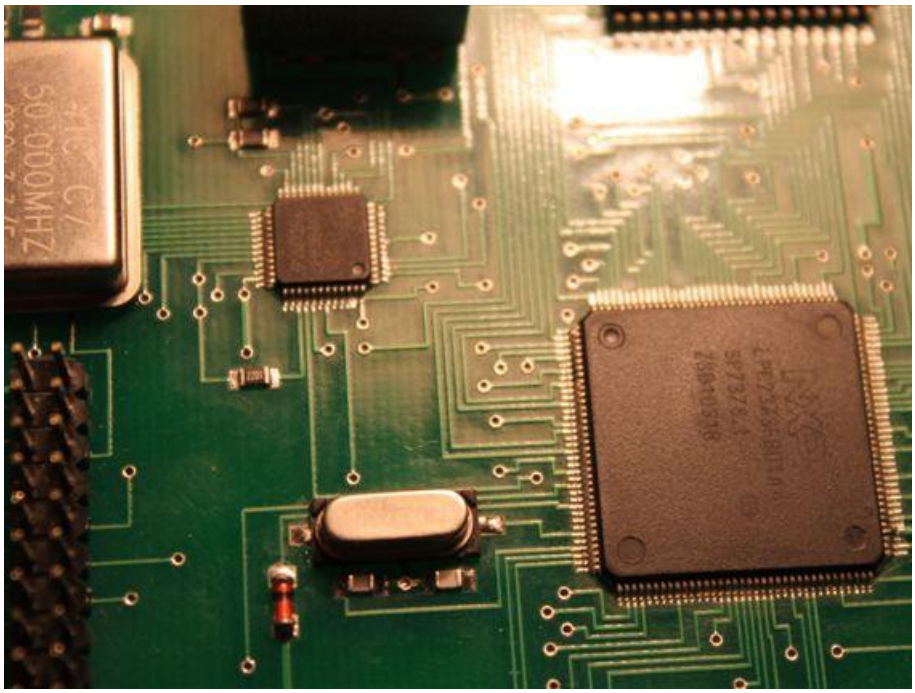
Andrzej Pawluczuk



Sieci komputerowe w praktyce elektronika rozważania o adresacji

Dla Elportal.pl

Nikogo nie trzeba przekonywać, że sieci komputerowe doprowadziły do prawdziwej światowej rewolucji. To, co jeszcze kilkadziesiąt lat temu było jedynie w sferze fantastyki, powoli stało się codziennością. Ogromna liczba komputerów została połączona ze sobą za pomocą „wynalazku" stworzonym w amerykańskiej DARPA (amerykańskiej agencji rządowej zajmującej się rozwojem zaawansowanych technologii na potrzeby wojska). Pozwoliło to na bardziej lub mniej swobodną wymianę informacji pomiędzy różnymi użytkownikami komputerów przyłączonych do sieci. Wiadomo, informacja jest „towarem" wręcz bezcennym. Powstanie sieci internetowej, bazującej właśnie na sieciach komputerowych, doprowadziło do przysłowiowej eksplozji nowych możliwości. Dzisiaj trudno sobie wyobrazić egzystencję bez wszystkowiedzącego wujka googla. Nieprzebrane zasoby sieci internetowej pozwalają pokonywać wiele barier, które jeszcze nie tak dawno były nie do przebycia. Doskonałym tego przykładem jest właśnie ten cykl artykułów (zatytułowany: *Domowa automatyka – system Infinity*). Naturalną cechą ludzi jest dążenie do wiedzy i rozwoju, co z kolei przekłada się na pomysły, koncepcje i rozwiązania. Nie byłoby to możliwe bez wynalazku, który narodził się w laboratoriach DARPA oraz słynnego ośrodka CERN w Szwajcarii. To właśnie tam postała koncepcja „World Wide Web". System ten zaprojektowano, aby zbierać zasoby ludzkiej wiedzy i umożliwić użytkownikom w odległych miejscach dzielenie się swoimi pomysłami oraz zgłębianie wszystkich aspektów wspólnego projektu. W przypadku, gdy dwa projekty tworzone były niezależnie od siebie, WWW pozwalała skoordynować pracę naukowców. W dzisiejszych czasach ta koncepcja wyewoluowała przede wszystkim do potężnej biblioteki zawierającej praktycznie wszystko (między innymi sieć internetowa stała się źródłem informacji dotyczących działania sieci komputerowych oraz źródłem programów narzędziowych, bez których pokonanie wielu barier nie byłoby możliwe).

