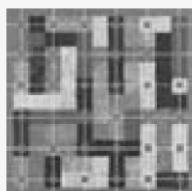


Uso de diagramas de flujo de datos

OBJETIVOS DE APRENDIZAJE

Al completar este capítulo usted podrá:

1. Comprender la importancia del uso de diagramas de flujo de datos (DFD) lógicos y físicos para describir gráficamente el movimiento de datos para humanos y sistemas en una organización.
2. Crear, usar y expandir los DFD lógicos para capturar y analizar el sistema actual por medio de niveles padre e hijo.
3. Desarrollar y expandir DFD lógicos que ilustren el sistema propuesto.
4. Producir DFD físicos con base en los DFD lógicos que desarrolle.
5. Comprender y aplicar el concepto de particionamiento de DFD físicos.



El analista de sistemas requiere la libertad conceptual que los diagramas de flujo ofrecen; éstos caracterizan gráficamente los procesos y flujos de datos en un sistema empresarial. En su estado original, los diagramas de flujo de datos describen con la mayor generalidad posible las entradas, los procesos y las salidas del sistema, las cuales corresponden a las del modelo de sistemas general que vimos en el capítulo 2. Podemos usar una serie de diagramas de flujo de

datos en capas para representar y analizar con detalle los procedimientos del sistema más amplio.

LA METODOLOGÍA DEL FLUJO DE DATOS PARA DETERMINAR LOS REQUERIMIENTOS HUMANOS

Para que los analistas de sistemas puedan comprender los requerimientos de información de los usuarios, deben ser capaces de conceptualizar la forma en que los datos se mueven a través de la organización, los procesos o la transformación por la que pasan los datos y las salidas de los mismos. Aunque las entrevistas y la investigación de datos “duros” proveen una narrativa verbal del sistema, una descripción visual puede cristalizar esta información para los usuarios y analistas de una manera útil.

Por medio de una técnica de análisis estructurado conocida como diagramas de flujo de datos (DFD), el analista de sistemas puede ensamblar una representación gráfica de los procesos de datos a través de la organización. Al usar combinaciones de sólo cuatro símbolos, el analista puede crear una descripción ilustrada de los procesos con el fin de elaborar una documentación sólida para el sistema.

Ventajas de la metodología del flujo de datos

La metodología del flujo de datos tiene cuatro ventajas importantes en comparación con las explicaciones narrativas sobre la forma en que se mueven los datos a través del sistema:

1. No hay que comprometerse demasiado pronto con la implementación técnica del sistema.
2. Permite comprender con más detalle la capacidad de interrelación de los sistemas y subsistemas.
3. Se puede comunicar el conocimiento del sistema actual a los usuarios por medio de diagramas de flujo de datos.
4. Se puede analizar un sistema propuesto para determinar si se han definido los datos y procesos necesarios.

Tal vez la mayor ventaja recaiga en la libertad conceptual que se obtiene al usar los cuatro símbolos (reconocerá tres de ellos del capítulo 2; la próxima subsección sobre las convenciones de los DFD aborda todos). Ninguno de estos símbolos especifica los aspectos físicos de la implementación. Los DFD se enfocan en el procesamiento de los datos o en la transformación de los mismos a medida que avanzan a través de varios procesos. En los DFD lógicos no hay distinción entre los procesos manuales o los automatizados. Tampoco se describen en forma gráfica los procesos en orden cronológico, sino que, en última instancia, se agrupan entre sí cuando un análisis posterior indique que es conveniente hacerlo. Se agrupan los procesos manuales; también los automatizados pueden asociarse entre sí. En una sección posterior veremos más detalles sobre este concepto, conocido como *particionamiento*.

Convenciones usadas en los diagramas de flujo de datos

Se utilizan cuatro símbolos básicos para graficar el movimiento de los datos en los diagramas: un cuadrado doble, una flecha, un rectángulo con esquinas redondas y un rectángulo con un extremo abierto (cerrado del lado izquierdo y abierto del lado derecho), como se muestra en la figura 7.1. Podemos describir en forma gráfica todo un sistema y numerosos subsistemas al combinar estos cuatro símbolos.




