance del límite de interés entre el contexto y el sistema.

c) Se suele representar como un círculo que encierra al sistema, y que deja afuera del límite de interés a la parte del contexto que no interesa al analista.

d) En lo que hace a las relaciones entre el contexto y los sistemas y viceversa. Es posible que sólo interesen algunas de estas relaciones, con lo que habrá un límite de interés relacional.

Determinar el límite de interés es fundamental para marcar el foco de análisis, puesto que sólo será considerado lo que quede dentro de ese límite.

Entre el sistema y el contexto, determinado con un límite de interés, existen infinitas relaciones. Generalmente no se toman todas, sino aquellas que interesan al análisis, o aquellas que probabilísticamente presentan las mejores características de predicción científica.

### **Rango:**

En el universo existen distintas estructuras de sistemas y es factible ejercitar en ellas un proceso de definición de rango relativo. Esto produciría una jerarquización de las distintas estructuras en función de su grado de complejidad.

Cada rango o jerarquía marca con claridad una dimensión que actúa como un indicador claro de las diferencias que existen entre los subsistemas respectivos.

Esta concepción denota que un sistema de nivel 1 es diferente de otro de nivel 8 y que, en consecuencia, no pueden aplicarse los mismos modelos, ni métodos análogos a riesgo de cometer evidentes falacias metodológicas y científicas.

Para aplicar el concepto de rango, el foco de atención debe utilizarse en forma alternativa: se considera el contexto y a su nivel de rango o se considera al sistema y su nivel de rango.

Refiriéndonos a los rangos hay que establecer los distintos subsistemas. Cada sistema puede ser fraccionado en partes sobre la base de un elemento común o en función de un método lógico de detección.

El concepto de rango indica la jerarquía de los respectivos subsistemas entre sí y su nivel de relación con el sistema mayor.

### **Subsistemas:**

En la misma definición de sistema, se hace referencia a los subsistemas que lo componen, cuando se indica que el mismo está formado por partes o cosas que forman el todo.

Estos conjuntos o partes pueden ser a su vez sistemas (en este caso serían subsistemas del sistema de definición), ya que conforman un todo en sí mismos y estos serían de un rango inferior al del sistema que componen.

Estos subsistemas forman o componen un sistema de un rango mayor, el cual para los primeros se denomina **macrosistema**.

### **Variables:**

Cada sistema y subsistema contiene un proceso interno que se desarrolla sobre la base de la acción, interacción y reacción de distintos elementos que deben necesariamente conocerse.

Dado que dicho proceso es dinámico, suele denominarse como variable, a cada elemento que compone o existe dentro de los sistemas y subsistemas.

Pero no todo es tan fácil como parece a simple vista ya que no todas las variables tienen el mismo comportamiento sino que, por lo contrario, según el proceso y las características del mismo, asumen comportamientos diferentes dentro del mismo proceso de acuerdo al momento y las circunstancias que las rodean.

### **Parámetro:**

Uno de los comportamientos que puede tener una variable es el de parámetro, que es cuando una variable no tiene cambios ante alguna circunstancia específica, no quiere decir que la variable es estática ni mucho menos, ya que sólo permanece inactiva o estática frente a una situación determinada.

### **Operadores:**

Otro comportamiento es el de operador, que son las variables que activan a las demás y logran influir decisivamente en el proceso para que este se ponga en marcha. Se puede decir que estas variables actúan como líderes de las restantes y por consiguiente son privilegiadas respecto a las demás variables. Cabe aquí una aclaración: las restantes variables no solamente son influidas por los operadores, sino que también son influenciadas por el resto de las variables y estas tienen también influencia sobre los operadores.

### **Retroalimentación:**

La retroalimentación se produce cuando las salidas del sistema o la influencia de las salidas del sistema en el contexto, vuelven a ingresar al sistema como recursos o información.

La retroalimentación permite el control de un sistema y que el mismo tome medidas de corrección en base a la información retroalimentada.

### **Feed-forward o alimentación delantera:**

Es una forma de control de los sistemas, donde dicho control se realiza a la entrada del sistema, de tal manera que el mismo no tenga entradas corruptas o malas, de esta forma al no haber entradas malas en el sistema, las fallas no serán consecuencia de las entradas sino de los procesos mismos que componen al sistema.

### **Homeostasis y entropía:**

La homeostasis es la propiedad de un sistema que define su nivel de respuesta y de adaptación al contexto.

Es el nivel de adaptación permanente del sistema o su tendencia a la supervivencia dinámica. Los sistemas altamente homeostáticos sufren transformaciones estructurales en igual medida que el contexto sufre transformaciones, ambos actúan como condicionantes del nivel de evolución.

La entropía de un sistema es el desgaste que el sistema presenta por el transcurso del tiempo o por el funcionamiento del mismo. Los sistemas altamente entrópicos tienden a desaparecer por el desgaste generado por su proceso sistémico. Los mismos deben tener rigurosos sistemas de control y mecanismos de revisión, reelaboración y cambio permanente, para evitar su desaparición a través del tiempo.

En un sistema cerrado la entropía siempre debe ser positiva. Sin embargo, en los sistemas abiertos biológicos o sociales, la entropía puede ser reducida o mejor aun transformarse en entropía negativa, es decir, un proceso de organización más completo y de capacidad para transformar los recursos. Esto es posible porque en los sistemas abiertos los recursos utilizados para reducir el proceso de entropía se toman del medio externo. Asimismo, los sistemas vivientes se mantienen en un estado estable y pueden evitar el incremento de la entropía y aun desarrollarse hacia estados de orden y de organización creciente.

### **Permeabilidad:**

La permeabilidad de un sistema mide la interacción que este recibe del medio, se dice que a mayor o menor permeabilidad del sistema el mismo será más o menos abierto.

Los sistemas que tienen mucha relación con el medio en el cuál se desarrollan son sistemas altamente permeables, estos y los de permeabilidad media son los llamados sistemas abiertos.

Por el contrario los sistemas de permeabilidad casi nula se denominan sistemas cerrados.

### **Integración e independencia:**

Se denomina sistema integrado a aquel en el cual su nivel de coherencia interna hace que un cambio producido en cualquiera de sus subsistemas produzca cambios en los demás subsistemas y hasta en el sistema mismo.

Un sistema es independiente cuando un cambio que se produce en él, no afecta a otros sistemas.

### **Centralización y descentralización:**

Un sistema se dice centralizado cuando tiene un núcleo que comanda a todos los demás, y estos dependen para su activación del primero, ya que por sí solos no son capaces de generar ningún proceso.

Por el contrario los sistemas descentralizados son aquellos donde el núcleo de comando y decisión está formado por varios subsistemas. En dicho caso el sistema no es tan dependiente, sino que puede llegar a contar con subsistemas que actúan de reserva y que sólo se ponen en funcionamiento cuando falla el sistema que debería actuar en dicho caso.

Los sistemas centralizados se controlan más fácilmente que los descentralizados, son más sumisos, requieren menos recursos, pero son más lentos en su adaptación al contexto. Por el contrario los sistemas descentralizados tienen una mayor velocidad de respuesta al medio ambiente pero requieren mayor cantidad de recursos y métodos de coordinación y de control más elaborados y complejos.

### **Adaptabilidad:**

Es la propiedad que tiene un sistema de aprender y modificar un proceso, un estado o una característica de acuerdo a las modificaciones que sufre el contexto. Esto se logra a través de un mecanismo de adaptación que permita responder a los cambios internos y externos a través del tiempo.

Para que un sistema pueda ser adaptable debe tener un fluido intercambio con el medio en el que se desarrolla.

**Mantenibilidad:**

Es la propiedad que tiene un sistema de mantenerse constantemente en funcionamiento. Para ello utiliza un mecanismo de mantenimiento que asegure que los distintos subsistemas están balanceados y que el sistema total se mantiene en equilibrio con su medio.

**Estabilidad:**

Un sistema se dice estable cuando puede mantenerse en equilibrio a través del flujo continuo de materiales, energía e información.

La estabilidad de los sistemas ocurre mientras los mismos pueden mantener su funcionamiento y trabajen de manera efectiva (mantenibilidad).

**Armonía:**

Es la propiedad de los sistemas que mide el nivel de compatibilidad con su medio o contexto.

Un sistema altamente armónico es aquel que sufre modificaciones en su estructura, proceso o características en la medida que el medio se lo exige y es estático cuando el medio también lo es.

**Optimización y sub-optimización:**

Optimización modificar el sistema para lograr el alcance de los objetivos.

Sub-optimización en cambio es el proceso inverso, se presenta cuando un sistema no alcanza sus objetivos por las restricciones del medio o porque el sistema tiene varios objetivos y los mismos son excluyentes, en dicho caso se deben restringir los alcances de los objetivos o eliminar los de menor importancia si estos son excluyentes con otros más importantes.

**Éxito:**

El éxito de los sistemas es la medida en que los mismos alcanzan sus objetivos.

La falta de éxito exige una revisión del sistema ya que no cumple con los objetivos propuestos para el mismo, de modo que se modifique dicho sistema de forma tal que el mismo pueda alcanzar los objetivos determinados.

**APORTES METODOLOGICOS****Jerarquía de los sistemas**

Al considerar los distintos tipos de sistemas del universo Kenneth Boulding proporciona una clasificación útil de los sistemas donde establece los siguientes niveles jerárquicos:

- Primer nivel, estructura estática. Se le puede llamar nivel de los marcos de referencia.
- Segundo nivel, sistema dinámico simple. Considera movimientos necesarios y predeterminados. Se puede denominar reloj de trabajo.
- Tercer nivel, mecanismo de control o sistema cibernético. El sistema se autorregula para mantener su equilibrio.



- Cuarto nivel, "sistema abierto" o autoestructurado. En este nivel se comienza a diferenciar la vida. Puede de considerarse nivel de célula.
- Quinto nivel, genético-social. Está caracterizado por las plantas.
- Sexto nivel, sistema animal. Se caracteriza por su creciente movilidad, comportamiento teleológico y su autoconciencia.
- Séptimo nivel, sistema humano. Es el nivel del ser individual, considerado como un sistema con conciencia y habilidad para utilizar el lenguaje y símbolos.
- Octavo nivel, sistema social o sistema de organizaciones humanas constituye el siguiente nivel, y considera el contenido y significado de mensajes, la naturaleza y dimensiones del sistema de valores, la transcripción de imágenes en registros históricos, sutiles simbolizaciones artísticas, música, poesía y la compleja gama de emociones humanas.
- Noveno nivel, sistemas trascendentales. Completan los niveles de clasificación: estos son los últimos y absolutos, los ineludibles y desconocidos, los cuales también presentan estructuras sistemáticas e interrelaciones.

### **Teoría analógica o modelo de isomorfismo sistémico:**

Este modelo busca integrar las relaciones entre fenómenos de las distintas ciencias. La detección de estos fenómenos permite el armado de modelos de aplicación para distintas áreas de las ciencias.

Esto, que se repite en forma permanente, exige un análisis iterativo que responde a la idea de modularidad que la teoría de los sistemas desarrolla en sus contenidos.

Se pretende por comparaciones sucesivas, una aproximación metodológica, a la vez que facilitar la identificación de los elementos equivalentes o comunes, y permitir una correspondencia biunívoca entre las distintas ciencias.

Como evidencia de que existen propiedades generales entre distintos sistemas, se identifican y extraen sus similitudes estructurales.

Estos elementos son la esencia de la aplicación del modelo de isomorfismo, es decir, la correspondencia entre principios que rigen el comportamiento de objetos que, si bien intrínsecamente son diferentes, en algunos aspectos registran efectos que pueden necesitar un mismo procedimiento.

### **Modelo procesal o del sistema adaptativo complejo:**

Este modelo implica por asociación la aplicación previa del modelo del rango.

Dado que las organizaciones se encuentran dentro del nivel 8, critica y logra la demolición de los modelos existentes tanto dentro de la sociología como dentro de la administración.

Buckley, categoriza a los modelos existentes en dos tipos:

- a) aquellos de extracción y origen mecánico, a los que denomina modelo de equilibrio;
- b) aquellos de extracción y origen biológico, a los que llama modelos orgánicos u homeostáticos.

Y dice:

"...el modelo de equilibrio es aplicable a tipos de sistemas que se caracterizan por perder organización al desplazarse hacia un punto de equilibrio y con posterioridad tienden a mantener ese nivel mínimo dentro de perturbaciones relativamente estrechas. Los modelos homeostáticos son aplicables a sistemas que tienden a mantener un nivel de organización dado relativamente elevado a pesar de las tendencias constantes a disminuirlo. El modelo procesal o de sistema complejo adaptativo se aplica a los sistemas caracterizados por la elaboración o la evolución de la organización; como veremos se benefician con las perturbaciones y la variedad del medio y de hecho dependen de estas".

Mientras que ciertos sistemas tienen una natural tendencia al equilibrio, los sistemas del nivel 8 se caracterizan por sus propiedades morfogénicas, es decir que en lugar de buscar un equilibrio estable tienden a una permanente transformación estructural. Este proceso de transformación estructural permanente, constituye el pre-requisito para que los sistemas de nivel 8 se conserven en forma activa y eficiente, en suma es su razón de supervivencia.

## **LAS ORGANIZACIONES COMO SISTEMAS**

Una organización es un sistema socio-técnico incluido en otro más amplio que es la sociedad con la que interactúa influyéndose mutuamente.

También puede ser definida como un sistema social, integrado por individuos y grupos de trabajo que responden a una determinada estructura y dentro de un contexto al que controla parcialmente, desarrollan actividades aplicando recursos en pos de ciertos valores comunes.

### **Subsistemas que forman la Empresa:**

- a) Subsistema psicosocial: está compuesto por individuos y grupos en interacción. Dicho subsistema está formado por la conducta individual y la motivación, las relaciones del status y del papel, dinámica de grupos y los sistemas de influencia.
- b) Subsistema técnico: se refiere a los conocimientos necesarios para el desarrollo de tareas, incluyendo las técnicas usadas para la transformación de insumos en productos.
- c) Subsistema administrativo: relaciona a la organización con su medio y establece los objetivos, desarrolla planes de integración, estrategia y operación, mediante el diseño de la estructura y el establecimiento de los procesos de control.

## **METODOLOGIA DE APLICACION DE LA T.G.S., PARA EL ANALISIS Y DISEÑO DE SISTEMAS**

Desde el punto de vista de la administración está compuesta de las siguientes etapas:

- a) Análisis de situación: es la etapa en que el analista toma conocimiento del sistema, se ubica en cuanto a su origen, objetivo y trayectoria.
  - 1. Definición de objetivo: el analista trata de determinar para que ha sido requerido ya que en general se le plantean los efectos pero no las causas.
  - 2. Formulación del plan de trabajo: el analista fija los límites de interés del estudio a realizar, la metodología a seguir, los recursos materiales y humanos que necesitará, el tiempo que insumirá el

trabajo y el costo del mismo. Esta etapa se conoce como propuesta de servicio y a partir de su aprobación se continúa con la metodología.

3. Relevamiento: el analista recopila toda la información referida al sistema en estudio, como así también toda la información que hace al límite de interés.

4. Diagnóstico: el analista mide la eficacia y la eficiencia del sistema en estudio. Eficacia es cuando el sistema logra los objetivos y eficiencia es cuando el sistema logra los objetivos con una relación costo beneficio positiva. Si un sistema es eficaz pero no eficiente el analista deberá cambiar los métodos del sistema, si un sistema no es eficaz el analista deberá cambiar el sistema y si un sistema es eficiente el analista sólo podrá optimizarlo.

5. Diseño: el analista diseña el nuevo sistema.

a) Diseño global: en el determina la salida, los archivos, las entradas del sistema, hace un cálculo de costos y enumera los procedimientos. El diseño global debe ser presentado para su aprobación, aprobado el diseño global pasamos al siguiente paso.

b) Diseño detallado: el analista desarrolla en detalle la totalidad de los procedimientos enumerados en el diseño global y formula la estructura de organización la cual se aplicara sobre dichos procedimientos.

6. Implementación: la implementación del sistema diseñado significa llevar a la práctica al mismo, esta puesta en marcha puede hacerse de tres formas.

a) Global.

b) En fases.

c) En paralelo.

7. Seguimiento y control: El analista debe verificar los resultados del sistema implementado y aplicar las acciones correctivas que considere necesarias para ajustar el problema.

## **EL SISTEMA DE CONTROL**

### **Concepto:**

Un sistema de control estudia la conducta del sistema con el fin de regularla de un modo conveniente para su supervivencia. Una de sus características es que sus elementos deben ser lo suficientemente sensitivos y rápidos, como para satisfacer los requisitos para cada función del control.

### **Elementos básicos:**

- a) Una variable; que es el elemento que se desea controlar.
- b) Los mecanismos sensores que son sencillos para medir las variaciones a los cambios de la variable.
- c) Los medios motores a través de los cuales se pueden desarrollar las acciones correctivas.
- d) Fuente de energía, que entrega la energía necesaria para cualquier tipo de actividad.
- e) La retroalimentación que a través de la comunicación del estado de la variable por los sensores, se logra llevar a cabo las acciones correctivas.

