

Диаграми на потоците от данни (ДПД).

1. Същност и компоненти на ДПД.

Когато изследват и описват една система, анализаторите търсят отговори на следните основни въпроса:

- Кой процес изграждат системата?
- Какви данни се използват при всеки процес?
- Кой данни се съхраняват?
- Какви данни влизат в и напускат системата?

Методологията, която търси отговор на тези въпроси и има за цел моделиране на системата от гледната точка на движението на данните се нарича Анализ на потоците от данни. Анализът на потоците от данни¹ (АПД) използва принципите на структурния анализ и изучава използването на данните във всеки процес или дейност. За целта той използва **диаграми на потоците** от данни, които графично показват връзката между процесите и данните и **речници на данните**, които формално описват данните за системата и местата на тяхното използване. Анализът на потоците от данни разглежда процесите в системата от гледната точка на данните: къде възникват, как се използват и променят и къде отиват, включително и “спирките” по пътя им от източника до получателя им.

Основното средство на анализа на потоците от данни са диаграми на потоците от данни (ДПД). ДПД са графично средство, използвано за описание и анализ на движението на данните в системата и включващо процесите, местата на съхранение, източниците и получателите на данни. Диаграмите на потоците от данни са основното средство и базата за развитие на останалите компоненти на АПД. Движението и промяната на данните през процесите от входа към изхода на системата може да бъде описано логически и независимо от физическите компоненти на системата (компютри, папки и т.н.) чрез логически диаграми на потоците от данни². За разлика от тях физическите диаграми³ на потоците от данни показват действителната реализация и движението на данните между хората, отделите, и работни места.

¹DataFlowAnalysis

²logicalDFD

³physicalDFD- диаграма, представяща реализацията на системата

Освен ДПД, анализът на потоците от данни използва:

- Речници на данни (Data Dictionary), които поддържат описание на компонентите, представени в ДПД;
- Диаграма на структурата на данните (Data Structure Chart) -графично описание на връзката между същностите в системата (хора, места, събития и т.н.) и набор от информация за всяка отделна същност;
- Графика на структурата на системата (Structure Chart) - средство за проектиране, което нагледно показва връзката между обработващите модули в информационната система. Описва йерархията на модулите и данните, които се обменят между тях. Включва анализ на входно-изходните преобразувания и анализ на транзакциите.

Средствата на методологията АПД могат да се прилагат както в процеса на определяне на системните изисквания, така и при проектиране на новата система. Съдържанието на речниците на данни се използва при проектиране на входа, изхода и БД. Диаграмите на структурите от данни са собено важни при проектирането на БД и разработката на софтуера. Диаграмите на потоците от данни служат като важно средство за определяне на рамките на изследваната система и проверка пълнотата на анализа на системата. Простите означения в диаграмите на потоците от данни са лесно разбираеми от потребителите и по този начин те могат да участват в процеса на изучаване на диаграмите. Това е предпоставка за своевременно откриване на евентуални грешки и пропуски на един по-ранен етап от създаването на новата система. ДПД предоставят общ език за комуникация с потребителите и добра база за пълно описание на системата и дефиниране на потребителските изисквания.



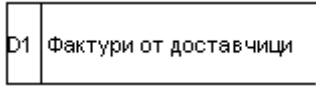
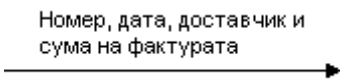
Анализът на потоците от данни позволява на проектантите да изолират области, представляващи интерес в организацията и да ги изучават. Пълното изследване на системата се осъществява посредством създаването на множество диаграми на потоците от данни, някои от тях представящи основните процеси, а други, навлизащи в по-големи детайли. Ако анализаторите искат да прегледат

цялата система, те могат да използват диаграмите от по-високо ниво. Когато се интересуват от някой специфичен процес в подробности, могат да използват диаграмите от по-ниско ниво.

Компоненти на диаграмите на потоците от данни

В теорията и практиката се използват няколко системи за означения на компонентите на диаграмите на потоците от данни. По известни са тези на Yourdon⁴ и GaneandSarson⁵. Логическите диаграми на потоците от данни могат да бъдат напълно представени с използване на 4 прости означения: външна единица, процес, поток от данни и хранилище на данни (вж табл.1).