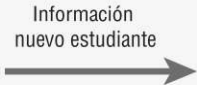

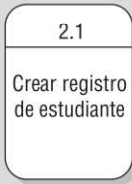

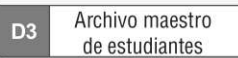
El cuadrado doble se utiliza para describir una entidad externa (otro departamento, una empresa, una persona o una máquina) que pueda enviar/recibir datos hacia/desde el sistema. La entidad externa, o simplemente entidad, también se conoce como origen o destino de los datos, y se considera externa al sistema que se está describiendo. Cada entidad se identifica con un nombre apropiado. Aunque interactúa con el sistema, se considera fuera de los límites de éste. Se debe denominar a las entidades con un sustantivo. Se puede utilizar la misma entidad más de una vez en un diagrama de flujo de datos para evitar cruzar las líneas de flujo de datos.

La flecha muestra el movimiento de los datos de un punto a otro; la cabeza de la flecha apunta hacia el destino de los datos. Los flujos de datos que ocurren al mismo tiempo se pueden describir mediante el uso de flechas paralelas. Como una flecha representa datos sobre una persona, lugar o cosa, también se debe describir con un sustantivo.

Se utiliza un rectángulo con esquinas redondas para mostrar la ocurrencia de un proceso de transformación. Los procesos siempre expresan un cambio o transformación en los datos; por ende, el flujo de datos que sale de un proceso *siempre* se identifica de manera distinta al flujo que entra al proceso. Los procesos representan el

FIGURA 7.1

Los cuatro símbolos básicos que se utilizan en los diagramas de flujo de datos, sus significados y ejemplos.

Símbolo	Significado	Ejemplo
	Entidad	
	Flujo de datos	
	Proceso	
	Almacén de datos	

trabajo que se realiza en el sistema y se deben denominar mediante el uso de uno de los siguientes formatos. Un nombre claro facilita la acción de entender lo que el proceso lleva a cabo.

1. Al denominar un proceso de alto nivel, asigne al proceso el nombre de todo el sistema. Por ejemplo, SISTEMA DE CONTROL DE INVENTARIO.
2. Al denominar un subsistema importante, use un nombre tal como SUBSISTEMA DE INFORME DE INVENTARIOS o SISTEMA DE CUMPLIMIENTO CON CLIENTES DE INTERNET.
3. Al denominar procesos detallados, use una combinación verbo-sustantivo-adjetivo. El verbo describe el tipo de actividad, como CALCULAR, VERIFICAR, PREPARAR, IMPRIMIR o AGREGAR. El sustantivo indica cuál es el resultado principal del proceso, como INFORME o REGISTRO. El adjetivo ilustra la salida específica que se produce, como PEDIDO PENDIENTE o INVENTARIO. Algunos ejemplos de nombres de procesos completos son CALCULAR IMPUESTO DE VENTAS, VERIFICAR ESTADO DE CUENTA DE CLIENTE, PREPARAR FACTURA DE ENVÍO, IMPRIMIR INFORME DE PEDIDOS PENDIENTES, ENVIAR CONFIRMACIÓN POR EMAIL AL CLIENTE, VERIFICAR SALDO DE TARJETA DE CRÉDITO y AGREGAR REGISTRO DE INVENTARIO.

Un proceso también debe recibir un número de identificación único que indique su nivel en el diagrama. Más adelante en el capítulo hablaremos sobre esta organización. Puede haber varios flujos de datos que entren y salgan de cada proceso. Examine los procesos que tengan sólo un flujo entrante y saliente para determinar si no hacen falta más flujos de datos.

El último símbolo básico que se utiliza en los diagramas de flujo de datos es un rectángulo con un extremo abierto, el cual representa a un almacén de datos. El rectángulo se dibuja con dos líneas paralelas que se cierran mediante una línea corta del lado izquierdo y cuyo extremo derecho está abierto. Estos símbolos se dibujan con la anchura suficiente como para permitir una leyenda de identificación entre las líneas paralelas. En los diagramas de flujo de datos lógicos no se especifica el tipo de almacenamiento físico. En este punto, el símbolo del almacén de datos muestra sólo un depósito de datos que permite examinar, agregar y recuperar los datos.

El almacén de datos puede representar un almacén manual como un archivero, o un archivo o una base de datos computarizada. Como los almacenes de datos representan a una persona, lugar o cosa, se denominan con un sustantivo. Los almacenes de datos temporales, como el papel de borrador o un archivo temporal de computadora, no se incluyen en el diagrama de flujo de datos. Hay que dar a cada almacén de datos un número de referencia único, como D1, D2, D3, por ejemplo.

CÓMO DESARROLLAR DIAGRAMAS DE FLUJOS DE DATOS

Podemos y debemos dibujar los diagramas de flujos de datos en forma sistemática. En la figura 7.2 se sintetizan los pasos involucrados para completar con éxito los diagramas de flujos de datos. En primer lugar, el analista de sistemas necesita conceptualizar los flujos de datos desde una perspectiva de arriba-abajo.

