

Projektowanie obiektowe oprogramowania

Architektura systemów (3)

Service Oriented Architecture

Wykład 15

Wiktor Zychła 2014

Spis treści

| | | |
|-----|---|----|
| 1 | Modele integracji danych | 2 |
| 1.1 | Integracja przez wymianę plików | 2 |
| 1.2 | Integracja przez wspólną bazę danych | 2 |
| 1.3 | Integracja przez usługi aplikacyjne (Web Services)..... | 2 |
| 1.4 | Integracja za pośrednictwem brokera/szyny usług | 2 |
| 2 | Message Oriented Architecture | 4 |
| 3 | Enterprise Service Bus | 6 |
| 3.1 | Szyna usług - podstawowe pojęcia..... | 6 |
| 3.2 | Przegląd implementacji | 9 |
| 3.3 | Przykład..... | 9 |
| 4 | Service Oriented Architecture | 10 |
| 5 | Command-Query Responsibility Segregation..... | 10 |
| 6 | Literatura | 12 |

1 Modele integracji danych

1.1 Integracja przez wymianę plików

- + Możliwe do implementacji w niejednorodnym środowisku
- - Konieczność ręcznej ingerencji użytkownika w proces czyni ten proces bezużytecznym tam gdzie oczekuje się automatyczności przepływów danych

1.2 Integracja przez wspólną bazę danych

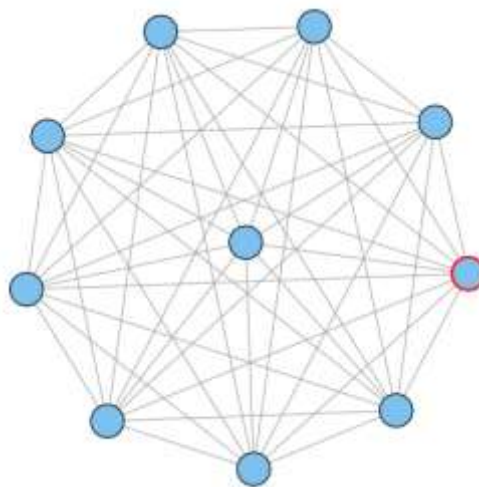
- + Integracja jest automatyczna i natychmiastowa
- - Duża struktura danych paraliżuje swój rozwój
- - W ten sposób nie da się integrować aplikacji pochodzących od różnych dostawców

Hohpe/Woolf:

If the technical difficulties of designing a unified schema aren't enough, there are also severe political difficulties. If a critical application is likely to suffer delays in order to work with a unified schema, then often there is irresistible pressure to separate. Human conflicts between departments often exacerbate this problem

1.3 Integracja przez usługi aplikacyjne (Web Services)

- + Możliwe do implementacji w niejednorodnym środowisku
- + Można integrować aplikacje różnych producentów
- - Właściwie jedyna wada – przypadłość „sieci połączeń”



1.4 Integracja za pośrednictwem brokera/szyny usług

- Połączenie pomysłu integracji przez usługi aplikacyjne z ideą wzorca Observer
- Każda aplikacja łączy się z podsystemem komunikacyjnym i publikuje lub subskrybuje powiadomienia
- Złożoność sieci połączeń jest liniowa, nie ma problemu „sieci powiązań” – każda aplikacja integruje się tylko z podsystemem komunikacyjnym
- Szyna dostarcza mechanizmów wiarygodnego dostarczania komunikatów



2 Message Oriented Architecture

Message-oriented architecture – model architektury zorientowany na komunikaty wymieniane między modułami. Ciekawie zaczyna robić się wtedy, kiedy mówimy o architekturze **systemów** ponieważ wtedy do wymiany komunikatów niezbędny staje się zewnętrzny pośrednik – **Message-oriented Middleware (MOM)**.

Gdyby każdy system typu MOM zbudowany był w taki sposób, że komunikacja z nim odbywałaby się w jego własnym „dialekcie”, to budowa rozwiązań integracyjnych byłaby utrudniona. Próbuje się więc budować abstrakcje, które gwarantują wymiennność implementacji infrastruktury oraz możliwość łączenia różnych rozwiązań typu MOM.

