

Platformy softwarowe dla rozwoju systemów intranetowych i internetowych

Maciej Zakrzewicz
Politechnika Poznańska, Instytut Informatyki
ul. Piotrowo 3a, 60-965 Poznań
mzakrz@cs.put.poznan.pl

Streszczenie

W ostatnich latach, Internet i intranet stały się dominującymi platformami dla budowy rozproszonych systemów informatycznych. Ich niekwestionowanymi zaletami są: skalowalność i przenaszalność systemów, łatwość obsługi aplikacji przez użytkowników oraz minimalne nakłady administracyjne. W artykule dokonano przeglądu podstawowych technologii intra- i internetowych, stanowiących wybranych budowy nowoczesnych aplikacji.

1 Wprowadzenie

Pod pojęciem Internetu rozumie się zbiór wielu milionów komputerów podłączonych do wspólnej, ogólnoswiatowej sieci komputerowej. Komputery łączone są m.in. w celu umożliwienia wymiany plików/dokumentów i zdalnego wykonywania programów. Istnieje szereg rozwiązań technologicznych, które ułatwiają realizację oprogramowania pracującego w sieci Internet. W ostatnich latach zauważa się gwałtowny wzrost liczby biznesowych zastosowań Internetu: od „pojawiania” się stron internetowych firmy w sieci aż do pełnej obsługi działalności przedsiębiorstwa (np. handel elektroniczny).

Internet posiada charakter otwarty: w praktyce każdy jego użytkownik posiada swobodny dostęp do informacji. Cecha taka, jakkolwiek pożądana, pierwotnie dyskwalifikowała tę technologię z grona rozważanych dla realizacji informatyzacji przedsiębiorstw. Sukces zastosowań Internetu spowodował jednak dążenie do budowy wielu „własnych internetów” - sieci typu intranet. Dziś termin „intranet” opisuje zamkniętą i bezpieczną sieć komputerową, wykorzystującą standardy technologiczne Internetu, budowaną przez przedsiębiorstwa na potrzeby własne.

Z informatycznego punktu widzenia, technologie internetowe nie wnoszą nowej jakości do tradycyjnych rozwiązań programistycznych. Jednak ich podstawowymi walorami są m.in.: względnie niskie koszty wdrożenia, łatwość wykorzystywania narzędzi przez użytkowników (jednakowe zasady obsługi), niezależność od stosowanych rozwiązań sprzętowych, szybki i nieskomplikowany dostęp do informacji, możliwości personalizacji aplikacji i treści.

Zastosowania intranetu dla automatyzacji procesów w przedsiębiorstwie mogą zostać sklasyfikowane w następujący sposób:

1. **Przekazywanie** informacji i dokumentów pomiędzy pracownikami przedsiębiorstwa.
2. **Przechowywanie** w postaci elektronicznej i udostępnianie dokumentów pracownikom przedsiębiorstwa.
3. **Generowanie** nowych dokumentów i udostępnianie ich pracownikom przedsiębiorstwa.
4. Udostępnianie **oprogramowania**, służącego do automatyzacji funkcji realizowanych przez pracowników przedsiębiorstwa.

W niniejszym artykule zostaną scharakteryzowane i porównane najważniejsze intra- i internetowe rozwiązania technologiczne, które stanowią podstawy konstrukcji nowoczesnych systemów informatycznych.

2 Przekazywanie informacji

Intranet dostarcza wielu rozwiązań umożliwiających użytkownikom wymianę informacji o charakterze tekstowym. Trzy najpopularniejsze z nich to: poczta elektroniczna (e-mail), tekstowe telekonferencje (IRC) oraz listy dyskusyjne. Poczta elektroniczna pozwala nadawcy na przesłanie wiadomości do tzw. elektronicznej skrzynki pocztowej odbiorcy, który w wygodnym dla siebie momencie odczyta jej treść. Tekstowe konferencje umożliwiają wielu równocześnie pracującym użytkownikom prowadzenie wspólnej konwersacji, w ramach której wymieniane są krótkie, zwykle jednozdaniowe wypowiedzi. Listy dyskusyjne są narzędziem, przy pomocy którego prowadzenie konwersacji przez wielu użytkowników nie wymaga ich równoczesnej obecności w sieci.