Para empezar un diagrama de flujo de datos, contraiga la narrativa (o historia) del sistema de la organización en una lista con las cuatro categorías de entidad externa, flujo de datos, proceso y almacén de datos. A su vez, esta lista ayuda a determinar los límites del sistema que va a describir. Una vez que haya compilado una lista básica de elementos de datos, empiece a dibujar un diagrama de contexto.

He aquí unas cuantas reglas básicas a seguir:

1. El diagrama de flujo de datos debe tener por lo menos un proceso y no debe haber objetos independientes o conectados a sí mismos.
2. Un proceso debe recibir por lo menos un flujo de datos entrante y debe crear por lo menos un flujo de datos saliente.
3. Un almacén de datos debe estar conectado con por lo menos un proceso.
4. Las entidades externas no se deben conectar entre sí. Aunque se comunican en forma independiente, esa comunicación no forma parte del sistema que diseñamos mediante el uso de DFD.

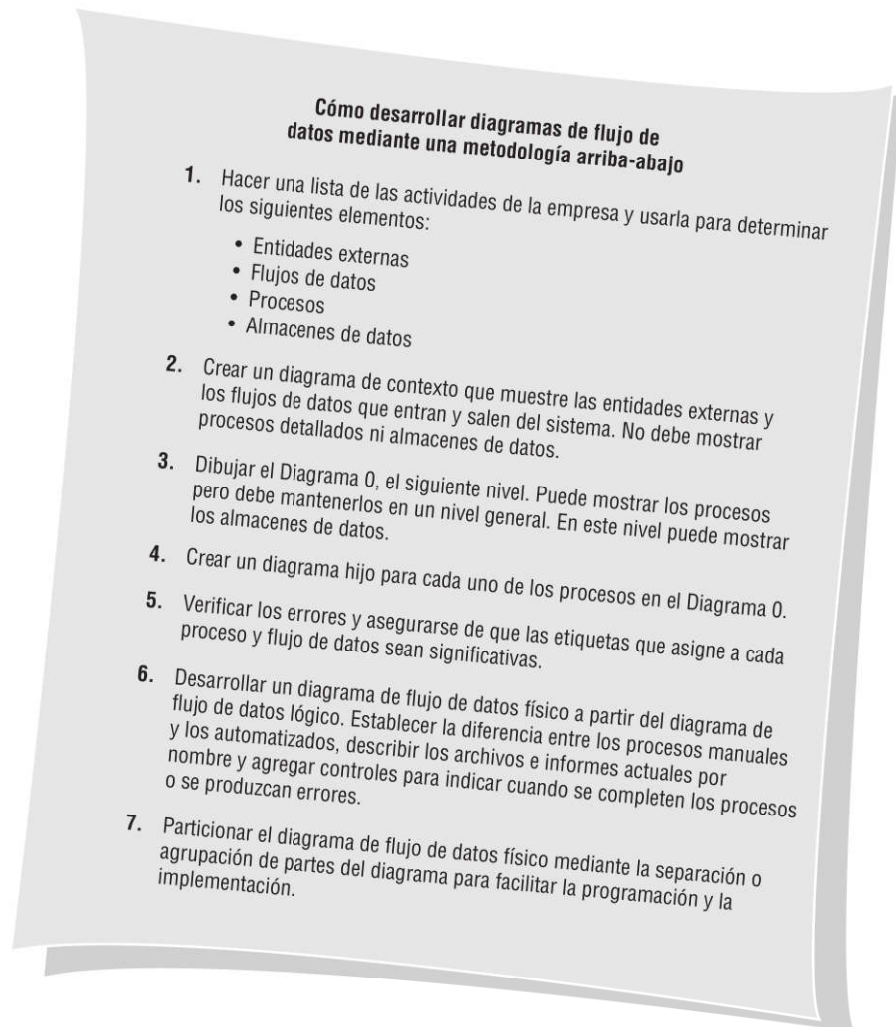
Creación del diagrama de contexto

Con una metodología arriba-abajo para crear un diagrama del movimiento de los datos, los diagramas avanzan de generales a específicos. Aunque el primer diagrama ayuda al analista de sistemas a comprender el movimiento de datos básico, su naturaleza general limita su utilidad. El diagrama de contexto inicial debe ser una vista general que incluya las entradas básicas, el sistema general y las salidas. Este diagrama será el más general, una verdadera vista panorámica del movimiento de datos en el sistema y la conceptualización más amplia posible del sistema.