Z punktu widzenia implementacji mechanizmów integracyjnych po stronie aplikacji-nadawców i aplikacji-odbiorców danych, istnieją dwie filozofie budowy abstrakcji

- **API Driven** (zorientowane na interfejs programistyczny) – podstawą abstrakcji mechanizmów integracyjnych jest API udostępniane przez mechanizm integracyjny. API jest ustandaryzowane, na zewnątrz gwarantuje stały kontrakt, a szczegóły wewnętrznej implementacji są nieistotne.

Przykład – JMS, Java Message.

http://en.wikipedia.org/wiki/Java_Message_Service

JMS jest zbiorem API, to znaczy że posiada abstrakcje takie jak połączenie, nadawca, odbiorca, komunikat wyrażone w języku obiektowej specyfikacji. JMS ma implementacje różnych „dostawców”, trochę tak jak nHibernate ma implementacje dostawców dialektów dla różnych baz danych. Wikipedia wymienia kilkanaście istniejących implementacji.

Plus abstrakcji typu API driven jest taki że raz napisany kod zawsze pozostaje niezmienny.

Minus jest taki, że JMS jako API jest częścią J2EE, więc nie występuje w środowiskach .NET, PHP itd. W efekcie, aby zaimplementować mechanizm integracji z dostawcą specyfikacji JMS w innej technologii niż J2EE i tak trzeba poznać szczegóły implementacji protokołu komunikacyjnego konkretnego dostawcy.

- **Protocol Driven** (zorientowane na protokół komunikacyjny) – podstawą abstrakcji mechanizmów integracyjnych jest gwarantowany protokół komunikacyjny między aplikacją a podsystemem MOM. Protokół jest gwarantowany, może być oparty o TCP albo o protokół wyższej warstwy, np. HTTP/HTTPS.

Przykład – AMQP Advanced Message Queuing Protocol.

http://en.wikipedia.org/wiki/Advanced_Message_Queueing_Protocol

AMQP jest otwartym standardem protokołu komunikacyjnego, opartego o TCP, który specyfikuje formaty ramek TCP w komunikacji klienta (aplikacji) z serwerem (brokerem)

MOM. Istnieją gotowe implementacje AMQP dla większości przemysłowych technologii wytwarzania oprogramowania, ale jeśli z jakiegoś powodu takiej implementacji brak – w oparciu o specyfikację protokołu zawsze można dostarczyć własnej implementacji. Wikipedia wymienia kilka (mniej niż 10) przemysłowych implementacji, w tym open-source.

Plus abstrakcji typu Protocol driven jest taki że jest rzeczywiście interoperacyjna, to znaczy aplikacje integrujące się z brokerem mogą być napisane w różnych technologiach/językach.

Specyfikacja ta nie posiada ewidentnych wad.

Współcześnie preferuje się specyfikacje typu Protocol Driven, specyfikacje typu API Driven są uznawane za przestarzałe z uwagi na problemy z interoperacyjnością.

3 Enterprise Service Bus

3.1 Szyna usług - podstawowe pojęcia

Enterprise Service Bus – usługa zewnątrzprocesowa typu MOM, dostarczająca narzędzi komunikacyjnych w rozległym środowisku aplikacyjnym. Szczegóły niżej.

Przykład – wiele niezależnych systemów bankowych, integracja przelewów i innych operacji finansowych – natychmiastowa aktualizacja między systemami.

W praktyce wzorzec ESB tak samo dobrze organizuje architekturę systemu, jak Event Aggregator organizował architekturę aplikacji.