### **Método de control:**

Es una alternativa para reducir la cantidad de información recibida por quienes toman decisiones, sin dejar de aumentar su contenido informativo. Las tres formas básicas de implementar el método de control son:

1.- Reporte de variación: esta forma de variación requiere que los datos que representan los hechos reales sean comparados con otros que representan los hechos planeados, con el fin de determinar la diferencia. La variación se controla luego con el valor de control, para determinar si el hecho se debe o no informar. El resultado del procedimiento, es que únicamente se informa a quién toma las decisiones acerca de los eventos o actividades que se apartan de modo significativo que los planes, para que tomen las medidas necesarias.

2.- Decisiones Programadas: otra aplicación de sistema de control implica el desarrollo y la implantación de decisiones programadas. Una parte apreciable de las decisiones de carácter técnico y una parte pequeña de las decisiones tácticas abarcan decisiones repetitivas y rutinarias. Diseñando el sistema de información de manera que ejecute esas decisiones de rutina, el analista proporciona a los administradores más tiempo para dedicarse a otras decisiones menos estructuradas.

Si se procura que el sistema vigile las órdenes pendientes y se programa las decisiones de cuáles pedidos necesitan mayor atención, se logrará un significativo ahorro de tiempo y esfuerzo.

3.- Notificación automática: en este caso, el sistema como tal, no toma decisiones pero como vigila el flujo general de información puede proporcionar datos, cuando sea preciso y en el momento determinado.

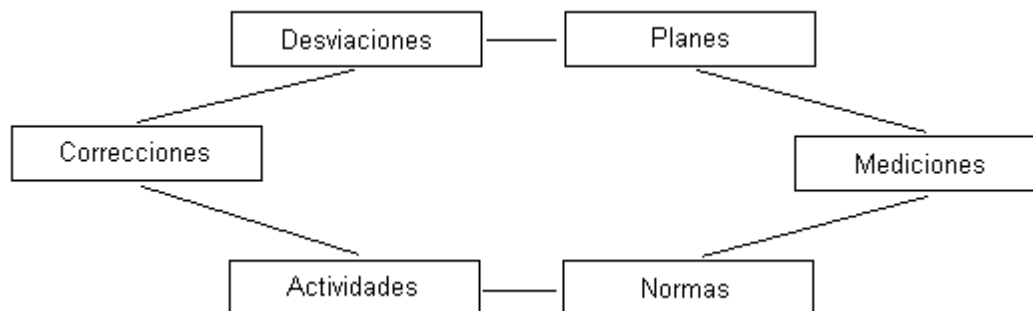
Las notificaciones automáticas se hacen en algunos criterios predeterminados, pero solo quienes toman las decisiones deben decir si es necesario o no emprender alguna acción.

## El Sistema de Control en las Organizaciones:

El control es uno de los cinco subsistemas corporativos (organización, planificación, coordinación y dirección son los restantes) los cuales son muy difíciles de separar con respecto al de control. De ello se desprende todo el proceso administrativo, debe considerarse como un movimiento circular, en el cual todos los subsistemas están ligados intrincadamente, la relación entre la planificación y el control es muy estrecha ya que el directivo fija el objetivo y, además normas, ante las cuales se contrastan y evalúan acciones.

Es necesario ver al control para determinar si las asignaciones y las relaciones en la organización están siendo cumplimentadas tal como se las había previsto.

### Gráfico del Sistema o Proceso de Control



Este gráfico representa el proceso de control como un sistema cerrado, es decir, que posee la característica de la retroalimentación o autorregulación. El movimiento es circular y continuo, produciéndose de la siguiente manera: se parte de la actividad o realidad a la cual debemos medir, con el auxilio o utilización de normas, efectuada la decisión comparamos los resultados de los planes, de esta manera la realidad quedará ajustada para el futuro. Se nota en este punto que no sólo la realidad puede ser ajustada, otras veces son los planes los que necesitan corrección por estar sensiblemente alejado de las actividades.

### BIBLIOGRAFIA CONSULTADA

Hermida, Jorge A. Ciencia de la administración. Ediciones Contabilidad Moderna S.A.I.C. Buenos Aires mayo de 1983.

## **ESTUDIOS DE FACTIBILIDAD**

En general los análisis de factibilidad más profundos, o los estudios de factibilidad, se completan durante la fase de diseño de sistemas, en general durante la consideración de la evaluación de las diferentes alternativas de solución propuestas. Los estudios de factibilidad consideran la factibilidad técnica, económica y operacional de cada alternativa, así como si el proyecto es o no apropiado dados los factores políticos y otros del contexto institucional.

### **Factibilidad operativa:**

Esta factibilidad comprende una determinación de la probabilidad de que un nuevo sistema se use como se supone. Deberían considerarse cuatro aspectos de la factibilidad operacional por lo menos.

- Primero, un nuevo sistema puede ser demasiado complejo para los usuarios de la organización o los operadores del sistema. Si lo es, los usuarios pueden ignorar el sistema o bien usarlo en tal forma que cause errores o fallas en el sistema.
- Segundo, un sistema puede hacer que los usuarios se resistan a él como consecuencia de una técnica de trabajo, miedo a ser desplazados, intereses en el sistema antiguo u otras razones. Para cada alternativa debe explorarse con cuidado la posibilidad de resistirse al cambio al nuevo sistema.
- Tercero, un nuevo sistema puede introducir cambios demasiado rápido para permitir al personal adaptarse a él y aceptarlo. Un cambio repentino que se ha anunciado, explicado y “vendido” a los usuarios con anterioridad puede crear resistencia. Sin importar qué tan atractivo pueda ser un sistema en su aspecto económico si la factibilidad operacional indica que tal vez los usuarios no aceptarán el sistema o que uso resultará en muchos errores o en una baja en la moral, el sistema no debe implantarse.
- Una última consideración es la probabilidad de la obsolescencia subsecuente en el sistema. La tecnología que ha sido anunciada pero que aún no está disponible puede ser preferible a la tecnología que se encuentra en una o más de las alternativas que se están comparando, o cambios anticipados en las prácticas o políticas administrativas pueden hacerse que un nuevo sistema sea obsoleto muy pronto. En cualquier caso, la implantación de la alternativa en consideración se convierte en impráctica.

Un resultado frecuente de hallazgos negativos acerca de la factibilidad operacional de un sistema es que éste no se elimina sino que se simplifica para mejorar su uso. Otras posibilidades son que los programas de relaciones públicas o de entrenamiento estén diseñados para enfocarse a sobreponerse a la resistencia a un nuevo sistema, o se desarrollan formas para hacer fases en el nuevo sistema en un largo periodo para que el cambio total, que traumatizaría a los usuarios u operadores, se convierta en una serie de pequeños cambios.

### **Factibilidad Técnica:**

El análisis de factibilidad técnica evalúa si el equipo y software están disponibles (o, en el caso del software, si puede desarrollarse) y si tienen las capacidades técnicas requeridas por cada alternativa del diseño que se esté considerando. Los estudios de factibilidad técnica también consideran las interfaces entre el sistema actual y el nuevo. Por ejemplo, los componentes que

tienen diferentes especificaciones de circuito no pueden interconectarse, y los programas de software no pueden pasar datos a otros programas si tienen diferentes formatos en los datos o sistemas de codificación; tales componentes y programas no son compatibles técnicamente. Sin embargo, puede hacerse una interface entre los sistemas no compatibles mediante la emulación, la cual son circuitos diseñados para hacer que los componentes sean compatibles, o por medio de la simulación, que es un programa de cómputo que establece compatibilidad, pero con frecuencia estas formas de factibilidad técnica no están disponibles o son demasiado costosas.

Los estudios de factibilidad técnica también consideran si la organización tiene el personal que posee la experiencia técnica requerida para diseñar, implementar, operar y mantener el sistema propuesto. Si el personal no tiene esta experiencia, puede entrenársele o pueden emplearse nuevos o consultores que la tengan. Sin embargo, una falta de experiencia técnica dentro de la organización puede llevar al rechazo de una alternativa particular.

### **Factibilidad Económica:**

Un proyecto es económicamente factible cuando sus ingresos son capaces de cubrir los gastos y generar un excedente adecuado para las condiciones de riesgo del proyecto.

$$\text{INGRESOS} \geq \text{EGRESOS} + \text{EXCEDENTE}$$

$$\text{INGRESOS} \geq \text{EGRESOS} + \text{EXCEDENTE} + \text{RIESGO}$$

Los estudios de factibilidad económica incluyen análisis de costos y beneficios asociados con cada alternativa del proyecto. Con análisis de costos/beneficio, todos los costos y beneficios de adquirir y operar cada sistema alternativo se identifican y se hace una comparación de ellos. Primero se comparan los costos esperados de cada alternativa, con los beneficios esperados para asegurarse que los beneficios excedan a los costos. Después la proporción costo/beneficio de cada alternativa se compara con las proporcionan costo/beneficio de las otras alternativas para identificar la alternativa que sea más atractiva e su aspecto económico. Una tercera comparación, por lo general implícita, se relaciona con las formas en que la organización podría gastar su dinero de modo que no fuera en un proyecto de sistemas.

Los costos de implementación incluyen comúnmente el costo remanente de la investigación de sistemas (para este propósito, los costos en los que ya se ha incurrido no son relevantes), los costos de hardware y software, los costos de operación del sistema para su vida útil esperada, y los costos de mano de obra, material, energía, reparaciones y mantenimiento. A través del análisis de costo/beneficio, la organización debe apoyarse en los conceptos tradicionales de análisis financiero y las herramientas como teoría del valor presente, análisis de costos diferenciales y análisis de flujos descontados.

Algunos costos y beneficios pueden cuantificarse fácilmente. Los beneficios que pueden cuantificarse con facilidad son de dos tipos generales: Ahorros en costos, tales como una disminución en costos de operación y aumentos en las utilidades directas. Como un ejemplo de lo último, un cliente podría haber contratado la suministración de pedidos de una cantidad conocida si la organización implanta un sistema que información que proporcione al cliente información

continúa acerca del estado de la producción en proceso de los embarques planeados de mercancía, de tal forma que a los clientes de dicho cliente pueda dárseles estimaciones exactas de cuándo estará disponible la mercancía.

Un problema importante con el análisis de costos/beneficio es la atención inadecuada de costos y beneficios intangibles. Éstos son aspectos de las alternativas de los nuevos sistemas que sí afectan los costos y utilidades y deberían evaluarse pero que los afectan en formas que no pueden cuantificarse fácilmente. Los factores intangibles con frecuencia están relacionados a la calidad de la información proporcionada por el sistema y a veces a formas sutiles en que esta información afecta a la empresa, tal como alternando las actitudes para que la información sea vista como un recurso.

Con frecuencia los diseñadores de sistemas no están a gusto basando sus recomendaciones en intangibles "vagos" que deben estimarse en forma contraria a lo que se llama "hechos Duros" de costos y beneficios fácilmente cuantificables; prefieren justificar sus recomendaciones con datos determinados objetivamente. Cuando se da mayor importancia a los costos y beneficios cuantificables que a los costos y beneficios intangibles, quizá haya una desviación contra el nuevo sistema porque la mayoría de los costos pueden cuantificarse de manera fácil, mientras muchos de los beneficios más importantes pueden ser intangibles y por lo tanto no se consideran correctamente.

Dos beneficios intangibles son el servicio a clientes y mejor información administrativa. Por ejemplo, los clientes pueden recibir información puntual y exacta acerca de los envíos, estados y otros informes más exactos, y nuevos servicios. Los cajeros electrónicos en los bancos que permiten a los clientes realizar operaciones 24 horas al día y que pueden resultar en un mayor número de clientes y utilidades para el banco, son un ejemplo de un servicio al cliente. Además, un nuevo sistema puede proporcionar una mejor imagen de la organización a sus clientes, vendedores, y empleados, que ayuda a atraer más clientes a que ayuda a retener a los empleados.

Los beneficios intangibles importantes pueden ser adquiridos de un nuevo sistema de información. Es cierto que el principal ímpetu al desarrollar un nuevo sistema puede ser la expectativa de información más exacta y a tiempo, un mejor formato de los informes, o informes que estén más enfocados a áreas particulares de problemas. Por ejemplo, los informes pueden recibirse más pronto después del cierre del periodo, o el nuevo sistema puede hacer que la información esté disponible con base en preguntas durante todo el tiempo. Además en muchos casos el nuevo sistema proporciona información que antes no estaba disponible, como información de los costos estándares o incrementos en los costos.

También puede haber menos beneficios intangibles obvios. Un nuevo sistema puede proporcionar mejor control sobre las operaciones de la organización, o puede ser que la auditoría sea más rápida o a un costo menor. Un beneficio intangible final es que la experiencia obtenida de la investigación de sistemas y del uso de un sistema de información más avanzado a menudo coloca a la organización en una mejor posición para tomar ventajas de desarrollos futuros en tecnología de computación y sistemas de información. Por ejemplo, es posible que la experiencia obtenida del desarrollo de una base de datos de personal tenga mucho valor si la organización decide implantar una base de datos financiera; no sólo estará afectando positivamente el diseño de la base de datos financiera, sino que también existirá una reducción en los costos de su desarrollo, que es un ahorro en costos hacia el siguiente proyecto de sistemas que debería considerarse como un beneficio proporcional por el proyecto actual.



La mayoría de los costos y beneficios intangibles de una alternativa afectan en forma indirecta las utilidades, pero esto es difícil de medir. La siguiente es una forma de cuantificar los costos y beneficios intangibles:

1. Identificar las causas y efectos directos. Por ejemplo, el efecto directo de computarizar tareas repetitivas puede ser que un nuevo sistema mejore los trabajos actuales y mejore la moral.
2. Identificar los efectos indirectos. Por ejemplo, una mejor moral puede resultar en cerca de 5% menos ausentismo y un 10% menos en el índice de rotación de empleados.
3. Estimar el impacto económico de los efectos indirectos para la vida estimada del sistema. Por ejemplo, una reducción en los retrasos de la programación y horas extras debidas a la reducción del ausentismo puede ahorrar casi \$2,000 al año, y una reducción en los costos de entrenamiento debidos a una reducción en la rotación de los empleados puede ahorrar hasta \$3,000 al año. El beneficio total (ahorro en costos) debido a una mejora en los empleos sería entonces \$5,000 al año o de \$20,000 para una vida estimada de 4 años del sistema.

Esta forma puede usarse para una gran variedad de costos y beneficios intangibles. Aunque arbitraria y subjetiva, es preferible a ignorar los intangibles. Esta forma puede describirse como “hacer tangibles los intangibles”.

Una forma alternativa es dejar sin cuantificar a los intangibles. Después, los usuarios y diseñadores de sistemas los estudian y llegan a un acuerdo acerca de la importancia relativa de lo cuantificado y de los costos y beneficios intangibles. Sin embargo, con frecuencia los costos y beneficios intangibles no se analizan completamente, y no se hace ningún intento para llegar a un acuerdo acerca de su importancia.

### **Factibilidad Legal:**

Al diseñar deben considerarse las implicaciones legales que surjan de los estatutos aplicables estatales, estatutos aplicables de la propia empresa, las reglas de la ley común, las agencias administrativas públicas, y de todos aquellos organismos públicos o privados que puedan tener alguna injerencia con el producto terminado de un Profesional de Sistemas de Información.

## TIPOS Y MODOS DE PROCESAMIENTO

Una de las tareas más importante que tiene el analista es poder determinar en que forma se pueden procesar los datos para hacer el sistema más eficiente y en que tipo de estructura física va a desarrollar su sistema.

En el caso de haber desarrollado bien el paso previo al diseño, es decir el estudio inicial, y análisis propiamente dicho, estas dos respuestas se deberían dar naturalmente siguiendo o estudiando las respuestas del usuario entrevistado y por su puesto estudiando detalladamente cada uno de los procesos y su interacción.

Según la tarea a desarrollar existen diferentes formas de procesar los datos:

### Modos de procesamiento.

Existen básicamente tres modos de procesar los datos que son en tiempo real o interactivo, batch o lotes y mixto.

#### Tiempo real

Es cuando los datos son cargados y procesados en el mismo momento. Se puede decir que es en forma interactiva, siempre y cuando exista un usuario que necesita llevarse una respuesta. Un ejemplo claro de esto es las cuentas bancarias en que los retiros son procesados en forma instantánea, deduciendo el monto retirado de la cuenta.

#### Por lotes o batch

Es en aquellos casos en que se acumula cierta cantidad de información antes de ser cargada y luego cuando hay un lote bastante considerable se procesa. Un ejemplo típico de esto es la liquidación de sueldos en que los datos del mes se acumulan en lotes y luego llegado la fecha de pago se procesan.

#### Mixtos

Son aquellos en que los datos se cargan en tiempo real y se procesan en forma diferida en el tiempo.

La principal pregunta es nítidamente, saber cual método emplear para procesar los datos para obtener información, ahí es donde la pregunta encuentra la respuesta en forma directa, pues es necesario saber en que momento se desea tener la información.

El problema se puede presentar en aquellos datos que pueden ser procesados tanto en **tiempo real**, como en forma **mixta** y aquellos que pueden ser por **lotes** o **mixta**. Nunca entre **lotes** y **tiempo real**, pues la naturaleza de los datos no permite la duda.

Las ventajas de cada método son fáciles de determinar:

En tiempo real se cuenta con información actualizada todo el tiempo, pero tiene como desventaja en contrapartida que necesitamos de recursos de hardware suficientes para ello y disponer en todo momento de ellos.

