Michał Małyszko

Jak działa Internet i co z tego wynika dla prób jego regulacji?

1. Wstęp

Celem niniejszego dokumentu jest opisanie w przystępny i zrozumiały sposób zasad działania sieci Internet oraz proponowanych narzędzi, które służyć mają jej ewentualnej regulacji.

Dokładniejsza znajomość zasad działania Internetu nie jest powszechna, co utrudnia prowadzenie merytorycznej dyskusji na temat regulacji sieci i skutków, jakie ta regulacja może spowodować. Nasze opracowanie ma na celu przedstawienie mechanizmów działania Internetu, a następnie pokazanie na tym tle proponowanych narzędzi regulacyjnych. Omawiając proponowane narzędzia regulacji Internetu, staramy się wyjaśnić mechanizm ich działania, a następnie opisujemy, jak wdrożenie danego narzędzia wpłynąć może na samą sieć i możliwości jej dalszego rozwoju.

Kiedy mówimy o proponowanych narzędziach regulacji Internetu, mamy na myśli wszelkie propozycje, które mają na celu kontrolowanie sposobu i zakresu korzystania z sieci Internet (np. poprzez ograniczenie dostępnych w niej treści, monitoring przesyłanych treści).

2. Czym jest Internet?

Wiele nieporozumień w dyskusjach na temat Internetu wynika z odmiennego rozumienia tego pojęcia przez dyskutujących. W naszym opracowaniu **patrzymy na Internet od strony technicznej, rozumiejąc go jako zbiór sieci umożliwiających transmisję danych pomiędzy podłączonymi do nich urządzeniami**. Opisujemy Internet od strony technicznej, ponieważ w naszym przekonaniu jest to potrzebne dla poprawnego zrozumienia jego obecnych cech oraz ewentualnych skutków wdrażania narzędzi regulacji Internetu. Narzędzia te opierają swoje działanie na rozwiązaniach technicznych. Aby móc poprawnie ja opisać i ocenić, powinniśmy więc poruszać się w tej płaszczyźnie.

Nie będziemy tu mówić o Internecie jako zjawisku społecznym czy kulturowym. Nie postrzegać też będziemy sieci przez pryzmat usług świadczonych z jej wykorzystaniem, co zdaje się być dość powszechnym potocznym rozumieniem pojęcia Internet („Internet to e-mail, strony WWW, komunikatory i programy do ściągania plików”).

Poniższy opis z założenia nie jest wyczerpujący i w wielu miejscach stosuje uproszczenia. Staramy się, by te uproszczenia nie fałszowały opisywanych mechanizmów działania sieci. Bardziej szczegółowe opis dołączamy jako załączniki do tego rozdziału.

W niektórych miejscach używamy analogii, by ułatwić zrozumienie działania sieci. Jesteśmy świadomi ograniczeń i zagrożeń wynikających z użycia analogii, jednakże uznaliśmy, że jest to wysoce efektywna metoda służąca zrozumieniu przekazywanej wiedzy.

2.1. Definicja Internetu

W naszym dokumencie posługujemy się prostą definicją Internetu[1], która podkreśla jego główną funkcję, którą jest przesyłanie danych.

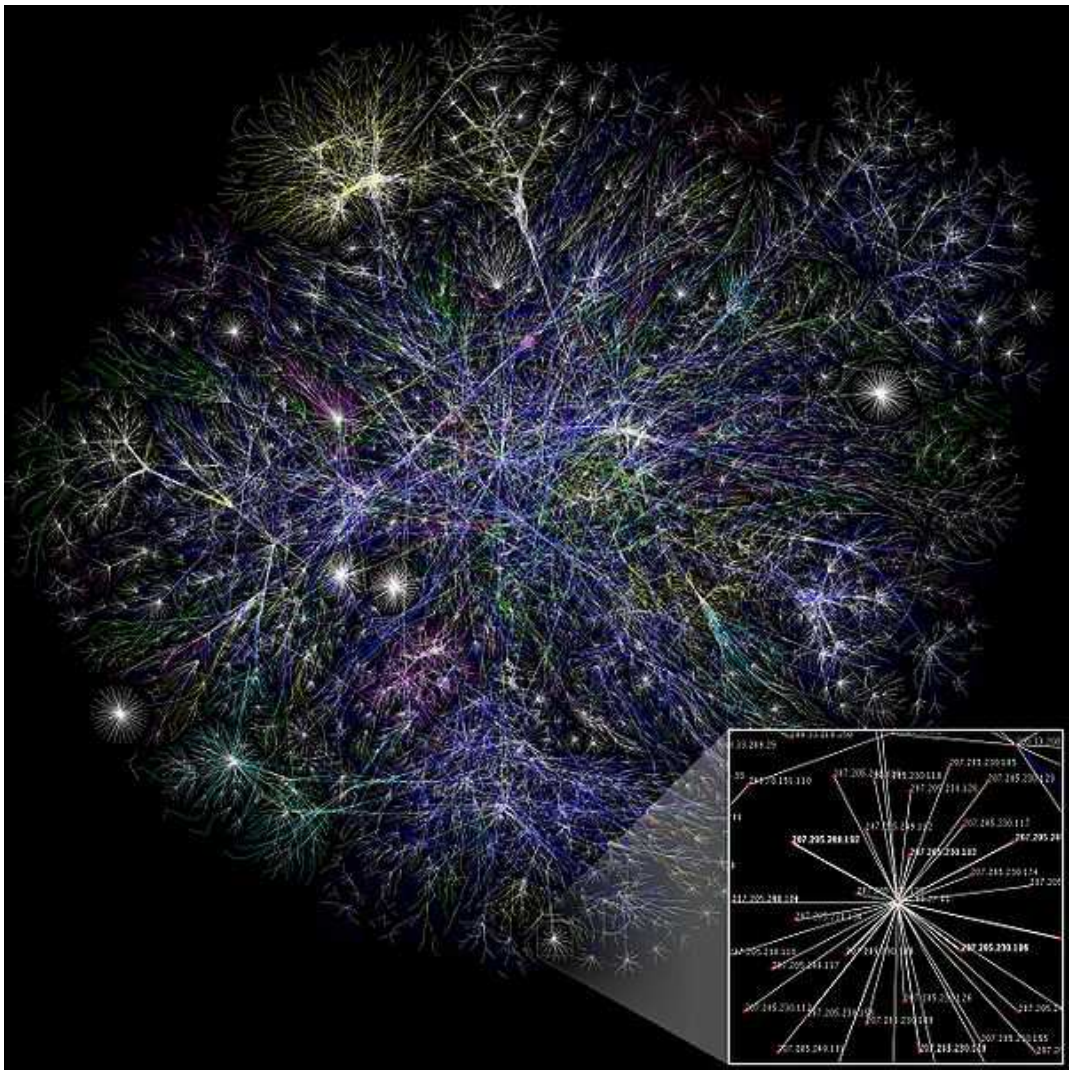
> **Internet to połączone ze sobą sieci, które przesyłają dane pomiędzy urządzeniami podłączonymi do tych sieci.**

Uszczegóławiając tę definicję, można powiedzieć, że:

> **Internet to połączone ze sobą tysiące sieci działające według odmiennych standardów, które dzięki temu, że potrafią „rozmawiać" ze sobą we „wspólnym języku" (protokole internetowym), potrafią przesłać dane z punktu A w dowolnej z tych sieci do punktu B w dowolnej z tych sieci.**

Powyższe definicje podkreślają dwie ważne cechy Internetu:

1) Internet jest „siecią sieci" – składa się z tysięcy różnorodnych połączonych ze sobą sieci,
2) głównym zadaniem Internetu jest transport danych pomiędzy nadawcą i odbiorcą.