Jak mówi znane porzekadło „apetyt rośnie w miarę jedzenia". Mając komputery przyłączone do ogólnoświatowej sieci, naturalną konsekwencją staje się budowanie małych urządzeń elektronicznych, które również będą przyłączone do tej samej sieci. Puszczając wodze fantazji, można stworzyć przykładowo układ pomiaru temperatury, którego dane pomiarowe będą mogły być zaprezentowane na ekranie komputera lub smartfonu. Wystarczy, że ten układ będzie potrafił wygenerować dane na potrzeby przeglądarki internetowej (a nie jest to skomplikowane). Również utworzenie odpowiednich poleceń do bazy danych (których interfejsy bazują na sieciach komputerowych) pozwoli na ich gromadzenie i w dalszej kolejności na realizację różnego typu analiz w komputerach. Analogicznie, komputer może stać się źródłem poleceń do różnorodnych urządzeń wykonujących odpowiednie funkcje sterujące. Powoli zaczynają znikać bariery techniczne na realizację różnych śmiałych pomysłów a pozostaje do pokonania jedynie bariera własnej pomysłowości. Może nadszedł czas na powiedzenie sobie, że jest coraz mniej rzeczy, których nie da się zrobić w warunkach hobbystycznych pracowni elektronicznych. Tak więc technologia z wielkiego świata komputerów zawitała pod strzechy małych mikrokontrolerów. Z ogromnego bogactwa różnych rozwiązań pozwalających na budowę sieci, w tym dokumencie mowa będzie przede wszystkim

o wariancie sieci określanym jako ethernet (a właściwie Fast Ethernet). Sieci ethernetowe to sieci oparte na przewodach elektrycznych. Typowy kabel ethernetowy zawiera cztery pary przewodów elektrycznych skręconych ze sobą (choć w rzeczywistości do transmisji wykorzystane są jedynie dwie pary przewodów) i głównie jest wykorzystywany do budowy lokalnych sieci komputerowych.

Zanim dojdzie do pierwszego starcia mikrokontrolera z siecią ethernetową, konieczne jest zapoznanie się z kilkoma istotnymi pojęciami, bez zrozumienia których stawianie kroków na nowym polu może być trudne. Ważnym elementem jest zrozumienie modelu **OSI** (ang. Open Systems Interconnection – model odniesienia łączenia systemów otwartych). Jego podstawowym założeniem jest podział systemów sieciowych na 7 warstw współpracujących ze sobą w ściśle określony sposób. Są to:

- Warstwa 7: aplikacji,
- Warstwa 6: prezentacji,
- Warstwa 5: sesji,
- Warstwa 4: transportowa,
- Warstwa 3: sieciowa,
- Warstwa 2: łącza danych,
- Warstwa 1: fizyczna

Generalnie zrozumienie idei podziału na w/w warstwy jest przydatne ale nie jest wymogiem absolutnie koniecznym. Warstwa aplikacji, jako warstwa najwyższa, zajmuje się specyfikacją interfejsu, który wykorzystują aplikacje do przesyłania danych do sieci. W przypadku sieci komputerowych aplikacje są zwykle programami, które komunikują się z siecią za pomocą obiektów nazywanych gniazdami (ang. socket) w architekturze klient-serwer. Serwer, to program, który serwuje możliwość nawiązania połączenia, jest uruchomiony i oczekuje na połączenie. Klient, to program, który inicjuje połączenie do oczekującego na połączenie serwera. Po nawiązaniu połączenia (w przypadku protokołu TCP) następuje wymiana danych pomiędzy klientem i serwerem. Komunikacja nigdy nie odbywa się bezpośrednio między tymi programami lecz za pośrednictwem wspomnianych gniazd, które są identyfikowane poprzez numer gniazda. Dane, które mają być przesłane, są uzupełnione informacjami opisującymi gniazdo. W niższej warstwie (prezentacji) przesyłane dane są przetworzone do postaci znormalizowanej. Najprościej jest to wyjaśnić na przykładzie przesyłania danych dwubajtowych. W pamięci RAM mikrokontrolera liczba zajmująca dwa bajty może być zapisana w ten sposób, że młodsza jej część będzie na młodszym adresie oraz starsza część będzie miała w pamięci adres o jeden większy. Nie jest to jedyne możliwe rozwiązanie. Równie dobrym rozwiązaniem może być odwrotna notacja (starsza część liczby będzie zapisana pod adresem mniejszym niż młodsza część). Zależy to wyłącznie od architektury danego mikroprocesora lub mikrokontrolera, dla których definiuje się pojęcie:

- big endian – jako forma zapisu danych, w której najbardziej znaczący bajt umieszczony jest jako pierwszy (pod adresem liczbowo mniejszym),

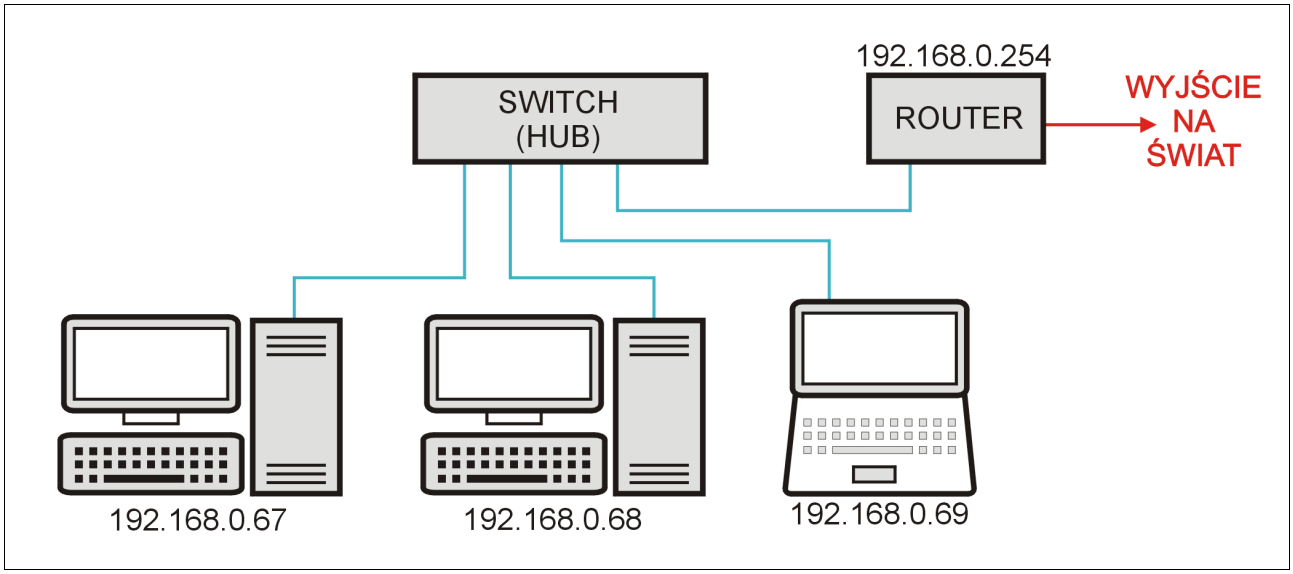
- little endian – jako forma zapisu danych, w której najmniej znaczący bajt umieszczony jest jako pierwszy.

By uniknąć niejednoznaczności przy przesyłaniu binarnych danych wielobajtowych, w komunikacji sieciowej używany jest wariant „big endian". Oznacza to, że programy w procesorach stosujących formę zapisu „little endian" muszą odwracać kolejność bajtów. Zagadnienie „wielkości endian" nie wyczerpuje funkcjonalności przewidzianej dla tej warstwy (odpowiada ona za kodowanie, konwersję danych, kompresję i dekompresję). Kolejna warstwa (sesji) odpowiedzialna jest za synchronizację przepływu danych pomiędzy nadawcą i odbiorcą. Niższa warstwa (transportowa) segmentuje dane w strumień i zapewnia połączenie między stacją źródłową oraz docelową. Do transmisji danych wykorzystywane są dwa protokoły komunikacyjne: TCP (ang. Transmission Control Protocol) oraz UDP (ang. User Datagram Protocol). W przypadku gdy do transmisji danych wykorzystany jest protokół TCP stacja docelowa po odebraniu segmentu wysyła potwierdzenie odbioru. W wyniku niedotarcia któregoś z segmentów realizowana jest retransmisja (ponowne przesłanie zagubionych danych). W przeciwieństwie do protokołu TCP, w protokole UDP nie stosuje się potwierdzeń, nie ma żadnej gwarancji, że wysłane dane dotarły do adresata. Następna niższa warstwa (sieciowa) zajmuje się przesyłaniem danych z uwzględnieniem fizycznej topologii sieci oraz może posługiwać się różnymi protokołami w warstwie sieci. W tym artykule opisany będzie protokół IPV4 (ang. Internet Protocol Version 4) – czwarta wersja protokołu komunikacyjnego IP przeznaczonego dla Internetu (obecnie powszechnie obowiązująca, choć na horyzoncie pojawia się rozwojowa wersja IPV6). Identyfikacja stacji w IPv4 opiera się na adresach IP, które są 32-bitową liczbą. W popularnej notacji, wartość tego adresu jest zapisywana jako cztery liczby z zakresu od 0 do 255 (wynikających z pojemności jednego bajta) rozdzielone znakiem kropki. Kolejna niższa warstwa (łącza danych) zajmuje się pakowaniem danych w ramki i przekazywaniem ich do warstwy fizycznej. W przypadku odbierania danych, warstwa ta wybiera z całego strumienia właściwe ramki i przekazuje odebrane dane do wyższej warstwy w celu ich przetworzenia. Najniższa warstwa (fizyczna) zapewnia właściwą obsługę wysyłania i odbierania sygnałów (elektrycznych, optycznych).

