

1. Internet

W latach sześćdziesiątych ubiegłego stulecia rozpoczęto prace nad zbudowaniem sieci komputerowej, która mogła by zapewnić komunikację między komputerami pracującymi w różnych sieciach lokalnych i znajdujących się w dużych odległościach od siebie. Pierwsze tego rodzaju sieci powstały w ośrodkach wojskowych, a następnie zaczęły szybko się rozwijać w ośrodkach naukowych i akademickich.

Internet to sieć komputerowa o zasięgu globalnym łącząca ze sobą setki tysięcy małych i dużych sieci lokalnych oraz pojedyncze komputery.

Dzisiaj Internet jest siecią o zasięgu globalnym i jest dostępny dla każdego użytkownika.

Ogólna struktura sieci Internet składa się z sieci lokalnych i pojedynczych komputerów oraz z szybkiej sieci szkieletowej. Sieć szkieletowa służy do przekazywania danych między sieciami lokalnymi.

Wśród komputerów pracujących w sieci można wyróżnić dwa typy nazywane serwerami i klientami. *Serwer* to komputer, który udostępnia swoje zasoby innym komputerom w sieci, np. udostępnia dane z sesji giełdowej. *Klient* to komputer, który korzysta z zasobów serwerów, np. domowy komputer, który łączy się poprzez sieć z serwisami informacyjnymi. (Nowoczesny system organizacji sieci zbudowany na podstawie tych założeń nosi nazwę modelu klient-serwer). Komputer podłączony do sieci Internet nazywa się też hostem.

Możliwości wykorzystania sieci Internet są bardzo różnorodne. Wśród nich można wyróżnić kilka najważniejszych. Noszą one nazwę usług. Do najważniejszych usług oferowanych w sieci Internet należą:

- WWW
- Poczta elektroniczna
- Grupy dyskusyjne
- IRC– (ang. Internet Relay Chat)
- FTP

(1.1) WWW

World Wide Web (WWW) powstał około 1990 roku. Za jego twórcę uważany jest Anglik Sir Tim Berners-Lee, pracownik CERN (ang. *European Organization for Nuclear Research* – Europejskie Centrum Badań Molekularnych w Szwajcarii). Pracował on razem z grupą programistów nad stworzeniem systemu wymiany informacji dla naukowców. Umożliwił komputerom znajdującym się

fizycznie w różnych miejscach wyświetlanie tekstu, który zawierał łącza (ang. *links* – po polsku odnośniki lub potocznie linki) do innych plików. Był to początek języka HTML (ang. *Hypertext Markup Language* – język hipertekstowego znakowania informacji), języka do tworzenia stron (zwanych też witrynami lub dokumentami) internetowych. Łącza stały się cechą zasadniczą witryn. Dzięki nim strony mogą zawierać informacje o wzajemnych relacjach, użytkownik może przechodzić z jednego dokumentu do drugiego. Tim Berners-Lee opracował także pierwowzory protokołu HTTP (ang. *Hypertext Transfer Protocol* – protokół przesyłania dokumentów HTML) i URL (ang. *Universal Resource Locator* – adres internetowy, unikalny identyfikator każdego dokumentu w sieci).

WWW jest tak zwanym systemem hipermedialnym umożliwiającym użytkownikowi dostęp interakcyjny do gromadzonych informacji.

Pojęcie systemu hipermedialnego stanowi rozszerzenie zwykłego systemu hipertekstowego.