Rys. 1. Sieć sieci – wizualizacja połączeń pomiędzy głównymi sieciami tworzącymi Internet
źródło: http://commons.wikimedia.org/wiki/File:Internet_map_1024.jpg, autor: Matt Britt

[1] Jesteśmy świadomi, że istnieje wiele definicji Internetu. Naszym celem w tym opracowaniu nie jest stworzenie nowej, wyczerpującej definicji. Nasz definicja ma na celu podkreślenie niedostrzeganej często różnicy między siecią a usługami świadczonymi z jej udziałem.

Jak napisaliśmy, Internet przesyła dane pomiędzy urządzeniami. **Istnieją dwa, zasadniczo odmienne w swojej funkcji, rodzaje tych urządzeń**:

1) **urządzenia końcowe**, **zlokalizowane na brzegu sieci**, a więc „klienci" Internetu, które używają sieci do wysyłania i odbieranie danych przeznaczonych dla nich samych:
   - są to urządzenia używane przez internautów w celu korzystania z Internetu, takie jak **komputery osobiste, telefony komórkowe, tablety** itp.;
   - są to też wszelkiego rodzaju urządzenia serwujące treści dla użytkowników, takie jak **serwery WWW** dostawców treści; oraz urządzenia koordynujące i udostępniające usługi, takie jak **serwery usług** komunikatorów internetowych, poczty elektronicznej, serwery DNS itp.
2) **urządzenia szkieletowe, kierujące i dokonujące transportu danych w Internecie, zlokalizowane wewnątrz sieci,** takie jak routery, serwery proxy itp.

2.2. Główne cechy sieci Internet

Internet powstał w latach 60. XX wieku w Stanach Zjednoczonych, by połączyć ze sobą pierwotnie kilka oddalonych ośrodków posiadających systemy komputerowe. Dzięki podstawowym założeniom przyjętym przy tworzeniu zasad transportowania danych w Internecie (tworzeniu protokołów komunikacyjnych dla Internetu), takim jak:

- otwarta specyfikacja Internetu,
- decentralizacja zarządzania siecią,
- wspólny język Internetu – protokół internetowy TCP/IP,
- przełączanie pakietów zamiast zestawiania kanału komunikacyjnego na czas transmisji,
- zasada *best-effort*,
- zasada *end-to-end* („głupia sieć" i inteligentne urządzenia na jej brzegach).

Internet stał się nieskończenie[2] skalowalnym systemem, co pozwala mu na ciągły wzrost, zapewnia wysoką niezawodność działania i odporność na uszkodzenia.

2.2.1. Otwarta specyfikacja Internetu

W odróżnieniu od wielu innych przedsięwzięć mających na celu stworzenie systemów służących komunikacji **Internet powstał jako projekt niekomercyjny**. W jego początkowy rozwój zaangażowane były agencje wojskowe, środowisko akademickie oraz National Science Foundation w USA. Internet uniknął „grzechu pierworodnego" większości komercyjnych projektów telekomunikacyjnych: dążenia właściciela projektu do pełnej kontroli nad tym, kto, w jaki sposób, kiedy, za jaką opłatą i z wykorzystaniem jakich urządzeń może komunikować się w danej sieci.

W Internecie, **dzięki prostocie i publicznej (otwartej) specyfikacji użytych protokołów komunikacyjnych, nastąpił bardzo szybki rozwój i wysoka konkurencja** producentów urządzeń, oprogramowania i dostawców usług. W jednym z pierwszych zamówień publicznych ogłoszonym w początkowej fazie budowy sieci z góry zamówiono trzy różne zestawy oprogramowania realizującego funkcje internetowe – chcąc zapewnić niezależność działania sieci od dostawców sprzętu i oprogramowania oraz ułatwić powszechną wymianę informacji według ustalonych standardów (tzw. interoperacyjność). W tym aspekcie Internet jest znacząco różny od innych, zamkniętych systemów telekomunikacyjnych (jak np. telefonia komórkowa GSM), w których chcąc stać się ich częścią, należy ponieść koszty związane z nabyciem licencji na

[2] Prawie nieskończenie: przyjęta liczba dostępnych adresów IP w pierwotnej wersji tego protokołu zostanie rozdysponowana prawdopodobnie do końca 2011 r.

wykorzystanie używanych rozwiązań technicznych. W przypadku Internetu usługi zapewnienia dostępu do sieci czy dowolne usługi świadczone przy pomocy samej sieci zaoferować może każdy podmiot, przy niskim koszcie wejście na ten rynek.

Szeroka konkurencja w usługach fizycznego dostępu do sieci, w hurtowych usługach dostępu do Internetu dla mniejszych sieci oraz w świadczeniu dowolnych usług wykorzystujących sieć, napędzana otwartością standardów, powoduje ciągły rozwój Internetu.

2.2.2. Decentralizacja zarządzania siecią

Internet nie posiada zarządzającego nim podmiotu – ani w kwestiach technicznych, ani biznesowych, ani w kwestii określania sposobu i jakości samego dostępu do sieci dla użytkowników końcowych. Każda z sieci stanowiących część Internetu jest niezależnym podmiotem. Każda samodzielnie decyduje o połączeniu się z innymi sieciami, zasadach i zakresie transportowania przez siebie danych czy też o warunkach podłączania do siebie użytkowników końcowych. Mimo to (a być może właśnie **dzięki** tej niezależności składających się na niego sieci!) Internet stale rośnie i rozwija się.

W Internecie nie znajduje się żaden centralny techniczny ośrodek odpowiedzialny za jego działanie (brak jest centralnego systemu sterującego). Nie da się więc zdestabilizować działania sieci poprzez atak na jakąś jej cześć czy też odcięcie tej części od reszty sieci. Ta niezawodność funkcjonowania Internetu podkreślana jest często zdaniem, wg którego Internet został zaprojektowany tak, aby przetrwać atak atomowy. I choć taki cel nie przyświecał twórcom sieci, to charakterystyczna dla Internetu autonomia działania poszczególnych jego części zapewniła mu bardzo wysoką odporność na ewentualne uszkodzenia.

Jak podaliśmy powyżej, nie istnieje jeden podmiot zarządzający Internetem, lecz dla zapewnienia poprawności i stabilności działania sieci **pewne elementy sieci są zarządzane centralnie**. W szczególności dotyczy to **przydzielania adresów internetowych (adresów IP)** oraz **zarządzania najwyższym poziomem systemu DNS (Domain Name System),** a więc systemem tłumaczącym nazwy domen (np. www.panoptykon.org) na ich adres internetowy (w tym przypadku: 174.121.79.98).