2.1 Poczta elektroniczna

Usługa poczty elektronicznej składa się z następujących typów komponentów:

1. Programów nazywanych *klientami poczty elektronicznej*, wykorzystywanych przez użytkowników.

2. Programów nazywanych *serwerami poczty elektronicznej*, stanowiących elementy infrastruktury informatycznej.

Użytkownik-nadawca wprowadza adres użytkownika-odbiorcy i treść wysyłanej wiadomości korzystając z programu klienta poczty elektronicznej. Następnie, program klienta porozumiewa się z programem serwera przy pomocy języka-protokołu SMTP (Simple Mail Transfer Protocol). Program serwera korzysta z własnych informacji konfiguracyjnych w celu przesłania wiadomości przez łańcuch innych serwerów, aż do serwera docelowego. Serwer docelowy przechowuje elektroniczną skrzynkę pocztową odbiorcy i w niej zapisywana jest otrzymana wiadomość.

Użytkownik-odbiorca korzysta z programu klienta poczty elektronicznej w celu odczytania wiadomości znajdujących się w skrzynce pocztowej. Tym razem, komunikacja odbywa się w języku-protokole POP3 (Post Office Protocol) lub IMAP (Internet Message Access Protocol). Wiadomość pobierana z serwera jest zapisywana na dysku komputera użytkownika i prezentowana na ekranie.

Usługa poczty elektronicznej posługuje się pojęciem *adresu poczty elektronicznej* (adresu e-mail) użytkownika. Adres ten pozwala na jednoznaczne kierowanie przesyłek pocztowych – w sieci Internet nie istnieje dwóch różnych użytkowników posiadających identyczne adresy poczty elektronicznej. Adres poczty elektronicznej składa się z dwóch części: nazwy własnej użytkownika (nazwa ta może powtarzać się pomiędzy systemami) – np. jkowalski – oraz z adresu domenowego komputera, na którym znajduje się skrzynka pocztowa użytkownika – np. alfa.com.pl. Obie części adresu są łączone w jeden ciąg znaków przy pomocy symbolu @, np. jkowalski@alfa.com.pl.

Popularne rozszerzenia technologii poczty elektronicznej umożliwiają przesyłanie dokumentów o charakterze innym niż tekstowy w postaci załączników do wiadomości.

2.2 Tekstowe telekonferencje

Usługa tekstowych telekonferencji (IRC – Internet Relay Chat) składa się z następujących typów komponentów:

1. Programów nazywanych *klientami IRC*, wykorzystywanych przez użytkowników do prowadzenia konwersacji.
2. Programów nazywanych *serwerami IRC*, stanowiących element infrastruktury informatycznej.

Pojedyncza dyskusja nazywana jest *kanalem IRC*. Użytkownicy wybierają kanał, w ramach którego będą prowadzić konwersację. Następnie, krótkie teksty wprowadzane przy pomocy klawiatury do programu klienta IRC są za pośrednictwem serwerów IRC natychmiast dostarczane na ekrany programów klientów IRC pozostałych użytkowników kanału.

2.3 Listy dyskusyjne

Usługa list dyskusyjnych nie wprowadza własnych rozwiązań komunikacyjnych do transmisji wiadomości, lecz wykorzystuje istniejącą technologię poczty elektronicznej. Zbiór adresów poczty elektronicznej użytkowników zainteresowanych konwersacją na wspólny temat nazywany jest *listą dyskusyjną*. Sama lista dyskusyjna również posiada adres poczty elektronicznej. Każda wiadomość wysłana przez uczestnika listy na jej adres jest automatycznie dystrybuowana do pozostałych uczestników listy.