El diagrama de contexto es el nivel más alto en un diagrama de flujo de datos y contiene sólo un proceso, el cual representa a todo el sistema. El proceso recibe el número cero. Todas las entidades externas se muestran en

FIGURA 7.2

Pasos para desarrollar diagramas de flujo de datos.



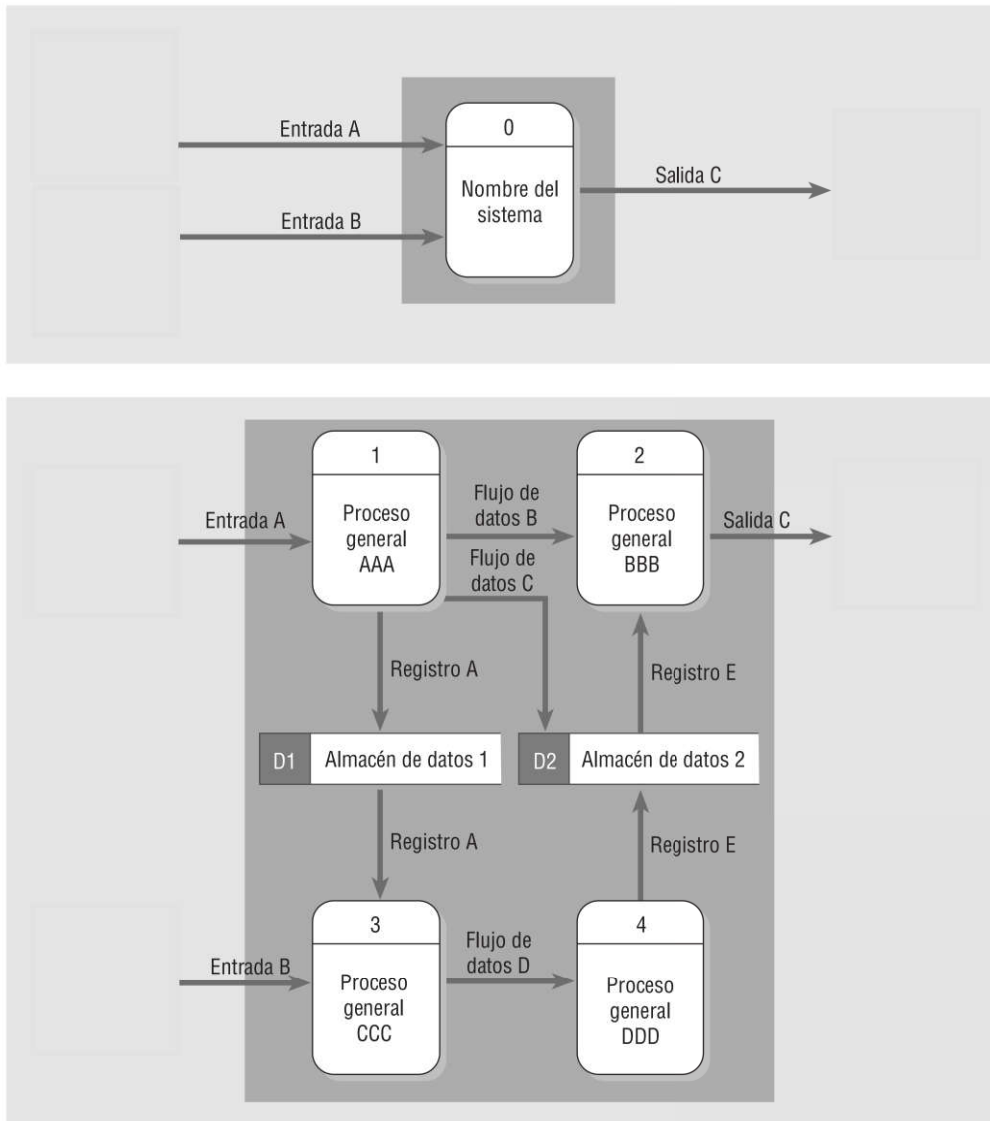
el diagrama de contexto, así como el flujo de datos principal que entra y sale de ellas. El diagrama no contiene almacenes de datos y es bastante simple de crear una vez que los analistas conocen las entidades externas y el flujo de datos que entra y sale de ellas.

