

Sieć z giełdowym systemem zawierania kontraktów na usługi przesyłowe – propozycja architektury

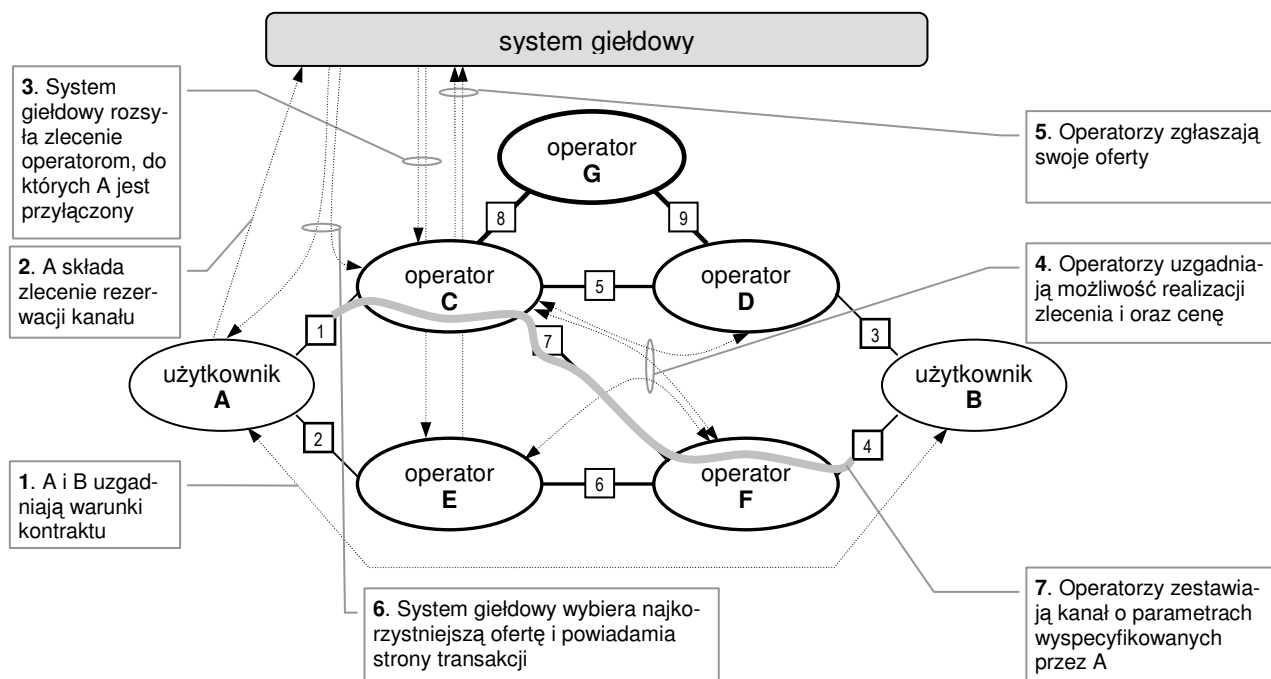
W artykule przedstawiono propozycję architektury systemu realizującego kontrakty przesyłowe i inne usługi o gwarantowanej jakości w sieci globalnej. Jako, że istniejące koncepcje rozwiązań technologicznych kwestii zapewnienia jakości można uznać za dojrzałe, autorzy skupili się na opracowaniu schematu negocjacji kontraktów w aspekcie ekonomicznym. Proponowane rozwiązanie zachowuje obecną, hierarchiczną strukturę relacji ekonomicznych w internecie, precyzując scenariusze negocjacji cen towarzyszące zawieraniu kontraktów dynamicznych (tj. na okres minut lub dłuższy). Operatorzy i użytkownicy, korzystając z mechanizmu rozproszonej giełdy usług, mają możliwość wyboru dostawcy usługi na najkorzystniejszych dla siebie warunkach. Obecna hierarchiczna struktura rynku oraz jego naturalny podział geograficzny zostały wykorzystane w proponowanej architekturze dla uzyskania skalowalności oraz możliwości kształtowania procesu negocjacji przez regulatora tak, by uzyskać korzystny społecznie wynik negocjacji. Opłacalność wdrożenia takiej architektury została wykazana dla prostych modeli ekonomicznych działania operatorów.

1. Wprowadzenie

Wzrastająca potrzeba wykorzystania internetu do realizacji usług opartych o transmisję z gwarancją jakości (*quality of service* – QoS) pozostaje niekwestionowanym faktem. Na przeszkodzie w realizacji świadczenia takich usług przesyłowych w skali globalnej stoją jednakże problemy natury i technologicznej, i ekonomicznej. Wyzwania wobec technologii świadczenia takich usług dotyczą głównie efektywnego przekazywania olbrzymiej liczby zróżnicowanych strumieni danych w węzłach sieci oraz zapewnienia sprawnej sygnalizacji dla całego wachlarza wykorzystywanych mediów i protokołów komunikacyjnych w niższych warstwach. Natomiast problemy ekonomiczne obejmują przede wszystkim podział zysków ze świadczenia usług pomiędzy zaangażowane podmioty w sposób umożliwiający inwestycje w adekwatne technologie transmisji, inżynierii ruchu i zapewnienia jakości – bez względu na to, czy miałyby być one wyrafinowane, czy prymitywne.

W dużym uproszczeniu można powiedzieć, że sieć globalna nowej generacji (*next-generation network* – NGN) będzie analogią do istniejącej sieci telefonicznej (*public switched telephone network* – PSTN): siecią z gwarancjami usług oraz z dojrzałymi regulacjami prawnymi zapobiegającymi patologiom. Jednakże, w odróżnieniu od PSTN, sama definicja usług w NGN powinna pozostawać otwarta, by stymulować innowacyjność. Stawia to dużo poważniejsze wyzwania technologiczne dla samej sieci. Z drugiej strony legislacja dotycząca przekazywania ruchu IP wydaje się nieadekwatna do potrzeb. Wskutek tego – i wbrew jasno sformułowanym strategiom regulatorów rynku – zamiast otwartej struktury sieci obserwuje się powszechnie rozwiązania zamknięte (*walled garden*), czyli wertykalną integrację usług, spotykaną u operatorów o dużej liczbie abonentów. Stan taki wynika naturalnie z układu sił i stanowisk regulatorów i operatorów: ci ostatni nie są skłonni mocno inwestować w infrastrukturę wobec perspektywy przymusowego udostępnienia jej operatorom alternatywnym, bez partycypowania w zyskach z usług dodanych sieci. Z kolei regulatorzy nie pracują nad bardziej zaawansowanym mechanizmem podziału zysków, który jednocześnie stymulowałby rozwój sieci otwartych, a pośrednio, nowych usług i całego rynku.

Wierzymy, że ścieżka rozwoju sieci nowej generacji powinna być wyznaczona przez dobrze skonstruowane mechanizmy wolnorynkowe. Podobnie, jak w szeroko rozumianym biznesie zrównoważony rozwój jest warunkowany ekonomiczną motywacją do działania dla firm wszystkich rozmiarów, również rozwój NGN potrzebuje uczestnictwa obecnych, ale też – obowiązkowo – nowych podmiotów. Taka gra rynkowa ma sumę dodatnią, gdyż daje możliwość tworzenia nowych usług. Podstawowym mechanizmem współdziałania podmiotów powinna stać się giełda tych usług. W dalszej części artykułu dokonuje się konkretyzacji rozważań dla usługi transmisji danych z gwarancjami jakości w globalnej w sieci IP.