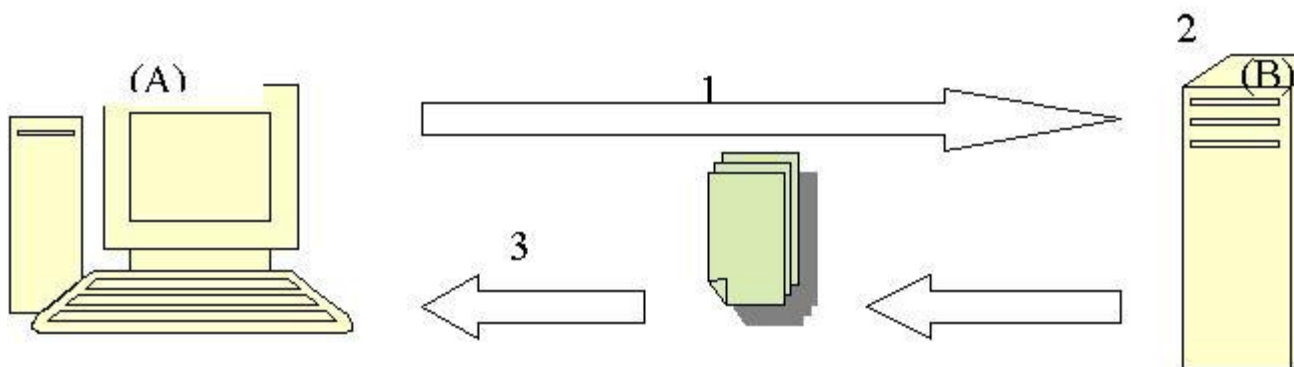
Oba te systemy są zbiorem dokumentów. Każdy z tych dokumentów ma jednak jedną charakterystyczną cechę: zawiera odnośniki (odsyłacze) do innych dokumentów w zbiorze, które to z kolei mogą zawierać odsyłacze do następnych dokumentów i tak dalej. Każdy odsyłacz zawarty w dokumencie jest związany ze wskazywaną przez użytkownika częścią dokumentu. Wybranie takiego odsyłacza powoduje uzyskanie dostępu do kolejnej informacji umieszczonej w systemie.

Na uwagę zasługuje fakt, że cały system usługi WWW jest specyficzną bazą danych. Stanowi on przykład rozproszonej bazy danych.

Ze względu na rozproszoną strukturę systemu hipermedialnego mogą zdarzać się sytuacje w których odsyłacze zawarte w dokumentach tego systemu są niepoprawne. Natrafiać na taką sytuację otrzymujemy najczęściej komunikat, o tym że dana strona nie jest dostępna.

(1.2) Zasada działania WWW

WWW działa w oparciu o pewien schemat, który wygląda zawsze tak samo bez względu na zastosowane oprogramowanie i technologie. Schemat działania WWW ilustruje rysunek 1:



Rysunek 1 Schemat działania World Wide Web oraz protokołu HTTP

Pierwszym krokiem jest otwarcie połączenia za pomocą protokołu transmisji TCP pomiędzy komputerem, na którym zainstalowana jest przeglądarka internetowa (A), a serwerem (B). Następne zadania są realizowane przez protokół HTTP, który jest położony w wyższej warstwie (w warstwie aplikacji według modelu warstwowego OSI) niż protokół TCP.

Protokół HTTP uczestniczy w następujących procesach (numery odpowiadają numerom na rysunku):

1. przeglądarka wysyła zlecenie do serwera WWW; każde zlecenie składa się z metody, identyfikatora zasobu oraz oznaczenia wersji protokołu; po wysłaniu pierwszej linii zlecenia przesyłane są inne informacje o serwerze WWW w formie nagłówków HTTP,
2. jeśli wysłane przez przeglądarkę zlecenie dotyczy dynamicznie generowanego dokumentu HTML, serwer przetwarza odpowiedni skrypt generując odpowiedni kod HTML,
3. po otrzymaniu zlecenia serwer odpowiada klientowi, wysyła nagłówki odpowiedzi – informacje o sobie i zwracanym dokumencie, informacje dotyczące oprogramowania serwera oraz ustawienia cookie; po wysłaniu nagłówków serwer wysyła pustą linię, a następnie właściwy dokument w formacie określonym przez jeden z nagłówków wysłanych przez przeglądarkę.

Gdy protokół HTTP zakończy wysyłanie danych, zamykane jest połączenie TCP.

(1.3) Format HTML

Język HTML jest językiem opisu struktury. Nie służy on do podawania szczegółów dotyczących wyglądu strony. Podaje on natomiast informacje o strukturze dokumentu, czyli wskazówki dla przeglądarki w jaki sposób ma go pokazywać. Stąd też ten sam dokument zapisany w HTML może być różnie prezentowany w zależności od zastosowanej przeglądarki.

HTML jest też językiem bezformatowym. Oznacza to że przejścia do nowego wiersza, czy wielokrotne spacje są tutaj ignorowane. Wymagana struktura dokumentu uzyskiwana jest za pomocą specjalnych elementów języka zwanych znacznikami.