En lotes la ventaja que tiene es que los datos pueden ser chequeados o depurados al momento de cargarlos y así obtener la información justa y necesaria a partir de datos seleccionados. Otra ventaja es que se utiliza mejor el equipamiento ya que se distribuyen las cargas a otros horarios en que no sea crítica la situación y por lo tanto podemos utilizar hardware mas barato. La principal desventaja es que la información no es actual.

Mixto ofrece un poco de cada uno de los descriptos anteriormente, si bien la información no puede ser seleccionada, en el momento de la carga pero puede ser procesada a elección y de esa forma obtener una ventaja respecto de la forma en tiempo real. Pero mantiene la desventaja de la falta de información instantánea.

La elección entonces será en base a la naturaleza de los datos y la importancia de contar con información en tiempo y forma que permita tomar las decisiones correctas. Es muy importante obtener una primera aproximación de lo que va a pasar ya que en esta etapa definimos el hardware a utilizar y no se puede improvisar luego sobre la marcha.

## **Tipos de procesamiento**

Está ligado a la distribución geográfica de la empresa o la distribución de oficinas de la misma, como así también a una filosofía de trabajo.

Los tipos de procesamiento los vamos a distinguir en:

Centralizado, cliente servidor y distribuido.

### **Centralizado:**

Hay una sola máquina capaz de procesar información y es comúnmente llamado servidor. Las estaciones de trabajo si bien pueden tener procesador (en caso de que la terminal sea un PC), este no procesa la información del sistema. Si sus trabajos locales. Esta metodología de trabajo esta basada en el hardware con que se cuenta o lo que piensa invertir. Ya que si hay un gran procesador que pueda satisfacer los requerimientos de performance del sistema es una forma viable de otra manera habrá que distribuir las cargas de alguna otra forma.

### **Cliente Servidor**

Con respecto a la definición de arquitectura cliente/servidor se pueden citar las siguientes:

- Cualquier combinación de sistemas que pueden colaborar entre si para dar a los usuarios toda la información que ellos necesiten sin que tengan que saber donde esta ubicada.
- Es una arquitectura de procesamiento cooperativo donde uno de los componentes pide servicios a otro.
- Es un procesamiento de datos de índole colaborativo entre dos o más computadoras conectadas a una red.
- El término cliente/servidor es originalmente aplicado a la arquitectura de software que describe el procesamiento entre dos o más programas: una aplicación y un servicio soportante.

- IBM define al modelo Cliente/Servidor. "Es la tecnología que proporciona al usuario final el acceso transparente a las aplicaciones, datos, servicios de cómputo o cualquier otro recurso del grupo de trabajo y/o, a través de la organización, en múltiples plataformas. El modelo soporta un medio ambiente distribuido en el cual los requerimientos de servicio hechos por estaciones de trabajo inteligentes o "clientes", resultan en un trabajo realizado por otros computadores llamados servidores".
- "Es un modelo para construir sistemas de información, que se sustenta en la idea de repartir el tratamiento de la información y los datos por todo el sistema informático, permitiendo mejorar el rendimiento del sistema global de información"

## Elementos principales

"Los elementos principales de la arquitectura cliente servidor son justamente el elemento llamado cliente y el otro elemento llamado servidor". Por ejemplo dentro de un ambiente multimedia, el elemento cliente sería el dispositivo que puede observar el vídeo, cuadros y texto, o reproduce el audio distribuido por el elemento servidor.

Por otro lado el cliente también puede ser una computadora personal o una televisión inteligente que posea la capacidad de entender datos digitales. Dentro de este caso el elemento servidor es el depositario del vídeo digital, audio, fotografías digitales y texto y los distribuye bajo demanda de ser una maquina que cuenta con la capacidad de almacenar los datos y ejecutar todo el software que brinda éstos al cliente.

C/S es una relación entre procesos corriendo en máquinas separadas

El servidor (S) es un proveedor de servicios.

El cliente (C) es un consumidor de servicios.

C y S Interactúan por un mecanismo de pasaje de mensajes:

Pedido de servicios

Respuesta

## Distribuido:

El hardware capaz de procesar se encuentra distribuido y cada sucursal procesa su información local y luego por algún medio de comunicación se transfieren los datos al sistema principal. Este método de comunicación no hace al tipo de procesamiento y puede ser cualquiera de los conocidos.

## Ventajas y desventajas:

Las **ventajas** del procesar en forma centralizada es que en todo momento se cuenta con una base de datos consistente y con datos actualizados (siempre se debe tener en cuenta el modo de procesamiento).

Las principales **desventajas** que tiene es que el servidor debería ser una máquina con gran capacidad de procesamiento. Ya que debe atender a todas las terminales y procesar toda la información. La otra es que debe existir un medio de comunicación permanente que vincule las sucursales con la casa central y el costo de la instalación se ve visiblemente incrementado.

La **ventaja** de procesar en forma distribuida es que al ser datos locales únicamente, el procesador central puede ser de menor envergadura. Otra ventaja es que las comunicaciones se hacen una sola vez al día, las variantes que se presentan son muchas.

La principal **desventaja** es que los datos en el procesador central no están actualizados en línea.

## TOPOLOGÍAS

Otra de las cosas a tener en cuenta cuando se diseña una instalación es la topología de red a emplear.

Con la evolución de hardware el costo del computadora básica ha bajado sensiblemente, con ciertos costos que tiene que pagar por ello como es el envejecimiento prematuro, etc.

Las aplicaciones que tiene trabajar en red es:

- compartir recursos de alto costo
- compartir datos (bases de datos distribuidas)
- correo electrónico
- acceso a hosts
- automatización de oficinas
- aplicaciones industriales

Una cualidad que caracteriza a una red es las distancias en donde esta se desarrolla, en este curso abarcaremos solamente las LAN, que son aquellas que cubren áreas de menos de 1000 metros.

Alta velocidad de transmisión, baja tasa de errores (menos de .001 ppm), y bajo retardo en la transmisión.

Existen diferentes tipos y las más usadas son las siguientes:

### BUS

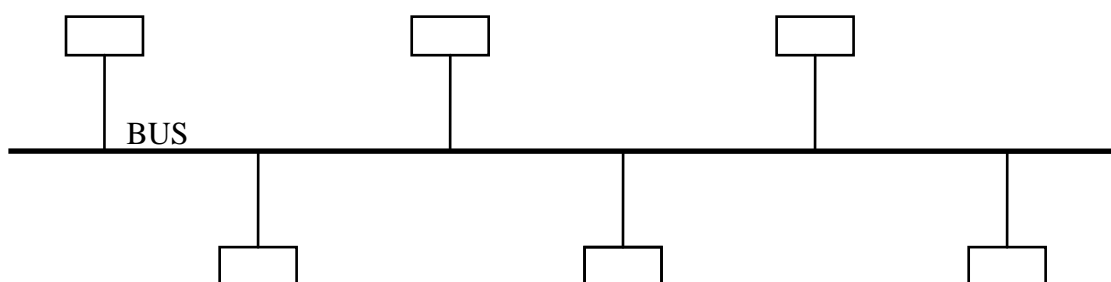
Está asociada directamente a una norma que es la IEEE 802.3, también conocida como Ethernet.

La velocidad de transmisión va entre los 10 y 100 Mb (IEEE 802.12).

Es la más usada por su bajo costo en relación a las prestaciones que presta.

Puede usar diferentes tipos de medios físicos lo que la hacen muy versátil. Fibra óptica, par trenzado, cable coaxial, etc.

Conceptualmente las computadoras conectadas comparten un BUS de comunicación y se lo disputan en forma contenciosa. El primero que lo toma es el que transmite.



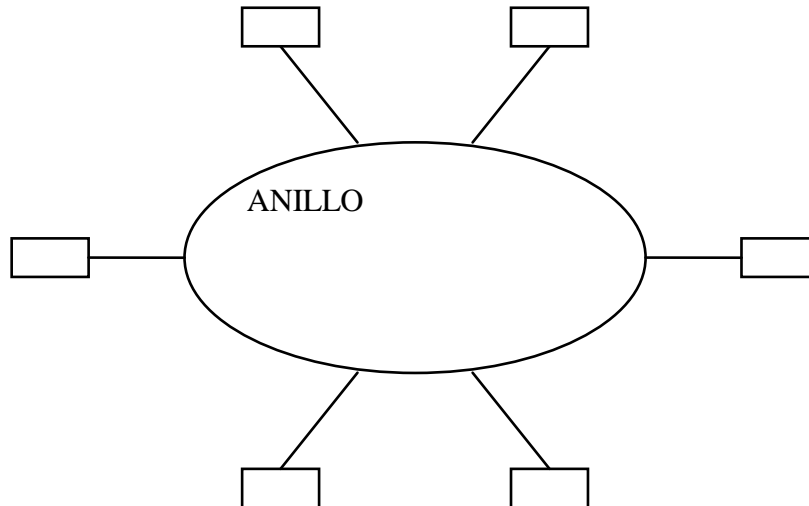
### ANILLO

Está asociada directamente a una norma que es la IEEE 802.5, también conocida como Token Ring.

La velocidad de transmisión va entre los 4 , 16 y 100 Mb(FDDI).

Puede usar diferentes tipos de medios físicos. Fibra óptica, par trenzado, cable coaxial (en caso de los 4Mb), etc.

Conceptualmente las computadoras conectadas a un anillo, en el que circula un Token o ficha y transmiten en el momento que disponen de ese Token, no pelen por el medio físico, es determinístico.



## ESTRELLA

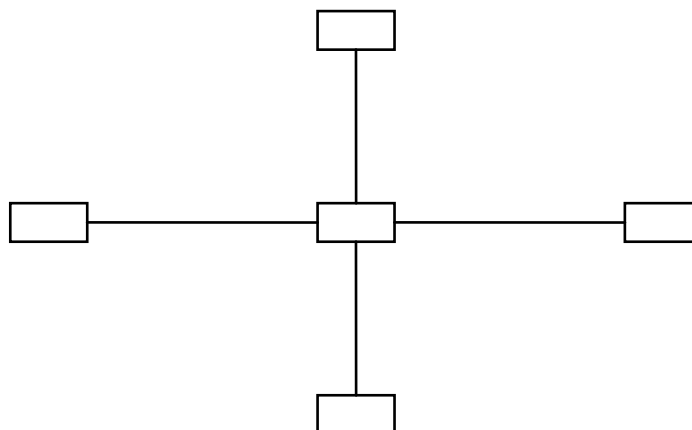
Antiguamente esta topología estaba asociada directamente a una norma que es la IEEE 802.6, también conocida como Arcnet.

La velocidad de transmisión va entre los 2 Mb.

Usa como medio físico cable coaxial de 75 Ohm.

Conceptualmente las computadoras conectadas a un procesador central, el cual administra las comunicaciones.

En la actualidad con el advenimiento de los cableados estructurados y los Hubs de comunicaciones se logra en la mayoría de los casos una topología de este tipo que es la más conveniente. Además las velocidades son las descritas para las otras topologías como así también los medios físicos



Las combinaciones de topologías son posibles y algunos ejemplos de ellos son:

ESTRELLA-ESTRELLA

ANILLO-ESTRELLA

ANILLO-ANILLO  
ANILLO-VERTEBRAL  
ARBOL DE BARRAS



## **SELECCION DEL HARDWARE Y SOFTWARE**

Cuando el analista-diseñador debe seleccionar el hardware y el software (en caso de que no sea desarrollo propio), se encuentra ante una gran variedad de productos y tecnologías, que hacen a la hora de decidir una tarea difícil y riesgosa en algunos casos.

### **SELECCION DEL HARDWARE**

En la selección del hardware deberíamos hacer conjugar una serie de cosas que hemos visto, definido y que a la hora de la aplicación se torna difícil, una de ellas es que el equipo computacional a menudo se adquieren sin un debido estudio de los requerimientos (diseño prematuramente físico) o ya estaban en la empresa al momento de solicitar nuestro servicio.

Cuando se esta ante el diseño de una instalación nueva el analista debe formularse una serie de preguntas, que determinan los factores de importancia para determinar las necesidades de tamaño y capacidad del equipo para poder así comparar sistemas distintos. Examinar los factores financieros implicados en la adquisición o alquiler del equipo.

El punto de partida para la selección se encamina en el requerimiento de tamaño y capacidad. Cada sistema en particular tiene sus propios requerimientos ya que un equipo puede ser apropiado para una tarea pero no para otra, esto específicamente se nota la diferencia cuando un sistema es administrativo y otro computacional.

La capacidad de un equipo generalmente es un factor determinante ya que dentro de la gran variedad de equipos que el mercado ofrece hoy no todos satisfacen la necesidad de nuestro sistema, ya sea por defecto o por exceso.

Dentro de las características mas comunes a evaluar se destacan las siguientes:

- Capacidad de memoria y posibilidad de crecimiento.
- Velocidad de procesamiento
- Características de los Buses y comunicaciones de datos.

Y entre las menos comunes se evalúa:

- Capacidad de almacenamiento masivo en disco interna.
- Cantidad de slots disponibles para tarjetas multiplexoras o canales SCSI.
- Monoprocesador o multiprocesador (en este último máximo de procesadores)

Un factor que se torna como determinante es que tipo de tecnología queremos que tenga la instalación que estamos diseñando, esto es CISC, RISC, microcanal, etc. Que define una particularidad a priori del sistema, esto es debido a que las pruebas comparativas (las más superficiales) en principio no tienen el mismo análisis y los valores no determinan nada.

Las necesidades de software y la cantidad de usuarios, dictan a menudo las necesidades del hardware (tamaño de los discos, velocidades de acceso a los mismos, memoria RAM, configuración de las comunicaciones, y los sistemas de almacenamiento masivo de respaldo o de seguridad, etc. ).

Una vez lograda una aproximación a lo necesitado, una forma útil de confrontar los requerimientos con lo disponible en el mercado es invitar a los proveedores a ofrecer su

equipamiento en base a las premisas antes establecidas, ya que esto además es útil en caso de que la compra o el alquiler sea de grandes proporciones, para que los proveedores no se sientan perjudicados, si no poseen la tecnología requerida, ofreciendo a cambio alternativas que muchas veces son válidas. Esto le sirve al analista para comprobar lo requerido con la folletería aportada por los mismos.

Es mas aún, algunas empresas con algunos datos de carga que va a tener el sistema, tienen la capacidad de simular en una máquina lo que sería el trabajo y así determinar una configuración aproximada.

## **Medición y evaluación del Hardware (Ver Benchmarks).**

### **Medición del Hardware**

Una vez determinados los requerimientos llega la etapa de prueba del hardware.

Una cosa importante en la etapa de prueba es ser sistemático, ya que hay que ser lo mas ecuánime posible en la medición de la respuesta.

Existen programas comerciales que se utilizan para medir performance de procesadores y discos, si bien estos son muy buenos para hacer una comparación, hay que manejar con mucho cuidado los resultados de los mismos. Muchas veces las empresas proveedoras de hardware diseñan los equipos conociendo los test por los que pasarán.

Lo ideal es hacer programas de prueba, que permiten someter el equipo a una mezcla bastante representativa de la carga proyectada.

Las pruebas más comunes son de velocidad del procesador, haciendo que este ejecute un número de instrucciones típicas ejecutando un número de programas en un ambiente de multiprogramación.

La misma prueba realizada en distintos equipos nos dará una clara idea de como se comporta cada procesador y el acceso a disco tanto para lectura, como para escritura.

Una vez concluidas las pruebas en donde en cada una se han realizado planillas con los resultados y estas firmadas por los presentes, se realiza una tabla comparativa, la que dará ya un resultado a tener en cuenta como un factor de mucha importancia.

### **Evaluación del Hardware**

Hasta ahora hemos obtenido resultados medidos del rendimiento del hardware, ya sea en tiempos de respuestas como en velocidad de procesamiento, pero otro factor a evaluar es la compatibilidad de los equipos que se van a adquirir y esto es así porque determinadas marcas de hardware con sistemas operativos propios cuando se van a instalar periféricos (discos, Impresoras, etc.), no los traen definidos como hardware conocido y la instalación, aunque posible, se hace engorrosa.

Además esto agrega al diseño del contrato un condimento diferente de que si se le compra todo a un mismo proveedor. Porque el analista deberá formular los pliegos licitatorios y contratos de manera tal que cada uno asuma las responsabilidades que le competen a la hora de un mal funcionamiento del equipo.

## **Mantenimiento y soporte**

El mantenimiento es el soporte que se le dará al equipo una vez concluida la etapa de garantía. Ya que se supone que la garantía cubre todos aquellos desperfectos que se produzcan en los equipos durante esa etapa.

Quizá sea el factor más importante a la hora de la elección del hardware, ya que las cláusulas que se formulen para el mantenimiento serán en definitiva las que le den peso al contrato. Esto es así porque las empresas buscan cubrirse con un buen contrato de mantenimiento y el proveedor es en este rubro donde encuentra los mejores réditos.

Existen dos tipos de mantenimiento correctivo y preventivo, las diferencias son obvias pero hay que aclarar que abarca cada uno de ellos y que tipo de cambios cubre cada uno.

Acá también es muy importante la participación del analista en la confección del contrato, ya que muchas empresas no dan mantenimiento a equipos que no vendieron o bien que no sean de su marca.

El soporte está mas bien ligado a cuanto tiempo la empresa proveedora del hardware soportará los equipos que hemos adquirido y que los mismos no estén discontinuados ya que eso se transforma en un factor de vital importancia a la hora de los problemas con los mismos.