Dibujo del Diagrama 0 (el siguiente nivel)

Podemos obtener más detalles de lo que permite el diagrama de contexto si “expandimos” los diagramas. Las entradas y salidas especificadas en el primer diagrama permanecen constantes en todos los subsiguientes. Sin embargo, el resto del diagrama original se expande en acercamientos que incluyan de tres a nueve procesos y muestren los almacenes de datos, junto con los nuevos flujos de datos de niveles inferiores. El efecto es como si tomáramos una lupa para ver el diagrama de flujo original. Cada diagrama expandido debe ocupar sólo una hoja de papel. Al expandir los DFD en subprocesos, el analista de sistemas empieza a llenar los detalles sobre el movimiento de los datos. El manejo de las excepciones se ignora durante los primeros dos o tres niveles en los diagramas de flujo de datos.

El Diagrama 0 es la expansión del diagrama de contexto; puede incluir hasta nueve procesos. Si incluimos más procesos en este nivel obtendremos un diagrama abarrotado de información que será difícil de comprender. Cada proceso se enumera con un entero, por lo general empezando a partir de la esquina superior izquierda del diagrama y avanzando hacia la esquina inferior derecha. En el Diagrama 0 se incluyen los principales almacenes de datos del sistema (que representan a los archivos maestros) y todas las entidades externas. En la figura 7.3 se ilustra de manera esquemática el diagrama de contexto y el Diagrama 0.

Como un diagrama de flujo de datos es bidimensional (en vez de lineal), tal vez quiera empezar en cualquier punto y avanzar hacia delante o hacia atrás por el diagrama. Si no está seguro de lo que podría incluir en cualquier punto, tome una entidad externa, proceso o almacén de datos distinto y empiece a dibujar el flujo a partir de él. Usted puede:

**FIGURA 7.3**

Los diagramas de contexto (superior) se pueden “expandir” en un Diagrama 0 (inferior). Observe el mayor detalle en el Diagrama 0.

1. Empezar con el flujo de datos proveniente de una entidad en el lado de entrada. Haga preguntas tales como: “¿Qué ocurre con los datos que entran al sistema?”, “¿Se guardan?”, “¿Constituyen la entrada para varios procesos?”.
2. Trabaje en sentido inverso desde un flujo de datos de salida. Examine los campos de salida en un documento o pantalla (este método es más sencillo si se crearon prototipos). Para cada campo en la salida pregunte lo siguiente: “¿De dónde proviene?” o “¿Se calcula o se guarda en un archivo?”. Por ejemplo, cuando la salida es un CHEQUE DE NÓMINA, el NOMBRE DE EMPLEADO y la DIRECCIÓN se ubicarían en un archivo de EMPLEADO, las HORAS TRABAJADAS estarían en un REGISTRO DE TIEMPO y se calcularían el SUELDO BRUTO y las DEDUCCIONES. Cada archivo y registro estaría conectado al proceso que produce el cheque de nómina.
3. Examine el flujo de datos que entra o sale de un almacén de datos. Pregunte: “¿Qué procesos colocan datos en el almacén?” o “¿Qué procesos utilizan los datos?”. Tenga en cuenta que un almacén de datos que se utilice en el sistema en el que usted esté trabajando puede ser producido por un sistema distinto. Por ende, desde su posición de ventaja, tal vez no haya ningún flujo de datos que entre al almacén de datos.
4. Analice un proceso bien definido. Examine los datos de entrada que necesita el proceso y la salida que produce. Después conecte la entrada y la salida a los almacenes de datos y las entidades apropiadas.
5. Tome nota de cualquier área confusa en donde no esté seguro de lo que se debería incluir o de la entrada o salida requerida. Al estar consciente de las áreas problemáticas podrá formular una lista de preguntas para las entrevistas de seguimiento con los usuarios clave.

Creación de diagramas hijos (niveles más detallados)

Cada proceso en el Diagrama 0 puede a su vez expandirse para crear un diagrama hijo más detallado. Al proceso que se expande en el Diagrama 0 se le conoce como el *proceso padre*, y al diagrama que resulta se le conoce como el *diagrama hijo*. La regla principal para crear diagramas hijos es el balanceo vertical; esta regla establece que un diagrama hijo no puede producir salida o recibir entrada que el proceso padre no produzca o reciba también. Todos los datos entrantes o salientes del proceso padre deben mostrarse como entrantes o salientes en el diagrama hijo.

El diagrama hijo recibe el mismo número que su proceso padre en el Diagrama 0. Por ejemplo, el proceso 3 se expande en el diagrama 3. Los procesos en el diagrama hijo se enumeran mediante el uso del número del proceso padre, un punto decimal y un número único para cada proceso hijo. En el diagrama 3 los procesos se enumerarían como 3.1, 3.2, 3.3 y así en lo sucesivo. Esta convención permite al analista rastrear una serie de procesos a través de muchos niveles de explosión. Si el Diagrama 0 describe los procesos 1, 2 y 3, los diagramas hijos 1, 2 y 3 se encuentran todos en el mismo nivel.