Rys. 1. Model sieci i podstawowy scenariusz ustanawiania kanału transmisyjnego

Rys. 1 przedstawia sposób modelowania sieci IP oraz zasadnicze elementy proponowanej architektury. Elipsy obrazują fragmenty sieci zarządzane przez ten sam podmiot – w szczególności przynależące do operatora bądź klienta. Sieci te są połączone łączami (odcinki), przy czym zawsze można wyróżnić punkt styku – miejsce, gdzie infrastruktura obu sieci jest fizycznie połączona. Zróżnicowana grubość linii na rysunku obrazuje wielkość uczestników rynku i przepustowość łączy w punktach styku. W szczególności, punkt styku może być ulokowany w punkcie wymiany ruchu (*internet exchange point* – IXP), gdzie stykają się więcej niż dwie sieci, często należące do operatorów różnej wielkości. Proponowany sposób realizacji połączeń QoS polega na zakupie na giełdzie odpowiedniego logicznego kanału przesyłowego pomiędzy właściwymi punktami styku.

Zgrubny scenariusz korzystania z rynku rezerwacji kanałów jest następujący:

1. Zestawienie kanału ma na celu umożliwienie komunikacji pomiędzy dwoma użytkownikami (A i B). Użytkownicy ci, w dowolny, ustalony sposób, uzgadniają:
 - i. *Parametry techniczne kontraktu:*
 - przepustowość kanału i maksymalne opóźnienie,
 - czas, w którym kanał będzie potrzebny,
 - listę alternatywnych punktów styku dla obu krańców kanału (tutaj: $\{1,2\} \times \{3,4\}$);
 - ii. *Parametry ekonomiczne kontraktu:*
 - maksymalną cenę za ustanowienie kanału i za jednostkę taryfikacji (czas trwania, ilość przesłanych danych, faktyczne wykorzystanie wg 95 percentyla itp.),
 - sposób podziału kosztów używania kanału pomiędzy podmioty,
 - który z podmiotów stanie się formalnym właścicielem kanału.
2. Uzgodniony w p. 1 użytkownik (A) składa do systemu giełdowego zlecenie rezerwacji kanału. Kontrakty zawarte w systemie giełdowym są wiążące – użytkownicy są zobowiązani prawem do uiszczenia opłat, a operatorzy – do realizacji zakontraktowanej usługi.
3. Zlecenie zostaje udostępnione operatorom przyłączonym do wyspecyfikowanych punktów styku z użytkownikiem-oferentem (C i E).
4. Powiadomieni operatorzy, w drodze negocjacji dwustronnych operatorami niezbędnymi do zestawienia połączenia (D i F), uzgadniają możliwości technologiczne i kwestie ekonomiczne; np. operator C może rozważyć warianty zestawienia kanału poprzez operatora D lub F.
5. Operatorzy składają do systemu giełdowego oferty zestawienia kanału.
6. System giełdowy wybiera najkorzystniejszą ofertę i doprowadza do zawarcia kontraktu oraz powiadamia kontrahentów (A i C).
7. Wszystkie zaangażowane strony dokonują odpowiedniej rekonfiguracji urządzeń sieciowych we własnej sieci i punktach styku tak, by zapewnić realizację kontraktu.

Proponowana architektura systemu charakteryzuje się asymetrią: początkowy wybór użytkownika-oferenta determinuje, którzy operatorzy będą współdziałać z systemem giełdowym. Mogą oni, niestety, wykorzystywać tę uprzywilejowaną pozycję podczas dwustronnych negocjacji z ich partnerami kwestii ekonomicznych kontraktu. Jednakże w sytuacji pełnej konkurencji na rynku żaden z nich (por. C i E na Rys. 1) miał takiej motywacji. Ponadto, proponowana asymetria ewidentnie poprawia skalowalność systemu, w przeciwieństwie do innych, konkurencyjnych rozwiązań [3].

Uargumentowanie zasadności i wyłożenie w szczegółach idei powszechnej giełdy usług przesyłowych zostanie przedstawione w dalszej części artykułu. Rozdział 2 prezentuje aktualny stan technologii i pokrewne, funkcjonujące rozwiązania komercyjne, a także stan badań nad NGN. Rozdział 3 przedstawia, jak naturalna hierarchia internetu przekłada się na hierarchiczność działania systemu giełdowego. Rozdział 4 przedstawia szczegóły techniczne systemu, w szczególności możliwość przetwarzania żądań w sposób zdecentralizowany. W rozdziale 5 zawarte są wstępne kalkulacje dowodzące opłacalności korzystania z systemu rezerwacji jako takiego, niezależnie od miejsca użytkownika w hierarchii. Możliwości zastosowania i rozwoju systemu oraz jego potencjalny wpływ na oblicze sieci zostały przedstawione w rozdziale 6.

2. Stosowane i opracowywane mechanizmy realizacji rynku usług przesyłowych z QoS

Jako, że proponowana architektura usług wymaga odpowiednich technologii zapewnienia jakości oraz mechanizmów organizujących wymianę towarową, warto przyjrzeć się obecnym dokonaniom i stanowi badań w obu dziedzinach.

Ugruntowaną technologią służącą zapewnianiu jakości transmisji w sieci jest ATM. Podobną funkcjonalność można również osiągnąć stosując łącznie odpowiednie rozwiązania w warstwie sieciowej: wymiarowanie, DiffServ, dopuszczanie połączeń. Można uznać, że oba podejścia są dojrzałe i – mimo swoich wad – nadają się w praktyce do realizacji kontraktów na transmisję z gwarancjami jakości. W rzeczywistości bardzo rozpowszechniona jest realizacja usług z gwarancjami w obrębie jednej domeny administracyjnej przy wykorzystaniu MPLS i DiffServ jednocześnie.

Kooperacja operatorów w celu wspólnego świadczenia usług z QoS jest od strony technicznej względnie łatwa, jeśli do różnicowania jakości używają tego samego rodzaju technologii (np. DiffServ). Natomiast zadanie zapewnienia jakości w sytuacji, gdy poszczególni operatorzy korzystają z różnych technologii, jest trudniejsze. Poświęcono mu kilka badawczych projektów europejskich (EuQoS, DAIDALOS, MESCAL, NETQoS); we wszystkich starano się wypracować technologię realizacji usług wymagających współpracy operatorów na całej ścieżce transmisji danych – najlepiej, wykorzystującą istniejące protokoły komunikacji pomiędzy systemami autonomicznymi (np. BGP). Projekty te nie skończyły się jednak upowszechnieniem opracowanych w nich technologii. Projekty uruchomione obecnie, w 7. Programie Ramowym (4WARD, AUTOI, CARMEN), skupiają się na zastosowaniu tzw. sieci nakładkowych (*overlay network*), czyli wykorzystujących podstawową infrastrukturę internetu do tworzenia własnej, topologii połączeń logicznych. Pozostaje jednak fundamentalny problem skutecznej współpracy sieci tworzących internet, zarówno dla realizacji sieci nakładkowych, jak i skutecznego zapewnienia jakości dla transmisji realizowanych w sposób bezpośredni.

Taka współpraca w skali globalnej nie zachodzi ze względów ekonomicznych, a nie technologicznych: podmioty na rynku nie mają czytelnej motywacji do współpracy. Jako kilka kontrprzykładów dla idei globalnej współpracy operatorów należy wymienić powszechne zjawiska: dyskryminacji w umowach peeringowych, wymuszania tras przesyłu przez droższych operatorów, czy tzw. *hot potato routing*.

Skuteczne wypromowanie środowiska i mechanizmów rynkowej rywalizacji dostawców o klientów poprzez zautomatyzowany system transakcyjny (giełdę usług) jest zagadnieniem delikatnym, dużo trudniejszym. Każda ze stron musi w uczestniczeniu w systemie znajdować swoje racje ekonomiczne. Przystąpienie do mechanizmu zautomatyzowanego wymaga zupełnej zmiany nastawienia uczestników: porównywanie ofert sprovedzone będzie do jednego kryterium, a segmentacja rynku czy uznaniowe zniżki – niemożliwe. Konfrontacja uczestników rynku jest o wiele bardziej namacalna, dziejąca się w zupełnie innej skali czasowej. Z drugiej strony, automatyzacja zawierania kontraktów kusi możliwościami ekspansji na rynku i zwiększonymi obrotami, rekompensującymi obniżone marże.