2.2.3. Internet jako metasieć

Specyfika rozwiązania użytego przy budowie Internetu polega na tym, że jego częściami mogą stać się (i stały się) całkowicie odmienne od siebie typy sieci. Każdy z tych typów działa z wykorzystaniem innych wewnętrznych i specyficznych dla siebie sposobów transportowania danych: każda sieć ma własny rodzaj „kontenera" służącego transportowaniu danych (w języku inżynierów są to różnego rodzaju pakiety, ramki, komórki itp.). Nie ma tu znaczenia, czy dany typ sieci jest stary czy nowy; czy powstał w celu obsługi połączeń telefonicznych, czy też do np. transmisji strumieni audio-wideo. Aby dana sieć mogła stać się częścią Internetu, wystarczy, by nauczyła się obsługiwać protokół TCP/IP, czyli nauczyła się transportować pakiety IP. Wracając do naszej analogii: przykładowo sieć kolejowa stała się częścią sieci transportu standardowych kontenerów (metasieci), bo nauczyła się transportować swoimi istniejącymi kontenerami (wagony kolejowe) standardowe kontenery używane w obrocie międzynarodowym.

„Nałożenie" na dany typ sieci dodatkowej funkcji polegającej na nauczeniu się przez nią transportowania pakietów IP możliwe jest do uzyskania w praktycznie dowolnej istniejącej sieci.

Dzięki temu Internet mógł rozwijać się tak dynamicznie, obejmując skokowo coraz większy obszar i zwiększając, czasami w dramatyczny sposób, swoją dostępność. Przykładowo umożliwienie dostępu do Internetu z wykorzystaniem standardowych linii telefonicznych i modemów instalowanych w komputerach (dostęp wdzwaniany, ang. *dial-up*) zwiększyło skokowo dostępność Internetu w Polsce. Inny, podobny przykład to upowszechnienie usługi dostępu do Internetu poprzez sieci telefonii komórkowej.

2.2.4. Wspólny język Internetu – protokół internetowy TCP/IP

Definiując Internet, napisaliśmy, że sieci będące częścią Internetu potrafią „rozmawiać" ze sobą we „wspólnym języku". Tym językiem jest **Internet Protocol Suite, potocznie określany jako TCP/IP**, od nazw dwóch najważniejszych protokołów wchodzących w jego skład: Transmission Control Protocol i Internet Protocol. Odpowiadają one wspólnie za poprawne działanie transportowania danych w sieci Internet.

2.2.4.1. Pakiet IP

Podstawowym pojęciem niezbędnym do zrozumienia działania Internetu jest **pakiet IP**.

Używając analogii z branżą logistyczną, można powiedzieć że **pakiet IP jest czymś na wzór zestandaryzowanego kontenera używanego do transportu towarów**. Dzięki standaryzacji formy kontenera jego transportem zająć mogą się zupełnie odmienne od siebie sieci: może on być przewożony tirem (sieć drogowa), samolotem (sieć lotnicza), pociągiem (sieć kolejowa) czy statkiem (sieć rzeczna i morska). Wszystkie te sieci „nauczyły się" transportować kontenery, używając swoich standardowych metod transportu, pojazdów i zasad kontroli ruchu. W ten sposób powstała jedna, globalna sieć transportu kontenerów, niezależna od lokalnej specyfiki różnorodnych sieci transportujących same kontenery. Powstała metasieć.

Internet jest właśnie taką metasiecią transportującą pakiety IP z wykorzystaniem różnorodnych istniejących typów sieci znajdujących się „pod spodem".

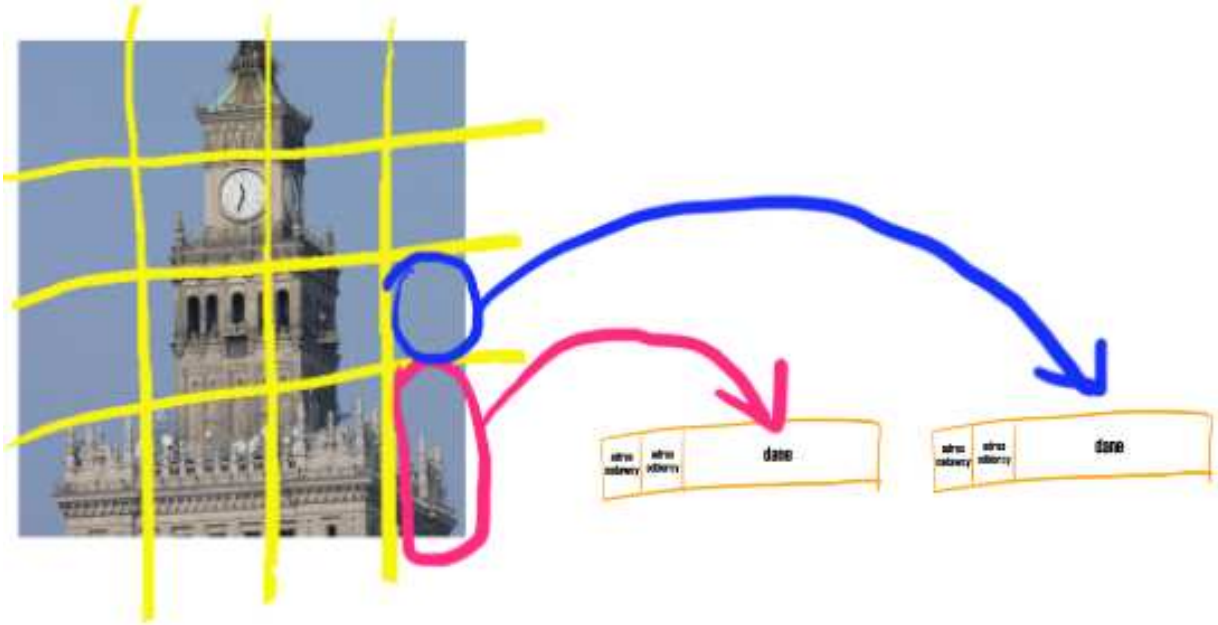
Pakiety IP zawiara w sobie przede wszystkim **adres nadawcy, adres odbiorcy** oraz **transportowane dane** (*payload*).

Rys. 2. Schemat pakietu IP

Internet, w uproszczeniu, to tysiące różnorodnych sieci, które przesyłają dane od nadawcy do odbiorcy przy pomocy zestandaryzowanych „kontenerów", którymi są właśnie pakiety IP.