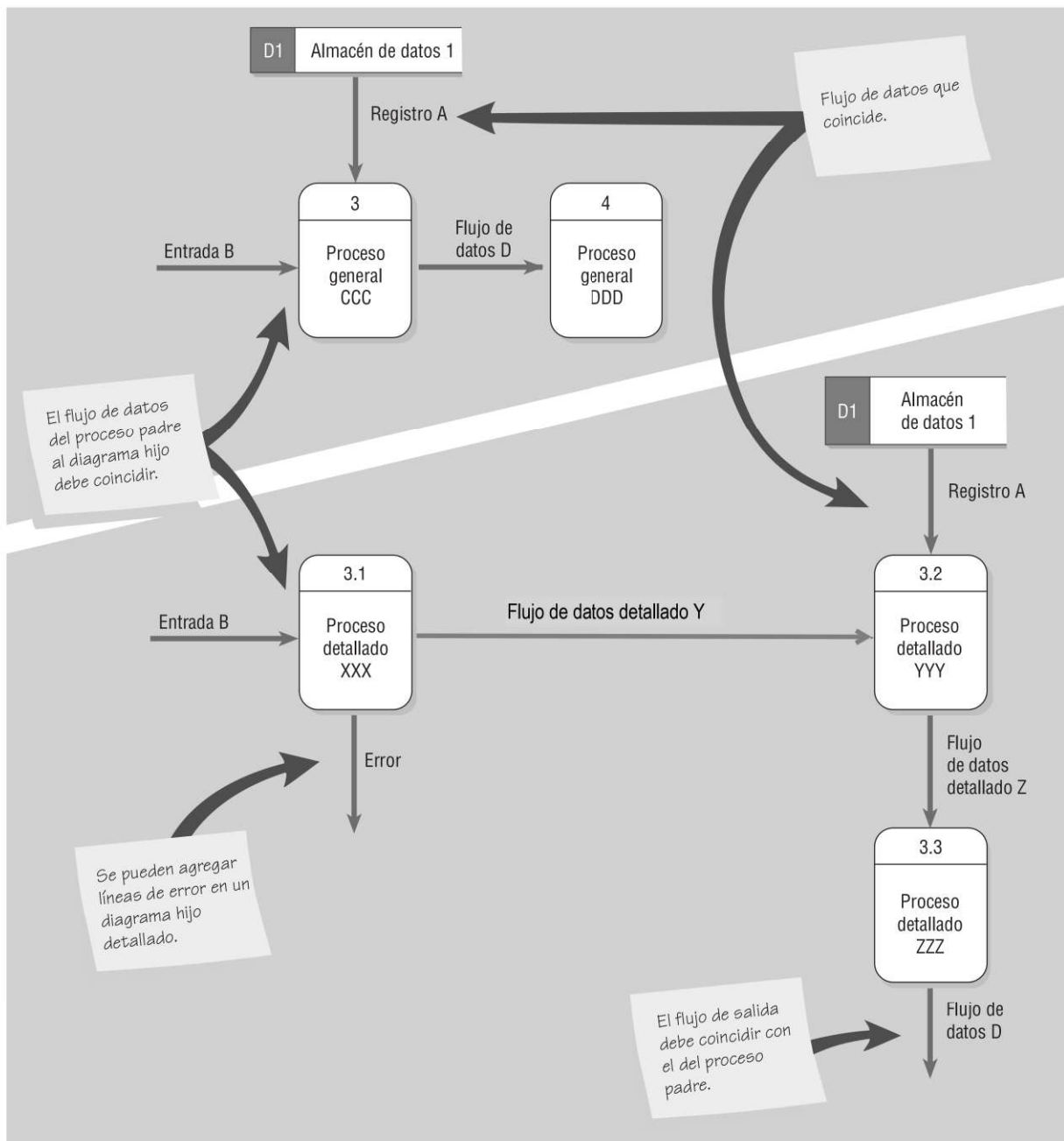
Por lo general, las entidades no se muestran en los diagramas hijos debajo del Diagrama 0. El flujo de datos que concuerda con el flujo padre se denomina *flujo de datos de interfaz* y se muestra como una flecha que entra o sale de un área en blanco del diagrama hijo. Si el proceso padre tiene un flujo de datos que lo conecta con un almacén de datos, el diagrama hijo puede incluir el almacén de datos también. Además, este diagrama de nivel inferior puede contener almacenes de datos que no se muestren en el proceso padre. Por ejemplo, se puede incluir un archivo que contenga una tabla de información tal como una tabla de impuestos, o un archivo que vincule dos procesos en el diagrama hijo. Los flujos de datos menores, como una línea de error, se pueden incluir en un diagrama hijo pero no en el padre.