Można wskazać nieliczne przykłady skutecznie funkcjonujących rynków usług teleinformatycznych. Jednym z nich jest MERKATO [7] – platforma transakcyjna, w której towarem jest pasmo na styku międzyoperatorskim, w konkretnej klasie DiffServ. Uczestnicy składają oferty kupna i sprzedaży w czasie rzeczywistym. Platforma funkcjonuje dla kilku punktów wymiany ruchu w Nowym Jorku i okolicach – działa więc w sposób skupiony i scentralizowany. Innym przykładem jest TelArena – platforma detalicznej sprzedaży połączeń do wybranych operatorów sieci telefonicznej stacjonarnej i ruchomej.

Należy zauważyć, że istotnym warunkiem nadającym rację bytu systemom giełdowym jest autentyczna liberalizacja rynku usług. Dominacja rynku przez jednego z graczy – co ma z reguły miejsce w Europie w przypadku ruchu IP – podważa sens wdrażania takiego systemu w postaci pierwotnej. Można jednakże zaprojektować go tak, by skłonić uczestników do zachowań racjonalnych dla nich samych, a jednocześnie składających się na realizację pewnego, zakładanego celu społecznego [5], [4].

Badania nad architekturą usług uwzględniającą racje poszczególnych podmiotów zaangażowanych na całej ścieżce transmisji były podejmowane w projektach europejskich oraz innych badaniach. Projekt M3I (Market Managed Multiservice Internet) skupiał się na technologii realizacji transmisji, której towarzyszyłby sygnał cenowy o wartości będącej kumulacją cen częściowych dla domen administracyjnych, przez które ruch jest przesyłany. Głównym zastosowaniem takiej technologii miało być zapobieganie przeciążeniom sieci poprzez skłonienie użytkowników do ograniczenia korzystania z niej w obliczu wzrostu ceny za usługę. Natomiast Co-COMBINE (Competition, Contents and Broadband for the Internet in Europe) był poświęcony analizie układu sił na rynku telekomunikacyjnym pod kątem istnienia rzeczywistych możliwości konkurowania ze sobą dostawców, a co za tym idzie, umożliwienia użytkownikowi wyboru spośród wielu ofert. Projekt nie skupiał się ani na realizacji kontraktów dynamicznych, ani na obsłudze połączeń z gwarancjami w sieci globalnej. Nieznane są losy komercjalizacji wyników obu projektów. Pokrewnym do M3I podejściem charakteryzowała się architektura z protokołem RNAP [12] – podobnym w działaniu do RSVP, ale umożliwiającym obliczanie i sygnalizowanie bieżących stawek za transmisję użytkownikom w czasie rzeczywistym.

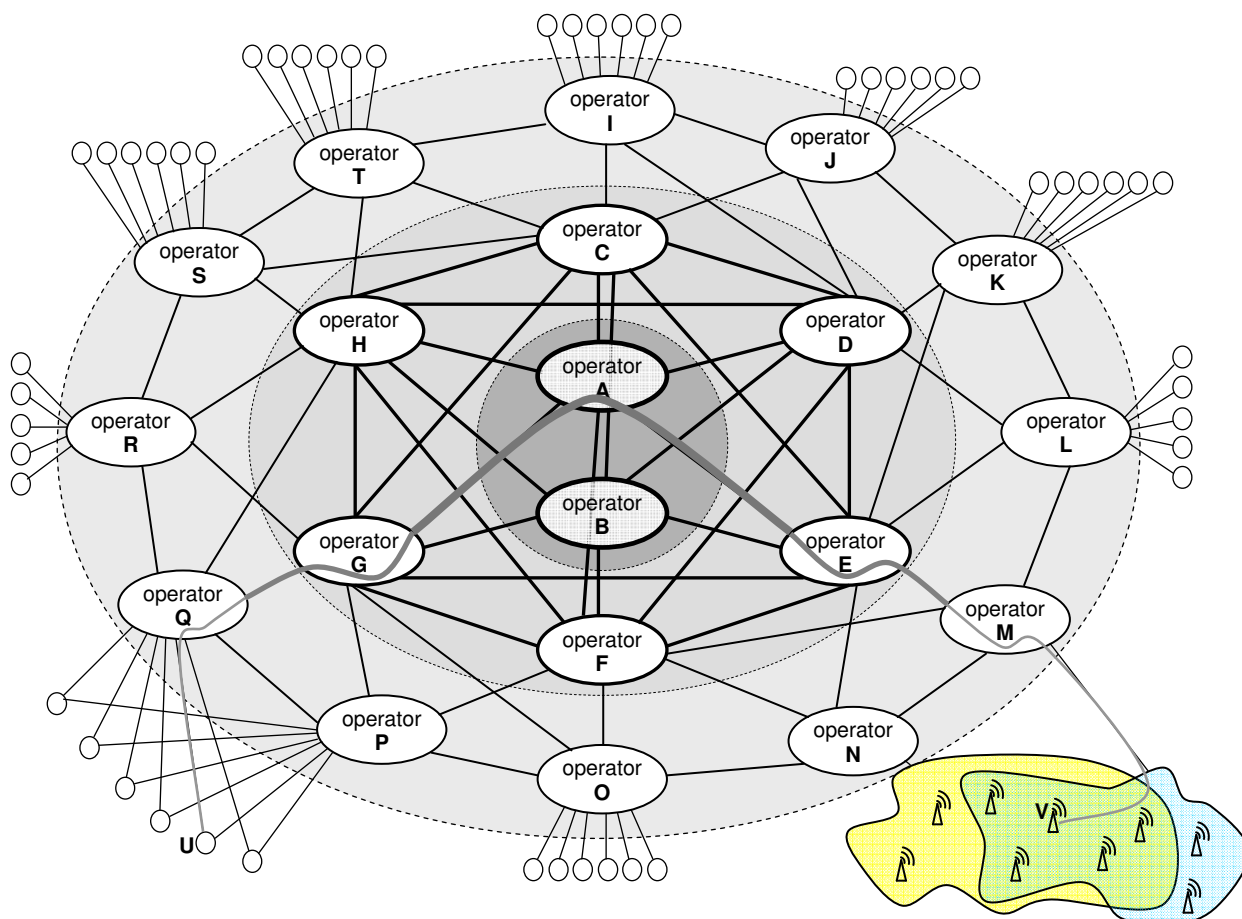
Badania nad aspektami ekonomicznymi internetu w przyszłym kształcie są obecnie kontynuowane w 7. PR w ramach projektu SMOOTHIT (w dużej części przez zespół odpowiedzialny koncepcyjnie za projekt M3I). W ramach prac ma powstać teoria oraz technologia zapewniająca korzyści zarówno dostawcom usług transportowych składających się na sieci nakładkowe, jak i użytkownikom takich sieci. Badania obejmują głównie aspekty ekonomiczne organizacji i działania takich sieci, oraz specyfikację technologii mającej służyć takiej efektywnej organizacji (tj. implementującej opracowaną teorię). Wątek badań nad ekonomią sieci spleta się zatem w 7PR z zagadnieniami opracowania technologii transmisji i organizacji takich sieci, reprezentowanymi jednakże przez dużo większą liczbę projektów.

Wzrastająca ważność danych transmitowanych w internecie uczyni zeń prędzej czy później medium równie ważne dla ludzkości jak energia elektryczna czy surowce naturalne. Dlatego zasadne jest myślenie o większej ingerencji regulatora w funkcjonowanie rynku usług sieciowych w celu zapewnienia nieprzerwanego dostępu do internetu – podobnie jak jest obecnie w przypadku rynku energii elektrycznej. Dlatego wartą odnotowania inicjatywą jest M3 [6] – model rynku wielotowarowego, wyrastający z prac autorów nad rynkiem energii, a mogącym mieć zastosowanie również dla rynku szeroko rozumianych usług telekomunikacyjnych. Cenną cechą modelu jest możliwość formułowania ofert kupna i sprzedaży produktów zawierających usługi składowe. Określone dla modelu zadanie optymalizacji jest sformułowane w sposób ciekawy, niosący znamiona wytycznych dla rynku energii.