Demás esta decir que el tipo de mantenimiento y soporte deben quedar por escrito rubricados en un contrato y que además diga cada cuanto tiempo se ajustará ese contrato, porque es sabido que el hardware a medida que transcurre el tiempo baja su costo y por ende el de los repuestos más comunes.

### **Servicio y tiempo de respuesta**

El servicio de mantenimiento es útil si está disponible cuando se lo necesita. Este concepto abarca entonces dos elementos: Servicio que se le da a un pedido de mantenimiento y tiempo de respuesta.

Es entonces necesario aclarar que lo que se va evaluar es que tipo de respaldo técnico (también conocido como soporte técnico) tiene la empresa adjudicada y en que lapso de tiempo están en condiciones de cumplir. Es necesario que la empresa tenga una oficina técnica en el medio y se diga como se van a recibir los llamados de socorro ante una eventualidad clasificados por orden de prioridad.

La empresa por su parte nos dará un listado del personal técnico que nos dará la atención y que posibilidades tiene de recambio de partes en tiempo y si no fuera así que técnica emplean para el reemplazo del equipo en caso de fallas de larga reparación. Además acompañara la información de con que rutina y procedimiento se harán los mantenimientos preventivos.

En definitiva la **SELECCION DEL HARDWARE** no es sólo la prueba del rendimiento de los equipos y su método de evaluación, es una combinación medida y detallada de todos los elementos que intervienen en este rubro en una instalación. Algunos autores ponen en este rubro bajo consideración las condiciones de compra de los equipos, (compra, alquiler y otras), las cuales consideran que no hacen a la selección del hardware sino mas bien a una cuestión técnico-financiera que es mas compleja de evaluar.

**NOTA:** en caso de ser equipos propios es necesario hacer todas las mediciones detalladas para saber si los equipos soportarán la carga del nuevo sistema.

## SELECCION DEL SOFTWARE

Muchos de los aspectos de interés señalados en la elección del hardware también se aplican a la selección del software.

La selección del software hoy en día dista mucho de ser lo que figura en los libros como una mera selección de programas comerciales que satisfagan la necesidad de la empresa. Es mucho más importante que eso.

Con la aparición de equipos personales potentes y la expansión definitiva de las redes de computadoras, intranet, etc., el sistema operativo pasa a ser algo susceptible de ser evaluado y esto es porque a diferencia del pasado donde las computadoras aceptaban sólo un sistema operativo propietario, las computadoras personales en general soportan una gran variedad de sistemas operativos y plataformas de red.

Necesariamente la selección del software de base debe ligarse íntimamente con la topología elegida y el hardware antes perfectamente detallado, es entonces prácticamente fácil de elegirlo.

En cambio una de las tareas más difíciles que hay es la elección del software una vez que se conocen los requerimientos del sistema, y es determinar si cierto paquete de software cumple con estos requerimientos.

La selección debe comenzar por saber si los paquetes están disponibles para el software de base elegido y si no ver a que solución de compromiso se llega. Esto deja un camino más corto para allanar ya que muchos de los vistos se decantarán.

Después de esta selección inicial es necesario ahondar mas las pruebas.

Si el software fue desarrollado a medida, es prácticamente lo mismo que si el desarrollo fuera propio, sólo es importante saber si son varias las empresas que se presentan lo siguiente:

- ¿quedan los programas fuentes para la empresa?.
- ¿que participación tiene el personal que va a hacer el mantenimiento propio de la empresa en el desarrollo?.
- ¿que capacitación vendrá acompañando al software?.
- Y otras de la misma naturaleza.

En cambio si se va a comprar un producto comercial ya desarrollado, los analistas evalúan que el mismo se adapten buena medida a los requerimientos que fueron determinados previamente en el análisis.

Lo fundamental a evaluar es también que requerimientos de hardware se requiere y que tipos de comunicación admite.

Que limitaciones tiene el software.

### Flexibilidad

Esta capacidad del software está referida además de cumplir con los requerimientos a todas aquellas otras cosas que el sistema es capaz de cumplir sin desarrollos adicionales. Básicamente la flexibilidad se remite a la emisión de reportes, almacenamientos, etc.

Tampoco es deseable tanta flexibilidad ya que corremos riesgos inherentes, como sobrecargar al equipo con excesiva carga, cuando las consultas o query, se hacen en forma indiscriminada y sin ningún control.

## **Auditoria y confiabilidad**

Algo muy importante de evaluar es la posibilidad que el sistema me brinda para hacer auditorias, sin necesidad de recurrir a software especializado.

Los auditores por lo general requieren que el sistema sea capaz de almacenar por si solo datos acompañados por fechas que sirvan para evaluar determinadas actividades y así emitir reportes.

La confiabilidad está referida a que el sistema debe obtener datos confiables, precisos y ciertos. Pero también se evalúa acá la seguridad que el sistema da a sus archivos y a los datos en forma individual para la protección contra tareas no autorizadas.

Por lo general se necesita un sistema de validación de claves para acceder al sistema y un registro de las mismas, independientemente de esto las claves para acceder a ciertos niveles de información.

Para esto se requiere de un sistema de seguridad que debe ser bien evaluado en todas sus partes y en lo posible exigir un sistema administrador separado para la seguridad.

## **Soporte técnico del proveedor**

Se toman las mismas consideraciones que se tuvieron a la hora de evaluar el hardware. Recordemos que el software si o si necesitará mantenimiento y entonces es muy importante dejar perfectamente determinado quien lo hará, como y con que frecuencia.

Como se proporcionarán las nuevas versiones y conocer la política de soporte de la empresa.

Una vez evaluado el software en su capacidad de producción, es necesario evaluar el plan de capacitación que acompaña.

Algo muy importante de evaluar y que no siempre se tiene en cuenta, es en aquellas empresas que ya contaban con un desarrollo de software, que grado de adaptabilidad tienen los usuarios al mismo.

NOTA: el software que a priori aparece como que no necesita soporte, en el corto plazo dejará de ser útil.

## ¿Que es un benchmark? (Comparador de rendimiento)

### 1- Introducción

Un benchmark es un conjunto de procedimientos (programas de computación) para evaluar el rendimiento de un ordenador. Hay cuatro categorías generales de pruebas de comparación :

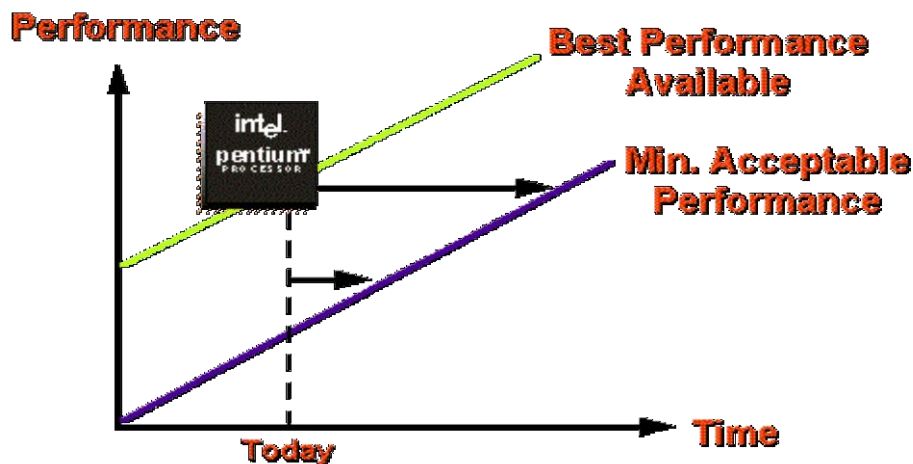
- Pruebas aplicaciones-base (**application-based**) las ejecuta y las cronometra.
- Pruebas playback (**playback test**), las cuales usan llamadas al sistema durante actividades específicas de una aplicación(Ej.: Llamados a gráficos o uso del disco) y las ejecuta aisladamente.
- Prueba sintética (**synthetic test**) , la cual enlaza actividades de la aplicación en subsistemas específicos.
- Prueba de inspección (**inspection tests**), la cual no intenta imitar la actividad de la aplicación, sino que las ejecuta directamente en los subsistemas específicos.

Los tests de aplicaciones base entregan la mejor forma de medir el rendimiento completo del sistema en el mundo real.

### 2- Usando el rendimiento para elegir el sistema correcto

#### ¿Porqué considerar el Rendimiento?

Juzgar el rendimiento de un sistema cuando se están tomando decisiones de compra es algo crítico a fin de retardar la obsolescencia y proteger su inversión.



Los procesadores Pentium ofrecen la vida útil más larga.

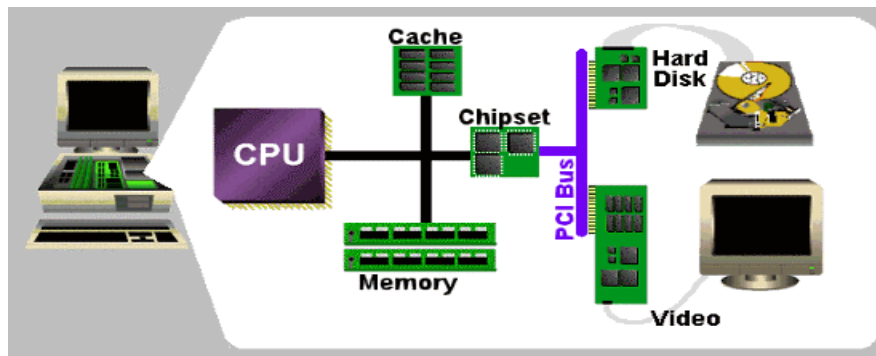
Ud. quiere obtener el máximo en la compra de su sistema. Parte de esto consiste en asegurarse que el sistema que elige le permite hacer lo que desea durante toda la vida del sistema. Esto

significa que no solo va a gestionar el software de hoy sino que dispone de los recursos necesarios para ejecutar los mas complejos programas del día de mañana.

Tal como se muestra en el diagrama de más arriba, para cualquier punto determinado en el tiempo se encuentra disponible una gama de procesadores, aquellos que ofrecen el máximo rendimiento posible y aquellos que ofrecen un mínimo nivel de rendimiento. Las flechas indican que comprando el de mayor rendimiento se obtiene una vida útil más larga.

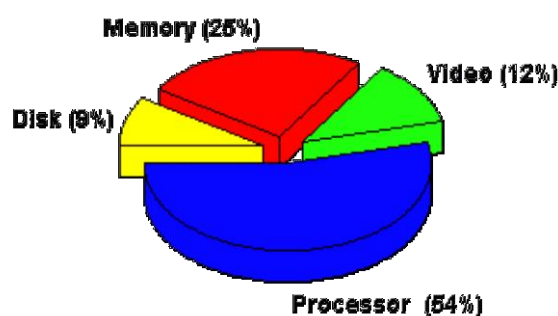
Es importante evaluar a una serie de aspectos relativos al rendimiento cuando se han de tomar decisiones de compra.

### 3- Factores del rendimiento de una computadora



El rendimiento global que puede Vd. obtener de su PC depende de cómo cada uno de sus componentes trabajan juntos para realizar una tarea. El efecto de cada componente individual puede variar dependiendo de en cuanto está interviniendo en ejecutar una aplicación particular. Como se muestra más abajo, el 54% del rendimiento del sistema depende de su procesador, o CPU, cuando se ejecutan aplicaciones Windows\* típicas. Otros componentes tales como memoria, vídeo y disco también juegan un papel. Una "conexión débil" en el sistema puede reducir significativamente la velocidad y el tiempo de respuesta que Vd. experimente.

Typical Windows\* Application Profile



Los porcentajes representan la utilización de los componentes cuando se ejecutan aplicaciones Windows\* típicas

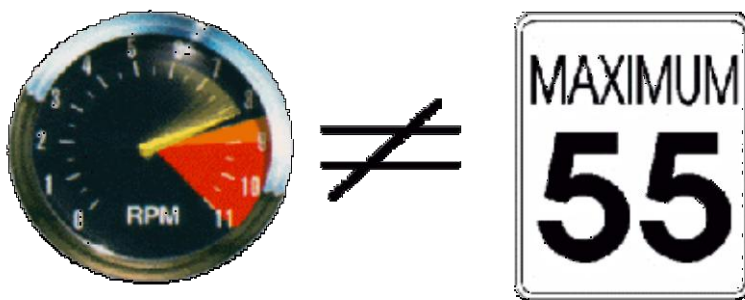
Por tanto, cuando elija un sistema, debe de empezar Vd. buscando cual es el procesador que le ofrece el mejor rendimiento para su software, a un precio razonable. Entonces puede Vd. empezar a analizar los otros componentes del sistema para asegurarse que se benefician de todas las ventajas de la potencia del procesador.

#### 4- ¿Qué es lo que tiene que tener en cuenta acerca del rendimiento ?

El decidir qué procesador Intel y qué sistema es el adecuado tiende a ser cada vez más difícil dado los nuevos productos PC que están apareciendo en el mercado. Una manera de diferenciar estos productos consiste en saber cómo su rendimiento se ve afectado con los trabajos de los que está Vd. interesado en obtener un mejor rendimiento, a fin de ganar tiempo de respuesta en las actividades que diariamente realiza, tales como crear un documento o trabajar sobre el último programa multimedia. También puede Vd. precisar de un rendimiento que le posibilite la opción de ejecutar el software cada vez más complejo de hoy día. Saber cómo el rendimiento del procesador y el sistema se puede ver afectado, le ayudará a realizar una elección mejor informada. Sin embargo, medir el rendimiento no es siempre algo simple. Para tomar la decisión correcta, necesita Vd. comprender qué es lo que está midiendo.

#### 5- MHz - Rendimiento

Un error frecuente es medir el rendimiento entre PCs estableciendo la comparación entre la velocidad del procesador en MHz o velocidad del reloj. Es algo así como utilizar RPM para medir la velocidad de su automóvil.



¿Miraría Vd. el tacómetro de su automóvil, que mide RPM, para determinar si está excediendo el límite de velocidad (y evitándose una elevada multa)? Probablemente no, dado que el tacómetro NO mide la velocidad del vehículo, sino que mide cuán rápido está rotando el motor.

Así como las RPM no indican cuán rápido está circulando su automóvil, los Gigahertzios no indican lo rápido que su hoja de cálculo se ejecutará. Es, por lo tanto, inadecuado comparar cuán rápido dos PCs realizarán una tarea comparando únicamente velocidades de reloj o Gigahertzios. Por ejemplo, un procesador Intel Core 2 Duo corriendo a 1.75 Gigahertzios supera fácilmente a un procesador Intel Pentium IV a 3.0 Gigahertzios. Estas es una de las razones por las que la industria creó las pruebas de rendimiento.

#### 6- ¿Qué son las pruebas comparativas?

Una prueba comparativa es un programa software que mide el rendimiento de una computadora, o las partes que la componen, ejecutando exactamente la misma cada vez. Estas pruebas no solamente pueden ayudarle a comprender el rendimiento de los diferentes sistemas a comparar, sino que también pueden ayudarle a evaluar el rendimiento de un sistema determinado bajo diferentes configuraciones hardware y software. Cada prueba evalúa un diferente tipo de trabajo. Algunas pruebas evalúan lo rápido que una computadora puede generar un documento. Algunas pruebas evalúan lo rápido que se dibuja un gráfico en la pantalla. Otras evalúan cómo de rápidas



se realizan determinadas operaciones matemáticas. Algunas pruebas pueden evaluar todas esas tareas al mismo. Para obtener resultados que tengan algún significado para Vd., use pruebas que reflejen el uso que hace Vd. de su computadora.

Hay dos niveles de pruebas: **COMPONENTE y de SISTEMA.**

- Las pruebas de componente evalúan solamente partes específicas de una computadora, tales como el procesador, el disco, la tarjeta gráfica, etc.
- Las pruebas de sistema evalúan el rendimiento global del sistema.

Hay dos tipos de pruebas: **APLICACION y SINTETICAS.**

- Pruebas de aplicación evalúan el rendimiento usando aplicaciones reales en la forma en que se utilizan.
- Las pruebas sintéticas usan programas software creados específicamente para evaluar el rendimiento.

Ambos tipos tienen ventajas y desventajas.

Hay muchas pruebas de rendimiento diferentes disponibles hoy día, pero no todas las pruebas han sido creadas de la misma forma. Crear una prueba requiere una considerable experiencia e inversión. La utilidad y exactitud de cualquier prueba depende fundamentalmente de la calidad de su implementación y las suposiciones que se hacen, por lo que puede darse que los resultados pueden variar de una prueba a otra.

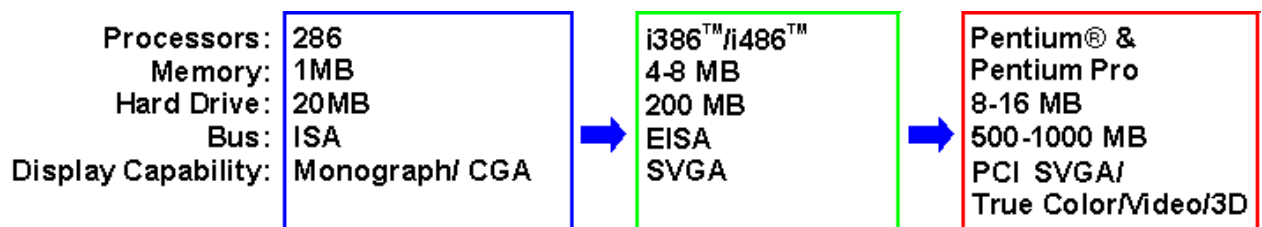
Es importante que, como hace normalmente con su software cuando se le queda obsoleto, consiga Vd. versiones actualizadas de las pruebas.