Znacznik ma postać nazwy ujętej w znaki mniejszości i większości. Znacznik NAZWA może więc w dokumencie HTML wyglądać tak:

```
<NAZWA>
```

Taki znacznik określa operację lokalną lub początek operacji formatującej.

Znacznik określający koniec operacji wygląda podobnie, z tym przed jego nazwą umieszcza się znak prawego ukośnika. Może więc wyglądać tak:

```
</NAZWA>
```

(1.4) Struktura dokumentu HTML

Każdy dokument zapisany w języku HTML jest podzielony na dwie główne części:

Nagłówka

Zawiera szczegółowe informacje o dokumencie; umieszczany jest między znacznikami <HEAD> i </HEAD>

i

Treści dokumentu

Czyli całej reszty; umieszczana jest między znacznikami <BODY> i </BODY>.

Podstawowa struktura dokumentu w języku HTML wygląda tak:

```
<HTML>
```

```
  <HEAD>
```

```
    <TITLE>
```

tutaj jest tytuł dokumentu

</TITLE>

</HEAD>

<BODY>

tutaj jest treść dokumentu

</BODY>

</HTML>

Znaczenie poszczególnych znaczników jest następujące:

Para <HTML>, </HTML> oznacza początek i koniec dokumentu HTML

Para <HEAD>, </HEAD> oznacza początek i koniec nagłówka

Para <BODY>, </BODY> oznacza początek i koniec treści dokumentu

Para <TITLE, </TITLE> w obrębie nagłówka oznaczają początek i koniec tytułu dokumentu.

Tekst dokumentu HTML może być zapisany tak jak powyżej, czyli z odpowiednimi wcięciami. Nic nie stoi jednak na przeszkodzie, aby ten dokument zapisać tak:

```
<HTML><HEAD><TITLE>tutaj jest tytuł dokumentu </TITLE></HEAD><BODY>tutaj jest treść dokumentu </BODY></HTML>
```

Z punktu widzenia przeglądarki stosowanie wcięć czy spacji nie ma znaczenia. Ta pierwsza forma jest jednak dużo czytelniejsza. Łatwiej w niej zorientować się w strukturze dokumentu i znaleźć ewentualne błędy.

Przykładowe znaczniki

BR

Znacznik
 nakazuje przeglądarce przejście do nowego wiersza. Pomiedzy wiersze nie jest wstawiany żaden odstęp.

P

Znacznik <P> definiuje akapit

H*i*

Znacznik <H*i*>, gdzie $i = 1, 2, \dots, 6$, oznacza początek nagłówka poziomu i . Na przykład tekst nagłówka najbardziej znaczącego poziomu jest ujęty w znaczniki <H1> , </H1>.

UL

Znacznik oznacza początek listy nieuporządkowanej, czyli takiej w której lista nie jest porządkowana według numerów. Każda pozycja na liście jest poprzedzana znacznikiem . Przeglądarka najczęściej poprzedza każdy element listy kropką.

IMG

Znacznik ten pozwala na dołączenie grafiki do dokumentu. Jego użycie może być następujące

```
<IMG SRC="moj_obrazek.gif">
```

Zastosowanie takiego znacznika spowoduje umieszczenie na stronie WWW w miejscu zastosowania znacznika obrazka zapisanego w pliku moj_obrazek.gif.

A

Za pomocą znacznika <A> realizowane są odsyłacze hipertekstowe. Znacznik ten definiuje tak zwaną kotwicę (ang. anchor). Aby użyć dowolny fragment dokumentu jako odsyłacza, wskazany fragment dokumentu należy ująć między znacznikami <A> , . Kotwica może wskazywać na połączenie do obiektu docelowego lub do innego miejsca w tym samym dokumencie. Jeśli połączenie ma być realizowane do innego dokumentu, to położenie obiektu definiowane jest przez parametr HREF.

Prosty odsyłacz hipertekstowy może wyglądać tak:

```
<A HREF="wstep.htm"> Wprowadzenie </A>
```

Umieszczenie takiego wpisu w dokumencie spowoduje zaznaczenie tekstu Wprowadzenie jako odsyłacza do dokumentu wstep.htm.

[NASTĘPNA](#)

2. Protokół FTP

Protokół FTP (ang. *File Transfer Protocol*) jest protokołem ogólnego przeznaczenia wykorzystywany w najbardziej rozpowszechnionych usługach przesyłania plików.