Myślenie o sieci przyszłości jako o konglomeracie usług prostych i złożonych, transportowych i dodanych (billing, uwierzytelnianie, hosting, streaming), możliwych do kupienia na jednym, wspólnym rynku jest też cechą charakterystyczną inicjatywy IPsphere [3], [8]. IPsphere zamierza opracować specyfikację technologii umożliwiającej kontraktowanie różnego rodzaju usług w sposób swobodny i dynamiczny. Taki mechanizm miałby stymulować innowację i powstawanie nowych usług w sieci oraz – w wyniku tego – zwiększone zapotrzebowanie na ruch u operatorów i, oczywiście, zwiększoną satysfakcję klientów (*a win-win solution*). Wizja sieci globalnej z dynamicznie kontraktowanymi usługami jest artykułowana indywidualnie również przez poszczególnych dostawców sprzętu i operatorów sieciowych [9], [11].

Z poczynionych obserwacji wynika, że ścieżka rozwoju internetu, uznana za najbardziej właściwą lub prawdopodobną, będzie polegać na upowszechnianiu się sieci nakładkowych. Sieci takie, swobodnie ewoluujące w swojej topologii, będą, jak można mniemać, świadczyły zbiór najbardziej aktualnie pożądaných usług w skali globalnej (dystrybucja streamingu, videokonferencje, gry, telemedycyna). Taka wizja może budzić obawy: 1) czy koszt budowy sieci nakładkowej rzeczywiście nie ograniczy innowacyjności, 2) czy nadal istotne są i w jaki sposób mają być potraktowane kwestie zapewnienia jakości, oraz 3) czy rzeczywiście wszyscy operatorzy znajdą podstawy ekonomiczne do współdziałania we wspólnym udostępnianiu infrastruktury dla sieci nakładkowych?

Proponowana architektura systemu rezerwacji, wyłożona poniżej w szczegółach, nie ucieka się do pojęcia sieci nakładkowej. Jej głównym założeniem jest możliwość samoczynnego, hierarchicznego organizowania się relacji biznesowych i idąca w ślad za nią, hierarchiczna struktura kontraktów przesyłowych. Hierarchia taka ma swoje silne źródło w obecnej topologii i technologii internetu.



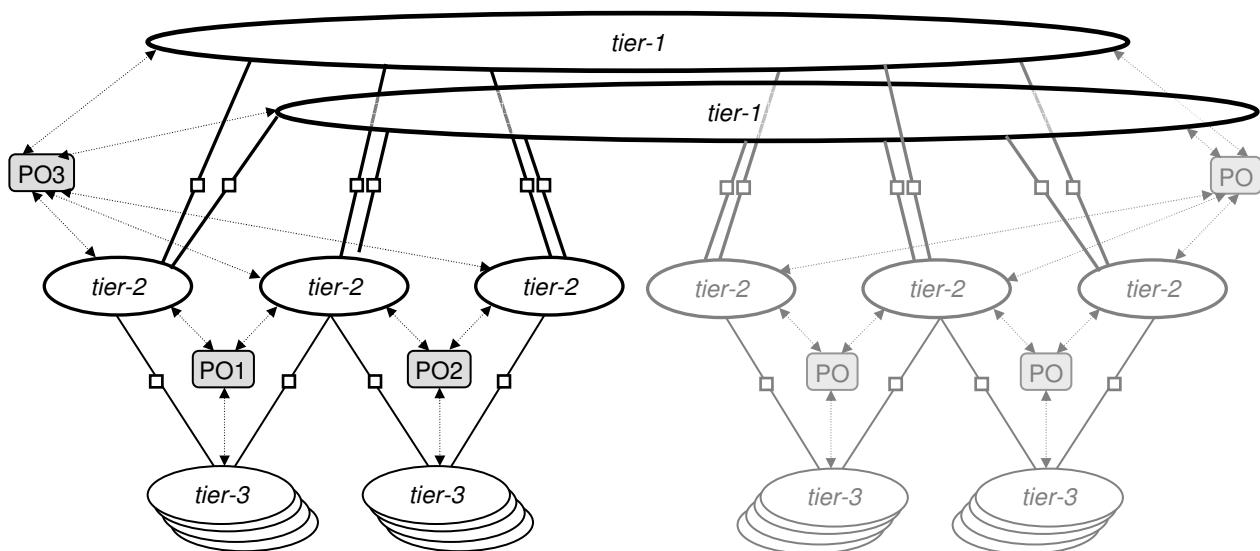
3. Hierarchiczność sieci globalnej wykorzystana w systemie giełdowym

Łączność w sieci globalnej zapewnia współpraca operatorów, powiązanych umowami typu *transit* (transmitowanie danych pod dowolny adres w sieci globalnej) oraz *peering* (transmitowanie pod dowolny adres w sieci partnera – z wzajemnością). Usługa tranzytowa jest zawsze płatna, usługa wymiany ruchu między systemami autonomicznymi (AS) jest zazwyczaj bezpłatna. Relacje, w jakich pozostają ze sobą operatorzy systemów autonomicznych, determinują ich pozycję w hierarchii:

- *Tier-1* – operatorzy, którzy są w stanie komunikować się z dowolnym punktem w internecie wyłącznie dzięki umowom o bezpłatnej wymianie ruchu;
- *Tier-2* – operatorzy, którzy są w stanie komunikować się ze znaczną liczbą AS dzięki posiadanym umowom peeringowym, ale do komunikacji z pozostałymi systemami potrzebują usługi tranzytowej świadczonej przez *tier-1*;
- *Tier-3* – operatorzy, którzy świadczą usługi klientom z reguły niebędącym systemami autonomicznymi; detaliści rynku usług przesyłowych.

W chwili obecnej działa dziewięciu operatorów typu *tier-1*. Tworzą oni pewnego rodzaju kartel, w którym próba jakiegokolwiek wyłomu skutkuje wykluczeniem buntownika z relacji peeringowych. Nie jest regułą, że operatorzy *tier-1* jednocześnie przenoszą najwięcej ruchu w sieci – istnieją bowiem operatorzy mający formalnie status *tier-2*, którzy z uwagi na swoją pozycję rynkową zamykają olbrzymią część swojego ruchu w formie peeringu, oraz mogą liczyć na bardzo korzystne warunki korzystania z tranzytu operatorów *tier-1*. Niezwykle pouczający jest również wykres topologii internetu wykreślonej dla systemów autonomicznych, wykonany przez CAIDA [1]. Wynika zeń m.in. że największe możliwości bezpośredniej transmisji do innych systemów autonomicznych ma operator Level 3: jest to 1845 systemów autonomicznych. Możliwości innych dużych operatorów (w zestawieniu jest ich około 50) stopniowo maleją, by osiągnąć wskaźnik rzędu 200 przyległych AS dla operatora z końca tej krótkiej listy.

Wdrożenie w tak ewidentnie zhierarchizowanej topologii systemu dynamicznej rezerwacji kanałów transmisji działającego sposób płaski (z jednakowym traktowaniem uczestników, bez względu na ich rolę) jest niemożliwe. Oznaczałoby bowiem, że operatorzy działający w rdzeniu sieci musieliby z olbrzymią wydajnością obsługiwać żądania dla każdego nowego kontraktu, powstające na brzegu sieci. W ustalonej technologii, np. MPLS, oznaczałoby to nieustanne tworzenie i usuwanie ścieżek, plus narzut obliczeniowy związany z kalkulacjami QoS w następstwie zawartego kontraktu.