## 7- Las pruebas comparativas modernas

Como medir el rendimiento de los PC de hoy: La tecnología actual suministra rápidos procesadores de 32-bits, mejor memoria, y subsistemas mejorados de display. Estos avances han coincidido con nuevos sistemas operativos, aplicaciones y características, lo que da como resultado una continuada necesidad de pruebas comparativas modernas para medir adecuadamente el rendimiento de los PCs de hoy y mañana.

Cuando se compra un PC, es crítico incorporar a su proceso de evaluación métodos actualizados para obtener el rendimiento. Esto se debe a que pruebas obsoletas basadas en antiguos entornos hardware y software, no pueden decirle todo lo que Vd. necesita saber.

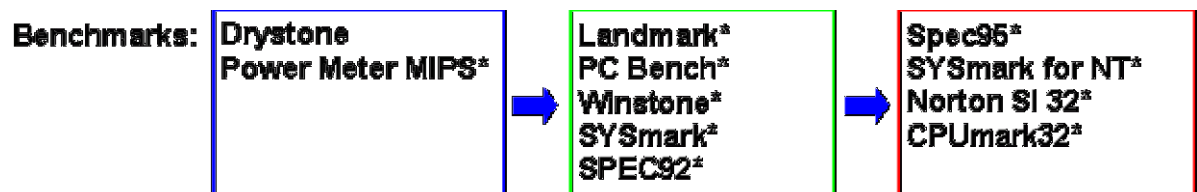
Ante todo, los procesadores y las plataformas son cada vez más sofisticados.



Al mismo tiempo, el software continua mejorando en características y prestaciones.



El resultado: una continua necesidad de nuevas pruebas comparativas para medir el rendimiento.



Pruebas como Dhrystone y PowerMeter MIPS\* podían considerarse adecuadas en los tiempos del procesador X86 cuando los sistemas se configuraban de una forma relativamente simple comparada con hoy día y cuando se ejecutaban aplicaciones DOS. Al mismo tiempo que la sofisticación ha aumentado, la industria se ha dirigido a crear nuevos tipos y herramientas de evaluación del rendimiento a fin de poder obtener resultados más ajustados.

Para explorar la evolución de las nuevas pruebas para PC y el porqué son tan importantes en su decisión de compra.

## 8- ¿Porqué las antiguas pruebas pueden ser inadecuadas?

Ese tipo de pruebas no tienen completamente en cuenta la diferencia de rendimiento entre el software de 16, 32 y 64 bits, ni son capaces de testear las posibilidades globales de las plataformas o los sistemas operativos de 64-bits de hoy día. Dos factores contribuyen principalmente al declive de este tipo de pruebas para medir el rendimiento. Ante todo, hay significativas diferencias entre el software para 64 bits, 32-bits y 16-bits. Y los avances en el hardware y los sistemas operativos están creando una gama de nuevas características del software que este tipo de pruebas antiguas sencillamente no tienen en cuenta.

### Diferencias entre software de 64-bits y 32-bits

La continua evolución a sistemas operativos de 64-bits está alentando a los vendedores de software a crear nuevas aplicaciones escritas en el modelo de programación para 64-bits. El uso de software de 64-bits tiene implicaciones importantes para el procesador de los PC's dado que el sistema usa el tiempo de proceso de forma diferente --y de forma más eficaz-- cuando ejecuta aplicaciones de 64-bits.

- El software de 32 y 16-bits necesita más tiempo en gestionar todo el entorno (tal como la gestión del segmento de memoria, por ejemplo). Esto proviene de los tiempos de programación para el x86 y el DOS, donde solamente era posible un direccionamiento de memoria de 64K al mismo tiempo.
- Comparativamente, el software para 32-bits típicamente precisa menos gestión de memoria, liberando al procesador para gastar más tiempo en actividades críticas, tales como tratamiento de instrucciones ALU y coma flotante. El gráfico de más abajo muestra el porcentaje del tiempo que el procesador ha gastado en diferentes instrucciones para aplicaciones de 32-bits en Windows\*95, así como para aplicaciones de 16-bits en Windows 3.1.

Esto significa que todas las aplicaciones de 32-bits se ejecutan más rápidamente que todas sus equivalentes de 16-bits, con lo que típicamente las aplicaciones de 32-bits deberán de tener una significativa mejora sobre la versión para un entorno DOS de 16-bits.

## 9- Nuevas características del Software

La mayoría de las pruebas antiguas pueden también fallar al analizar las nuevas características ofrecidas por el software más recientes como:

- Multitarea

La mayoría de los sistemas operativos de 32-bits y 64-bits dividen el tiempo de proceso en tareas múltiples --muchas de las cuales pueden ejecutarse de forma continuada en segundo plano-- tales como impresión o conmutar entre distintas aplicaciones. Este tipo de trabajo en modo múltiple requiere una gran potencia de proceso.

- Aplicaciones más grandes

El mayor número de características de una aplicación obligan a consumir un mayor espacio en todos los subsistemas de memoria. Esto aumenta los requerimientos de rendimiento para discos, la memoria y la arquitectura de la gestión de memoria.

- Incremento del tamaño de los datos

El aumento de aplicaciones que hacen un uso intensivo de vídeo y gráficos, así como velocidades más rápidas de red, requiere operaciones de entrada/salida más eficaces para un ancho de banda más alto del sistema.

- Multitarea / Multiproceso

La multitarea tiene una mayor eficacia ejecutando algunas aplicaciones modernas. La evaluación del rendimiento de aplicaciones multitarea y sistemas multiprocesador requiere técnicas especiales.

- Mejor protección de memoria

Los nuevos sistemas operativos como Windows 2003, Linux, Unix propietarios, dotan al sistema de una mayor fiabilidad para aislar las aplicaciones, lo que es por aplicaciones aislantes desde el uno al otro y el sistema operativo -- otra característica que requiere adicional capacidad de proceso.

- Nuevos recursos del sistema

Estos Sistemas Operativos que aparecen incluyen nuevos Interfaces de Programación de Aplicaciones (APIs) que hacen más fácil a los desarrolladores de software integrar las más nuevas posibilidades en sus aplicaciones. Estas incluyen mejores gráficos 3D, protocolos de telefonía (TAPI), y soporte de datos más complejos tales como audio y vídeo, que precisan de una mayor potencia de computación.

**En resumen, estas pruebas antiguas solamente podrán mostrar el rendimiento del software antiguo. Las pruebas comparativas modernas son necesarias para medir el rendimiento del software moderno.**

## 10- ¿Porqué se necesitan pruebas comparativas modernas?

La gente quiere que sus computadoras duren el mayor tiempo posible. Por esta razón tiene sentido comprar máquinas que tengan el mayor rendimiento disponible. Pero esto precisa de pruebas comparativas diseñadas para las aplicaciones de hoy, a fin de poder evaluar el rendimiento de proceso.

A continuación le indicamos en qué le pueden ayudar las pruebas comparativas modernas a medir el rendimiento de proceso:

- Adecuar apropiados estándares de rendimiento que tomen en cuenta las diferentes características del software de 64-bits y las nuevas posibilidades de multitarea, video y 3D que las viejas pruebas no tienen en cuenta.
- Una medida más precisa de las diferencias de rendimiento relativo entre dos PCs distintos, permitiéndole aumentar la vida útil de un PC con la vista puesta en el futuro.
- Ayudarle a prevenir que tenga que aumentar su inversión antes de tiempo, mostrándole la verdadera diferencia de rendimiento entre dos sistemas.

## **11- Las nuevas Pruebas Comparativas que se pueden utilizar en la actualidad.**

### **- Pruebas a nivel de procesador**

- SPECint\*95, SPECfp\*95, SPECfp\*0x

- consorcio de la industria SPEC

Una prueba CPU/memoria basada en aplicaciones reales

- Norton SI 32\*

- Prueba sintética de 32-bits de Symantec basada en aplicaciones Windows

- CPUmark32\*

- Prueba sintética de 32-bits de Ziff-Davis basada en aplicaciones de 32-bits .

- Indice iCOMP® 2.0

Actualización del actual 'Intel Comparative Microprocessor Performance Index',  
haciendo énfasis en pruebas de 32-bits

### **- Pruebas a nivel de Sistema**

- SYSmark\*32 para Windows 95\* .

Una prueba para aplicaciones Windows 95 •Disponible

- SYSmark para Windows NT\* , XP, 2003, 2008

Una prueba de 32-bits a nivel de sistema de BAPCo

Aplicaciones reales Windows NT, algunas multitarea

Prueba a nivel de sistema solo para 32-bits

Otras áreas del rendimiento del PC necesitan ser exploradas e incorporadas a las pruebas comparativas de la industria para que se pueda disponer de una correcta visión de las capacidades del sistema. Estas áreas incluyen la multitarea, vídeo, gráficos 3D y sonido, comunicaciones y reconocimiento de voz.

Intel está involucrándose activamente, junto con muchos consorcios y organizaciones de la industria, en ayudar a desarrollar estas pruebas. Como resultado, podrá Vd. tener un mejor indicador para estar informado en la decisión de compra de los PCs de hoy y en el futuro.

## **12- Bibliografía**

Sitios de Internet : Intel,SEI, ZD-NET.-

## Normalización de Archivos

¿Qué es normalización?

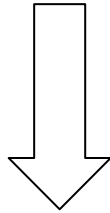
Normalización es un proceso que clasifica relaciones, objetos, formas de relación y demás elementos en grupos, en base a las características que cada uno posee. Si se identifican ciertas reglas, se aplica una categoría; si se definen otras reglas, se aplicará otra categoría.

Estamos interesados en particular en la clasificación de las relaciones BDR. La forma de efectuar esto es a través de los tipos de dependencias que podemos determinar dentro de la relación. Cuando las reglas de clasificación sean más y más restrictivas, diremos que la relación está en una forma normal más elevada. La relación que está en la forma normal más elevada posible es que mejor se adapta a nuestras necesidades debido a que optimiza las condiciones que son de importancia para nosotros:

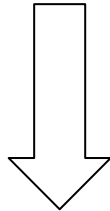
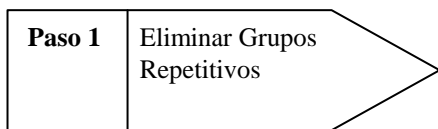
- La cantidad de espacio requerido para almacenar los datos es la menor posible;
- La facilidad para actualizar la relación es la mayor posible;
- La explicación de la base de datos es la más sencilla posible.

# NORMALIZACION

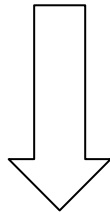
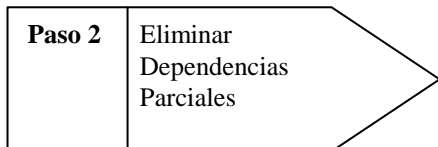
Vista del usuario



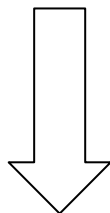
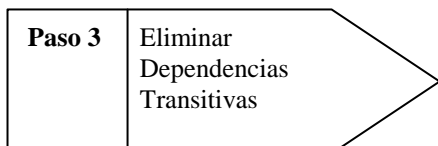
Relación sin normalizar



Relaciones en Primera  
Forma Normal



Relaciones en  
Segunda Forma  
Normal



Relaciones en Tercera  
Forma Normal

<b>Id. Vendedor:</b> 5000 <b>Nombre Vendedor:</b> Daniel Lillo <b>Area de Venta:</b> San Rafael				
Nro de Cliente	Nombre Cliente	Nro. Almacén	Ubicación Almacén	Ventas
18475	Delta Systems	4	Mitre 580	8
18300	A&R Refrigeración	3	Pellegrini 8	5
18450	Azcarate y Marra	8	Avellaneda 45	4

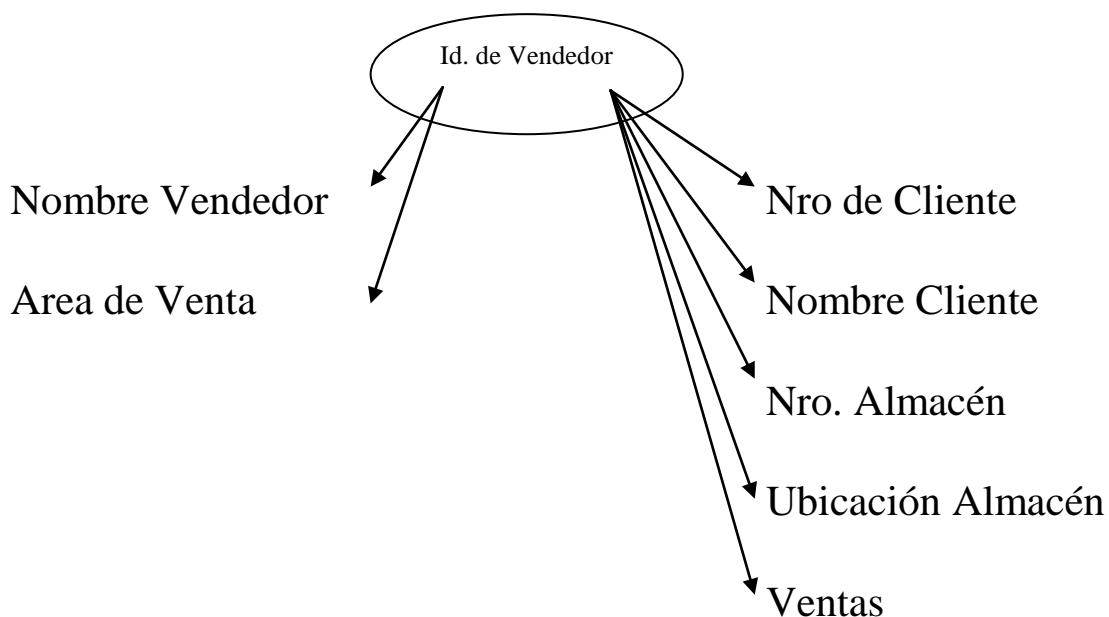
Cada Vendedor va a tener una de estas listas

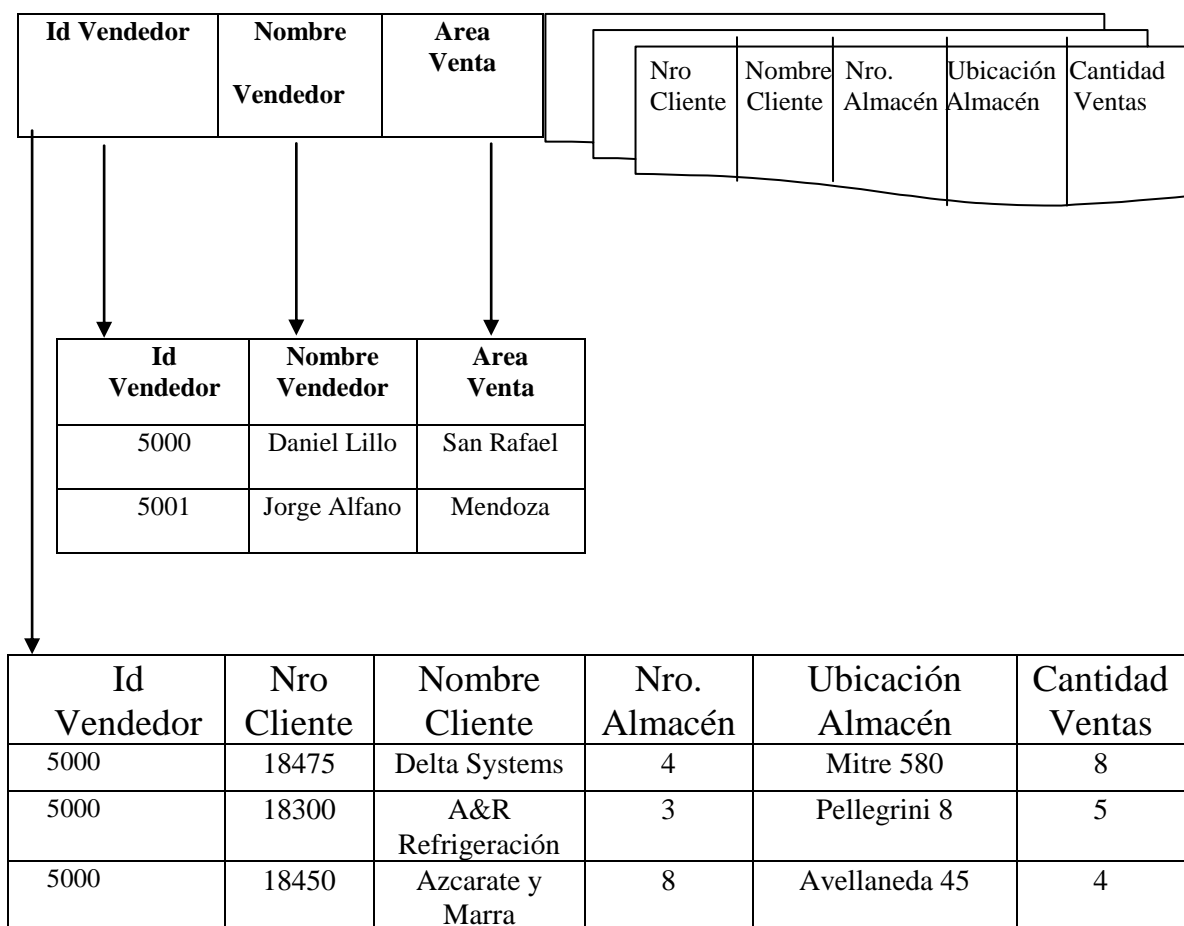
### Primera Forma Normal

Una relación se encuentra en 1era FN si no contiene estructuras repetitivas.