Protokół FTP umożliwia:

- przesłanie dowolnego pliku
- określenie właściciela pliku
- określenie ograniczenia dostępności
- ukrycie szczegółów poszczególnych systemów operacyjnych

Jest to też jeden z najstarszych protokołów wykorzystywanych w sieci Internet. Powstał wcześniej niż protokoły TCP czy IP.

Najważniejszą cechą protokołu FTP, jest to że może on być wykorzystywany do kopiowania dowolnego pliku z jednego komputera na inny.

(2.1) Polecenie FTP

Interfejs dostępny dla użytkownika korzystającego z z FTP może być różny w zależności od implementacji, gdyż standard protokołu FTP nie definiuje interfejsu użytkownika a jedynie sposób w jaki oprogramowanie jednego komputera współpracuje z oprogramowaniem drugiego komputera.

Najpopularniejszy jest interfejs opracowany dla systemu UNIX BSD. Obejmuje on ponad 50 poleceń.

Po podaniu polecenia *open* protokół FTP tworzy połączenie TCP z odległym komputerem. Połączenie to nosi nazwę połączenia sterującego. Połączenie sterujące służy tylko do przekazywania poleceń, a nie do przesyłania danych.

Po ustanowieniu połączenia sterującego oprogramowanie FTP wymaga podania danych pozwalających na identyfikację użytkownika. Należy więc podać identyfikator oraz hasło.

Po uzyskaniu autoryzacji można już zlecać wykonywanie poleceń w tym przesyłanie plików.

Wszystkie polecenia przesyłane są połączeniem sterującym. Jeśli użytkownik zamierza zamknąć połączenie sterujące musi wydać polecenie *close*.

Close nie kończy jednak działania programu FTP.

(2.2) Autoryzacja w FTP

Możliwość autoryzacji użytkowników w FTP jest bardzo cenna, gdyż pozwala na kontrolę dostępu do zasobów serwera FTP.

Wiele jednak serwerów FTP udostępnia swoje zasoby dowolnemu użytkownikowi. Aby uniknąć konieczności nadawania wszystkim użytkownikom, którzy zamierzają skorzystać z zasobów serwera FTP kont, wprowadzono jedno specjalne konto. Kont to ma identyfikator *anonymous*. Identyfikator ten pozwala na dostęp do zasobów serwera dowolnemu użytkownikowi. Czasami zdarza się, że z kontem *anonymous* związane jest hasło dostępu. Najczęściej jako hasło dostępu dla konta *anonymous* należy podać swój adres e-mail.

(2.3) Typy plików

Program FTP obsługuje dwa podstawowe rodzaje plików. Są to

- binarne
- tekstowe

Korzystając z protokołu ftp trzeba ustalić przed przesłaniem pliku tryb transmisji (czyli typ przesyłanego pliku)

Tryb tekstowy służy do przesyłania podstawowych plików tekstowych podzielonych na wiersze.

Tryb binarny służy do przesyłania wszystkich innych rodzajów plików.

W takim trybie należy więc przysyłać pliki zawierające obrazki, muzykę, programy czy dokumenty Worda.

(2.4) Połączenia w FTP

Jak już wspomnieliśmy wcześniej po wydaniu polecenia *open* protokół FTP otwiera połączenie TCP, którym przesyłane są komendy interfejsu FTP. Połączenie to nazywane jest połączeniem sterującym.

Natomiast pliki nie są przesyłane połączeniem sterującym. Dla każdego pliku ustanawiane jest oddzielne połączenie i tym połączeniem przesyłany jest plik. W celu uniknięcia konfliktów między połączeniem sterującym a połączeniem dla danych każde z nich wykorzystuje inny port.

W trakcie sesji FTP połączenie sterujące pozostaje stałe i nie zmienia się. Połączeni dla danych zmieniają się natomiast często. W trakcie sesji są najczęściej wielokrotnie otwierane i zamykane.