Dlatego proponujemy, by system rezerwacji kanałów działał w sposób hierarchiczny, zgodnie z przykładem zilustrowanym na Rys. 2. Użytkownicy końcowi (np. indywidualni) zamierzają zestawić pomiędzy sobą kanał komunikacji. Jeden z nich występuje ze stosownym żądaniem do właściwego operatora (pomi-

Rys. 3. Obsługa żądań przez wiele niezależnych procesorów ofert

nięto fazę wyboru tego operatora przez system giełdowy). Operator taki, działający na wyższym poziomie agregacji ruchu, powinien już posiadać odpowiednio zwymiarowany kanał łączący go z operatorem, który umożliwi zestawienie kanału dla użytkowników końcowych. Kanał zbiorczy powinien być tak zwymiarowany, by jego utrzymanie było opłacalne wobec aktualnych statystyk zgłoszeń na rezerwację kanałów pomiędzy opisywanymi operatorami. Jeśli ruch na wybranej trasie (liczba zgłoszeń) jest duży, opłaca się utrzymywać stałą

Rys. 2. Funkcjonowanie systemu w obecnej, hierarchicznej strukturze zależności w sieci globalnej

pojemność kanału. Jeśli natomiast ruch jest mały, być może opłaca się utworzyć lub przewymiarować kanał w reakcji na każde, konkretne zgłoszenie od użytkowników końcowych. Ustanowienie takiego kanału dla ruchu zagregowanego odbywa się analogicznie, jak dla scenariusza w niższym szczeblu hierarchii: operatorzy występują do swoich dostawców usług z żądaniem rezerwacji. Sytuacja powtarza się do czasu, aż operator obsługujący żądanie dysponuje odpowiednim łączem fizycznym, i dokonuje rezerwacji pasma na rzeczywistych urządzeniach.

Można szacować, że na najwyższym szczeblu hierarchii (dla *tier-1*) liczba osiągalnych przez największego operatora (tj. AT&T) systemów autonomicznych wynosi łącznie około 7000 **Błąd! Nie można odnaleźć źródła odwołania.** Przy założeniu, że każdy osiągalny AS utrzymuje za pomocą tego operatora kanał logiczny z wszystkimi innymi, otrzymujemy liczbę około 25 mln jednoczesnych połączeń. Dokładność szacowania tej liczby jest niewielka; nie wydaje się, by każdy z 7 tysięcy AS miał potrzebę komunikowania się z wszystkimi pozostałymi naraz. Jeśliby jednak taka potrzeba zachodziła, wystarczy rozważyć stosowanie tego samego mechanizmu agregacji wewnątrz systemów autonomicznych, grupując np. wszystkie kanały skierowane do konkretnej lokalizacji danego operatora.

4. Decentralizacja procesu przetwarzania żądań w systemie

Skuteczne i wydajne funkcjonowanie proponowanego systemu giełdowego zależy w dużej mierze od topologii sieci globalnej oraz, dopasowanej do niej, architektury systemu giełdowego. Rozważmy infrastrukturę niezbędną do realizacji żądania zawarcia kontraktu, według Rys. 1, z uwzględnieniem hierarchicznej struktury sieci globalnej, przedstawionej na Rys. 2. Została ona przedstawiona na Rys. 3. Na system giełdowy składają się w proponowanej architekturze niezależne, dedykowane jednostki zwane procesorami ofert (PO). Procesory ofert mają formę serwisów utrzymywanych przez podmiot, lub podmioty, niereprezentujące interesów żadnego z graczy na rynku, i działające według jednolitego, uznanego przez wszystkich graczy, algorytmu. Pełnią one rolę pośredniczącą i są gwarantami uczciwej realizacji uzgodnionych przez wszystkich zasad funkcjonowania giełdy. Lokalizacja jest znana zarówno użytkownikowi, jak i wszystkim operatorom, z których usług korzysta dany użytkownik. Tę informację użytkownik sam, w swoim dobrze pojętym interesie przekaże tym operatorom, by mogli być oni powiadamiani o jego żądaniach.

W momencie przyjęcia żądania, na podstawie listy alternatywnych punktów styku wyspecyfikowanych w żądaniu, PO ustala listę operatorów, którym należy przekazać to żądanie. Żądanie, przekazywane do konkret-

nego operatora, jest pozbawiane alternatywnych punktów styku z innymi operatorami, wyspecyfikowanych przez użytkownika. Tak spreparowane żądanie trafia do ustalonego przez operatora miejsca obsługi, gdzie jest przetwarzane. Każdy z powiadomionych operatorów, jako informację zwrotną, przekazuje swoją ofertę na realizację usługi. PO ocenia otrzymane oferty i wybiera tę najkorzystniejszą dla klienta. Kryteria wyboru klienta mogą być przekazane przezeń każdorazowo wraz z żądaniem, lub też mogą spoczywać zapisane w profilu klienta w PO.

W opisywanej architekturze zasięg rozgłaszania żądania klienta jest zdeterminowany topologią sieci. Oznacza to, że nie obejmuje on operatorów, którzy nie dysponują bezpośrednim fizycznym przyłączem do klienta. Takie ograniczenie konserwuje obecny hierarchiczny obraz sieci, ale dzięki temu jest skalowalne. Na forach naukowych pojawiały się wprawdzie inicjatywy stworzenia globalnego, rozproszonego systemu aukcyjnego, w którym dynamicznie zrzeszający się, mali operatorzy, oferowaliby usługi transportowe na znaczne odległości po cenach konkurencyjnych do *tier-1*. Jednak w opinii ekspertów mocno osadzonych w realiach funkcjonowania rynku, cena jednostkowa przepustowości zestawionej w taki sposób i tak okazałaby się wyższa niż obecnie oferowana przez oligopolistyczny układ dużych operatorów tranzytowych. Natomiast skalowalność takiego rozwiązania byłaby zdecydowanie mniejsza i możliwa do poprawienia jedynie kosztem suboptymalności rozwiązań.

Wcześniej zaprezentowany schemat dla relacji dostawca-klient jest powielany również w wyższych warstwach hierarchii, z dotychczasowym dostawcą wcielającym się w rolę klienta wobec swojego dostawcy usług hurtowych. Częstość zgłaszania żądań przez dostawcę swojemu PO (a więc i częstość negocjacji kontraktów) zależy od jego własnej strategii ekonomicznej wobec rodzaju żądań otrzymywanych od klientów, posiadanych punktów styku z dostawcami hurtowymi i strategii ekonomicznych tych dostawców. Strategia działania operatora nie jest w jakikolwiek sposób narzucona. Wobec zgłoszonego żądania klienckiego, operator sprawdza, czy jest w posiadaniu kanału transmisyjnego do operatorów obsługujących alternatywne punkty styku wyspecyfikowane przez klienta. Jeśli tak, to oferuje cenę wynikającą z uzgodnionych wcześniej z wybranym operatorem opłat za kanał transmisyjny, plus koszt transmisji wewnątrz sieci własnej i operatora. Jeśli natomiast nie ma możliwości zaoferowania klientowi usług (nie zarezerwował kanału, albo pojemność kanału jest za mała, albo zaoferowana cena byłaby ewidentnie niekonkurencyjna), wówczas może albo odmówić świadczenia usługi, albo na bieżąco zestawić nowy kanał bądź renegotjować parametry istniejącego.

Zestawienie kanału lub renegotjacja wymaga natychmiastowego zgłoszenia operatorom hurtowym zapotrzebowania na nowy kanał. Jego parametry będą wynikać z szacunków operatora co do przyszłego ruchu na właśnie zestawianym kierunku. Oczywiście, najmniej wyrafinowanym zachowaniem jest zestawienie takiego kanału tylko na czas realizacji kontraktu żadanego przez użytkownika. Prowadziłoby ono do całkowitej utraty skalowalności: operatorzy *tier-1* musieliby de facto rezerwować zasoby dla każdego połączenia między użytkownikami końcowymi, a całym system rezerwacji stałby się kosztownym i spowalniającym przekąźnikiem żądań detalicznych aż na szczyt sieciowej hierarchii. Tak się jednak nie stanie z dwóch powodów. Po pierwsze, operatorzy duzi nie będą zainteresowani małymi i krótkotrwałymi kontraktami, tj. będą oferować za nie ceny nie do przyjęcia przez operatorów mniejszych, nie wspominając już o użytkownikach końcowych. Po drugie, w obronie przed ewentualnym przeciążeniem serwisów realizujących funkcjonalność systemu, PO mogą dokonywać wstępnej priorytetyzacji żądań, faworyzując zlecenia „duże” (znaczna przepustowość, dystans, czas trwania).