Świat mikrokontrolerów, w stosunku do komputerów, wygląda raczej jak ubogi krewny. Ograniczone zasoby, głównie wielkości pamięci RAM, zmuszają do znaczących uproszczeń wynikających z przedstawionych wyżej funkcjonalności poszczególnych warstw. Maksymalna wielkość ramki w sieci ethernet wynosi 1536 (600 hex) oktetów, toteż przy założeniu, że w przepływie danych w ramach modelu OSI nie następuje fragmentacja i defragmentacja przesyłanego strumienia danych, musi istnieć limit wielkości jednorazowo przesyłanych danych. Takie ograniczenie jest do zaakceptowania. W komputerze jednym zleceniem możliwe jest przesłanie przykładowo bloku danych o wielkości 1MB (obsługa sieci w niższych warstwach podzieli blok danych na mniejsze części). W przypadku mikrokontrolerów takie rozwiązanie staje się kłopotliwe, gdyż większość z nich nie dysponuje wewnętrzną pamięcią RAM o wymaganej wielkości. Gdyby nawet zaistniała konieczność

przesłania danych o dużej objętości, to w warstwie aplikacji można dokonać wymaganej fragmentacji strumienia. W mikrokontrolerach właściwym rozwiązaniem będzie pofragmentowanie bloku danych w warstwie aplikacji na takie części, by ich wielkości nie przekroczyły dopuszczalnej wielkości ramki, czyli przykładowo przesłanie obrazka, który ma objętość kilkadziesiąt kilobajtów nie będzie przesłane jednym zleceniem, tylko przykładowo w kawałkach po 1 kilobajt – nastąpiła defragmentacja w warstwie aplikacji.

Kolejnym ważnym elementem jest wyjaśnienie pojęcia sieci lokalnych. Stanowi to strukturę komunikacyjną łączącą urządzenia na określonym obszarze. Warto zauważyć, że aspekt fizycznego obszaru jest istotny, ale nie najważniejszy. Z reguły, w sensie fizycznego obszaru, sieć lokalna dotyczy ograniczonego terenu. W sensie komunikacyjnym, sieć lokalna stanowi zbiór urządzeń elektronicznych należących do tej samej sieci. Przykładową taką sieć pokazuje rysunek 1, gdzie występuje kilka urządzeń (komputerów), urządzenie określane jako switch oraz urządzenie określane jako router.