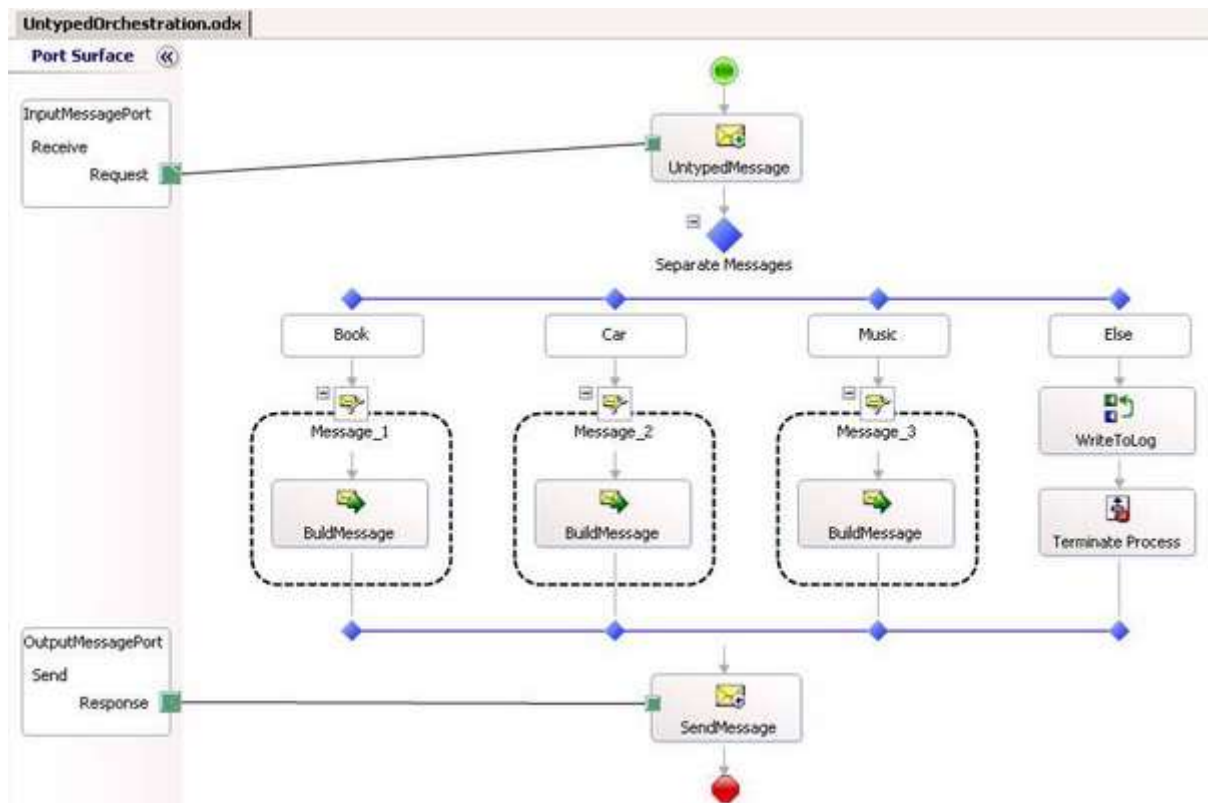
Publish/Subscribe – wzorzec wymiany komunikatów w szynie usług, w którym wybrane aplikacje pełnią rolę publikatorów komunikatów (powiadomień), a inne pełnią rolę subskrybentów tych komunikatów. Na tym atomowym wzorcu oparte są pozostałe wzorce złożone – Request/Reply i Saga.

Request/Reply – wzorzec wymiany komunikatów w szynie usług, w którym aplikacja zadaje zapytanie do szyny usług, a jedna (zwyczajowo) z aplikacji jest zarejestrowana jako dostawca takich danych i odpowiada na żądanie. Należy zauważyć, że R/R implementuje się z wykorzystaniem wzorca P/S – R/R dla aplikacji A i B wygląda następująco:

- B subskrybuje komunikatu typu „Żądam danych typu X”
- A subskrybuje komunikat typu „Odpowiedź na zapytanie o dane typu X”
- A publikuje komunikat typu „Żądam danych typu X”
- B otrzymuje komunikat i go przetwarza; publikuje komunikat typu „Odpowiedź na zapytanie o dane typu X” i w adresata wstawia A
- (opcjonalnie) Szyna filtruje komunikat tak żeby dostał go tylko A
- A otrzymuje odpowiedź

Saga (inaczej *long running transaction*; **transakcja długoterminowa**) (także: **proces biznesowy**) – wzorzec wymiany komunikatów w szynie usług, oparty o złożony proces biznesowy, którego koordynatorem jest szyna. Proces zwykle składa się z atomowych kroków, najważniejsze z nich to komunikacja z portami wejścia/wyjścia.