Para pasar a 1era FN se deben realizar los siguientes pasos:

- Selección de una clave primaria
- Identificación de estructuras repetitivas
- Por cada estructura repetitiva debemos crear una nueva relación con los atributos de la estructura repetitiva más la clave primaria de la relación principal. Luego elegir identificador primario para la nueva relación secundaria.
- Reformular la relación principal quitando las estructuras repetitivas ya derivadas.





Id.\_Vend, N\_Vend, A\_Ventas, {N\_Cliente, Nom\_Clie, N\_Alm, Ubic, C\_Ventas}

**1FN1:** Id\_Vend, N\_Vend, A\_Ventas

**1FN2:** Id\_Vend, N\_Cliente, Nom\_Clie, N\_Alm, Ubic, C\_Ventas



## Segunda Forma Normal

Una relación se encuentra en 2da FN cuando cada atributo no-clave se corresponde con la totalidad de la clave y además está en 1era FN. Una relación que tiene clave simple (no compuesta) está automáticamente en 2da FN (también puede no tener más atributos que las claves).

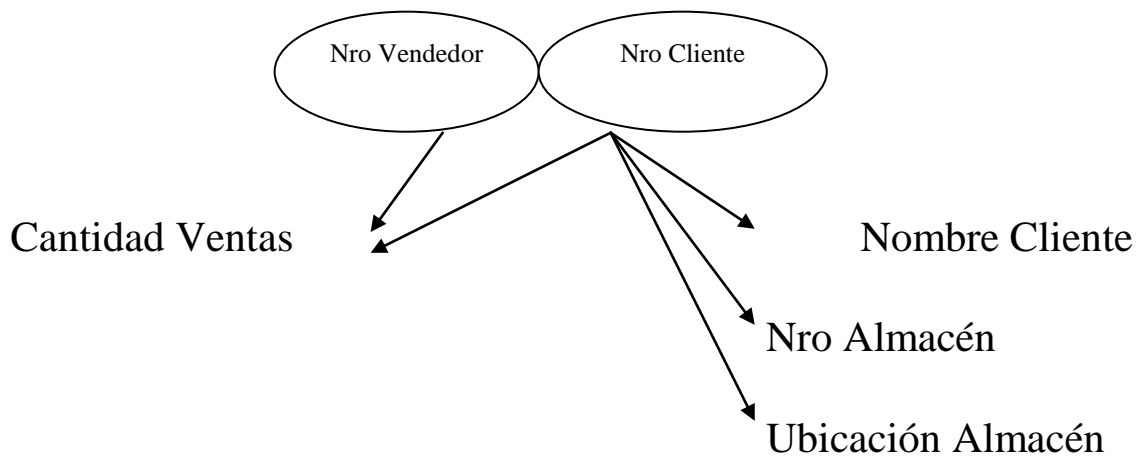
Para pasar a 2da FN se deben realizar los siguientes pasos:

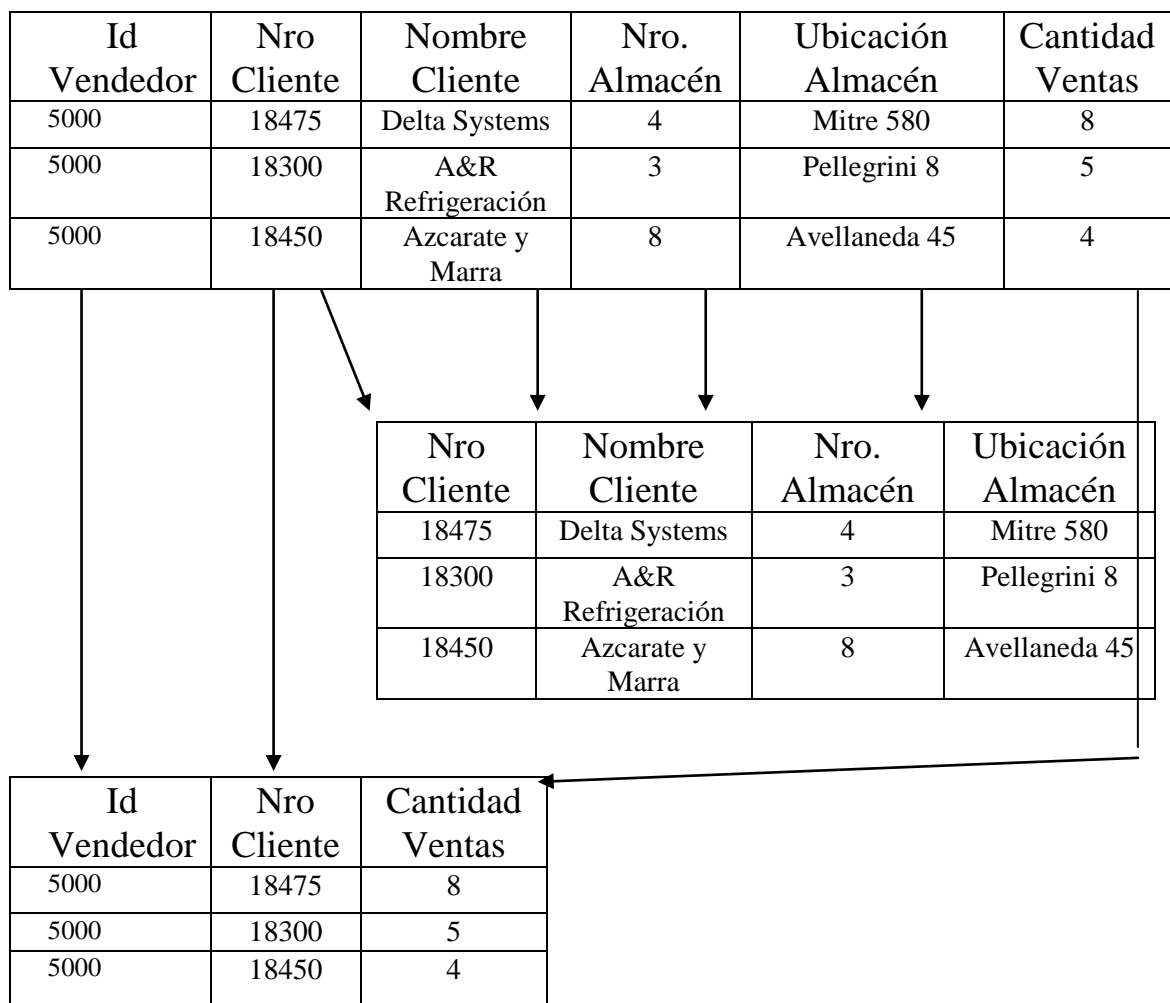
Comprobar por cada atributo no-clave, el grado de correspondencia con la clave (parcial o total).

Por cada atributo que tenga una dependencia parcial de la clave, hay que derivar una nueva relación con el atributo no-clave y el subconjunto de la clave de la cual depende.

A ésta nueva relación hay que escogerle clave primaria.

Reformulamos la relación original descartando estos atributos.





Id\_Vend, N\_Cliente, Nom\_Clie, N\_Alm, Ubic, C\_Ventas

**2FN1:** Id\_Vend, N\_Cliente, C\_Ventas

**2FN2:** N\_Cliente, Nom\_Clie, N\_Alm, Ubic

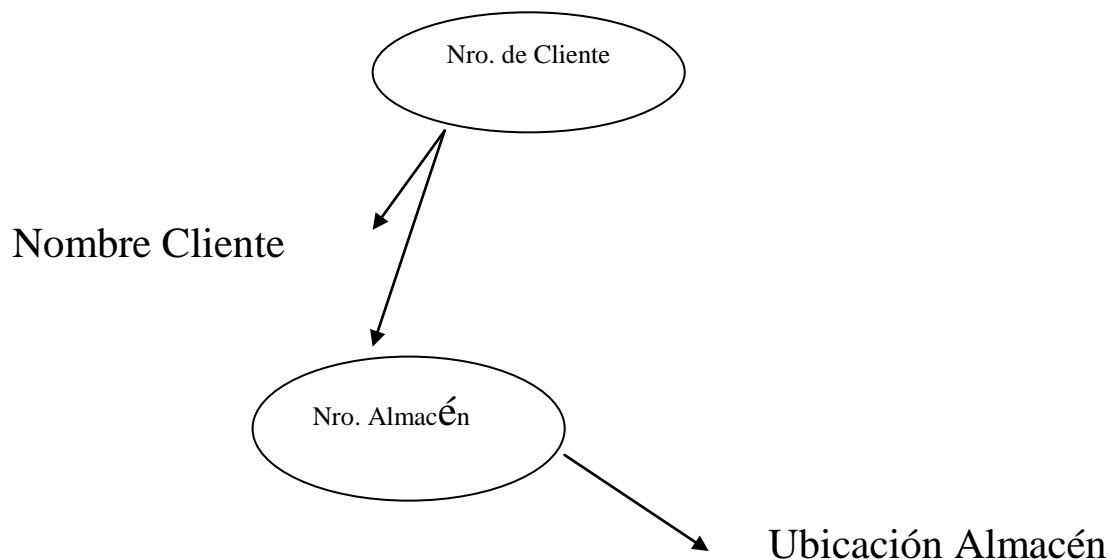
## Tercera Forma Normal

Una relación se encuentra en 3era FN si no existe interdependencia entre atributos no clave.

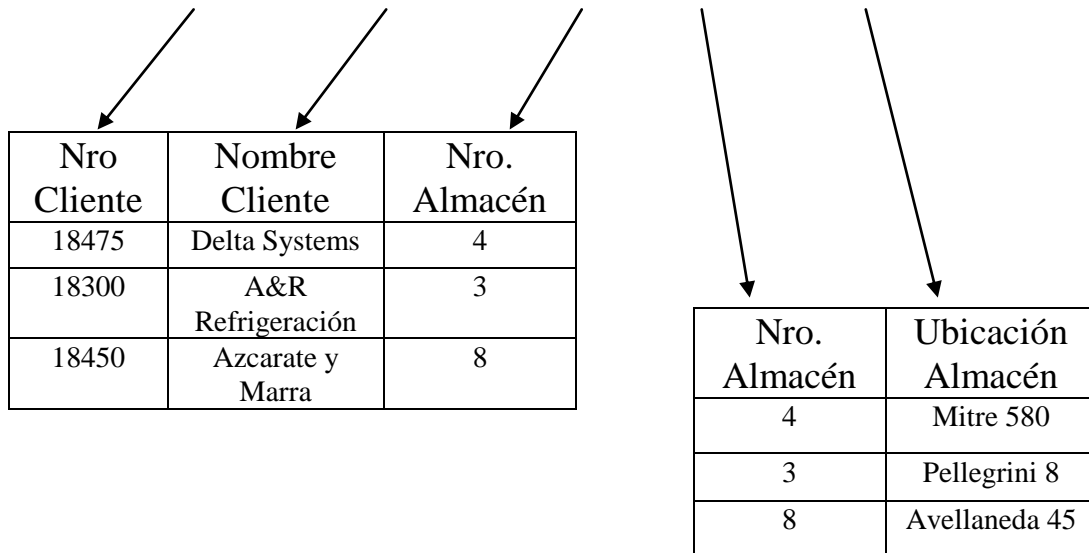
La interdependencia puede ser funcional o aritmética.

Para pasar a 3ra FN se deben realizar los siguientes pasos:

- Identificar interdependencias entre atributos no-clave y clasificarlos en aritméticos o funcional relacional (relación de datos).
- Aritméticas: se eliminan los atributos dependientes.
- Funcionales: se deriva una nueva relación, elegimos clave primaria para esta nueva relación. Luego eliminamos de la relación original los que son atributos noclave en la derivada.



Nro Cliente	Nombre Cliente	Nro. Almacén	Ubicación Almacén
18475	Delta Systems	4	Mitre 580
18300	A&R Refrigeración	3	Pellegrini 8
18450	Azcarate y Marra	8	Avellaneda 45



N\_Cliente, Nom\_Clie, N\_Alm, Ubic

**3FN1:** N\_Cliente, Nom\_Cliente, N\_Alm

**3FN2:** N\_Alm, Ubic

# Normalización de base de datos

## Indice

[1. Descomposición y Normalización](#)

[2. Dependencia](#)

[3. Normalización](#)

[4. Primera Forma Normal](#)

[5. Segunda Forma Normal](#)

[6. Tercera Forma Normal](#)

[7. Cuarta Forma Normal](#)

## 1. Descomposición y Normalización

Siempre que un analista de sistemas de base de datos arma una base de datos, queda a su cargo descomponer dicha base en grupos y segmentos de registros. Este proceso es la descomposición; el mismo es necesario independientemente de la arquitectura de la base de datos -relacional, red o jerárquica-. Sin embargo, para la base de datos relacional, la acción correspondiente puede dividirse y expresarse en términos formales y se denomina normalización a la misma.

La normalización convierte una relación en varias sub-relaciones, cada una de las cuales obedece a reglas. Estas reglas se describen en términos de dependencia. Una vez que hayamos examinado las distintas formas de dependencia, encontraremos procedimientos a aplicar a las relaciones de modo tal que las mismas puedan descomponerse de acuerdo a la dependencia que prevalece. Esto no llevará indefectiblemente a formar varias subrelaciones a partir de la única relación preexistente.

## 2. Dependencia

Significado :

Antes de entrar en el tópico principal de dependencia, vamos a rever algunos conceptos acerca de los individuos y acerca de las tuplas que los describen en la base de datos relacional (BDR). Restringiremos la discusión a la BDR, si bien la misma se aplica igualmente a las otras arquitecturas.

Los individuos tienen muchos atributos que pueden ser de interés a diferentes personas en diferentes momentos. Nuestro problema actual es con una sola aplicación o conjunto de aplicaciones: solamente son de interés algunos de los atributos.

Los símbolos aplicables a la relación han sido introducidos previamente.

- R es una tupla general o vector que describe a un individuo;
- R es una relación, una matriz o un conjunto de vectores que pertenecen a la población de interés.
- U es el universo consistente en todas las posibles descripciones individuales, obtenido mediante una combinación exhaustiva de los valores a atributos.

La tupla general toma la siguiente forma

$R = (a, b, c, \dots, n)$  La pertenencia con respecto a relaciones, tuplas y universos se indica mediante. Con respecto a los atributos:

- A es el símbolo del nombre de un atributo
- a es el símbolo de un valor del atributo.

Dominio (A) es el dominio para el atributo cuyo nombre es A.

Campo de aplicación

Estamos interesados en relaciones dependientes entre atributos de los individuos en una o varias poblaciones. Consideramos a los atributos D, E, y F. La dependencia es una relación funcional tal que los valores de una (o más de una) de las variables determina y fija el valor de las otras variables en la relación dependiente. Consideramos el caso en el que E y F dependen de D. Esto se describe más brevemente en forma simbólica:

$$e = e(d) \qquad f = f(d)$$

Existen tres tipos distintos de dependencia.

- Total uno-uno-sinónimo
- Completa - subtupla
- Transitiva - múltiple.

La dependencia es una relación funcional que penetra en el universo de posibilidades. La dependencia no puede deducirse solamente de los datos de nuestra, ya que éstos son necesariamente incompletos, sino que debe ser inherente al comportamiento del sistema. Por ejemplo, si los datos revelan que cada uno de nuestros proveedores tiene exactamente una planta y

que todas estas plantas están en diferentes ciudades, podemos asumir una dependencia total entre proveedor, planta y ciudad. Es decir, dada una ciudad, la misma está asociada con un proveedor; y dado este proveedor estará asociado con una ciudad. En la práctica, solamente cuando un nuevo proveedor se incorpore con una planta en la misma ciudad que uno de nuestro antiguos proveedores, resultará claro que no existe dicha dependencia total, Esto no podría ser deducido a partir de los datos previos.

### Dependencia Total

Consideremos los atributos  $x$  e  $y$ . Cada valor de  $x$  tiene uno y solo un valor de  $y$  asociados a él; e inversamente, dado un valor de  $y$  existe solamente un valor de  $x$  asociado a éste. Se trata de una función unitaria de una variable tanto en sentido directo como inverso y por o tanto se denomina dependencia total. Otra forma de expresar lo mismo es decir que  $x$  e  $y$  son sinónimos; ambas expresiones son equivalentes.

### Ejemplo con clave

Si una de las variables es al mismo tiempo la clave, como consecuencia todo valor de ambas variables es único en cualquier tupla de la relación. Por ejemplo, consideremos un archivo de personal donde cada uno de los empleados es identificado de tres maneras.