2.2.4.2. Podział danych na części

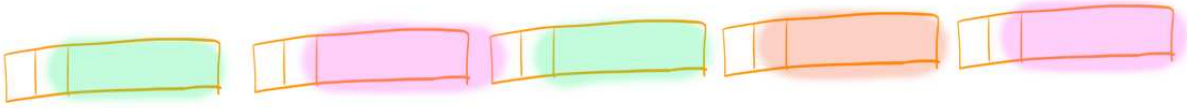
Specyfiką protokołu TCP/IP jest dzielenie transportowanych danych części. Dane, które chce przesłać nadawca, dzielone są na małe porcje, które po zapakowaniu w pakiety IP wysyłane są siecią Internet do odbiorcy. Odbiorca, po otrzymaniu pakietów IP, wydobywa z nich te małe porcje danych i łączy je w oryginalne dane, które nadał nadawca.



Rys. 3. Podział danych na porcje i pakowanie ich w kolejne pakiety

Podział treści na części jest efektywnym sposobem radzenia sobie z problemami w komunikacji. Jeśli np. w wyniku zakłóceń do odbiorcy nie dotarła cała wysłana treść, w celu poprawnego zakończenia transmisji wystarczy ponownie wysłać odbiorcy tylko te pakiety IP, które nie dotarły do niego w ramach pierwotnej transmisji.

Użycie małych pakietów jako elementu transportującego dane w Internecie pozwala także na **współdzielenie sieci przez wiele transmisji zachodzących w tym samym czasie**: pakiety należące do wielu transmisji „przeplatają się" na łączach i w urządzeniach sterujących ich ruchem, jako że intensywność pojedynczej transmisji danych jest zazwyczaj zmienna w czasie[3].



Rys. 4. Pakiety niosące różne podzielone dane (odmienne kolory) „przeplatają się" na jednym łączu

[3] Tzw. *bursty traffic* – transmisja o zmiennym natężeniu, która polega na tym, że od czasu do czasu następuje większy ruch transmisyjny, a pomiędzy tymi okresami transmisja jest na znacząco niższym lub zerowym poziomie. Przykładem może być przeglądanie stron WWW, kiedy najintensywniejsza transmisja zachodzi podczas pobierania treści strony WWW na komputer użytkownika, po czym następuje przerwa w transmisji (użytkownik zapoznaje się z treścią strony) aż do następnej transmisji zainicjowanej przez użytkownika poprzez zażądanie ściągnięcia następnej strony WWW.

2.2.4.3. Przełączanie pakietów – routery

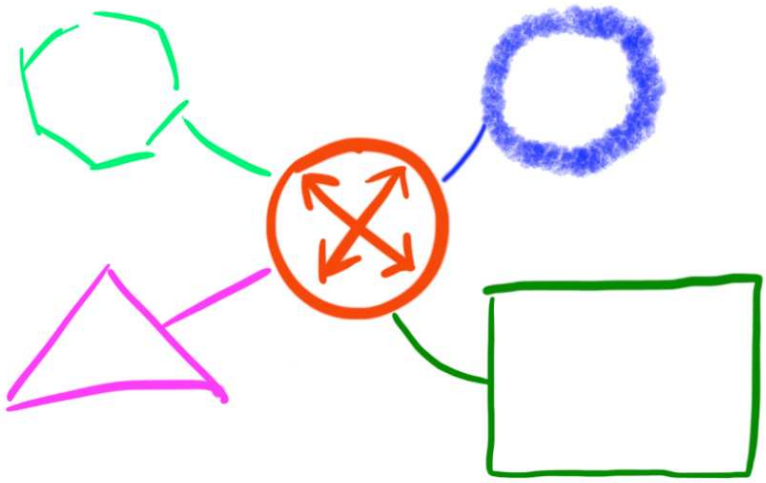
Kluczową rolę w transportowaniu pakietów IP w Internecie pełnią urządzenia nazywane **routerami**. **Routery** zapewniają podstawową usługę niezbędną dla funkcjonowania Internetu – **zajmują się przełączaniem pakietów IP**.

Przełączanie pakietów (routing, zwany też trasowaniem) to proces polegający na:

1) przyjęciu przez router pakietu IP nadchodzącego z dowolnej sieci, do której router jest podłączony,
2) analizie adresu odbiorcy pakietu IP i podjęciu decyzji o najlepszej dalszej drodze pakietu IP na podstawie swojej wiedzy o strukturze i połączeniach otaczających go sieci,
3) wysłaniu pakietu IP w dalszą podróż łączem, które zostało wybrane przez router.

Routery umożliwiają łączenie ze sobą całkowicie odmiennych technicznie sieci. Jak pisaliśmy poprzednio, każda sieć będąca częścią Internetu „nauczyła się" przenosić pakiety IP, umieszczając je wewnątrz swoich specyficznych elementów transportujących dane (wewnątrz swoich pakietów, ramek, komórek itp.). **Routery są „poliglotami" potrafiącymi rozmawiać z wieloma typami sieci w ich „językach"**. Dzięki temu zapewniają możliwość transportowania pakietów IP pomiędzy różnymi sieciami.

Powracając do analogii z kontenerami, **router przyrównać można do stacji przeładunkowej** „podłączonej" do np. paru systemów autostrad, linii kolejowych i posiadającej swój port lotniczy odwiedzany przez różne linie lotnicze. Stacja ta odbiera kontenery z tirów, pociągów i samolotów, odczytuje adresy odbiorców, a następnie na podstawie swojej wiedzy o geograficznej lokalizacji danego odbiorcy, wysyła każdy kontener w dalszą podróż odpowiednim środkiem transportu konkretnego przewoźnika. Stacja przeładowuje kontenery (pakiety IP) pomiędzy dowolnymi środkami transportu różnych przewoźników (sieciami).



Rys. 5. Router (czerwony okrąg w środku) podłączony do odmiennych typów sieci

Aby efektywnie przełączać pakiety IP (wysyłać je dalej we właściwym kierunku), **router wie, jak wygląda pewnie fragment Internetu dookoła niego**. Na podstawie tej wiedzy buduje tzw. tablicę routingu – zbiór informacji o osiągalności adresów IP poprzez łącza podłączone do routera wraz ze wskazaniem preferowanych łączy. Proces przełączania pakietów następuje właśnie na podstawie zapisów w tablicy routingu.

Jeśli router zajmuje się przełączaniem pakietów wewnątrz jednej sieci, potrzebuje orientować się tylko w jej topologii: wiedzieć, którym łączem ma wysłać pakiet IP przeznaczony dla adresata mającego dany adres IP. Jego tablica routingu jest niewielka.

Jeśli router zajmuje się przełączeniem pakietów między sieciami, otrzymuje on od nich informacje o osiągalności ich adresów IP oraz często adresów IP sieci znajdujących się za nimi itd. Tablica routingu takiego routera zazwyczaj jest znaczących rozmiarów.

Odporność na awarie połączeń. Połączenia pomiędzy sieciami mogą ulegać awariom. W takim przypadku router świadomy awarii danego połączenia zaczyna kierować pakiety (które uprzednio kierowałby tym łączem) inną, mniej preferowaną drogą, zgodnie z zapisami w swojej tablicy routingu. Ten przykład pokazuje sytuację, w której przed awarią połączenia cześć pakietów IP kierowana byłaby jedną trasą, a w trakcie awarii – drugą. Mimo awarii połączenia wszystkie pakiety dotarłyby do odbiorcy, choć różnymi drogami.