Przykładowa saga: komunikat od konsultanta do spraw sprzedaży z wnioskiem o kredyt dla klienta trafia do szyny. Jeśli kwota kredytu jest mniejsza niż 1000\$, szyna tworzy komunikat typu „aprobata wniosku” i przekazuje nowy komunikat do serwera departamentu wypłacania kredytów. Jeśli kwota kredytu jest większa lub równa 1000\$, szyna przekazuje ten komunikat dalej, do serwera departamentu kredytowego, gdzie komunikat oczekuje w systemie na akceptację kierownika działu. Szyna oczekuje na odpowiedź nie dłużej niż 48h. Jeśli w tym czasie nadchodzi odpowiedź pozytywna – szyna konwertuje ją na komunikat typu „aprobata wniosku” i przekazuje nowy komunikat do serwera departamentu wypłacania kredytów. Jeśli odpowiedź jest odmowna lub upłynęło więcej niż 48h, szyna tworzy komunikat typu „odmowa” i przekazuje zwrótnie do systemu, w którym zobaczy go konsultant.



Rysunek 1 (za <http://www.codeproject.com/Articles/13277/Configuring-BizTalk-Orchestrations-to-handle-un-ty>)

(Uwaga! Diagram nie odpowiada treści przykładowego procesu!)

Asynchroniczna komunikacja – dlaczego rozproszonych usług integracyjnych nie można zaimplementować tak samo jak wzorca Observer? Dlatego że subskrybenci mają różne tempo przetwarzania powiadomień. Sto powiadomień od A do B i C może być przez B przetwarzane w ciągu 1 sekundy, a przez C w ciągu 1 godziny. Szyna musi wiarygodnie koordynować przekazywanie komunikatów i wymaga wsparcia jakiegoś repozytorium trwałego, np. relacyjnej bazy danych lub usługi kolejki wiadomości.