3 Przechowywanie i udostępnianie dokumentów

Jednym z flagowych zastosowań intranetu jest łatwe udostępnianie użytkownikom dokumentów przechowywanych w systemach komputerowych. W tym zakresie najpopularniejszymi z dostępnych technologii są: transfer plików (FTP), grupy dyskusyjne (News), a przede wszystkim WWW (World Wide Web). Transfer plików pozwala na zdalne pobieranie plików/dokumentów ze zautomatyzowanej składnicy. Grupy dyskusyjne są alternatywą dla omawianych wcześniej list dyskusyjnych lecz w przeciwieństwie do list przechowują wiadomości uczestniczące w dyskusji. Usługa WWW umożliwia wygodny dostęp do sieci specjalnie przygotowanych dokumentów multimedialnych, umieszczonych w zautomatyzowanej składnicy.

3.1 Transfer plików

Usługa transferu plików (FTP – File Transfer Program) składa się z następujących komponentów:

1. Programu nazywanego *klientem FTP*, wykorzystywanego przez użytkownika do sterowania pobieraniem plików.

2. Programu nazywanego *serwerem FTP*, stanowiącego element infrastruktury informatycznej, odpowiedzialnego za przechowywanie i przesyłanie plików.

Przy pomocy programu klienta FTP użytkownik przyłącza się do programu serwera FTP i wysyła do niego żądania przesłania wskazanych plików. Serwer FTP odczytuje pliki z dysku i wysyła je siecią komputerową do użytkownika. Program klienta FTP zapisuje pobrane pliki na dysku lokalnym. Komunikacja pomiędzy programem klienta i serwera odbywa się w języku-protokole FTP (File Transfer Protocol).

3.2 Grupy dyskusyjne

Usługa grup dyskusyjnych (News, Usenet News, Network News) składa się z następujących komponentów:

1. Programów nazywanych *klientami News*, wykorzystywanych przez użytkowników do wysyłania i odczytywania wiadomości uczestniczących w dyskusji.
2. Programów nazywanych *serwerami News*, stanowiących element infrastruktury informatycznej, odpowiedzialnych za przechowywanie i udostępnianie wiadomości tekstowych.

Przy pomocy programu klienta News, użytkownik łączy się z serwerem News, wybiera temat dyskusji (tzw. grupę), z którego następnie pobiera wybrane wiadomości. Komunikacja pomiędzy programem klienta News a serwera jest realizowana przy użyciu języka-protokołu NNTP (Network News Transfer Protocol), a pobierane wiadomości są wyświetlane w postaci tekstowej na ekranie komputera użytkownika.

3.3 WWW

Usługa WWW (World Wide Web) jest zbudowana z następujących komponentów:

1. Programów nazywanych *przeglądarkami WWW*, wykorzystywanych przez użytkowników do pobierania i prezentacji dokumentów multimedialnych.
2. Programu nazywanego *serwerem WWW*, stanowiącego element infrastruktury informatycznej, odpowiedzialnego za przechowywanie i udostępnianie dokumentów.

Za pomocą przeglądarki WWW użytkownik przesyła do serwera WWW żądanie pobrania dokumentu. Żądanie to wyrażane jest w formie tzw. *adresu URL* (Uniform Resource

Locator), zawierającego adres IP lub domenowy serwera WWW, katalog, w którym znajduje się dokument, wreszcie nazwę pliku dokumentu, np. `x.com.pl/nowe/demo.html`. Po otrzymaniu żądania, serwer pobiera plik dokumentu z dysku i wysyła go do przeglądarki WWW użytkownika. Przeglądarka WWW wyświetla dokument na ekranie komputera.

W celu ujednoczenia formatów dokumentów udostępnianych poprzez usługę WWW, wprowadzono język opisu dokumentów HTML (Hypertext Markup Language). Językiem tym muszą posługiwać się twórcy dokumentów przechowywanych przez serwer WWW. Istnieje wiele narzędzi edytorskich, które ułatwiają wykorzystywanie języka HTML (np. nawet program Microsoft Word posiada funkcję zapisu dokumentu w tym języku).

Dokumenty WWW mogą zawierać obrazy graficzne, sekwencje wideo, informację dźwiękową, itp. Ponadto, dokumenty te łączone są ze sobą, dzięki czemu dostęp użytkownika do innych dokumentów o podobnej tematyce jest ułatwiony.