Los procesos se pueden o no expandir, dependiendo de su nivel de complejidad. Cuando un proceso no se expande, se dice que es funcionalmente primitivo y se le denomina *proceso primitivo*. Hay que escribir lógica para describir estos procesos; veremos esto con detalle en el capítulo 9. En la figura 7.4 se ilustran los niveles detallados en un diagrama de flujo de datos hijo.

Comprobación de errores en los diagramas

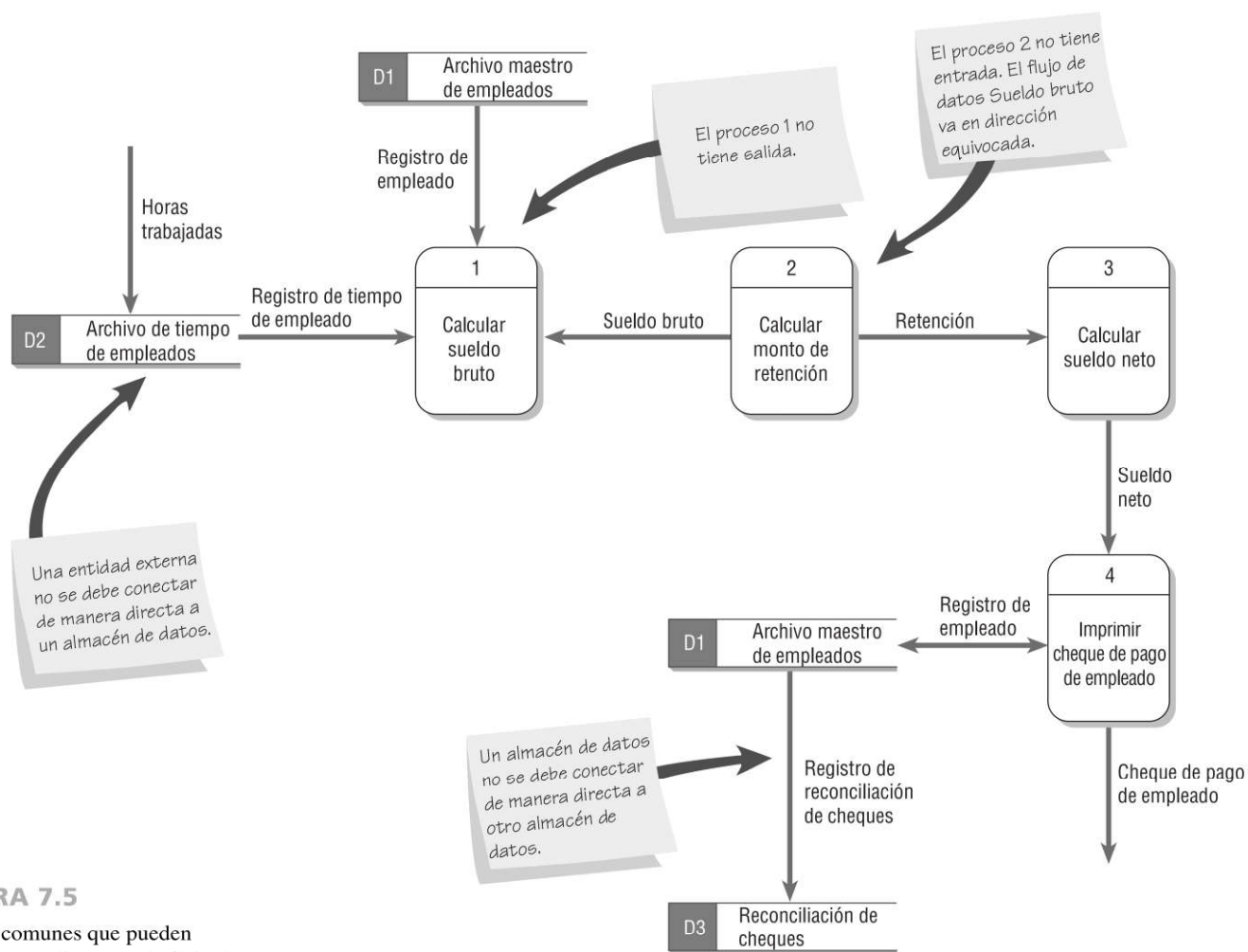
Varios de los errores que se cometen al dibujar diagramas de flujo de datos son:

1. Olvidar incluir un flujo de datos o apuntar una flecha en dirección equivocada. Un ejemplo es dibujar un proceso que muestra a todos sus flujos de datos como entradas o a todos como salidas. Cada proceso transforma los datos: debe recibir datos de entrada y producir datos de salida. Por lo general este tipo de error ocurre cuando el analista olvida incluir un flujo de datos o coloca una flecha apuntando en dirección equivocada. El proceso 1 en la figura 7.5 sólo contiene una entrada debido a que la flecha de SUELDO BRUTO apunta en dirección equivocada. Este error también afecta al proceso 2, CALCULAR MONTO A RETENER, al cual también le falta un flujo de datos que representa la entrada para las tasas de retención y el número de dependientes.
2. Conectar almacenes de datos y entidades externas directamente entre sí. No se pueden conectar los almacenes de datos y las entidades entre sí; se deben conectar sólo mediante un proceso. Un archivo no se interconecta con otro archivo sin la ayuda de un programa o una persona que mueva los datos, por lo cual ARCHIVO MAESTRO DE EMPLEADOS no puede producir directamente el archivo RECONCILIACIÓN DE CHEQUES. Las entidades externas no trabajan directamente con archivos. Por ejemplo, no es conveniente que un cliente hurgue por el archivo maestro de clientes. Por lo tanto, el EMPLEADO no puede crear el ARCHIVO DE TIEMPO DE EMPLEADOS. Dos entidades externas conectadas en forma directa indican que desean comunicarse entre sí. Esta conexión no se incluye en el diagrama de flujo de datos a menos que el sistema facilite la comunicación. La producción de un informe es una instancia de este tipo de comunicación. Sin embargo, debe haber un proceso interpuesto entre las entidades para producir el informe.
3. Etiquetar de manera incorrecta los procesos o el flujo de datos. Inspeccione el diagrama de flujo de datos para asegurar que cada objeto o flujo de datos esté identificado en forma apropiada. Un proceso debe indicar el nombre del sistema o debe usar el formato verbo-sustantivo-adjetivo. Cada flujo de datos se debe describir con un sustantivo.
4. Incluir más de nueve procesos en un diagrama de flujo de datos. Al tener muchos procesos se produce un diagrama sobrecargado de información que puede confundirnos al tratar de leerlo y entorpece la comunicación en vez de mejorarla. Si hay más de nueve procesos involucrados en un sistema, agrupe algunos de los procesos que trabajen en conjunto para formar un subsistema y colóquelos en un diagrama hijo.

**FIGURA 7.4**

Diferencias entre el diagrama padre (superior) y el diagrama hijo (inferior).

- Omitir el flujo de datos. Examine su diagrama en busca de flujo lineal; es decir, un flujo de datos en el que cada proceso sólo tiene una entrada y una salida. Excepto en el caso de los diagramas de flujo con datos de diagramas hijos muy detallados, el flujo de datos lineal ocurre raras veces. Por lo general su presencia indica que faltan flujos de datos en el diagrama. Por ejemplo, el proceso CALCULAR MONTO DE RETENCIÓN necesita el número de dependientes que tiene un empleado y las TASAS DE RETENCIÓN como entrada. Además, no se puede calcular el SUELDO NETO sólo con base en la RETENCIÓN, y el CHEQUE DE PAGO DEL EMPLEADO no se puede crear sólo a partir del SUELDO NETO; también hay que incluir un NOMBRE DE EMPLEADO así como las cifras de nómina actuales y del año a la fecha, además del MONTO DE RETENCIÓN.
- Crear una descomposición (o expansión) desbalanceada en los diagramas hijos. Cada diagrama hijo debe tener el mismo flujo de datos de entrada y salida que el proceso padre. La excepción a esta regla es la salida menor, como las líneas de error que se incluyen sólo en el diagrama hijo. El diagrama de flujo de datos de la

**FIGURA 7.5**

Errores comunes que pueden ocurrir en un diagrama de flujo de datos (ejemplo de nómina).

figura 7.6 está dibujado en forma correcta. Observe que, aunque el flujo de datos no es lineal, podemos seguir con claridad una ruta directamente desde la entidad de origen hasta la entidad de destino.