Specyfikacje formatów wymiany danych – zwyczajowo szyny usług do integracji aplikacji niejednorodnych technologicznie wykorzystują nośnik XML, wsparty opcjonalnie walidacją formalną XSD i mechanizmami autentykacji/autoryzacji opartymi o XmlDsig (patrz poprzedni wykład).

```

<?xml version="1.0" encoding="utf-8"?>
<xs:schema xmlns="http://ito.vulcan.edu.pl" targetNamespace="http://ito.vulcan.edu.pl"
xmlns:xs="http://www.w3.org/2001/XMLSchema"
elementFormDefault="qualified" attributeFormDefault="unqualified">
  <xs:complexType name="IT0JednostkaSamodzielna">
    <xs:sequence>
      <xs:element minOccurs="0" maxOccurs="1" name="Adres" type="IT0Adres"/>
      <xs:element minOccurs="0" maxOccurs="1" name="JednostkiSkladowe"
type="ArrayOfIT0JednostkaSkladowa"/>
      <xs:element minOccurs="1" maxOccurs="1" name="ID" type="xs:int"/>
      <xs:element minOccurs="0" maxOccurs="1" name="Nazwa" type="xs:string"/>
      <xs:element minOccurs="0" maxOccurs="1" name="NazwaZUS" type="xs:string"/>
      <xs:element minOccurs="0" maxOccurs="1" name="Kod" type="xs:string"/>
      <xs:element minOccurs="0" maxOccurs="1" name="Regon" type="xs:string"/>
      <xs:element minOccurs="0" maxOccurs="1" name="NIP" type="xs:string"/>
      <xs:element minOccurs="0" maxOccurs="1" name="Patron" type="xs:string"/>
      <xs:element minOccurs="0" maxOccurs="1" name="Siedziba" type="xs:string"/>
      <xs:element minOccurs="0" maxOccurs="1" name="Opis" type="xs:string"/>
      <xs:element minOccurs="0" maxOccurs="1" name="Uwagi" type="xs:string"/>
      <xs:element minOccurs="1" maxOccurs="1" name="DataOd" nillable="true"
type="xs:dateTime"/>
      <xs:element minOccurs="1" maxOccurs="1" name="DataDo" nillable="true"
type="xs:dateTime"/>
      <xs:element minOccurs="1" maxOccurs="1" name="PlatnikVAT" type="xs:boolean"/>
      <xs:element minOccurs="1" maxOccurs="1" name="PlatnikFGSP" type="xs:boolean"/>
      <xs:element minOccurs="0" maxOccurs="1" name="Rodzaj" type="xs:string"/>
      <xs:element minOccurs="0" maxOccurs="1" name="OrganProwadzacy" type="xs:string"/>
      <xs:element minOccurs="0" maxOccurs="1" name="OrganRejestrujacy" type="xs:string"/>
      <xs:element minOccurs="1" maxOccurs="1" name="IdOrganProwadzacy" nillable="true"
type="xs:int"/>
      <xs:element minOccurs="1" maxOccurs="1" name="IdOrganRejestrujacy" nillable="true"
type="xs:int"/>
      <xs:element minOccurs="1" maxOccurs="1" name="S4" type="xs:int"/>
      <xs:element minOccurs="0" maxOccurs="1" name="IdOsw" type="xs:string"/>
      <xs:element minOccurs="1" maxOccurs="1" name="MojA" type="xs:boolean"/>
      <xs:element minOccurs="1" maxOccurs="1" name="IsDeleted" type="xs:boolean"/>
      <xs:element minOccurs="1" maxOccurs="1" name="DataDodania" type="xs:dateTime"/>
    </xs:sequence>
  </xs:complexType>
</xs:schema>

```

Port wejścia – szyny są usługami zewnątrzprocesowymi; typowa szyna usług jest równocześnie usługą systemową (demonem systemowym). To oznacza, że komunikat trzeba do szyny jakoś dostarczyć. Port wejścia jest abstrakcją źródła danych, a typowe tzw. *adaptery* (implementacje konkretne) obejmują: relacyjne bazy danych, system plików czy usługi aplikacyjne. Na przykład więc port wejścia używający adaptera – systemu plików przyjmuje komunikat w postaci pliku pojawiającego się w określonym folderze. Aplikacja aby wysłać komunikat do szyny tworzy plik XML o określonej składni w tymże wybranym folderze, a szyna aktywnie podejmuje i przetwarza komunikat.

Mówimy że port wejścia pracuje w trybie **pull**, kiedy szyna musi aktywnie monitorować jakiś zasób w celu stwierdzenia, że pojawiają się tam dane. W trybie pull pracują adaptery – system plików i relacyjna baza danych. Mówimy że port wejścia pracuje w trybie **push**, kiedy usługa przekazania komunikatu jest wzbudzana „na żądanie”, bez potrzeby monitorowania zasobu. Tak działa adapter – usługa aplikacyjna – dane są przekazane do szyny przez wywołanie jakiejś jej usługi aplikacyjnej.

Port wyjścia – abstrakcja miejsca, w które szyna dostarcza dane. Typowe adaptery są podobne jak w przypadku portów wejścia. Na przykład port wyjścia używający adaptera – usługi aplikacyjnej dostarczy komunikat do aplikacji w ten sposób, że wywoła jakąś określoną usługę aplikacyjną (web service) po stronie aplikacji.