4 Generowanie i udostępnianie dokumentów

W ramach przedstawionej powyżej usługi WWW, użytkownik może pobierać z serwera dokumenty uprzednio przygotowane. Często jednak ręczne tworzenie takich dokumentów jest niemożliwe lub zbyt kosztowne. Przykładowo, udostępnianie w intranecie informacji finansowych z Giełdy Papierów Wartościowych wymagałoby praktycznie zatrudnienia dodatkowej osoby, tworzącej nowy, aktualny dokument co piętnaście minut. W związku z tym, stworzono szereg rozwiązań softwarowych, uzupełniających architekturę WWW o mechanizmy dynamicznego tworzenia dokumentów. Wydaje się, że najbardziej charakterystycznymi są: CGI (Common Gateway Interface), serwlety Java oraz JSP (Java Server Pages). Mechanizm CGI umożliwia programistom tworzenie prostych lecz mało wydajnych programów, które po uruchomieniu przez serwer WWW budują kompletny dokument w języku HTML. Serwlety Java są odmianą tej technologii, pozbawioną wad na poziomie wydajności, a ponadto przenaszalną pomiędzy różnymi platformami sprzętowo-systemowymi. Rozwiązanie JSP stosuje pomysł samomodyfikujących się dokumentów HTML.

4.1 CGI

W pierwszym historycznie rozwiązaniu problemu dynamicznej generacji dokumentów WWW, programista przygotowuje samodzielny program, który potrafi konstruować kompletny dokument w języku HTML. Program ten jest zapisywany na dysku serwera WWW, a następnie, na żądanie użytkownika, uruchamiany przez serwer. Wygenerowany

dokument jest przekazywany wprost do przeglądarki użytkownika. W praktyce, użytkownikowi trudno jest rozpoznać, czy otrzymany dokument został stworzony ręcznie przez autora, czy też wygenerowany przez program.

Często wymienianą wadą programów CGI jest konieczność ich kosztownego uruchamiania za każdym razem, kiedy użytkownik przesyła stosowne żądanie oraz zatrzymywania po zakończeniu generowania dokumentu.

4.2 Serwlety Java

Java to nowoczesny, liczący sobie dopiero 5 lat, język programowania. Serwlety to programy tworzone w tym języku i służące do generowania dokumentów w języku HTML. Zasady działania serwletów są zbliżone do zasad przedstawionych dla CGI, z tą jednak różnicą, że serwlet pracuje nieprzerwanie w środowisku serwera WWW, w wyniku czego żądania użytkowników są przez niego obsługiwane bardziej wydajnie. Dodatkową zaletą serwletów jest niezależność ich programisty od konkretnego typu serwera WWW, na którym będą pracowały. Rozszerzenie tej architektury może polegać na wzbogaceniu serwera WWW o szereg dodatkowych funkcji, przydatnych programistom, jak np. mechanizmów bezpieczeństwa, niezawodności, komunikacji z bazami danych – tak rozbudowane serwery WWW nazywane są *serwerami aplikacji* (Application Servers).

4.3 JSP

Pewnym utrudnieniem dla programistów tworzących zarówno programy CGI, jak i serwlety Java jest to, że programy te muszą generować kompletne dokumenty WWW. Często jednak pod pojęciem dynamicznej generacji rozumiemy modyfikację tylko fragmentu, a nie całości dokumentu. Rozwiązanie JSP (Java Server Pages) proponuje ręczne tworzenie dokumentów w języku HTML, w których umieszczane są specjalne „wstawki”, podmieniane na dynamiczną zawartość podczas dostarczania dokumentu użytkownikowi. Wstawki te zapisywane są w języku programowania Java. Technologia JSP cechuje się większą wydajnością pracy programistów, którzy nie muszą już być odpowiedzialni na szatę graficzną dokumentów.

5 Udostępnianie oprogramowania w intranecie

Intra- i Internet oferują również mechanizmy udostępniania tradycyjnego oprogramowania automatyzującego funkcjonalność przedsiębiorstwa. Udostępnianie oprogramowania polega na przesyłaniu kodu programu na komputer użytkownika, a następnie jego lokalnym

wykonaniu. Dzięki temu użytkownicy mogą korzystać z oprogramowania, które nie zostało wcześniej zainstalowane na ich komputerach. Flagową technologią w tym kontekście są tzw. *applety Java*. Applet Java jest niewielkim programem napisanym w języku Java, umieszczonym na dysku serwera WWW. Na żądanie użytkownika posługującego się przeglądarką WWW, applet jest wysyłany przez serwer do klienta i uruchamiany lokalnie na jego komputerze, w ramach zasobów przeglądarki WWW.