Zauważmy, że żądanie zestawienia kontraktu ma zawsze tę samą postać, niezależnie od miejsca w hierarchii, gdzie jest przetwarzane. W niższych warstwach hierarchii żądania takie zgłaszane są częściej i dotyczą mniejszych i zestawianych na krótszy okres, przepustowości. W warstwach wyższych, gdzie uwidaczniają się efekty agregacji ruchu, żądania zgłaszane są rzadziej, lecz dotyczą one większych przepustowości i dłuższego czasu zestawienia kontraktu. Zatem algorytm działania PO jest uniwersalny. Zauważmy również, że nic nie stoi na przeszkodzie, aby funkcjonalność systemu rozłożyć na wiele PO, dedykowanych do obsługi konkretnego fragmentu topologii, bądź to w sensie poziomu hierarchii (PO3), bądź w sensie geograficznym (PO1 i PO2 na Rys. 3). Nic nie stoi na przeszkodzie, by operatorzy-usługodawcy rejestrowali się (i uaktualniali swoje profile cenowe) w więcej niż jednym PO. Nic też nie stoi na przeszkodzie, by operator-klient kierował swoje żądania do różnych PO, pod warunkiem, że konkretne żądanie będzie wysłane tylko do jednego PO.

Procesory ofert są jedynie niezbędnym fragmentem infrastruktury systemu, nieingerującym zupełnie w ekonomię zawierania kontraktów. Mogą więc należeć do różnych podmiotów, którzy w przypadku skrajnym mogliby konkurować między sobą o klientów. Procesory ofert mogą działać niezależnie dopóki algorytm wyboru najkorzystniejszej decyzji nie uwzględnia realizacji celów społecznych, które mógłby narzucić regulator rynku. W takim przypadku system przestaje być zdecentralizowany, a PO nie mogą już działać niezależnie. Stopień autonomii PO i jego związek z realizacją tak zdefiniowanych celów wykracza poza ramy tego artykułu.

Proponowana architektura dobrze obsługuje łatwo agregujące się kanały transmisji: liczne i łączące podobne geograficznie lokalizacje. Niestety, połączenia sporadyczne, pojedyncze i zestawianych na „nietypowych” relacjach angażują w taki sam sposób operatorów *tier-1*, co połączenia duże. A nawet gorzej: ekonomia tych połączeń wymagałaby bowiem ich każdorazowego zestawiania i zrywania, gdy ustanie potrzeba komunikacji. Wydaje się, że przy obecnie bardzo rozwiniętej topologii rdzenia sieci (np. Level(3) dysponuje 5 tys. łączy pomiędzy swoimi 600 routerami rdzeniowymi w USA – zob. [10]) zestawienie pełnej siatki kanałów pomiędzy wszystkimi punktami styku lub przynajmniej punktami obecności byłoby technicznie wykonalne.

5. Opłacalność ekonomiczna systemu

Posługując się uproszczonym modelem, można wykazać opłacalność stosowania przedstawianej architektury na ustalonym szczeblu hierarchii. Załóżmy, że usługa transmisyjna polega na wydzieleniu pasma do dyspozycji użytkownika, na określonym odcinku. Zauważmy, że usługa ta nie musi być realizowana w warstwie 3. sieci: do jej świadczenia można wykorzystać równie dobrze ATM lub techniki agregowania i priorytetyzacji w technologii Ethernet. Wydzielone pasmo, o ile nie jest wykorzystywane przez użytkownika, może służyć transmisji w klasie domyślnej, bez gwarancji (*best effort* – BE). Użytkownicy-klienci operatora, których jest n_0 , korzystają z aplikacji elastycznych (stosunkowo niewrażliwych na spadek jakości transmisji); poza tym n_1 użytkowników korzysta również z aplikacji nieelastycznych: dla nich spadek jakości transmisji, rozumiany jako wzrost opóźnienia lub współczynnika strat wiąże się z dużym dyskomfortem. Korzystanie przez użytkownika z aplikacji nieelastycznej niech będzie modelowane – analogicznie do ruchu połączeń głosowych w telefonii – przez proces Poissona z częstością λ i o średnim czasie trwania równym h .

Dla odpowiednio dużej liczby użytkowników, liczba aktualnie aktywnych aplikacji elastycznych i nieelastycznych osiąga pewną średnią wartość. Wynika z niej natężenie ruchu na rozpatrywanym odcinku, z niego zaś użyteczności jednostkowe obu typów aplikacji, odpowiednio $u_0(c_0)$ i $u_1(c_1, b_1)$, gdzie c_0 oznacza pasmo przypadające aplikacji elastycznej, c_1 – pasmo przypadające aplikacji nieelastycznej, a b_1 – prawdopodobieństwo niezrealizowania usługi z powodu braku zasobów. Użyteczności zależą również od liczby użytkowników, tutaj potraktowanej jako stała i nieuwzględnionej we wzorach. Użyteczności są naturalnie funkcjami rosnącymi względem c , aczkolwiek nieliniowo; dodatkowo użyteczność aplikacji nieelastycznej maleje względem b_1 . Można utożsamiać użyteczności z maksymalną ceną, jaką użytkownik jest skłonny zapłacić za korzystanie z aplikacji, a więc również z maksymalnym przychodem operatora w sytuacji rynku zmonopolizowanego. W rozpatrywanym modelu nie zostały uwzględnione kwestie podziału tej użyteczności pomiędzy dostawcę a odbiorcę usług. Założono, że wynikają one z faktycznej konkurencji na rynku: zatem cena danej usługi będzie zapewne niższa niż jej użyteczność, a wyższa niż koszt realizacji. W odniesieniu do kosztów poczyniono podobne założenie: w modelu traktowane są one jako stałe, ustabilizowane w wyniku konkurencji operatorów znajdujących się o stopień wyżej w hierarchii i świadczących hurtowe usługi przesyłowe operatorowi z rozpatrywanego scenariusza.

Rozważmy **wariant A** funkcjonowania operatora, czyli obecny stan rzeczy, w którym operatorzy nie współdziałają na rzecz zagwarantowania użytkownikom jakości usług transmisyjnych. Wówczas oba rodzaje aplikacji korzystają z klasy BE (bez gwarancji). Mamy średnio $n_1\lambda h$ aktywnych sesji aplikacji nieelastycznych oraz n_0 użytkowników aplikacji elastycznych. Jeśli przyjąć sprawiedliwy podział dostępnego pasma C (co w praktyce może mieć miejsce, zważywszy że coraz więcej aplikacji nieelastycznych korzysta z protokołu TCP), wówczas pasmo dostępne aplikacji wynosi $c = C/(n_0+n_1\lambda h)$, a łączna użyteczność klientów operatora wynosi

$$n_0u_0(c) + n_1\lambda hu_1(c, 0).$$

Następnie rozważmy **wariant B** funkcjonowania operatora, w którym zostaje wydzielone dla aplikacji nieelastycznych pasmo C_1 , oraz zostaje przyjęta pewna strategia dopuszczania połączeń dla tych aplikacji, w której gwarantuje się połączeniu stałą przepustowość c_1 . Wówczas można zrealizować równocześnie $N = C_1/c_1$ połączeń. Prawdopodobieństwo niezrealizowania połączenia (tzw. blokowania) można wyznaczyć posługując się wzorem B Erlanga, jednak w przypadku ogólniejszym, gdy N jest małe (przyjmuje się, że poniżej 200), stosowniej jest korzystać ze wzoru Engseta. Załóżmy, że liczba użytkowników i kanałów operatora jest wystarczająco duża, by stosować wzór B Erlanga i oznaczymy prawdopodobieństwo blokowania wyznaczone w ten sposób przez $P(n_1\lambda h, N)$. Średnia liczba aplikacji korzystających z wydzielonego kanału wynosi $n_1\lambda h$, więc dla klasy BE pozostaje $C - n_1\lambda hc_1$ pasma. Pasma to jest dodatkowo współdzielone z aplikacjami nieelastycznymi, które nie uzyskały dostępu do wydzielonego kanału. Łączna użyteczność wynosi w tym przypadku