Mówimy że port wyjścia działa w trybie **pull**, kiedy aplikacja-subskrybent musi aktywnie monitorować jakiś zasób, żeby zorientować się że szyna ma dla niej jakieś dane. Na przykład port wyjścia używający adaptera – usługa aplikacyjna po stronie szyny, wymaga od aplikacji regularnego monitorowania stanu portu (wywoływania określonej metody) w celu stwierdzenia czy pojawiły się jakieś dane. Mówimy że port wyjścia działa w trybie **push**, jeśli szyna danych aktywnie przekazuje komunikat do subskrybenta. Na przykład port wyjścia używający adaptera – usługa aplikacyjna po stronie subskrybenta, po pojawieniu się komunikatu po prostu wywoła aktywnie wskazaną usługę aplikacyjną.

Oczywista obserwacja – porty wejścia z adapterami w trybie push są wydajniejsze i pozwalają na niemal on-line przekazywanie powiadomień. Jedyne problemy są takie, że to nie zawsze technicznie jest możliwe, na przykład tryb push dla portów wyjścia nie jest możliwy jeśli subskrybent nie jest aplikacją serwerową, posiadającą własne usługi aplikacyjne; z kolei tryb pull zawsze jest możliwy na portach wyjścia, bo szyna jest usługą serwerową.

3.2 Przegląd implementacji

Usługi ESB mają dziesiątki mniejszych i większych implementacji. Trochę wyobrażenia można nabrać przeglądając (niepełny!) rejestr na wiki:

http://en.wikipedia.org/wiki/Comparison_of_business_integration_software

| Product | Developer | Version | Release Date | Pricing | License | Category |
|--|------------------------------------|------------|-----------------|--|---------|----------------------------------|
| Apache Camel | Apache Software Foundation | 2.10.0 | March 2013 | Free/Commercial support available | Yes | Apache Software License |
| Apache ServiceMix | Apache Software Foundation | 3.0 | November 2010 | Free / Commercial support available | Yes | Apache Software License |
| AMN Open | CS&Business | 3.0 | November 2009 | starting from \$12,000 for standard edition (1) | No | proprietary |
| Arta ESB | Progress Software | 5.0 | 2007 | | No | proprietary |
| Aviatrix Software Enterprise Data Integrator | Aviatrix Software | 8.0 | NAQ 2013 | Varies | No | proprietary |
| Automation Anywhere Integration Pack | Automation Anywhere | 8.1 | Feb 2011 | MSMC | No | proprietary |
| BizTalk Server | Microsoft | 2010 | March 2013 | Enterprise Edition: \$18,000 per core; Standard Edition: \$2,400 per core; Speech Edition: \$600 per core; Developer Edition - expected release November 1st 2013 - not \$50 per seat, also available under MSDN license (2) | No | proprietary |
| Microsoft Server | Microsoft | 2010 | 2013 | | No | proprietary/SaaS |
| Flow Software | Flow Software Ltd | 2.3.0 | May, 2010 | Free Community Edition, and Enterprise license | No | proprietary |
| FINES ESB - Enterprise ServiceMix | Finisar | 4.2 | 2007 | | Yes | based on Apache Software License |
| Informatica Power Center | Informatica | 8.3 | October 2007 | Varies: 50,000 - 100,000 generally | No | proprietary |
| JBoss Enterprise Service Bus (ESB) | JBoss, a division of Red Hat, Inc. | 4.9 | August 2013 | Free / Commercial support available | Yes | LGPL |
| JBoss Enterprise 6.0 Platform | JBoss, a division of Red Hat, Inc. | 6.1 | February 2011 | Free / Commercial support available | Yes | LGPL |
| JBoss | JBoss | 2.0 | May 2005 | | Yes | GPL |
| Magi (an Integration Platform) | Magi Software Enterprises | 1.0 | May 2012 | | No | proprietary |
| Link 5 | Twinkl2.com | 5.0 | Feb 2014 | Free beta | No | proprietary |
| OpenAdaptor | The Software Conservancy | 1.0.0 | February 2011 | Free | Yes | subset of MIT |
| Mule ESB | MuleSoft | 3.2.1 | September 2012 | | Yes | LGPL |
| OpenESB | OpenESB Community | 2.0 | May 2008 | Supported by LogiCity Inc. http://logicity.com/support | Yes | LGPL |
| OpenLink Software | OpenLink Software | 4.5 | 2009 | \$50 per node user (1) | Yes | Dual GPL or proprietary |
| Oracle BPEL Process Manager (1) | Oracle Corporation | 10.1.2.0.2 | 21 January 2006 | \$2,000 per processor (1) | No | proprietary |
| Oracle Enterprise Service Bus (1) | Oracle Corporation | 10.1.1.1 | 7 | \$2,500 (1) | No | proprietary |
| PEARL ESB | ORQ Consortium | 3.0.3 | July 2011 | Free / Commercial support available | Yes | LGPL |
| Boon CMB | Progress Software | 8.4 | 2011 | | No | proprietary |