Router – gateway. Każde urządzenie końcowe w swojej konfiguracji umożliwiającej mu komunikację z siecią Internet posiadać musi adres IP routera, na który to adres posyłać ma wszystkie pakiety IP wysyłane przez to urządzenie. Ten **router stanowi bramę (*gateway*) do Internetu dla urządzenia końcowego**. Tak jak pisaliśmy, to routery posiadają wiedzę o topologii sieci i to one podejmują decyzję, którą drogą posłać dany pakiet IP w dalszą podróż do odbiorcy. Router będący gatewayem jest jedynym urządzeniem szkieletowym, przez które przechodzi cała komunikacja urządzenia końcowego z Internetem. Z tego powodu **router ten jest miejscem, w którym prowadzić można ewentualną analizę całości komunikacji urządzenia końcowego lub nadzór nad nią**.

Asymetria ruchu pakietów w Internecie. Jak napisaliśmy, każdy router podejmuje decyzję o posłaniu otrzymanego pakietu IP w dalszą drogę na podstawie tablicy routingu. Routery zajmujące się przełączaniem pakietów pomiędzy sieciami operatorów posiadają rozległe tablice routingu, w których dla danego docelowego adresu IP widnieje wiele możliwych dróg dotarcia. Mechanizm wyboru drogi, którą router pośle dalej pakiet, bazuje zazwyczaj na „długości" analizowanej drogi. Mechanizm ten może być zmieniany przez operatora sieci tak, by odpowiadał przyjętym wewnętrznym zasadom kierowania ruchem, tzw. polityce routingu (wynikającym np. z kosztów utrzymania łączy międzyoperatorskich).

Decyzję o każdym kolejnym fragmencie trasy, którą zostanie wysłany pakiet IP, podejmuje każdy kolejny router, do którego trafia tenże pakiet. Statystycznie, po drodze od nadawcy do odbiorcy znajduje się średnio kilka sieci, a każda z nich posiada swoją politykę routingu. **Powoduje to asymetrię ruchu w Internecie: trasa pokonywana przez pakiet z punktu A do B nie jest w większości przypadków taka sama, jak trasa pokonywana przez pakiet z punktu B do A.**

2.2.4.4. Brak wiedzy o przesyłanej treści

W kontekście routerów na sieć Internet można spojrzeć w ten sposób, że składa się ona z:

1) części transportowej – urządzeń przełączających pakiety (routerów),
2) urządzeń końcowych (*host*) komunikujących się ze sobą przy pomocy części transportowej (serwery, komputery użytkowników, telefon komórkowy podłączony do sieci).

Dla efektywnego działania Internetu znaczenie ma szybkie przełączania pakietów w części transportowej, co zostało zdefiniowane już przy projektowaniu zasad działania sieci (więcej o tym piszemy w punkcie *Zasada end-to-end*). **Dlatego też każdy router przełącza pakiety IP, biorąc**

pod uwagę wyłącznie adres odbiorcy pakietu. Routery nie analizują danych przenoszonych wewnątrz pakietów.

Bardzo istotne dla zrozumienie działania Internetu jest to, że Internet nie jest „świadomy" danych (treści), które są z jego wykorzystaniem transportowane, ani aplikacji, które z niego korzystają.

Jak napisaliśmy, routery zajmują się przełączaniem pakietów, a nie ich analizowaniem. Routery nie są świadome usług działających w oparciu o pakiety IP (takich jak np. poczta elektroniczna, telefonia internetowa, strony WWW) – nie jest to niezbędne dla zapewnienia ich poprawnego działania. Routery zajmują się jedynie przełączaniem pakietów, nie odróżniając, czy dany pakiet IP niesie w sobie np. zapytanie o stronę WWW czy fragment e-maila.

2.2.4.5. Adresy IP

Każde urządzenie podłączone do Internetu mieć przypisany do siebie **adres IP**[4], aby móc wysyłać i odbierać pakiety IP. Adres IP to unikatowy w skali Internetu adres zapisany jako ciąg liczb. Przykładowy adres IP to np. 174.121.79.98.

Każde urządzenie w Internecie może komunikować się z każdym innym urządzeniem, wysyłając pakiety IP oznaczone swoim adresem IP nadawcy i zaadresowane na adres IP odbiorcy.

Adresy IP są przypisywane zarówno urządzeniom końcowym użytkowników Internetu, urządzeniom dostawców usług (np. serwerom), jak i urządzeniom operatorów sieci (np. routerom). Każde urządzenie, które odbiera, nadaje lub przełącza (jak router) pakiety IP, musi mieć przyznany adres IP, aby można było „zaadresować" pakiety kierowane do tego urządzenia. Tak samo adres IP urządzenia wysyłającego pakiet IP służy identyfikacji go jako nadawcy.

Dla poprawności działania Internetu **konieczne jest zachowanie unikatowości przyznawanych urządzeniom adresów IP.** Dwa urządzenia w Internecie nie mogą mieć takiego samego adresu IP, ponieważ w takim przypadku nie wiadomo byłoby, do którego z tych urządzeń należałoby dostarczyć zaadresowany na taki adres IP pakiet IP.

Nad zachowaniem unikatowości przyznawanych adresów IP czuwają Regionalne Rejestry Internetowe. Przyznają one zakresy adresów IP Lokalnym Rejestrom Internetowym (LIR – często są nimi operatorzy sieci), a LIR przypisuje adresy IP swoim klientom: następnym, mniejszym operatorom sieci lub bezpośrednio użytkownikom końcowym.

Dla potrzeb naszego opracowania ważne jest **rozróżnienie publicznych i prywatnych adresów IP**:

1) **publiczne adresy IP** to unikatowe adresy IP, z których i do których możliwe jest słanie pakietów IP w globalnym Internecie;
2) **prywatne adresy IP** to określone, specjalne zakresy adresów IP, z których i do których nie jest możliwe słanie pakietów IP w globalnym Internecie, a które mogą komunikować się lokalnie między sobą.

Prywatne adresy IP powstały pierwotnie dla celów używania ich w sieciach niepołączonych z Internetem. Obecnie adresy te wykorzystywane są często przez dostawców Internetu oraz w domowych sieciach użytkowników prywatnych – najczęściej z użyciem mechanizmu NAT

[4] Adresy IP przydzielane są do urządzeń na stałe lub na dany okres. Przy odnowieniu przydziału okresowego adres IP może pozostać taki sam lub też zmienić się.

(*Network Address Translation* – mechanizm translacji adresów sieciowych). W obu wymienionych przypadkach adresy prywatne przypisywane są urządzeniom wewnątrz sieci „prywatnej" (sieci operatora lub sieci użytkownika domowego), natomiast w procesie ich komunikacji z globalnym Internetem pośredniczy urządzenie posiadające publiczny adres IP (proxy serwer, router).

Z punktu widzenia Internetu prywatne adresy IP nie są widoczne – widziany jest tylko publiczny adres IP routera / serwera proxy, za którym znajdują się („są schowane") urządzenia z przydzielonymi prywatnymi adresami IP.