$$n_0u_0(c_0) + (1 - P)n_1\lambda hu_1(c_1, P) + n_1\lambda Phu_1(c_0, 0),$$

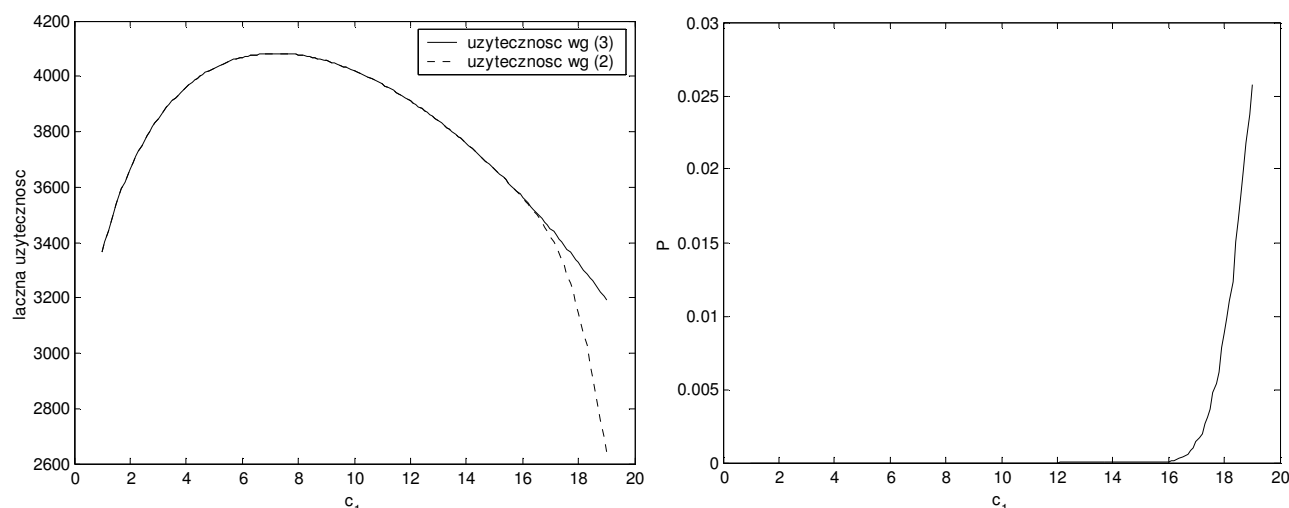
gdzie $c_0 = (C - n_1\lambda hc_1)/(n_0 + n_1P\lambda h)$ i oznacza pasmo przypadające na wszystkie aplikacje elastyczne i aplikacje nieelastyczne, które nie mogły skorzystać z kanału dedykowanego.

Operator, wprowadzając usługę transmisyjną z gwarancjami, staje przed zadaniem optymalnego doboru jakości świadczenia tej usługi (zmienna decyzyjna c_1) oraz jej dostępności (zmienna decyzyjna C_1), przy wartościach pozostałych parametrów traktowanych jako ustalone. Zauważmy, że przyjęcie $C_1 = C$ stanowi potencjalnie najlepszy wybór, gdyż pozwala na świadczenie usługi z gwarancjami jakości jak największej liczbie klientów, kosztem ograniczenia komfortu korzystania z klasy BE – ale nie zaprzestania świadczenia tej usługi w całości. Do przyjęcia takiej strategii jest konieczne, by użyteczność aplikacji nieelastycznej była istotnie większa niż elastycznej przy takim samym przydzielonym pasmie – wówczas rugowanie ruchu aplikacji elastycznych będzie opłacalne. Użyteczność aplikacji nieelastycznych jest również uwarunkowana prawdopodobieństwem blokowania – nie może być ono zbyt duże. Prawdopodobieństwo to wynika z wyboru pasma dostępnego pojedynczej aplikacji nieelastycznej, c_1 . Jeśli wykres funkcji użyteczności względem c_1 spłaszcza się powyżej pewnej wartości c_1' , wówczas przyjmowanie $c_1 > c_1'$ jest nieekonomiczne. Lepiej jest zmniejszyć c_1 , redukując tym samym prawdopodobieństwo blokowania i podnosząc wartość użyteczności aplikacji.

Powyższe rozważania prowadzą do wniosku: jeśli w sytuacji wyjściowej (bez różnicowania jakości) aplikacje nieelastyczne posiadają potencjał wzrostu użyteczności przy zwiększonym przydziale pasma, i jeśli jest on większy niż dla aplikacji elastycznych, wówczas operator powinien zwiększyć przydział pasma dla aplikacji nieelastycznych. Natomiast wartość pasma przydzielonego pojedynczej aplikacji elastycznej wynika z kształtów funkcji użyteczności aplikacji obu typów oraz liczby klientów. Zadanie optymalizacji ma więc jedną zmienną decyzyjną, c_1 . Jeśli przyjąć dodatkowe założenia: że liczba klientów jest duża, prawdopodobieństwo blokowania – małe (czemu sprzyja duże n_1), a zależności pomiędzy funkcjami użyteczności nie skłaniają do rugowania ruchu BE, wówczas zadanie optymalizacji przyjmie postać:

$$\max_{c_1} (n_0 u_0((C - n_1\lambda hc_1)/n_0) + n_1 \lambda h u_1(c_1, 0)), \quad c_1 \in \langle 0, C/n_1 \rangle$$

Jeśli, dla przykładu, przyjmiemy jednakowe, typowe, logarytmiczne funkcje użyteczności: $u_0 = \ln((C - n_1\lambda hc_1)/n_0 + 1)$, $u_1(c_1, 0) = x \ln(c_1 + 1)$, gdzie użyteczność aplikacji nieelastycznej jest x -krotnie większa od elastycznej przy takim samym pasmie, to otrzymujemy wklęsłe zadanie optymalizacji, z rozwiązaniem $c_1^* = [x(C + n_0) - n_0]/(n_0 + n_1\lambda hx)$. Można pokazać, że dla $x > 1$ rozwiązanie to odpowiada większemu przydziałowi pasma dla aplikacji nieelastycznej niż przypadkowi opisanemu przez (1), a co za tym idzie uzasadnia proponowane postępowanie (tzn. przydział pasma aplikacjom nieelastycznym kosztem elastycznych) jest opłacalne. Warto przy tym zauważyć, że dla $x = 1$, czyli przypadku granicznego, $c_1^* = (C + n_0 - n_0)/(n_0 + n_1\lambda h) = C/(n_0 + n_1\lambda h) = c$. Dla zobrazowania potencjalnych zysków ze stosowania priorytetyzacji ruchu, jeśli przyjmą $n_0 = 4000$, $n_1 = 2000$, $\lambda h = 0,1$, $C = 4000$, $x = 5$, wówczas osiągnięta łączna użyteczność wynosi w przypadku (1) 3345, a w zadaniu (3) – 4083, przy $c_1^* = 7,2$. Dla takiej wartości c_1 prawdopodobieństwo blokowania jest w tym przykładzie praktycznie zerowe; prawdopodobieństwo to osiąga istotne wartości z dala od optymalnej wartości c_1^* . Łączną użyteczność obliczaną wg (2) oraz (3) oraz prawdopodobieństwo blokowania w funkcji c_1 przedstawiono na Rys. 4. Dla wariantu z (2) przyjęto $u_1(c_1, P) = x \ln(c_1 + 1) - c_1 P$, co oznacza, że użyteczność spada proporcjonalnie do doświadczanego ryzyka blokowania, a dodatkowo spadek ten jest tym większy, im większa żądana przepustowość wchodzi w grę. Taki sposób uwzględnienia spadku jakości usług jest wzorowany na [2], gdzie za miarę jakości przyjęto opóźnienie transmisyjne. Podobną marginalizację oddziaływania blokowania na optymalną wartość c_1 uzyskiwano dla szerokiego zestawu wartości parametrów przykładu.