- Su nombre
- Su número de seguridad social
- Su número de empleado

Los tres pueden representar una dependencia total. Tanto el número de seguridad social como el número de empleado identifican al individuo en forma única. El número de seguridad social atañe a la población completa de trabajadores de los Estados Unidos. El número de empleado se aplica solamente al personal de una empresa en particular. El nombre puede no ser totalmente único y la dependencia total existe solamente cuando cada empleado tiene un nombre único.

Si el número de empleado es al clave de la relación, el número de seguridad social es sinónimo de aquel. Podemos en consecuencia decir que el número de seguridad social, el campo no clave, es totalmente dependiente de la clave, y es una clave candidata.

Si los nombres de todos nuestros empleados son únicos, también pueden, ser claves candidatas. Sin embargo puede existir alguna duplicación, dos personas llamadas John Smith, por ejemplo. Dado que esta es una posibilidad, no puede establecerse una dependencia total con

respecto total con respecto al nombre. Puede incorporarse a la firma un nuevo empleado y este puede tener el mismo nombre que uno de nuestros empleados actuales.

**Ejemplo con estado** Consideremos una relación que contiene información sobre estado en dos formas :

- Una identificación de estado con dos letras, tal como CA para California.
- Una designación con un número de dos dígitos tal como 12 para California.

Estas dos formas de información sobre estado ilustran una dependencia total. Debe notarse sin embargo que muchas tuplas pueden contener la misma identificación de Estado, dado que muchos de nuestros clientes pueden provenir de California. En consecuencia resulta claro que la dependencia total no significa unicidad.

#### Dependencia Completa

El concepto de dependencia completa se aplica solamente cuando:

- Tenemos más de dos variables, y
- Una variable dependiente depende de dos o más variables independientes.

Consideramos una relación que abarca las variables P, Q y R. Supongamos que P es la variable dependiente. Si el valor de P está determinado por una función de Q y R combinados, se trata de una dependencia completa. Esto es, el valor de P no depende únicamente ni de Q ni de R.

Vamos a repetir esto simbólicamente. El valor de P es completamente dependiente de los valores de q y r.

$$p = p(q,r)$$

#### Ejemplo con orden de compra

Como un ejemplo de dependencia completa, consideremos el caso de una orden de compra. Supongamos que esta orden de compra describe mediante tres variables que son de interés para nosotros:

- El número de orden de compra (PON) designa la orden completa;
- El número de parte de pieza designa una de las partes ordenadas por el pedido;
- La cantidad de piezas es el número de unidades de dicha pieza requerida para satisfacer el pedido.



Los pedidos describen en consecuencia una orden por medio de varias partes diferentes, y para cada una distinta asociada. El sistema contable ve varios pedidos diferentes. La misma parte puede aparecer en distintos pedidos y, cuando ello sucede, puede estar asociadas distintas cantidades con la misma parte.

Un tupla de la base de datos relacional contendrá un PON un número de parte y una cantidad. La cantidad es completamente dependiente del PON y del número de parte. Resulta claro que el número de pedido no es suficiente para determinar la cantidad (todas las partes de un determinado pedido no tiene la misma cantidad). Análogamente, un número de parte no es suficiente para determinar la cantidad ordenada, dado que diferentes pedidos pueden requerir distintas cantidades de dicha parte. Por lo tanto, en nuestro ejemplo, la cantidad no es dependiente solamente del PON o del número de parte; es completamente dependiente de ambos.

Puede imaginarse, aunque no es muy probable el caso de que cada vez ordenados una parte la ordenamos solamente por una cantidad como una docena, o tres gruesas o cualquier otro valor fijo. Si esto ocurre para todas las partes y para todos los pedidos de nuestro sistema, en consecuencia no existirá dependencia completa. En efecto podemos decir que hay dependencia total entre cantidad y número de partes - condición improbable-.

Hemos examinado anteriormente un ejemplo académico y las variables profesor, clase y sección. Tenemos en este caso una dependencia completa de profesor respecto de clase y sección. Si en nuestra facultad está establecido existirá dependencia completa. Esto existiría que un profesor enseñe siempre a todas las secciones de una clase particular - una condición no muy factible con un curso de 20 secciones-.

Dependencia transitiva

La dependencia transitiva se aplica a tres o más variables. Consideremos el caso de solo tres variables y llamémoslas S, T y V.

Diremos que S es la variable independiente si los valores de S determinan tanto a T como a V, y se simbolizará así:

$$S \twoheadrightarrow T; S \twoheadrightarrow V$$

Sin embargo, sería deseable encontrar una relación más restrictiva o definida.

Tenemos dependencia transitiva cuando S determina a T y V, pero los valores de V pueden considerarse siempre como dependiendo de los valores de T. Esto puede escribirse como

$$S \rightarrow T; T \rightarrow V$$

o alternativamente como

$$V = V(T); T = T(S) \quad V = V(T(S))$$

### Reducción

Si podemos manejar las dependencias transitivas, podremos reducir el espacio total requerido para almacenar los datos. Varios valores de S pueden generar un único valor de T. De modo similar, pueden existir varios valores de T asociados solamente con un valor de V. La separación de estas relaciones permite conservar espacios. Esto puede observarse mejor con respecto al ejemplo que se describe más abajo.

### Ejemplo

Consideramos un ejemplo que asocia cursos con departamento y con escuela. En consecuencia, canto será dictado por el departamento de música en la escuela de Artes y Ciencias; hidráulica será dictada por ingeniería civil en la Escuela de Ingeniería; impuestos será dictado por el departamento contable en la Escuela de Administración.

### Llamemos

- S al curso
- T al departamento
- V a la escuela

Por lo tanto

$$S \rightarrow T \rightarrow V$$

la descomposición consiste en la asociación de un curso con un departamento en una relación. Otra relación identifica a cada departamento con una escuela. Esta segunda relación es necesariamente menor tanto en grado como en cardinalidad y aquí reside el ahorro de espacio.

## 3. Normalización

¿Qué es normalización?

Normalización es un proceso que clasifica relaciones, objetos, formas de relación y demás elementos en grupos, en base a las características que cada uno posee. Si se identifican ciertas reglas, se aplica una categoría; si se definen otras reglas, se aplicará otra categoría.

Estamos interesados en particular en la clasificación de las relaciones BDR. La forma de efectuar esto es a través de los tipos de dependencias que podemos determinar dentro de la relación. Cuando las reglas de clasificación sean más y más restrictivas, diremos que la relación está en una forma normal más elevada. La relación que está en la forma normal más elevada posible es que mejor se adapta a nuestras necesidades debido a que optimiza las condiciones que son de importancia para nosotros:

- La cantidad de espacio requerido para almacenar los datos es la menor posible;
- La facilidad para actualizar la relación es la mayor posible;
- La explicación de la base de datos es la más sencilla posible.

#### **4. Primera forma normal**

Para que una relación esté en primera forma normal (1 FN), debe ser solamente una relación propia, una matriz  $m$  por  $n$ , donde:

- Ninguna celda de la matriz está vacía;
- El valor  $n$  cualquier columna está definido por el dominio para dicho atributo.
- Cada tupla tiene una clave que la identifica en forma unívoca, pero dicha clave no significa orden.

La aplicación determina la relación

Para que una relación sea normalizada en pasos adicionales, debe encontrarse en la primera forma normal. Colocar los datos en la primera forma normal está a cargo del diseñador de la aplicación. Estos datos se encuentran disponibles de alguna manera inicialmente. Si la aplicación existe en forma manual, o ha sido anteriormente computarizada pero no todavía como relación, el diseñador reorganiza los datos de modo de conformar una matriz 1FN.

La segunda inicial más importante es la dimensión de la relación ¿cuántos componentes existen en la tupla o cuántas columnas en la tabla? ¿De qué manera se compara esto con el número de campos en el documento fuente?.

- Información descriptiva para el usuario
- Nombres de atributos.

La información convertida queda formada en tuplas. La próxima pregunta es cuantas tuplas representarán a la formación en esta forma. Debe notarse que el número de partes ordenadas varía de una factura o pedido a otro.

orden	vendedor	fecha	via	orden wetco
de cliente		de la orden		
M0007	2-14	3/12/17	ups	1922447

Venta a flores associates expedido a 108 8 avenue brooklyn, n.y. CP 11215

cantidad		precio parte		descripcion	monto	
Ordenada	despachada	Pendiente				
2	2		3.50	018719	camisa	7.00
2	2		.35	020428	guia	.70
1	1		.70	020808	rodillo motor	.70
1	0	1	.25	020811	rodillo libre	0.00
1	1		6.00	020819	humidrum	8.00

Dado que una tupla debe tener un número fijo de componentes, necesitamos una tupla en primera forma normal para cada parte de cada pedido. Sin embargo, la información que se encuentra en la parte superior del formulario, y que se llena a máquina, es la misma para todas las partes ordenadas más abajo. Por lo tanto cada tupla consiste en una parte de datos que son variables y datos del pedido que se duplican para cada parte ordenada.

### Grafo de Dependencia

Una vez que los datos han sido puestos en primera forma normal, resulta conveniente descomponer la relación en un número de relaciones más pequeñas, cada una en forma normal superior, de modo de optimizar el almacenamiento y usar sus funciones. Para esto resulta necesario reconocer las dependencias existentes. Un grafo exhibe los distintos tipos de dependencias que existen, y enfatizan que hemos investigado completamente cada dependencia.

El grafo simple no está diseñado para mostrar dependencias. Para hacer utilizable a este grafo, se agregan colores pueden expresarse en blanco y negro mediante distintos tipos de líneas. Discutiremos estos tipos de líneas en términos de la dependencia que cada uno representa. En las figuras que siguen las formas gráficas aparecen a la izquierda y se utilizan para constituir un grafo completo. A la derecha se puede observar una forma simbólica para describir dependencias únicas.

### Dependencia única

En la figura vemos un arco que conecta dos vértices A y B. A es la cola y B es la cabeza de la "flecha". Esto significa que B depende de A. Es decir dado un valor de A podemos predecir de A. Es decir, dado un valor de A podemos predecir cuál será el valor de B.

### Dependencia total

La dependencia total se define como una dependencia bilateral o simétrica. Es decir, si C depende de D, en consecuencia D será dependiente en forma similar de C. Esto se expresa en la figura mediante una arista (sin una flecha) que une C y D. Para enfatizar la dependencia total, se usa una línea doble o una línea más gruesa. Esto representa una medida de seguridad para verificar que el usuario no dibuje un arco e inadvertidamente omita la flecha. Simbólicamente se utiliza una doble flecha.

### Dependencia completa

La variable G depende en forma completa de otras dos variables E y F, lo cual puede ilustrarse como se ve a la izquierda de la figura. Pero así no es representada adecuadamente la dependencia completa, ya que el valor de G no depende de E o F, independiente, sino que depende de ambos valores. Por lo tanto en el centro de la figura A, vemos una forma mejor; la arista que une E y F no intenta demostrar una dependencia entre E y F, por lo tanto se dibuja en líneas de trazos; a partir del centro de esta línea de trazos, se dibuja un arco dirigido hacia G para indicar que G depende de ambas variables E y F.

### Dependencia transitiva

Supongamos que dos variables, K y L, dependen de J. Si puede verificarse que L depende en forma primaria de K, existiría una dependencia transitiva. Mostramos a la izquierda de la figura B que L depende de J o de K. Más apropiado es el grafo del centro de la figura B, donde podemos ver que L está definida por K la cual, a su vez, está determinada por los valores de J.

Simbólicamente indicamos una dependencia transitiva de L respecto de J mediante una flecha de trazos desde J a L, como puede verse a la derecha de la figura B.

### Ejemplo

En la figura B se presenta un grafo de dependencia hipotético. En el mismo se dibujan las relaciones de dependencia entre atributos para una aplicación de remuneración. EMPNO y DEPTNO están subrayadas en la figura para expresar que ambas son partes de una clave compuesta para la relación. Una línea gruesa conecta EMPNO a EMPNOM para indicar que si nombre de empleado y existe una dependencia total.

Varios atributos dependen directamente del número de empleados:

- TITL es el título de la tarea del empleado
- PAYLVL es un carácter que indica el nivel de sueldo del empleado.
- HORAS representa el número de horas que el empleado ha trabajado la presente semana.
- PAYRT está apuntado a PAYLVL indicando que el régimen de pago es transitivamente dependiente del nivel de pago.

La línea de trazos que une PAYRT y HORAS indica que ambas participan en una dependencia completa por la cual el receptor es PAYAMT, el valor pagado para esta semana.

A la derecha de la figura, encontramos los atributos que dependen del número de departamento. Obsérvense la dependencia total entre número y nombre del jefe del mismo (MGRO y MGRNM).

Hay solamente un atributo que es completamente dependiente de ambas partes de la clave compuesta, es decir, el número de proyecto, PROJNO.

## 5. Segunda Forma Normal

Una relación está en segunda forma normal (2FN) solamente si todos los atributos son dependientes en forma completa de la clave.

### Descripción De La Segunda Forma Normal (2 F<sub>n</sub>)

Su nombre ya nos indica el hecho de que la segunda forma normal es por lo general el próximo paso de normalización y descomposición. Para ser accesible a la normalización, y poder ser puesta en segunda forma normal, la relación debe poseer las siguientes propiedades:

- Debe estar en primera forma normal
- Debe tener una clave compuesta.

La consecuencia inmediata de los requerimientos expresados más arriba es que cualquier relación en primera forma normal que tiene una clave simple, está automáticamente en segunda forma normal. Comencemos con un ejemplo en forma de tabla de una relación consistente en 17 atributos, que se presenta en la figura. La misma se encuentra en primera forma normal y tiene una clave compuesta que consiste en dos atributos P y Q. Estos están subrayados en la figura para mostrar que sirven como clave. La tupla de relación puede también escribirse linealmente en forma simbólicamente:

$$R = (A, B, C, D, E, F, G, H, I, L, M, N, O, \underline{P}, \underline{Q})$$

El próximo paso es crear un grafo de dependencia, presentando aquí como figura. Debe notarse que este grafo se crea examinando conocimientos y atributos para determinar como participan y relacionan entre ellos.

No resulta suficiente analizar la matriz de relación, la cual puede hacernos creer que existe una dependencia debido a que la muestra de la cual se ha extraído dicha relación es pequeña. Si somos inducidos a error por los datos existentes y construimos una dependencia donde esta no existe, se planteará un problema. Cuando lleguen nuevos datos que contradigan la dependencia, deberá dejarse de lado el esquema completo.

Supongamos en consecuencia que el grafo que se puede observar en la figura ha sido derivado en forma funcional y que expresa correctamente las dependencias. Resulta claro a partir de este grafo que los atributos que parten de P son dependientes solamente de este. De un modo similar los que parten de Q dependen solamente de este último. Solamente aquellos que parten de la línea de trazos que conecta a P y Q tienen dependencia completa de ambos. Esta es la guía para la descomposición.

#### Descomposición

La figura contiene 3 sub-árboles, la base de nuestra descomposición. Definimos una sub tupla general en base a cada sub-árbol y en consecuencia:

$$P' = (P, A, B, C, E, H, K)$$

$$Q' = (Q, F, G, J, N)$$

$$PQ = (P, Q, D, I, L, M, O)$$

Aquí la raíz de los sub-árboles de la izquierda y la derecha. P y Q, se convierte en la clave de sus respectivas sub tuplas; ambos. P y Q forman la clave compuesta para la sub tupla PQ.

#### Proyección

El próximo paso es proyectar la relación R sobre cada una de estas sub tuplas para formar tres nuevas relaciones, y en consecuencia.

$$P' = \text{proyectar } R(P')$$

$$Q' = \text{proyectar } R(Q')$$

$$PQ = \text{proyectar } R(PQ)$$



Las relaciones así formadas nos dan tres nuevas sub-relaciones. Una subrelación es la relación que deriva de una relación mayor. Las subrelaciones ilustradas en la figura están correlacionadas por medio de los componentes de sus claves. La clave compuesta P y Q de la relación original R, es también la clave de la sub-relación PQ. P y Q tienen a P y Q respectivamente como claves. La línea de trazos en la figura indica que Q está correlacionada con PQ por medio de la componente Q y P está correlacionada con PQ por medio de P.

Para restablecer la relación original R debemos juntar estas tres subrelaciones en algún orden, indicado simbólicamente como:

$$R = \text{juntar } P [\text{juntar } PQ, (Q)] (P) = \text{juntar } Q[\text{juntar } PQ P(P)] (Q).$$

## Grafos

La nueva sub-relación que se ve en la figura se presenta en forma de grafo en la figura siguiente. Existe una considerable analogía entre estas figuras y la figura anterior. Lo importante es la diferencia. En PQ existe una línea de trazos que conecta los componentes de la clave compuesta P y Q en el centro de la figura. Los arcos parten del centro de esta línea de trazos hacia todos los componentes de P y Q, los cuales son dependientes en forma completa de ambos, es decir de P y Q. Una línea de puntos conecta P en la relación PQ a P de la relación P. Esto representa la correspondencia entre ambas veces P. Una línea de puntos conecta de un modo similar Q en PQ a Q en Q para indicar una correspondencia similar.

## Efectos

El efecto de esta descomposición puede no resultar inmediatamente claro. Debemos insistir en que ninguna relación correcta debe contener tuplas duplicadas. La relación original R contiene muchas subtuplas duplicadas P' y Q'. Las mismas han sido eliminadas durante la descomposición. Esto facilita en forma extraordinaria la actualización y otras importantes operaciones que afectan a estas relaciones, las cuales serán aclaradas en los ejemplos que siguen.