Każdy pakiet wysłany z urządzenia używającego prywatnego adresu IP i zaadresowany na publiczny adres IP, po dostarczeniu go do routera lub serwera proxy, wysyłany jest ponownie przez router (lub serwer proxy) z użyciem jego publicznego adresu IP jako adresu nadawcy. Następnie, po otrzymaniu odpowiedzi na tak wysłany pakiet, router (lub serwer proxy) przekazuje tę odpowiedź na prywatny adres IP urządzenia, które zainicjowało tę komunikację. Można powiedzieć, że router (serwer proxy) anonimizuje tożsamość urządzeń używających prywatnych adresów IP.

Ważną **konsekwencją używania prywatnych adresów IP jest znacząco utrudniona identyfikacja urządzenia** korzystającego z adresacji prywatnej przez kogoś, kto chciałby sprawdzić, z którym dokładnie urządzeniem wymieniał dane (np. które urządzenie przesłało mu szkodliwą treść lub przeprowadziło atak). O tym, które urządzenie otrzymało który prywatny adres IP, wiedzieć może administrator, który je przyznał: operator używający adresacji prywatnej we własnej sieci albo użytkownik końcowy używający takich adresów np. w swojej sieci WiFi, do której podłączył kilka urządzeń. Wiedza ta jest użyteczna tylko wtedy, kiedy fakt przyznania adresu IP jest zapisywany i przechowywany. Jest to wiedza umożliwiająca odnalezienie urządzenia, które dokonało jakiegoś działania w sieci, pod warunkiem że router (serwer proxy), za pomocą którego to urządzenie komunikowało się z Internetem, zapisywał dodatkowe dane o transmisji pakietów od/do tego urządzenia, w której pośredniczył.

Dodatkowo, dane urządzenie może korzystać ze **stałego lub zmiennego adresu IP**. W przypadku używania prywatnych adresów IP przyznawane są one często na żądanie urządzenia, kolejno z dostępnej w danej chwili puli nieprzyznanych adresów (np. w momencie włączenia komputera użytkownika, w momencie podłączania się do sieci WiFi). **To samo urządzenie użytkownika może otrzymać inny prywatny adres IP za każdym „włączeniem" się do sieci**.

Najczęściej stosowanym mechanizmem przyznawania adresów IP urządzeniom końcowym jest DHCP (Dynamic Host Configuration Protocol), w którym to mechanizmie adres IP przyznawany jest na określony okres (np. 24 godziny), a następnie urządzenie występuje o ponowne przyznanie adresu IP. DHCP przyznaje ponownie ten sam lub inny adres IP, w zależności od przyjętej przez administratora zasady. Tak więc **nawet długie pozostawanie urządzenia w stanie podłączenia do Internetu nie gwarantuje, że będzie ono dostępne w tym okresie pod tym samym adresem IP**.

Mechanizm DHCP używany jest w sieciach przyznających urządzeniom zarówno prywatne, jak i publiczne adresy IP.

Reasumując, aby zidentyfikować urządzenie podłączone w danym okresie do Internetu, niezbędne jest:

1) jeśli używa stałego, publicznego adresu IP – znanie tego adresu IP;
2) jeśli używa zmiennego, publicznego adresu IP – podanie właściwemu operatorowi adresu IP i okresu, o który pytamy (przy założeniu, że operator prowadzi czasową ewidencję przyznawanych adresów IP);

3) jeśli używa adresu prywatnego – podanie właściwemu operatorowi publicznego adresu IP routera (serwera proxy), z którego / do którego wysyłane była analizowane przez nas pakiety IP wraz z czasem i dokładnym opisem analizowanego przez nas wydarzenia [przy założeniu, że operator prowadzi czasową ewidencję przyznanych adresów IP oraz zapisuje szczegóły połączeń i transmisji nawiązywanych przez urządzenia korzystające z danego routera (serwera proxy)].

2.2.4.6. Bezpołączeniowość

W odróżnieniu od klasycznych sieci, takich jak np. sieć telefoniczna, komunikacja w Internecie jest komunikacją bezpołączeniową. Oznacza to, że **przed rozpoczęciem transmisji danych z punktu A do B w Internecie nie zostaje zestawiony dedykowany, wydzielony kanał transmisyjny między tymi punktami**.

W Internecie komunikacja odbywa się bez uprzedniego zestawiania kanału transmisyjnego. Urządzenia podłączone do sieci wysyłają pakiety IP do najbliższego znanego sobie routera, wiedząc, że pakiety te będą przełączane przez kolejne routery, aż dotrą do odbiorcy. Brak zestawienia kanału skutkuje też brakiem blokowania zasobów sieci na czas trwania transmisji.

W sieciach telefonicznych dla każdej transmisji (rozmowy) zestawiany jest dedykowany kanał, w którym odbędzie się rozmowa. Po jej zakończeniu rozmowy kanał ten zostaje rozłączony. Dokładny zestawienia i rozłączenia kanału oraz „adresy" (numery telefoniczne) abonentów z niego korzystających są zapisywane w bazach operatorów, którzy brali udział w zestawieniu kanału. Dzięki temu możliwe są rozliczenia między operatorami oraz obciążenie ceną rozmowy abonenta/-ów. Sieci telefoniczne dla celów przekazywania sobie informacji (tzw. sygnalizacji) o potrzebie zestawiania kanałów utrzymują oddzielne, dedykowane połączenia międzysieciowe. Dzięki nim możliwy jest dokładny billing.

W Internecie nie możliwe jest użycie tak rozumianego billingu, ponieważ komunikacja pomiędzy urządzeniami jest komunikacją bezpołączeniową oraz nie istnieją łącza sygnalizacyjne, które przekazywałyby dane o „rozmowach" pomiędzy urządzeniami z danymi adresami IP.

2.2.4.7. Zasada *best effort*

W związku z przyjętymi rozwiązaniami przy projektowaniu podstawowych protokołów Internet nie gwarantuje żadnych parametrów związanych z dostarczaniem pakietów IP. Sieć nie gwarantuje nawet tego, że pakiet zostanie dostarczony do odbiorcy. Nie wiadomo też, w jakim czasie pakiet dotrze od nadawcy do odbiorcy (może być przesłany bardzo szybko lub bardzo wolno, w zależności od chwilowego obciążenia fragmentów sieci). Trasa przesyłania pakietów także nie jest z góry określona i może się zmieniać w czasie, a nawet podczas transmisji.

Sieć „podejmuje próbę" dostarczenia pakietu w czasie i taką drogą, jaka będzie możliwa w danym momencie. Ta cecha sieci to właśnie zasada *best effort* (dosłownie: 'dołożenia starań'; sieć zrobi, co w jej mocy, aby dostarczyć pakiet, lecz nie gwarantuje niczego).