Rys. 4. Łączna użyteczność oraz prawdopodobieństwo blokowania w systemie w funkcji przepustowości przydzielonej pojedynczej aplikacji nieelastycznej

Strategia operatora ulega zmianie w sytuacji, gdy z użytkowaniem kanału o gwarantowanych parametrach związane są koszty. Tak jest w przypadku dzierżawienia kanału przez operatora, czyli w mniejszym lub w większym stopniu dla wszystkich uczestników rynku, poza *tier-1*. Rozważmy dwa podstawowe rodzaje kosztów: utrzymania kanału o określonej przepustowości oraz zestawienia bądź zmiany parametrów kanału. Dla ustalonej, niezmiennej w czasie, częstości zgłoszeń λ operator ma do wyboru dwie strategie: albo utrzymywać zarezerwowany na stałe kanał o określonej przepustowości (wariant C), albo zmieniać jego parametry i dopasowywać przepustowość do aktualnej liczby aktywnych aplikacji nieelastycznych (wariant D). Przy ustalonym λ , jakakolwiek strategia pośrednia, polegająca np. na okresowej zmianie parametrów kontraktu, będzie nieoptymalna, gdyż spowoduje powstawanie dodatkowych kosztów. Bilans użyteczności i kosztów dla **wariantu C** można zapisać, rozwijając (2) do postaci

$$n_0 u_0(c_0) + (P-1)n_1 \lambda h u_1(c_1, P) + n_1 \lambda P h u_1(c_0, 0) - k N c_1,$$

gdzie k jest kosztem wynajęcia jednostki przepustowości w jednostce czasu. Natomiast w **wariancie D** mamy

$$n_0 u_0(c_0) + (P-1)n_1 \lambda h u_1(c_1, P) + n_1 \lambda P h u_1(c_0, 0) - k n_1 \lambda h c_1 - s n_1 \lambda (1-P),$$

gdzie s jest kosztem pojedynczej zmiany rezerwacji. Koszt dzierżawy przepustowości zależy od stopnia uśrednienia ruchu na dzierżawionym łączu, przy założonym prawdopodobieństwie blokowania P . Dostawcy realizujący kontrakty na pojedyncze połączenia mają $n_1 \lambda h \ll N$. Natomiast dostawcy realizujący kontrakty hurtowe lub znaczną liczbę kontraktów indywidualnych mają $n_1 \lambda h$ zbliżone do N . Założmy, że koszt s zestawienia kontraktu jest niezależny od wielkości operatora i liczby zgłoszeń. Wówczas, przy założonym odpowiednio małym P , wariant C jest opłacalny dla $n_1 \lambda > N c_1 / (h c_1 - s/k)$. W przeciwnym przypadku opłaca się wariant D.

Z racji małej intensywności ruchu i nieefektywnego uśredniania, użytkownikom indywidualnym i małym operatorom opłaca się każdorazowe żądanie zestawienia kanału. Po osiągnięciu pewnego stopnia agregacji ruchu, lepszą strategią jest rezerwacja kanału na stałe. W praktyce jednak częstość zgłoszeń podlega zmianom, według co najmniej dobowej okresowości. Dlatego nawet operatorom stosującym wariant C opłaca się okresowa zmiana parametrów kontraktu. Odbywa się ona jednak w zupełnie innej skali czasowej niż u operatorów małych. Zatem raczej ekonomiczne wszystkich uczestników rynku gwarantują skalowalność systemu: na wyższych szczeblach hierarchii operatorzy muszą, co prawda, zarządzać większą liczbą punktów styku oraz fizycznych połączeń między nimi, ale żądania rekonfiguracji połączeń są rzadsze.

Przedstawiony model ekonomiczny może wydawać się nieadekwatny do zaprezentowanego wcześniej scenariusza rezerwacji pasma, gdzie uczestnik rynku, aby poznać cenę usługi, każdorazowo, za pośrednictwem systemu giełdowego, angażuje potencjalnych dostawców w wycenę. W praktyce tak być nie musi: dostawca może przechowywać swój profil w systemie na potrzeby kolejnych, podobnych zapytań, i uaktualniać go, jeśli zmieniły się jego własne koszty świadczenia usługi.

6. Możliwości wdrożenia i rozwoju systemu

W poprzednich rozdziałach zaproponowano sposób agregacji kontraktów umożliwiający skalowalność architektury. Opłacalność zawierania kontraktów została wykazana przy założeniu niemonopolistycznego zachowania operatorów – wówczas w szacunkach można przyjąć, że parametry ścieżek transmisji i koszty korzystania z nich są w przybliżeniu podobne dla każdego operatora. Jednak aby rynek usług umożliwił taką konkurencję, powinien zrzeszać możliwie dużą liczbę i operatorów, i klientów, uczestniczących nie tylko formalnie, ale mających rzeczywistą możliwość wyboru sposobu realizacji usługi – co jest determinowane topologią połączeń międzyoperatorskich. Dla przytaczanego już przykładu MERKATO warunki te zachodzą, aczkolwiek w ograniczonym zasięgu terytorialnym. Natomiast w Polsce, mimo nominalnie dużej liczby operatorów telekomunikacyjnych (ok. 600), możliwości wyboru dostawcy usługi ma tylko niewielka ich część, a i tak warunki, w jakich jest on dokonywany, trudno nazwać wolnorynkowymi. Taki stan rzeczy jest spowodowany płytkością polskiego rynku – skutkującą zarówno drzewiastą topologią połączeń jak i niskim umotywowaniem dostawców do konkurowania między sobą. Stan idealny – i utopijny – polegałby na obowiązkowym uczestniczeniu wszystkich podmiotów w globalnej giełdzie usług oraz na towarzyszącej mu, adekwatnie rozwiniętej topologii połączeń międzyoperatorskich.

Po stworzeniu odpowiednich ram prawnych, prezentowany system mógłby stać się narzędziem regulatora rynku usług, umożliwiającym takie zdefiniowanie prawideł wyboru dostawcy usługi, by osiągać różne cele

społeczne. Mogą one promować rozwój operatorów konkurencyjnych wobec dominującego, przydział pasma w celu maksymalizacji korzyści społecznej, optymalizacji wykorzystania infrastruktury itp. System giełdowy mógłby jednocześnie służyć jako miejsce stosowania ewentualnych dopłat i dotacji, przydzielanych według określonego klucza.

Proponowana architektura systemu nie precyzuje konkretnych technologii agregacji przepływów i gwarantowania jakości usług w ramach kontraktu. Obecnie najbardziej rozpowszechnioną i adekwatną do powszechnego zastosowania wydaje się być technologia T-MPLS. Niewykluczone jest jednak stosowanie i innych rozwiązań, jak ATM czy Ethernet z agregacją etykiety VLAN, pod warunkiem, że rzeczywista jakość usługi nie będzie odbiegać od zakontraktowanej. Najpowszechniejszym sposobem zapewnienia tych gwarancji jest połączenie technologii różnicowania jakości z technologią dopuszczania połączeń. Nie są to z reguły rozwiązania tanie; warto jednakże zauważyć, że wystarczy stosować je tylko w tych fragmentach sieci, które w istotny sposób wpływają na jakość usług i są wąskimi gardłami infrastruktury.

Bibliografia

- [1] CAIDA
- [2] Courcoubetis *Pricing Communication Networks: Economics, Technology and Modelling*
- [3] IPsphere Forum
- [4] Johari *Parameterized supply function bidding: equilibrium and welfare*
- [5] Karpowicz, *Characterization of scalar strategy mechanisms for efficient re-resource allocation: Supply side case,*
- [6] M3
- [7] MERKATO
- [8] Nolle *A New Business Layer For IP Networks*
- [9] Sen *Challenges in Evolving to a Converged Network*
- [10] Spring *Measuring ISP Topologies With Rocketfuel*
- [11] Stuart *Core Network Transformation*
- [12] Wang RNAP: A Framework for Congestion-Based Pricing and Charging for Adaptive Multimedia Applications