Możliwe są nawet proste i tanie implementacje wykorzystujące np. mechanizmy Pub/Sub w WCF, podsystem Workflow Foundation czy wprost usługi kolejkowania wiadomości (MSMQ) opakowane jakąś implementacją adapterów dla wybranych portów wejścia wyjścia.

3.3 Przykład

Interaktywny przykład najprostszej możliwej implementacji szyny integracyjnej opartej o wieloplatformowy system MOM typu open-source, **RabbitMQ**. Omówimy pojęcia:

- Kolejki (Queue)
- skrzynki nadawczej (Exchange),
- Powiązań (Binding)
- Routing typu „fan”
- Routing typu „direct”
- Routing typu „topic”

<http://rabbitmq.com>

4 Service Oriented Architecture

Service Oriented Architecture (SOA) to sposób myślenia o architekturze systemu, w którym poszczególne podsystemy dostarczają dobrze określonych funkcjonalności biznesowych (gromadzenie danych, wyszukiwanie danych, przetwarzanie danych, mechanizmy wydruków, usługi komunikacyjne itp.) ale zachowują dużą autonomię jako osobne “aplikacje” w rozumieniu klasycznych stosów aplikacyjnych.

Architektura SOA zwykle oznacza więc też Szynę Integracyjną (lub inny component typu MOM), model komunikacji oparty o XML/JSON i/lub jakiś mechanizm przetwarzania długotrwałych transakcji.

Architekturę typu SOA zwykle wybiera się dla dużych, heterogenicznych systemów, często wytwarzanych w różnym czasie i przez różnych dostawców.