## Ejemplo de inventario

Vamos a utilizar ahora un ejemplo práctico para demostrar la normalización. En la figura se observa una parte de la matriz de relación PW.

Pueden verse los nombres de los atributos simbólicos y sus significados, pero no sus valores. Las columnas no aparecen en ningún orden en particular. Debe observarse la clave compuesta que distingue cada tupla, que abarca el número de pieza y el número de depósito PNO y WNO.

## Arbol de Dependencia

El medio para descomponer la relación es el árbol de dependencia que se ve en la figura. Este árbol ha sido construido solamente teniendo en cuenta la dependencia completa, y no muestra las dependencias total o transitiva, que se describe más adelante, si es que las mismas existen.

Como podíamos esperar, aparecen tres sub-árboles. El sub-árbol de la izquierda, con raíz PNO, contiene los atributos que se aplican solamente a la pieza o parte. El sub-árbol de la derecha con raíz WNO describe cada depósito. El sub-árbol del centro corresponde a las partes y al depósito, y describe la cantidad de partes disponibles en el depósito, QOH, y el número de cajón o estante, BIN (o algún otro parámetro de ubicación), donde dichas partes pueden ser halladas.

El próximo paso es definir tres tuplas generales para cada sub-árbol,

$$P = (PNO, DESC, PR, UNIT)$$
$$W = (WNO, WAD, FUE)$$
$$P/W = (PNO, WNO, BIN, QOH)$$

La descomposición consiste en proyectar la relación PW sobre cada una de estas tuplas para obtener tres nuevas sub-relaciones:

$$P = \text{proyectar } PW(P)$$
$$W = \text{proyectar } PW(W)$$
$$P/W = \text{proyectar } PW(P/W)$$

La descomposición en la figura muestra las tres relaciones como matrices; la línea de trazos indica como se vinculan las relaciones.

## Efecto

Discutiremos ahora algunas de las ventajas obtenidas mediante la descomposición. Si estas relaciones se utilizan para el control de inventario, nuestra preocupación será cuantas piezas de cada tipo están disponibles en un depósito en particular. Cuando se retiran piezas o se reciben nuevos envíos la cantidad disponible, QOH será la variable de cambio. La actualización consiste en poner al día sub-relación P/W la cual ahora contiene solamente malos componentes en lugar de los nuevos P/W.

Existe una tupla  $P$  en la sub-relación de pieza o parte,  $P$ , para cada parte y una tupla  $W$ , en la sub-relación  $W$ , para cada depósito y estos últimos probablemente no serán muchos. Consideremos la facilidad de efectuar cambios en un depósito en particular. Si un atributo de uno de los depósitos varía entraremos en  $W$  para efectuar el cambio solamente en una tupla. En la primera forma normal para  $PW$  teníamos que encontrar todas las tuplas en las cuales el valor de  $WNO$  es el particularmente deseado, y efectuar el mismo cambio en cada una de ellas. Si dicho depósito almacenaba 100 partes, como consecuencia debía variar 100 tuplas de  $PW$ . El procedimiento de actualización se aplica también a las descripciones de partes. Si el precio de alguna parte o pieza cambia, este cambio es independiente del depósito en el cual se almacena dicha parte. Solamente se efectúa un cambio en  $P$  a diferencia de los muchos que hubieran sido requeridos para  $PW$ .

## 6. Tercera forma normal

Una relación se encuentra en tercera forma normal (EFN) si no existen transitividades entre sus atributos y si ya se encuentra en 2 FN.

### Descripción

Una relación  $R$  a poner en tercera forma normal debe estar en la segunda forma normal. Es muy común que  $R$  sea una sub-relación; la relación original estaba en primera forma normal (para ponerla en segunda forma normal fue descompuesta en varias sub-relaciones). Estas son ahora candidatas a una descomposición adicional.

Recordamos que las propiedades de la segunda forma normal (2Fn) son:

- Tenemos una matriz  $m \times n$  con un valor determinado para cada componente de cada tupla.
- Cada valor es obtenido a partir de un dominio propiamente definimos
- Cada valor contiene una clave, ya sea simple o compuesta
- Cada componente no clave es dependiente en forma completa de su clave.

En consecuencia es evidente que tenemos, o bien una clave simple, o una clave compuesta de la cual todos los componentes no clave son dependientes en forma completa.

El objeto de esta fase es determinar todas las dependencias transitivas; la descomposición producirá a continuación sub-relaciones para las cuales no existirán dependencias transitivas -la definición de la tercera forma normal (EFN)-.

Una dependencia transitiva abarca como mínimo tres componentes. Si los componentes fueran más, la dependencia múltiple puede derivarse en varias dependencias atransitivas de tres componentes solamente dada una. Por lo tanto dirigiremos nuestra atención a una dependencia transitiva simple de tres componentes. Tal dependencia puede expresarse como:

$$Q \rightarrow A \rightarrow B$$

En la cual se dice que B depende de A y que A depende de Q. La transitividad existe debido a que el valor de B depende en la última instancia del valor de Q.

La dependencia transitiva es degenerada si cualquiera de las dependencias anteriores es total. Esto es, podemos prever que la relación de Q a A es muchos-unos, donde varios valor único de A. Dado un valor tal Q el valor de A queda determinado. La inversa no se aplica y en consecuencia no existe una dependencia total: dado un valor de A el valor correspondiente de Q no queda determinado a menos de que se trate de una dependencia total.

El ahorro que surge de colocar la relación en tercera forma normal aparece a raíz de la granularidad del dominio involucrado. Se puede prever que:

$$\text{num dominio (Q)} > \text{num dominio (A)} > \text{num dominio (B)}$$

#### Determinación de la dependencia transitiva

Si el grafo utilizado para llevar la relación a la segunda forma normal es completo en termino de las transitividades existentes, no resulta necesario un grafo adicional. El grafo para convertir a la segunda forma normal requiere solamente que todas las dependencias completas y parciales sean conocidas. Supongamos que no hemos establecido todas las dependencias transitivas. Se presenta una situación simple en la figura anterior donde A, B y C son dependientes de Q. Si suponemos que existe una dependencia entre A, B y C son dependientes de Q. Si suponemos que existe una dependencia entre A y B debemos confirmarlo en forma funcional.

Una dependencia total entre A y B en el grafo de la figura puede representarse como se ve en la figura el arco desde A a B no muestra una dependencia de B respecto de A inversamente el arco a partir de B hacia A muestra una dependencia de A respecto de B; los arcos a partir de Q a A y a B nos muestra la dependencia de cada una de éstas respecto de Q. Esto puede observarse nuevamente en la figura, donde una doble arista entre A y B indica la bi-direccionalidad de esta

dependencia. El hecho de que Q apunte a esta arista nos muestra que cada una de las variables A y B es claramente dependiente de aquella.

Como ejemplo sea Q el número PO, A el número de parte o pieza y B el nombre de parte, A y B son totalmente dependientes y cada uno dependen de Q.

#### Transitividad simple

Para la dependencia transitiva unilateral, la variable independiente apunta a la variable dependiente, tal cual se presenta en el figura donde B depende de A. El arco entre B y Q ha sido eliminando; la dependencia implícita de B respecto de Q resulta obvia.

Si se presenta la dependencia inversa, debe grafificarse como se ve en la figura.

#### Descomposición

Dada una sub-relación con una o más dependencias transitivas, la descomposición consiste en partir la relación en una o más de una sub-relación, donde la variable intermedia aparezca como variable dependiente en una y como variable independiente en la otra.

Caso simple Tenemos:

$$Q \twoheadrightarrow A \twoheadrightarrow B$$

$$Q \twoheadrightarrow C$$

Dado que ambas, A y C dependen directamente de Q deben conservarse en una sub-relación Q, con clave Q.:

$$Q \twoheadrightarrow A; Q \twoheadrightarrow C$$

Debe separarse la relación directa remanente, y colocarla en su propia sub-relación A' con la A:

$$A \twoheadrightarrow B$$

Los grados de Q' y A'. Aquí la componente A relaciona Q' con A, a es la clave simple de A'. Si bien A no es la clave de Q' es le medio de relacionar un valor de Q en Q' con un valor de B en A' y se llama por lo tanto la clave externa de Q'. Para crear Q' y A' debemos utilizar las subtuplas generales Q' y A' denifidas en consecuencia:

$$Q' = (Q, A, C)$$

$$A' = (A, B)$$

donde el subrayado indica una clave.

Estos deben proyectarse sobre Q para obtener las sub-relaciones:

$$Q' = \text{proyectar } Q(Q')$$

$$A' = \text{proyectar } Q(A')$$

### Caso Compuesto

Las dependientes transitivas múltiples han sido investigadas y exhibidas. Tenemos en consecuencia.

$$Q \twoheadrightarrow C$$

$$Q \twoheadrightarrow A \twoheadrightarrow B1$$

$$Q \twoheadrightarrow A \twoheadrightarrow B2$$

$$Q \twoheadrightarrow A \twoheadrightarrow B3$$

La descomposición separa nuevamente todas estas variables directamente dependiente de la clase original en una subtupla.  $Q'' = (Q, A, C)$

Las variables restantes son todas dependientes directa o totalmente de A o C y se reorganizan de un modo similar.  $A'' = (A, B1, B2, B3); C'' = (C, D)$

Deben construirse tres sub-relaciones por proyección:

$$Q'' = \text{proyectar } Q(Q'')$$

$$A'' = \text{proyectar } Q(A'')$$

$$C'' = \text{proyectar } Q(C'')$$

Aquí Q'', A'' y C'' aparecen como sub-árboles. Las mismas se relacionan por medio de la clave externa de Q'' es decir A'' y C''; esto se muestra mediante la línea de puntos entre A'' y A'' y entre C'' y C''. Nos podemos mover directamente entre las dos figuras sin la intervención de pasos simbólicos, utilizando solamente manipulaciones gráficas.

### Descomposición Gráfica

Hemos discutido el enfoque simbólico. Dado un grafo 2FN. Debemos seleccionar en primer término los nodos apuntados por la raíz que no sean hojas. Los mismos se convierten en raíces de sus propios sub-árboles, A'' y C''. Estos sub-árboles son eliminados de Q dejando en Q'' solamente los nodos A'' y C'', que son las raíces de A''; y C''.

## Ejemplo de orden de compra

Examinaremos solamente una pequeña porción de la relación orden de compra que ha sido convertida en un grafo de dependencia. Para esta porción de la relación compra PP, tenemos:

- Las partes se compran utilizando el número de parte, PNO;
- Un vendedor, VNDR está asociado a cada parte;
- Cada vendedor tiene una clasificación de forma de pago, PAYCLS.

Por lo tanto PAYCLS representa si el vendedor debe cobrar dentro de los 10 días, 30 días, 60 días, etc. La acción para convertir la relación.

Tenemos aquí una relación transitiva que puede ser representada en consecuencia:

PNO ---> PAYCLS

Sabemos que la variable intermedia, el vendedor VNDR, es el que determina el tipo de pago de modo tal que

PNO ---> VNDR --> PAYCLS

para poner esta relación en la tercera forma normal, la misma se descompone en dos sub-relaciones. Las dos sub-relaciones PV y VP, se forman por proyección a partir de la relación original PP de modo tal que:

PV = proyectar PP (PNO, VNDR); PV = proyectar PP (VNDR, PAYCLS).

La relación PV relaciona partes con vendedores.

La identificación del vendedor, VNDR es la clave externa par PV. La misma se utiliza para entrar en la relación VP, en la cual es la clave primaria.

Debe notarse que, para el mantenimiento, si cambia la clase de pago solamente cambiara una entrada o tupla en VP y ninguna en PV. Para el caso de PP hubiera cambiado muchas tuplas.

## Ejemplo de inventario

Presentamos ahora una porción de un ejemplo de inventario, al cual corresponde el grafo parcial. Tenemos en este caso:

- PNO es un número de parte
- PNM es el nombre de parte y tiene dependencia total con el número de parte
- PREC es el costo de UNITS multiplicado por el número de partes
- PCL es la clase de parte, la cual da el tipo de parte en términos de su peso y de su forma.

- WHN es el número de depósito donde está almacenada la parte.
- WHLOC Es la ubicación del deposito
- FUE es la categoría de seguro de incendio del depósivto.

Resulta claro a partir del grafo que el número de parte determina la clasificación de la parte, la cual a su vez determina parcialmente el deposito donde está almacenada dicha parte. Usaremos esta dependencia transitiva, que está circundada con línea de trazos gruesos, para descomponer la relación en su tercera forma normal:  $PNO \text{ ---> } WHN$ ;  $PNO \text{ ---> } PCL \text{ ---> } WHN$

La variable intermedia, clase de parte, PCL, es el medio de que disponemos para descomponer el grafo. Se deja como ejercicio hallar las proyecciones y la relaciones resultantes.

#### Ejemplo bancario

Consideremos parte de un ejemplo de banco donde cada depositante tiene un número de cuenta que lo identifica. El depositante recibe una línea de crédito. Puede extraer dinero hasta dicho valor. La parte no utilizada de crédito puede ser retirada cuando lo desee. Vemos que la línea de crédito LNCR es funcionalmente dependiente del número de cuenta CUET; el valor ya extraído DEBIT es también dependiente del número de cuenta. El valor de crédito disponible en este momento, DISP, es dependiente en forma completa de ambos, LNCR y DEBIT.

Parecería que lo lógico es descomponer el grafo y volver a presentarlo. En base a esto, P tiene como clave el número de cuenta CUENT. Debemos entrar en P para obtener LNCR y DEBIT. Estas son claves externas para P; las mismas forman la clave compuesta para entrar en Q y hallar el valor de la variable completamente dependiente DISP.

Esto funcionaría, pero hay una forma más simple de resolver el problema. El valor de crédito disponible en la actualidad es simplemente la diferencia entre la línea de crédito y el debido corriente. Todo lo que tenemos que hacer es ejecutar una sustracción. La relación original no necesita contener DISP. dado que éste se calcula simplemente durante el procesamiento. Por lo tanto podemos sencillamente omitir Q.

Transitivas múltiples.



Establecemos de entrada la condición simple de que Z sea dependiente en forma transitiva de Q. Si existe más de una variable intermedia de dependencia, la transitiva no será completa hasta que se especifiquen todas dichas variables. Es decir, si bien empezamos con la condición de transitividad,

$$Q \longrightarrow Z,$$

la condición completa podría ser,  $Q \longrightarrow X \longrightarrow Y \longrightarrow Z$

Ninguna condición intermedia  $Q \longrightarrow X \longrightarrow Z \longrightarrow; Q \longrightarrow Y \longrightarrow Z$

sería suficiente para descomponer la original de la figura.

## 7. Cuarta forma normal

### Dependencias multivaluadas

La tercera forma normal toma en cuenta la dependencia transitiva y provee una reducción óptima universal, excepto para los casos infrecuentes de dependencia multivaluadas. Ha quedado claro en épocas recientes que es posible una reducción adicional en este caso, y esto es lo que se lleva a cabo mediante la cuarta forma normal.

Existe una dependencia multivaluada cuando un valor de una variable está siempre asociado con varios valores de otra u otras variables dependientes que son siempre las mismas y están siempre presentes. Esto se ilustra mejor con el ejemplo presentado en la figura. La relación FAB describe tejidos. La variable independiente (con respecto a las dependencias (multivaluadas) es el número de tejido FABNO. Con el se encuentra asociados un modelo (o patrón) y un color. En la figura, el tejido 345 vienen en dos modelos y entres combinaciones de modelo y color. En este caso se aplica el grafo de dependencia. Para hacer mas clara que esta es una dependencia multivariable, una cabeza doble de flecha apunta desde FABNO o PATRN y también desde FABNO a COLOR.

La ineficiencia en el registro de información y se resulta clara al examinar las dos nuevas relaciones. La primera de éstas, FABPAT lista el número de tejido contra el modelo; en el segundo caso, FABCOL, lista el número de tejido contra las combinaciones de color. Dado que la regla es que todas las combinaciones de las variables dependientes multivaluadas deben prevalecer, resulta simple reconstruir la relación FAB a partir de las dos sub-relaciones que resultaron.

**Descomposición** Para poner una relación o sub-relación en la cuarta forma normal debe poder aplicarse lo siguiente:

- Debe estar en la tercera forma normal.
- Deben existir una o mas multidependencias.

Después de construir el grafo de dependencia, el próximo paso es ejecutar proyecciones utilizando la variable independiente y una de las variables multidependientes.

FABPAT = proyectar FAB (FABNO, PATRN)

FABCOL = proyectar FAB (FABNO, COLOR)

El resultado son nuevas sub-relaciones que han sido utilizadas para ahorrar espacio y permitir una más fácil actualización.

#### Ejemplo de profesor y texto

Consideremos otro ejemplo. Los cursos dictados en una escuela corresponden a un número de curso. Asociada a cada número de curso se encuentra la descripción del mismo. Para cada curso existe una selección de textos y una selección de profesores. Puede darse cualquier combinación de texto y profesor.

El grafo de dependencia. El mismo nos muestra una dependencia total entre el número de curso y la descripción del curso. Existe una multidependencia entre texto y número de curso, y también entre profesor y número de curso.

Para descomponer la sub-relación en sus relaciones más pequeñas, se efectúan tres proyecciones. Las sub-relaciones resultantes.