6 Bezpieczeństwo połączeń sieciowych

Stosowanie sieci komputerowych (w tym intranetu) w przedsiębiorstwie narzuca pewne rygory bezpieczeństwa połączeń. Podstawowymi problemami w tym kontekście są:

1. Ochrona przed nielegalnym podsłuchem komunikacji sieciowej.
2. Ochrona przed nielegalną modyfikacją danych przesyłanych pomiędzy serwerem a użytkownikiem.
3. Nielegalne podszywanie się osoby nieupoważnionej pod osobę pracownika.

Dostępny dziś rozwiązaniami, pozwalającymi utrudnić lub wykryć wystąpienie każdego z powyższych zagrożeń są: algorytmy szyfrujące, algorytmy kryptograficznych sum kontrolnych i podpisy elektroniczne. Metody te nie są zwykle implementowane przez programistów, lecz dostępne w standardowych pakietach, znajdujących się na wyposażeniu serwerów i przeglądarek (np. SSL, PCT).

6.1 Algorytmy szyfrujące

Algorytmy szyfrujące przekształcają treść dokumentów do postaci nieczytelnej. Do odczytania zaszyfowanego dokumentu konieczne jest jego rozszyfrowanie. Zarówno podczas szyfrowania, jak i rozszyfrowywania, użytkownik musi znać specjalne hasła, nazywane *kluczami*. W zależności od liczby stosowanych kluczy, algorytmy szyfrowania dzielą się na dwie grupy:

1. Symetryczne – do zaszyfrowania i do rozszyfrowania służy ten sam klucz. Wadą tego typu algorytmów jest trudność bezpiecznego przekazania klucza przez nadawcę wiadomości do jej odbiorcy. Przykłady algorytmów symetrycznych: DES, IDEA, CAST, Skipjack, RC2, RC4.
2. Asymetryczne – stosują dwa różne klucze. Wiadomość jest szyfrowana przy użyciu klucza nazywanego *publicznym*, natomiast jej rozszyfrowanie możliwe jest wyłącznie przy pomocy innego klucza – *prywatnego*. W związku z tym klucz szyfrujący (publiczny) może być swobodnie przekazywany nadawcy wiadomości, gdyż jego ewentualne

nielegalne przejście nie stwarza zagrożenia dla bezpieczeństwa. Przykłady algorytmów asymetrycznych: RSA, PKCS, DSS.

6.2 Algorytmy kryptograficznych sum kontrolnych

Kryptograficzne sumy kontrolne umożliwiają wykrycie nielegalnej modyfikacji przesyłanych danych. Idea polega na wyznaczeniu dla wysyłanej wiadomości wartości pewnej funkcji matematycznej – nazywanej *sumą kontrolną*. Wiadomość wraz z sumą kontrolną jest przesyłana do odbiorcy, który ponownie wylicza sumę kontrolną i porównuje jej wartość z przesłaną przez nadawcę. Różnica sum kontrolnych oznacza, iż podczas transmisji wiadomość uległa zmianie. Ponieważ do wyliczenia kryptograficznej sumy kontrolnej konieczna jest znajomość tajnego klucza, to ewentualny włamywacz nie może jej odpowiednio zmodyfikować po wykonaniu modyfikacji wiadomości. Przykłady algorytmów kryptograficznych sum kontrolnych: MD2, MD4, MD5, SHA.