Dzięki rozdzieleniu połączeń dla danych i poleceń sterujących można:

- łatwiej dokonać implementacji
- wysłać polecenia w trakcie przesyłania pliku (np. domagające się przerwania transmisji pliku)
- korzystać z warunku końca pliku

[NASTĘPNA](#)

3. Poczta elektroniczna

Jest to serwis zapewniający wymianę wiadomości pomiędzy użytkownikami Internetu. Wiadomości mogą mieć postać albo tekstów (w znakach ASCII lub w kodowaniu narodowym - czyli zawierającym np. polskie znaki diakrytyczne) albo załączników binarnych. Szczególnie obecność załączników otwiera przed użytkownikami poczty elektronicznej nieograniczone możliwości – w załączniku może znajdować się dowolny plik binarny – np. pliku z edytora typu .doc, plik z obrazkiem typu .jpg albo plik z wiadomością głosową lub muzyką lub filmem.

Użycie poczty elektronicznej składa się więc z dwu elementów:

- odbieranie poczty – poczta dla nas przychodzi do naszej skrzynki pocztowej na serwerze; żeby móc ją przeczytać musimy ją odebrać z serwera przy pomocy protokołu POP3 lub IMAP4; należy pamiętać, że odebranie poczty oznacza przeważnie przeniesienie wszystkich wiadomości z serwera na nasz komputer – na serwerze te wiadomości przestają istnieć; na komputerze klienckim możemy odczytać poszczególne wiadomości przy użyciu programu pocztowego.
- wysyłanie poczty – polega na stworzeniu wiadomości na kliencie i przesłaniu jej dalej do serwera dostawcy usług, który przesyła wiadomość dalej do serwera, gdzie jest skrzynka odbiorcy wiadomości. Przesłanie polega na nawiązaniu przez program pocztowy klienta sesji wykorzystującej protokół SMTP z programem poczty elektronicznej na serwerze.

(3.1) MIME

Protokół MIME służy do przesyłania różnych typów danych takich jak czysty tekst, tekst formatowany, rysunki, dźwięk, obraz, dokumenty HTML itp.

Dzięki standardowi MIME możemy wysyłać informacje różnego typu. Mogą one być odebrane nawet przez osoby, które pracują z terminalami nie obsługującymi grafiki. W takim wypadku standard MIME wyświetli u użytkownika końcowego tylko tekstowy fragment naszej wiadomości pocztowej.

W standardzie MIME możemy również dołączać linki do innych dokumentów znajdujących się na przykład na serwerach FTP.

(3.2) POP

Aby list wysłany protokołem SMTP mógł być odebrany, komputer odbiorcy musi być dostępny. Dlatego też poczty nie wysyła się na stacje klienckie, ale do specjalnych serwerów pocztowych czyli hostów SMTP. Hosty te przechowują pocztę. W momencie, kiedy chcemy przeczytać naszą pocztę musimy połączyć się z serwerem pocztowym. Do tego celu służy protokół POP3. Wykorzystuje on jako warstwę transportową protokół TCP/IP. Serwer POP3 oczekuje na połączenie pod numerem

portu 110.

Standardowo w protokole POP3 identyfikator użytkownika oraz hasło przesyłane są otwartym tekstem. Powoduje to znaczne zagrożenie dla bezpieczeństwa systemu. Dlatego też istnieje również możliwość korzystania z bezpiecznego uwierzytelniania konta. Metoda ta bazuje na standardzie MD5.

(3.3) IMAP4

IMAP4 jest protokołem umożliwiającym dostęp do wiadomości pocztowych znajdujących się na serwerze pocztowym (bez konieczności ściągania ich z serwera) oraz zarządzania nimi bezpośrednio na serwerze pocztowym.

Protokół ten zapewnia więc możliwość takiego manipulowania skrzynkami pocztowymi znajdującymi się na serwerze, jak gdyby były one na stacji lokalnej klienta oraz możliwość synchronizacji skrzynek lokalnych i skrzynek znajdujących się na serwerze.

Najważniejsze zalety protokołu IMAP4 to:

- możliwość wykonywania operacji na wiadomościach bez konieczności ściągania ich z serwera;
 - przeglądanie wiadomości i załącznika bez pobierania ich z serwera;
 - praca w trybie off-line;
 - synchronizacja skrzynek lokalnych i skrzynek na serwerze.