2.2.4.8. Zasada *end-to-end*: głupia sieć i mądre urządzenia końcowe

Zasada *end-to-end* była jedną z podstawowych reguł przyjętych podczas projektowania Internetu. Jest ona odpowiedzią na pytanie, gdzie najlepiej implementować różne funkcje w systemach komunikacyjnych. **Zasada *end-to-end* głosi, że najlepiej działają te systemy, które jak najwięcej funkcji oddają pod kontrolę urządzeń końcowych, a nie implementują je w samym systemie**. Stąd właśnie pochodzi potoczne określenie „głupiej sieci" (*dumb network*) i inteligentnych urządzeń końcowych.

Przykładem zastosowania zasady *end-to-end* jest architektura protokołów IP i TCP:

1. Protokół IP jest prostym, „głupim" protokołem, którego zadaniem jest umożliwienie routerom przełączenia pakietów na ich drodze pomiędzy od nadawcy do odbiorcy; nie oferuje on żadnej bardziej zaawansowanej funkcjonalności.
2. Natomiast protokół TCP jest inteligentnym protokołem transportowym zapewniającym wykrywanie błędów w transmisji pakietów IP, retransmisję zagubionych pakietów, kontrolę szybkości wysyłania pakietów itp.
3. Prosty i przez to efektywny protokół IP zaimplementowany jest w sieci, a inteligentny i skomplikowany protokół TCP implementowany jest w urządzeniach końcowych.

Zasada *end-to-end* jest gwarancją prostoty i efektywnego działania Internetu. **Umożliwia ona ciągłą innowację poprzez brak usztywnienia samej sieci zbyt specyficznymi funkcjonalnościami realizowanymi w jej wnętrzu (sieć pozostaje prosta)**. Zastosowanie tej zasady umożliwia powstawanie nowych usług bez konieczności zmiany architektury, oprogramowania i konfiguracji Internetu. Dla wprowadzenia usług telefonii internetowej, czy też jeszcze bardziej wymagających transmisji wideo w Internecie, nie było konieczne przebudowanie czy rozszerzenie funkcjonalności samej sieci. **Nowe aplikacje i nowe zaawansowane funkcje realizowane są nie przez sieć, lecz przez urządzenia końcowe (*end-hosts*), podczas gdy sam Internet pozostaje tą samą, prostą siecią.**

„Inteligencja" znajduje się na brzegu sieci, podczas gdy ona sama ma po prostu transportować pakiety. Dzięki zasadzie *end-to-end* brzeg sieci (urządzenia użytkowników, serwery dostawców usług, same usługi i ich zakres) może swobodnie ewoluować, nie martwiąc się o zachowanie kompatybilności z siecią, z której korzysta.

Dzięki przyjętej filozofii budowy Internetu stworzono prostą, wydajną sieć, która obsługuje obecnie olbrzymie wielkości transmisji w skali globalnej. Udaje jej się to właśnie dlatego, że jest prosta: nie zajmuje się analizą i priorytetyzowaniem trafiających do niej pakietów. Zajmuje się po prostu ich przesyłaniem. Każdy może wykorzystać tę uniwersalną sieć transportową pakietów IP do przesyłania dowolnych treści służących dowolnym usługom.

2.2.5. System DNS

Jak już pisaliśmy, pewne elementy Internetu są zarządzanie centralnie. Jednym z takich elementów jest system DNS.

Przypomnijmy: dane w Internecie przesyłane są za pomocą pakietów IP pomiędzy urządzeniami posiadającymi adresy IP. Aby nawiązać kontakt z danym urządzeniem, konieczna jest znajomość jego adresu IP. O ile na początku funkcjonowania Internetu adresów IP nie było zbyt wiele i użytkownicy mogli je zapamiętać, a aplikacje mogły posiadać ich listę, o tyle obecnie zadanie to jest już niewykonalne.

W celu ułatwienia komunikacji w Internecie stworzono system **DNS (Domain Name System)**, który **przy pomocy łatwych do zapamiętania nazw (domen) uwalnia użytkowników i aplikacje od konieczności znajomości adresów IP**. Dzięki systemowi DNS, chcąc wejść na stronę WWW np. portalu Onet, wystarczy, że wpiszemy w oknie przeglądarki adres „onet.pl". Bez systemu DNS musielibyśmy wpisać tam adres IP serwera tego portalu.

DNS jest systemem hierarchicznym, zarządzanym centralnie. Na poziomie każdego kraju zadaniem zarządzania domeną krajową zajmuje się rejestr narodowy (w Polsce jest to NASK). Rejestry te stoją na straży poprawności działania systemu DNS, w szczególności dbają o to, by zmiany adresów IP przypisanych do domen dokonywane były wyłącznie przez uprawnione podmioty/osoby.

Jak działa DNS:

1. Zapytanie o adres IP wysłane np. przez przeglądarkę internetową, w której użytkownik wpisał adres www.costam.pl, przekazywane jest pomiędzy serwerami DNS tak długo, aż dotrze do serwera, który zna na nie odpowiedź.
2. Ten serwer zwraca informację o adresie IP przypisanym do domeny www.costam.pl do przeglądarki.
3. Przeglądarka może dzięki temu wysłać zapytanie o stronę WWW na podany jej adres IP.

Podkreślić tu należy, że system DNS jest dodatkowym elementem sieci Internet, ułatwiającym użytkownikom korzystanie z niej. Jego istnienie nie jest konieczne do korzystania z Internetu, ale w związku z popularnością usług WWW i używania domen **DNS stał się jednym z krytycznych systemów zapewniających poprawną pracę Internetu.**

Większość urządzeń podłączonych do Internetu, takich jak np. komputery użytkowników końcowych, oprócz przyznanego im adresu IP **otrzymuje także adres IP jednego lub paru serwerów DNS**, z których powinny korzystać. Przykładowo operatorzy ISP podają swoim klientom adresy utrzymywanych przez siebie serwerów DNS. **Pamiętać tu należy, że użytkownik zazwyczaj może zmienić w swoim urządzeniu adresy IP serwerów DNS**, co – jak później opiszemy – ma znaczący wpływ na skuteczność niektórych proponowanych narzędzi regulacji Internetu.

3. Praktyczne aspekty działania Internetu – w skrócie

1. Internet jest siecią złożoną z tysięcy oddzielnych sieci połączonych za sobą.
2. Wspólne dla nich wszystkich jest używania jednego języka – protokołu internetowego (TCP/IP).
3. Internet jest metasiecią – służy do przesyłania pakietów IP wykorzystując do tego celu różnorodne sieci znajdujące się „pod" Internetem.
4. Sieć jest „głupia" – zajmuje się przesyłaniem pakietów, dzięki czemu robi to szybko i efektywnie.
5. „Mądrość" i zaawansowane funkcje zlokalizowane są w urządzeniach końcowych (komputery, serwery itp.) i aplikacjach na nich uruchamianych.

Dzięki schematowi „głupia sieć – mądre urządzenie końcowe", czyli zasadzie *end-to-end*, Internet jest otwartą platformą dla innowacji, nowych aplikacji i zastosowań.

Użytkownik Internetu

1. Za każdym razem, kiedy użytkownik Internetu podłącza się do sieci, jego urządzenie otrzymuje adres IP:
   - publiczny lub prywatny,

- stały lub zmienny.