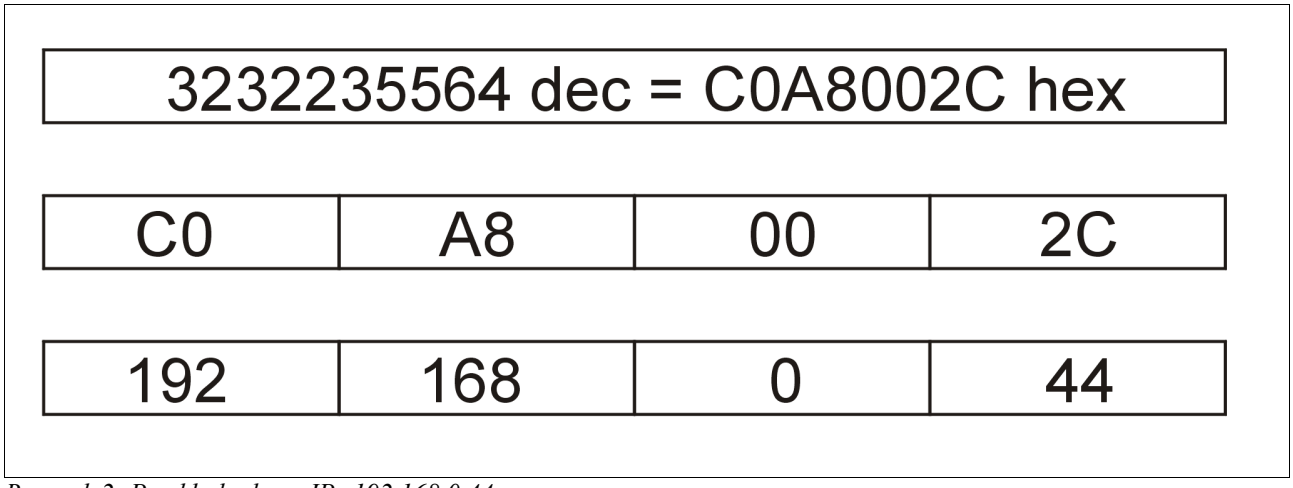


Rysunek 1: Sieć lokalna

Switch (dawniej hub) jest urządzeniem, którego zadaniem jest umożliwienie kilku urządzeniom dostęp do wspólnej magistrali komunikacyjnej. W dawnych czasach komputery były łączone jednym przewodem koncentrycznym. Z punktu widzenia elektrycznego odpowiada to połączeniu równoległemu. Rozwiązanie to miało wiele wad, gdzie najważniejszą była wysoka awaryjność. Obecnie powszechnym rozwiązaniem jest stosowanie odpowiedniej skrętki. Cechą charakterystyczną tego rozwiązania jest brak możliwości łączenia równoległego przewodów. Implikuje to stworzenie odpowiedniego urządzenia (switch lub hub), którego zadaniem jest swoiste zwielokrotnienie strumienia przesyłanych danych. Każdy switch (hub) obsługuje pewną liczbę wejść/wyjść. Dane odebrane z jednego wejścia (wysłane przykładowo przez komputer o adresie 192.158.0.67 na rysunku 1) są przesłane do wszystkich innych urządzeń przyłączonych do switcha (huba). Efektem tego, z punktu widzenia komunikacyjnego, jest to, że dane wysłane z jednej

stacji są odbierane przez każdą inną stację przyłączoną do switch'a (huba). Hub, jako pewnego rodzaju koncentrator, działa w pierwszej warstwie modelu OSI (warstwie fizycznej). Oznacza to, że przenosi sygnały z jednego wejścia na każde inne. Z kolei switch jest urządzeniem, które w swoim działaniu uwzględnia pierwszą warstwę modelu OSI. Na podstawie istniejącego ruchu jest w stanie dokonywać optymalizacji przesyłania ramek ethernetowych (przekazuje ramki wyłącznie do docelowego segmentu sieci bez niepotrzebnego powielania danych na każde inne wyjście).

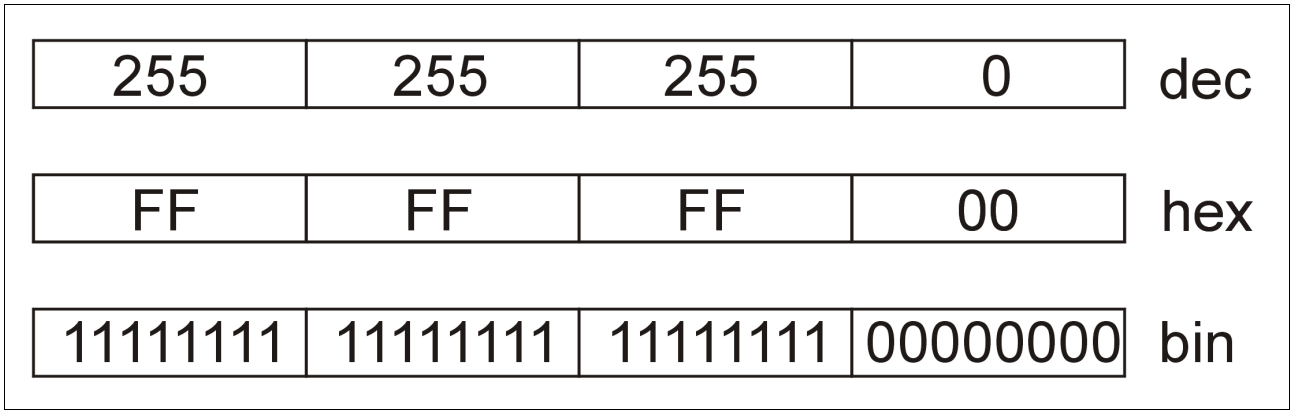
Często zdarza się, że lokalna sieć wymaga konieczności przesyłania danych poza własną sieć lokalną. By stworzyć taką możliwość, w sieci lokalnej musi znaleźć się urządzenie określane jako router. Jego zadaniem jest umożliwienie przesyłania danych do innej sieci. Z punktu widzenia sieci lokalnej, router musi być jej elementem, być osiągalnym z punktu widzenia sieci lokalnej. Oprócz tego by być fizycznie przyłączonym do wspólnej struktury komunikacyjnej, router musi mieć adres IP należący do puli adresowej sieci lokalnej. Wspomniany tutaj adres IP jest 32-bitowym identyfikatorem urządzenia. Właściwie należy napisać, że adres IP dotyczy interfejsu sieciowego, gdyż nie stanowi wielkiego problemu zbudowanie urządzenia, które miałoby dwa niezależne wejścia/wyjścia sieciowe. Jednak ograniczając się do budowy prostych urządzeń, które będą miały jeden interfejs sieciowy, można postawić znak równości pomiędzy urządzeniem a interfejsem sieciowym. Wspomniany adres IP w popularnej notacji jest zapisywany w postaci czterech liczb z zakresu od 0 do 255 rozdzielonych znakiem kropki. Rzutując liczbę 32-bitową na strukturę 4-bajtową i zapisując zawartość każdego bajtu w postaci dziesiętnej (rozdzielając je znakiem kropki) uzyskuje się powszechnie stosowaną notację adresów IP. Z punktu widzenia percepcji ludzkiej, łatwiej jest zapamiętać 4 niewielkie liczby niż jedną bardzo dużą, toteż przyczyn popularności wspomnianej notacji należałoby poszukiwać w aspektach psychologicznych (koncepcję tą ilustruje rysunek 2).