[NASTĘPNA](#)

4. Rodzaje dokumentów WWW

Wśród dokumentów WWW możemy wyróżnić trzy rodzaje. Są to dokumenty:

- statyczne
- dynamiczne
- aktywne

(4.1) Dokumenty statyczne

Dokument statyczny to „standardowy” dokument WWW zapisany w języku HTML raz na zawsze. Nie zmienia się on więc w czasie swego cyklu życia.

Zaletą tego typu dokumentów jest przede wszystkim ich prostota. Mogą one być tworzone przez niezbyt zaawansowanych użytkowników. Jeśli go przetestujemy, to mamy gwarancję, że pozostanie on poprawny do końca swego życia. Z takim dokumentem łatwo też poradzi sobie nasz serwer WWW oraz przeglądarka.

Z dokumentem statycznym wiążą się też pewne wady. Do zasadniczych należy ... niezmienniczość zawartych w nim informacji. Tak to właśnie jest, że to co jest zaletą może być jednocześnie wadą.

Zmiany w treści dokumentu statycznego wymagają za każdym razem ingerencji autora. Dokumenty takie są więc nieelastyczne, a wprowadzanie do nich zmian wiąże się z pewnymi kosztami, stratą czasu i możliwością wprowadzenia przypadkowych błędów.

(4.2) Dokumenty dynamiczne

Dokumenty dynamiczne cechują się tym, że mogą pokazywać aktualne informacje bez konieczności dokonywania zmian w tych dokumentach.

Mogą więc służyć do prezentowania informacji, która często się zmienia – takich jak na przykład wyniki sesji giełdowych, danych o pogodzie itp.

W jaki sposób działają dokumenty dynamiczne? Tak naprawdę są to dokumenty statyczne, ale przygotowywane „na bieżąco”. Po odebraniu przez serwer żądania przesłania dokumentu dynamicznego, serwer najpierw uruchamia specjalny program, który na podstawie aktualnych danych tworzy dokument HTML (statyczny). Z punktu widzenia przeglądarki dokumenty dynamiczne i statyczne są więc nierozróżnialne.

Do głównych zalet dokumentów dynamicznych należy ich duża elastyczność.

Nie ma jednak róży bez kolców. Dokumenty dynamiczne wymagają większych nakładów zarówno od autora tych stron (wiedza, doświadczenie) jak i od serwera WWW, który musi wykonać więcej operacji przed przesłaniem strony do klienta.

Strony dynamiczne muszą być znacznie uważniej opracowane i lepiej przetestowane niż strony statyczne. Ich ściągnięcie z serwera zajmuje też zazwyczaj więcej czasu niż stron statycznych.

Wadą tych stron jest też i to, że po przesłaniu ich do przeglądarki pozostają już niezmiennione. Użytkownik nie może więc pracować z nimi interaktywnie. Jeśli chce coś zmienić (np. układ danych), to musi ściągnąć z serwera kolejną stronę.

(4.3) Dokumenty aktywne

Brak interaktywności strony dynamicznej może być usunięty poprzez zastosowanie stron aktywnych. Dokument aktywny nie jest, w przeciwieństwie do pozostałych typów, w pełni określony przez serwer. W jego skład wchodzi program, który może obliczać i wyświetlać wartości w trakcie przeglądania dokumentu przez użytkownika

Tak więc zawartość strony aktywnej może się cały czas zmieniać, a decyzję o jej wyglądzie podejmuje użytkownik.

Dokument aktywny może też , co jest bardzo ważne, w sposób ciągły sięgać do swoich źródeł informacji i uaktualniać dane wyświetlane (np. notowania giełdowe).

Za te udogodnienia płacimy jednak większymi nakładami poniesionymi na stworzenie dokumentu oraz większymi kosztami jego prezentacji (przeglądarka musi umieć więcej i być zainstalowana na wydajniejszym systemie). Dodatkowo dokument aktywny stanowi potencjalne zagrożenie dla bezpieczeństwa systemu, gdyż może on również udostępniać informacje zgromadzone na komputerze przeglądarki i w pewnym sensie może przejąć kontrolę nad częścią jego zasobów.

[NASTĘPNA](#)

Podsumowanie

Internet stał się w ciągu ostatnich dwudziestu lat ważnym środkiem komunikacji międzyludzkiej. Jego popularność ciągle rośnie, a wraz z nią liczba użytkowników. Najważniejszymi usługami wykorzystywanymi w sieci Internet są usługi WWW, ftp oraz e-mail.