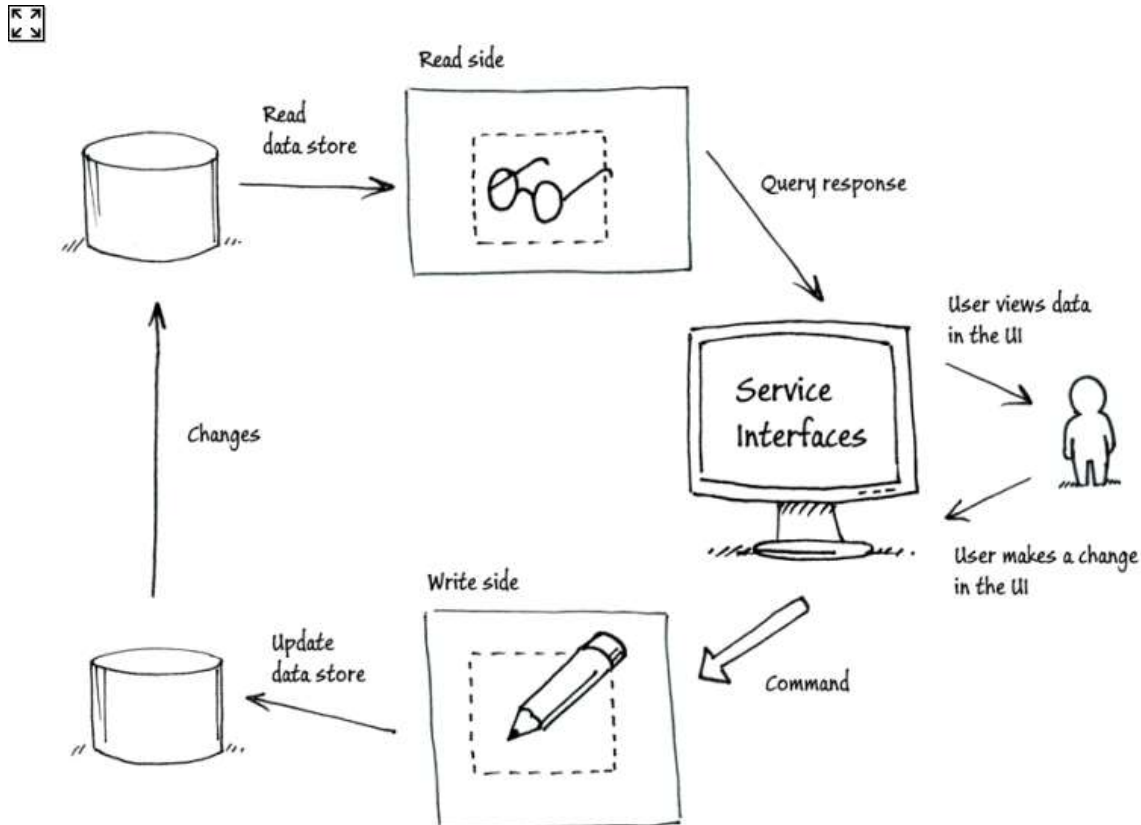
Zaletą architektury SOA jest duża podatność na autonomiczne zmiany w ramach każdego z podsystemów, w tym całkowita wymiana pewnych podsystemów na inne/nowsze, bez utraty spójności systemu..

Wadą SOA jest duży nakład początkowy, niezbędny chociażby do zapewnienia odpowiedniej standaryzacji poszczególnych usług oraz implementacji całej infrastruktury typu MOM.

Współcześnie rozważa się również tzw. **architekturę mikrousług** (Fowler: microservices, <http://martinfowler.com/articles/microservices.html>), w której podobne pojęcia jak w SOA pojawiają się na poziomie architektury pojedynczej aplikacji.

5 Command-Query Responsibility Segregation

Command-Query Responsibility Segregation – filozofia budowy systemu transakcyjnego, w którym rola odczytu i zapisu (modyfikacji) jest wyraźnie koncepcyjnie rozdzielona.



Rysunek 2 Za <http://msdn.microsoft.com/en-us/library/jj591573.aspx>

Podstawowe spostrzeżenia:

- W dużej części systemów dane głównie się ogląda – rozrzut między odczytami a zapisami to nawet 1000:1
- Zapis jest najwygodniejszy do struktury silnie znormalizowanej, odczyt jest najwygodniejszy ze struktury mocno zdenormalizowanej (idealnie: całkowicie spłaszczonej). To mogą być nawet dwa różne magazyny danych (np. baza relacyjna do zapisu i baza typu noSQL do odczytu)
- W praktyce do implementacji architektury CQRS można użyć systemu kolejkowego – każda operacja po stronie zapisu powoduje powstanie komunikatu (Command), który jest przetwarzany używany do zaktualizowania danych. Następnie w miejscu aktualizacji danych generuje się dodatkowe komunikaty o zmianach (Changes), które również za pomocą systemu kolejkowego przesyła się do części systemu odpowiedzialnej za generowanie widoków dla użytkowników.

System składa się wtedy z części do przetwarzania zapisów/modyfikacji, systemu kolejkowego i części do odczytu. Mając tak rozdzielone odpowiedzialności można łatwiej dostosować system do rzeczywistego obciążenia.

6 Literatura

1. Hohpe, Woolf – Enterprise Integration Patterns, <http://www.eaipatterns.com/> - podręcznik wzorców integracyjnych, szczegółowo omawia wzorzec ESB i szereg jego podwzorców
2. A. Rotem-Gal-Ol – Wzorce SOA
3. Videla, Williams – RabbitMQ in Action