Таблица 1. Компоненти на ДПД

Компонент на ДПД	Означение
Процес (Process)	
Външна единица (External Entity)	
Хранилище на данни (DataStore)	
Поток от данни (DataFlow)	

Процес е действие, което преобразува входни потоци от данни в изходни в съответствие с определен алгоритъм. Процесите могат да бъдат хора, процедури или устройства, които използват или обработват информация. Процесът се представя с правоъгълник, който съдържа числов идентификатор и наименование. Името на процеса представлява описание на процеса и се състои обикновено от отглаголно съществително (изчисляване, проверяване, определяне, създаване и т.н.) и съществително име, например: "Въвеждане на

⁴Shelly, Gary B., Henry J. Rosenblatt, Systems Analysis and Design 9th edition, Cengage Learning, 2011 – стр. 201

⁵Пак там

данни за клиентите"; "Изготвяне на информация за текущите разходи"; "Проверка на кредитоспособността на клиента".

Всички процеси трябва да бъдат именувани по такъв начин, че да отразяват спецификата на дейностите, включени в процеса. Имената "Покупка", "Продажба" са прекалено общи за да могат да се използват в диаграмите на потоците от данни. Вместо това могат да се използват имена като "Подготвяне на заявка за покупка", "Редактиране на заявки за продажби".

За идентифициране на процесите се използват и подходящи числови идентификатори, които отразяват йерархичната им съподчиненост. (1, 1.1, 1.1.1, 2.1 и т.н)

Външните единици (източник или получател на данни, external entity) могат да бъдат хора, програми, организации или други обекти, които взаимодействат със системата, но са извън нейните рамки. Те генерират информация към системата или получават такава от нея. Определянето на даден обект като външна единица означава, че той се намира извън границите на изследваната система. В някои системи от означения се използват термините Source (източник) или Destination (получател), Sink, Terminator. Информацията в системата трябва да произлиза от външни единици.

Потокът от данни представлява данни, които се движат в определена посока от източник към получател под формата на документ, писмо, телефонен разговор и др. Потокът от данни е пакет от информация. Той трябва винаги да е насочен към или да произлиза от процес. На диаграмата се представя чрез насочена стрелка, показваща посоката на движение на данните и етикет с описание на предаваната информация. Единият край на потока от данни трябва винаги да е процес, докато другият може да бъде външна единица, хранилище на данни или друг процес.

Имената на потоците от данни трябва да отразяват информацията, представляваща интерес за анализатора, а не документа, на който се базира тя. Например една фактура може да съдържа различна по характер информация. Анализаторите се интересуват от данните във фактурата, които са важни за конкретния процес. Това може да е номера на фактурата, датата на издаване,

общата сума, а не самият документ сам по себе си. Данните, които влизат в даден процес претърпяват промени. Затова изходните потоци от данни се именуват по различен начин от входните.

Хранилище на данни е мястото, където данните в системата се съхраняват и към което се обръщат процесите. Хранилищата от данни могат да бъдат компютъризирани или некомпютъризирани устройства. На диаграмите се идентифицират чрез буквата "D" и число. Името на хранилището на данни се избира от съображения за максимална информативност за проектанта. Хранилищата в най-общия случай се явяват прототип на бъдещите бази от данни и описанието на съхраняващите се в тях данни трябва да бъде обвързано с информационния модел на системата

2. Създаване на диаграми на потоците от данни

2.1. Контекстна диаграма на потоците от данни (ContextDataFlowDiagram).

Първата стъпка в процеса на определяне на системните изисквания е насочена към изучаване на общите характеристики на системата. Моделирането на системата чрез ДПД започва със създаването на контекстната диаграма на потоците от данни⁶. Тя представя системата като един процес и дефинира всички потоци от данни от и към нея без да описва вътрешната ѝ работа. Въпреки че контекстната диаграма не се използва пряко при проектирането, тя е полезна с това, че определя рамките на системата и може да се използва при обсъждането с потребителите и мениджърите за да се одобри обхвата на новата система. Всичко, което попада извън процеса, идентифициран в контекстната диаграма, не е обект на изследване. Начинът, по който функционират външните единици е извън контрола на системата и затова не се изучава в детайли. Ако те обаче влияят по някакъв начин върху системата в качеството си на източници или получатели на данни системата трябва да се разработят средства за взаимодействие с тях.

2.2. Йерархична графика на процесите (HierarchyProcessChart)

⁶Нарича се още и диаграма от нулево ниво -Level-0 Diagram

След контекстната ДПД се създава йерархичната графика на процесите (ЙГП). ЙГП представя в структуриран йерархичен вид всички действия, които генерират, модифицират или използват информация в рамките на системата. Структурирането на процесите в ЙГП не винаги съвпада с организационната структура на компанията, тъй като организационната структура може да описва дейности, които са извън обхвата на изследваната система.