6.3 Podpisy elektroniczne

Zadaniem podpisu elektronicznego jest zagwarantowanie odbiorcy, że przesyłana wiadomość pochodzi od osoby zapisanej jako nadawca, a nie od podszywającego się pod nią oszusta. Podpis elektroniczny jest formą cyfrowej informacji, której nie jest w stanie wygenerować inny użytkownik, niż prawdziwy nadawca wiadomości. Idea podpisu elektronicznego polega głównie na zastosowaniu asymetrycznego algorytmu szyfrowania i użyciu klucza prywatnego nadawcy w taki sposób, aby zaszyfrowana nim część wiadomości mogła być rozszyfrowana tylko przy pomocy klucza publicznego tego nadawcy (odwrotnie niż w przypadku omówionego szyfrowania treści wiadomości). Skuteczne użycie klucza publicznego nadawcy przez odbiorcę wiadomości jest dowodem na to, że wiadomość ta istotnie została zaszyfrowana przy użyciu klucza prywatnego nadawcy – a klucz ten znany jest wyłącznie nadawcy, co dowodzi jego tożsamości.

7 Zagadnienia planowania i budowy intranetu

Przedsiębiorstwo, które podjęło decyzję o budowie intranetu powinno rozważyć następujące zagadnienia, które wpływają na koszt i jakość wdrożenia:

1. Aspekt sprzętowy – gdzie będą zlokalizowane serwery?
2. Wybór technologii systemowych – które z technologii intra- i internetowych są najodpowiedniejsze do realizacji założonej strategii?
3. Inwestycje:

- w serwery i zapewnienie odpowiedniej przepustowości posiadanych sieci komputerowych (np. na potrzeby przesyłania danych multimedialnych)
 - w narzędzia służące do tworzenia oprogramowania intranetowego i dokumentów WWW
 - wynajęcie konsultantów, odpowiedzialnych za koordynację przedsięwzięcia
 - koszty pielęgnacji wdrożonego oprogramowania (konieczność usuwania ewentualnych usterek i dostosowywania systemu do zmieniającej się rzeczywistości)
 - w zabezpieczenia sprzętowe i programowe, zapewniające bezpieczne wykorzystywanie zasobów sieciowych.
4. Wybór przeglądarki (Microsoft vs. Netscape) – wybór wyłącznie jednej pozwala zaoszczędzić 20% pracy programistów (badania Zona Research)
 5. Integracja z istniejącymi systemami i serwerami baz danych – w jakim zakresie nowe oprogramowanie może wykorzystywać posiadaną infrastrukturę informatyczną?

8 Podsumowanie

Technologie intra- i internetowe stały się dominującymi technikami informatycznymi, stosowanymi do automatyzacji procesów w przedsiębiorstwach. W artykule przedstawiono najbardziej charakterystyczne i najważniejsze wg autora technologie softwarowe. Mimo, iż zakres ich zastosowań jest ograniczony wyłącznie wyobraźnią projektanta, to jednak na zakończenie przedstawimy kilka popularnych propozycji:

1. Zarządzanie zasobami ludzkimi: książki telefoniczne/informacje kontaktowe o pracownikach, rozpowszechnianie materiałów informacyjnych/ogłoszeń/gazetek, terminarze zajęć, itp.
2. Handel i marketing: cenniki, opisy produktów, rozpowszechnianie raportów o sprzedaży, itp.
3. Systemy informacyjne: asysta techniczna i centra zgłoszeniowe, multimedialne szkolenia, itp.
4. Zarządzanie i finanse: rozpowszechnianie informacji finansowych, telekonferencje, monitoring prasy, itp.
5. Obsługa klientów: składanie i obserwacja realizacji zamówień, rejestrowanie problemów, udostępnianie katalogów/cenników, itp.

9 Bibliografia

- [1] Exploring Java, P. Niemeyer, J. Peck, O'Reilly & Associates, Inc., 1996
- [2] Hypertext Transfer Protocol -- HTTP/1.0, T. Berners-Lee, R. Fielding, H. Frystyk,
Internet Draft
- [3] Internet Standards and Protocols, D.C. Naik, Microsoft Press, 1998
- [4] Java, A. van Hoff, S. Shaio, O. Starbuck, Helion, 1996
- [5] Java™ 2 Platform: Technology for the Enterprise, L. Perlstein, <http://java.sun.com/>
- [6] The WWW Common Gateway Interface Version 1.1, D.R.T. Robinson, Internet Draft