Rysunek 2: Przykład adresu IP=192.168.0.44

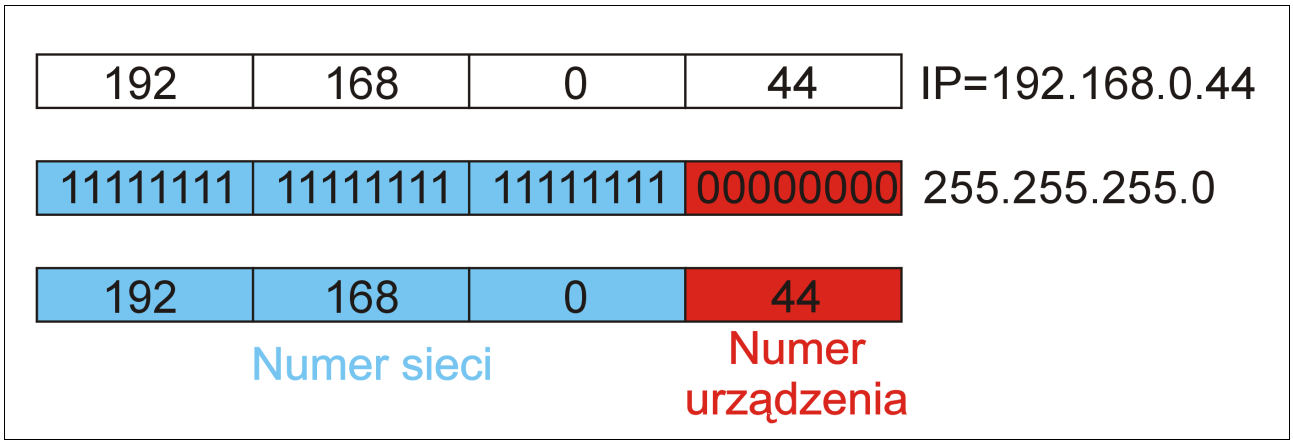
Każdy adres IP urządzenia przyłączonego do lokalnej sieci komputerowej musi być unikalny (co raczej jest zrozumiałe i nie wymaga szczegółowego

wyjaśnienia) oraz należeć do odpowiedniej puli adresowej. W rzeczywistości, adres IP niesie w sobie dwie informacje: adres (numer) sieci oraz adres (numer) urządzenia w tej sieci. Łącznie oba te numery tworzą liczbę zapisywaną na 32 bitach. Jej podział na dwie części nie jest określony w sposób sztywny, toteż do opisu sieci lokalnej konieczne jest dodanie kolejnej informacji, jaką jest maska podsieci. Jej znaczenie jest ściśle związane z adresem IP, toteż zwyczajowo ma ona identyczną konwencję zapisu. W większości sieci lokalnych maski podsieci wynosi 255.255.255.0. Rozpatrując ten adres w postaci binarnej uzyskuje się informację, która dzieli typowy adres IP na część identyfikującą numer sieci i część identyfikującą urządzenie w obrębie tej sieci (ilustruje to rysunek 3).