Препоръчително е броят на процесите от едно ниво на декомпозиция, подчинени на един процес от предишното ниво, да бъде между 4 и 6, като не е желателно да надхвърля 8. ЙГП служи за база за разработване на останалите ДПД като определя нивата на декомпозиция, броя на диаграмите и процесите, които се включват във всяка диаграма.

2.3. Диаграма на потоците от данни от първо ниво (Level 1 DFD).

Следващата стъпка е да се опише системата на първо ниво от йерархичната графика на процесите. Диаграмата от първониво представлява детайлизиране на контекстната диаграма. В нея се представят основните подсистеми и взаимодействието между тях.

При създаване на диаграмите на потоците от данни от първо ниво може да се спазва следната примерна последователност на работа:

1. В края на диаграмата се разполагат всички външни единици и асоциираните към тях входни и изходни потоци от данни;
2. За всеки поток от данни се идентифицира процесът, който получава или генерира данните. Потоците от данни от и към външните единици се асоциират с тези процеси. Това трябва да са процеси от първо ниво на ЙГП;
3. Добавят се вътрешните хранилища от данни. Те се свързват с процесите посредством вътрешни потоци от данни.
4. Добавят се вътрешни процеси, които не са били включени на втора стъпка.
5. Добавят се вътрешни потоци от данни, които свързват процесите.

6. Диаграмата се проверява за коректност и пълнота.

2.4. Създаване на ДПД от следващи нива.

Създаването на диаграмите от следващите нива се определя от ЙГП. Възможно е обаче да се пристъпи към разработката на ДПД от следващи нива преди да е окончателно довършена ЙГП. В този случай декомпозирането на процесите от ДПД се отразява в структурата на ЙГП.

След като се създаде диаграмата от първо ниво е необходимо да се реши кои от процесите се нуждаят от допълнително описание в диаграма от второ и следващо ниво. Ако процесът има кратко и специфично наименование и има не повече от 6 до 8 входа и изхода, той по всяка вероятност не се нуждае от допълнително разбиване. Процесите се **нуждаят от декомпозиране** ако:

- Включват изпълнението на няколко задачи, изискващи движение на данни между различни хора или места. Многократните действия, изпълнявани от един човек или на едно работно място не е необходимо да се декомпозират;
- Имат повече от 6-8 входни или изходни потоци от данни;
- Името на процеса е много дълго или общо, като “Приемане на поръчки и отразяване на продажби” или “Отдаване под наем на МПС”

За да има **съвместимост** между диаграмите от по-ниско ниво и тези от по-високо ниво трябва да се спазват следните изисквания:

- Всички потоци от данни, които се появяват в диаграмата от по-високо ниво, описваща процеса, трябва да бъдат включени в диаграмата от по-ниско ниво;
- Нови потоци от данни и нови хранилища на данни се добавят само ако се използват в рамките на процеса, за свързване на процеси, представени за пръв път на това ниво (вътрешни процеси);
- Потоците от данни и хранилищата на данни, които възникват в рамките на процеса трябва да бъдат показани. Хранилищата на данни и потоците от данни, които са специфични

само за даден процес не бива да се показват докато процесът не се разбие, т.е. не се покаже неговата вътрешна работа;

- Не трябва да има противоречие между описанието на диаграмата от по-високо и по-ниско ниво;
- С навлизането на по-ниско ниво на декомпозиция се добавят все повече детайли.

Тъй като сложността на системите варира в много широки граници, няма препоръчителен брой нива за диаграми на потоците от данни. Общото правило е, че процесите се декомпонират, докато се изследват напълно детайлите на системата и начинът, по който функционира.

3.Правила при създаването на диаграми на потоците от данни

При създаване на ДПД трябва да се спазват следните основни правила:

- Всеки процес трябва да има поне един вход и изход;
- Потоците от данни винаги произтичат или са насочени към процес. Другият им край може да е хранилище на данни, външна единица или друг процес. Следователно невалидни потоци от данни са тези между две външни единици, между две хранилища от данни и между външна единица и хранилище;
- Всеки компонент на диаграми на потоците от данни трябва да има кратък и ясен етикет, който дава точно описание на неговото функциониране в системата;
- Потоците от данни не трябва да се разделят на два или повече различни потоци. Същото правило важи и за обединяването на два или повече потока от данни;
- Всеки поток от данни, който напуска даден процес, трябва да се базира на данни, които са входни за процеса;
- Входни трябва да бъдат само данните, необходими за изпълнението на процеса;
- Даден процес трябва да бъде независим от останалите в системата, той трябва да зависи само от собствения си вход и изход.