2. Urządzenie użytkownika otrzymuje (lub ma wpisane na stałe) adresy serwerów DNS, do których kierowane będą zapytania o adresy IP przypisane do domen.
3. W trakcie korzystania z Internetu:
   - wszystko, co wysyła i odbiera użytkownik, transportowane jest w pakietach IP;
   - sieć, która transportuje pakiety, jest „głupia" (prosta i nieświadoma zawartości pakietów IP);
   - cała „mądrość" usług, pilnowanie poprawności dostarczania pakietów itp., „dzieje się" w samych aplikacjach uruchomionych na urządzeniach podłączonych do sieci.

4. Obecne cechy Internetu wpływające na narzędzia regulacyjne

4.1. Dominacja małej liczby aplikacji

Internet jest siecią transportującą dowolne dane. Może więc być wykorzystany przez różnorodne typy aplikacji oferujące dowolne usługi. Pomimo tej otwartości od początków historii Internetu daje się zaobserwować dominację stosunkowo niewielkiej liczby aplikacji, które używane są przez internautów. Obecnie najpopularniejsze usługi w Internecie to:

- korzystanie ze stron WWW za pomocą przeglądarek internetowych,
- poczta elektroniczna (e-mail),
- komunikatory i telefonia internetowa (VoIP),
- wymiana plików *peer-to-peer*.

4.2. Matrioszki, czyli usługi w usługach

W ostatnich latach obserwowalny w Internecie jest proces:

- zwiększania liczby świadczonych usług przez pojedyncze aplikacje,
- udostępniania danego typu usługi przez inna usługę (usługa w usłudze).

Pierwszy z tych procesów to sytuacja, w **której aplikacje dedykowane początkowo do jednego celu zaczynają oferować funkcje realizowane przez inne aplikacje**. Przykładem są tu komunikatory tekstowe (*instant messangers*), takie jak Gadu-Gadu, Tlen, które zaczęły po pewnym czasie umożliwiać także komunikację głosową (VoIP), do której istniały oddzielne aplikacje.

Drugi proces związany jest głównie z rozwojem przeglądarek internetowych i obsługiwanych przez nie standardów (takich jak HTML, Javascript, Flash). **Coraz więcej usług zaczyna być świadczonych poprzez strony WWW**. Dobrym przykładem tego procesu jest dostęp do poczty elektronicznej poprzez przeglądarkę internetową. Jest to nadal usługa WWW – użytkownikowi przesyłana jest strona internetowa, lecz jednocześnie jest to usługa dostępu do konta poczty elektronicznej, którą to usługę standardowo świadczyły wyspecjalizowane programy. Co szczególnie ważne, korzystały one innych protokołów komunikacyjnych niż protokół używany przez przeglądarki. W klasycznym scenariuszu aplikacja typu czytnik e-mail uruchomiona na urządzeniu użytkownika komunikowała się z serwerem poczty np. w protokole POP i SMTP w celu odbioru i wysyłania e-maili. W podanym powyżej przykładzie tego typu komunikacja nie zachodzi. Usługa dostępu do skrzynki e-mail świadczona jest za pośrednictwem usługi WWW – **jest to usługa w usłudze**. Ten sam proces zaczyna dotyczyć komunikatorów, grup newsowych, gier. Z punktu widzenia implementacji narzędzi regulacyjnych stanowi to dodatkową trudność w identyfikacji i potencjalnej kontroli danego typu usługi.

4.3. Szyfrowanie

Coraz więcej komunikacji w Internecie odbywa się z użyciem szyfrowania. Oznacza to, że zawartość pakietów IP (tzw. *payload*) jest szyfrowana przez nadawcę i odszyfrowana przez odbiorcę.

Najczęściej szyfrowanie występuje obecnie w sytuacjach:

- połączenia pomiędzy przeglądarką internetową uruchomioną na urządzeniu końcowym a serwerem dostawcy usługi (np. dostęp do systemów bankowości internetowej),
- połączenia pomiędzy programem pocztowym a serwerem poczty elektronicznej,
- dostępu do serwerów proxy,
- zdalnego dostępu do zasobów firmowych przez pracowników.

Stosowany w tych przypadkach algorytmy szyfrowania są wystarczająco silne, by zapewnić wymagany (np. przez banki) poziom bezpieczeństwa. Moc obliczeniowa urządzeń końcowych jest obecnie tak duża, że mogą one efektywnie stosować silne metody szyfrowania. Dla omawianych narzędzi regulacji oznacza to praktyczną niemożność poznania treści zaszyfrowanej transmisji w Internecie.

Istnieje także inna, mniej obecnie rozpowszechnioną sfera szyfrowania: szyfrowanie poczty elektronicznej. Na rynku dostępne są darmowe, silne algorytmy szyfrujące przeznaczone dla poczty elektronicznej. Najpopularniejszym takim systemem jest PGP (*Pretty Good Privacy*), oparty na mechanizmie publicznych i prywatnych kluczy szyfrujących. Niektóre organizacje i użytkownicy wyczuleni na punkcie prywatności korespondencji już obecnie używają tego typu szyfrowania, choć przyznać należy, że są na razie w mniejszości. Jednak w przypadku zwiększenia zainteresowania tą metodą (wywołanego np. nagłośnieniem informacji o „podsłuchiwaniu" korespondencji e-mail przez służby do tego uprawnione) użytkownicy w szybkim tempie zainstalują na swoich urządzeniach końcowych oprogramowanie szyfrujące.

4.4. Tunelowanie

Tunelowanie to sposób przesyłania pakietów między urządzeniami z wykorzystaniem Internetu, w którym danymi (*payload*) pakietów IP przesyłanych w Internecie są inne pakiety. Internet wykorzystywany jest w tym przypadku nie do transportu samych danych generowanych przez jakąś aplikację, lecz do transportu całych pakietów.

Tunelowanie stosowane jest np. do bezpiecznego (szyfrowanego) połączenia dwóch sieci przedsiębiorstwa poprzez publiczną sieć Internet. W obu sieciach przedsiębiorstwa konfigurowane są urządzenia (najczęściej routery) stanowiące zakończenia tunelu. Każdy pakiet z sieci A przeznaczany dla sieci B przesyłany jest tunelem poprzez zapakowanie go w pakiet/-y IP po jednej stronie tunelu i wypakowanie go z pakietu/-ów IP po drugiej stronie.

Tunelowanie używane jest także np. przy korzystaniu z serwerów proxy, które stają się w takim przypadku realnym gatewayem dla klienta, który z nich korzysta. Router będący oryginalnym gatewayem dla danego urządzenia końcowego (router dostawcy Internetu) przekazuje pakiety pomiędzy urządzeniem końcowym a serwerem Proxy, nie będąc świadomym, że pakiety te niosą w sobie właściwe pakiety IP służące komunikacji urządzenia końcowego z Internetem. Ta metoda komunikacji używana jest szczególnie często w przypadkach, gdy dostawca Internetu stosuje narzędzia ograniczające dostęp do wybranych zasobów i usług w Internecie.