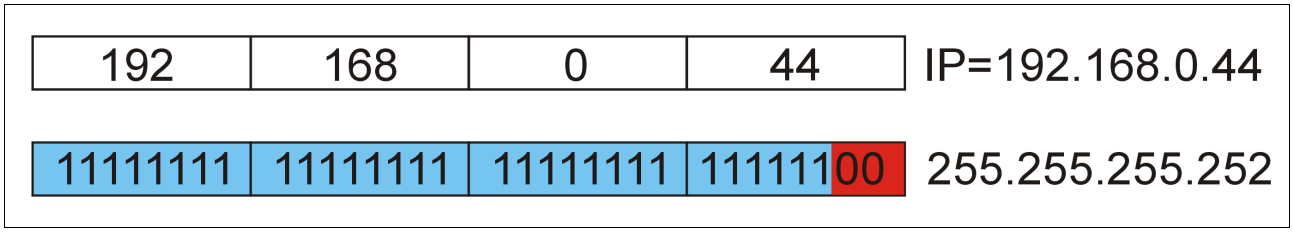
Rysunek 3: Maska podsieci 255.255.255.0

Uwzględniając maskę podsieci w stosunku do adresu IP urządzenia, wyodrębniany jest numer sieci oraz numer urządzenia w obrębie sieci. Część adresu IP urządzenia wydzielona przez bity o wartości 1 w masce podsieci identyfikuje numer sieci. Pozostała część adresu określa urządzenie w obrębie tej sieci. Rozumiejąc znaczenie maski podsieci, której zadaniem jest wyodrębnienie numeru sieci z adresu IP urządzenia, definicję sieci lokalnej należy uważać jako zbiór urządzeń przyłączonych do wspólnej struktury komunikacyjnej (jako okablowanie wraz z urządzeniami typu switch lub hub) mających identyczny numer sieci (identyczny ciąg bitów wynikający z nałożenia maski podsieci na adres IP urządzenia).



Rysunek 4: Podział adresu IP

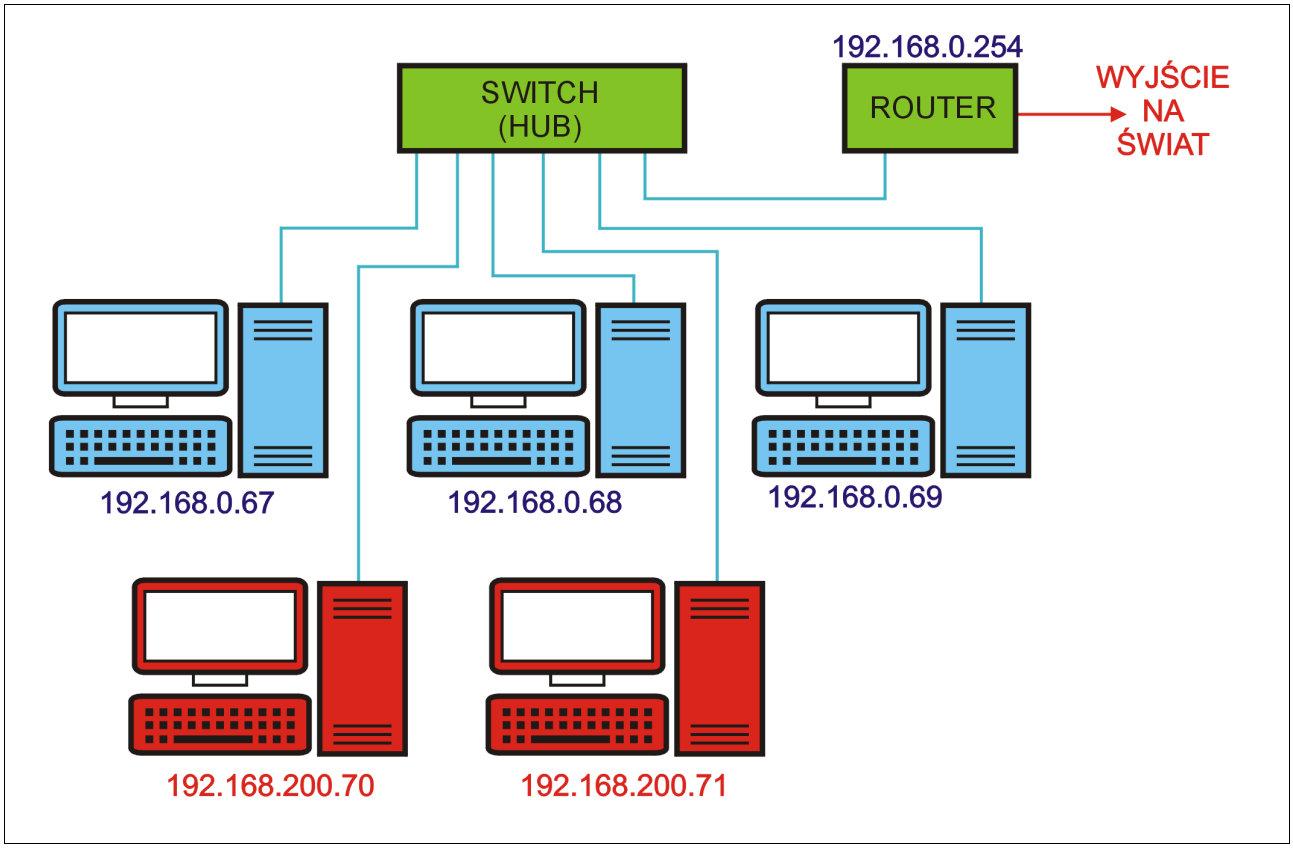
W przypadku maski podsieci zaprezentowanej na rysunku 4, wszystkie urządzenia o adresie IP=192.168.0.x (gdzie x jest dowolną liczbą z zakresu od 0 do 255) należą do tej samej sieci lokalnej. Z powyższego wynika istotna informacja: w przypadku sieci, dla których maska ma wartość jak w przykładzie, do tej samej sieci może należeć maksymalnie 256 urządzeń, gdyż tyle różnych kombinacji można zapisać na 8 bitach (w masce podsieci jest 8 bitów mających wartość 0). Generalnie nie ma obowiązku, by maska podsieci zawierała 0 ma ostatniej, najmłodszej pozycji (chodzi o 0 w zapisie 255.255.255.0). Przykładowo maska podsieci 255.255.255.252 opisuje sieć lokalną, do której może należeć jedynie 4 urządzenia (rysunek 5).



Rysunek 5: Sieć lokalna z maską 255.255.255.252

W przypadku, gdy urządzenie ma adres IP=192.168.0.44 z maską podsieci 255.255.255.252, sieć lokalną mogą stworzyć jedynie urządzenia o następujących adresach IP: 192,168.0.44, 192.168.0.45, 192.168.0.46 oraz 192.168.0.47 (tylko takie mają różne kombinacje na dwóch najmłodszych bitach).

Do dokładnego zrozumienia istoty sieci lokalnych warto rozpatrzeć jeszcze następujący przykład pokazany na rysunku 6, gdzie maska podsieci ma typową wartość (255.255.255.0).



Rysunek 6: Dwie sieci lokalne

Pomimo, że okablowanie sieci jest spięte jednym urządzeniem switch, rysunek przedstawia dwie niezależne sieci lokalne (jedna o adresacji 192.168.0.x, druga o adresacji 192.168.200.y). Dla pierwszej sieci (192.168.0.x) istnieje możliwość „wyjścia na świat", gdyż router, który to umożliwia należy do tej sieci lokalnej. Druga sieć lokalna (192.168.200.y) składa się z dwóch elementów i komunikacja może zachodzić jedynie między tymi dwoma komputerami. W tej sieci nie występuje żadne urządzenia umożliwiające przekierowanie danych z sieci 192.168.200.y do sieci 192.168.0.x (zakładając, że każdy komputer nie realizuje funkcjonalności związanej z przekierowaniem danych do innych sieci). Tu warto zauważyć, że zmiana maski podsieci na inną wartość (255.255.0.0) w każdym urządzeniu sieciowym (komputerze i routerze) spowoduje, że wszystkie komputery z rysunku 6 stworzą jedną sieć lokalną o adresie 192.168.x.y. Przykład ten pokazuje, że idea sieci lokalnej ma charakter bardziej logiczny niż fizyczny. Pewne szczegóły konfiguracyjne sieci (adresacja IP i maska podsieci) umożliwiają lub nie na przepływ informacji od jednego urządzenia do drugiego (zakładając istnienie „drożności" przewodów łączących). Jest oczywistym, że fizyczne rozdzielenie urządzeń uniemożliwi przesyłanie czegokolwiek pomiędzy nimi (choćby poprzez niezależne okablowanie), gdyż stworzy to niezależne sieci lokalne. Ta cecha daje ogromną elastyczność, gdyż w milionach domów egzystują maleńkie sieci komputerowe o adresie 192.168.0.x (z maską podsieci 255.255.255.0), które nie wpływają wzajemnie na siebie.