

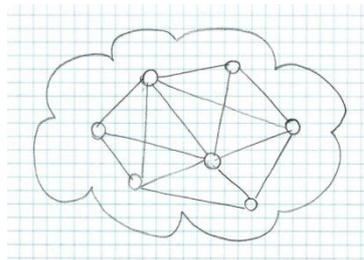
La rete internet

Un protocollo deve avere caratteristiche costanti per poter mettere in comunicazione dispositivi diversi. Esiste un'azienda, l'IEEE, che emette i parametri standard che ciascun protocollo deve possedere. Ogni aspetto delle connessioni viene precisato: dalle tensioni allo spessore dei fili, dai colori delle guaine alle velocità, ecc. Alcuni esempi sono lo standard 802.3, che regola le connessioni LAN e lo standard 802.11, che regola le connessioni wi-fi.

Finora abbiamo parlato di connessioni locali, cioè tra computer che si trovano a breve distanza tra loro. Se invece dobbiamo connettere computer molto distanti? Dobbiamo usare la linea telefonica. La complessità aumenta poiché è necessario mettere in comunicazione non solo singoli dispositivi lontani, ma intere reti LAN distanti anche migliaia di chilometri. Tuttavia la complessità non è data dalla distanza. Gli impulsi elettrici che trasportano le informazioni si muovono infatti a metà della velocità della luce, quindi possono percorrere in un secondo 150000 chilometri. La difficoltà da superare è il fatto che la grande quantità di computer connessi crea un'enormità di strade percorribili differenti. Le strade si intersecano formando numerosi incroci e le informazioni rischiano di perdersi se non vengono instradate nel modo corretto. Ogni incrocio si chiama ROUTER e ha proprio il compito di instradare correttamente le informazioni.

Ciascuna rete LAN si connette alla più grande rete (web) mediante un Access Point. Le odierne chiavette internet e gli smartphone hanno invece connessioni "mobili" BTS (Base Station Controller).

Ecco lo schema "a nuvola" della rete internet:



Mentre un tempo l'attività principale svolta su internet era scaricare informazioni, oggi gli utenti possono, grazie ai social network e a siti come wikipedia, caricare grandi quantità di dati.

Ora, come accade in una rete LAN, anche in una rete non locale è necessario identificare il mittente e il destinatario delle informazioni. Ci chiediamo se sia possibile allora utilizzare gli indirizzi MAC delle reti LAN su quelle mondiali. La risposta è no. Gli indirizzi MAC sono associati a postazioni fisse dei computer, come accade in un'aula di informatica. I byte che percorrono una rete mondiale provengono da punti che cambiano continuamente e che addirittura possono essere mobili sul territorio. Gli indirizzi con cui i dispositivi vengono identificati in internet sono associati alla posizione mobile degli stessi. Tali indirizzi sono detti IP (Internet Protocol address) e sono costituiti da quattro byte. Quattro byte equivalgono a $2^{32} = 4 \times 10^9$ indirizzi differenti. La rete internet presenta un numero notevolmente maggiore di dispositivi connessi rispetto ad una LAN. Allora perché, è lecito chiedersi, il numero di indirizzi componibili è inferiore rispetto a quelli di una rete LAN? Perché internet è nato prima rispetto alle reti locali e quattro miliardi di indirizzi sembrava un numero enorme. Infatti si è manifestata molto presto la necessità di trasmettere dati da una parte all'altra del mondo.

Quanti indirizzi IP possiede un dispositivo? Uno per ogni sua interfaccia fisica, cioè per ciascuna modalità con cui può navigare in internet. Ad esempio uno smartphone ha un indirizzo IP per la connessione UMTS, uno per quella Wi-Fi e uno per il Bluetooth.

Ecco un esempio di indirizzo MAC costituito da sei byte:

00 : 12 : AB : 01 : FC : 3A

Eccone uno IP:

192.168.1.2

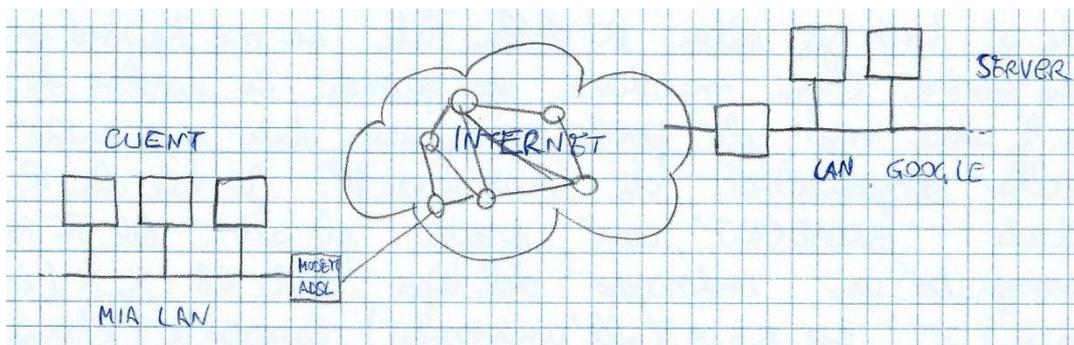
Come è nato Internet? Facciamo un breve excursus nella storia della più importante rete globale esistente. Le radici di Internet sono da ritrovarsi nella rete ARPANET della difesa americana. Il progetto era nato nel 1960. Successivamente l'università di Berkeley fondò una propria rete sulla scia di ARPANET. Tale seconda rete prese il nome di ARPA e constava di 2^{32} indirizzi IP. Venne messa in rete una macchina, il server web, che metteva a disposizione grandi quantità di informazioni. Chi si collegava a tale rete (inizialmente, le altre università) poteva scaricare i contenuti del server web. A partire da questo momento la rete inizia ad allargarsi sempre di più e, dopo vari rinnovamenti e ridenominazioni diviene l'odierna rete Internet. Inizialmente la connessione a Internet non era cosa semplice come è oggi grazie all'interfaccia a finestre. Era necessario conoscere almeno le basi della programmazione per poter sfruttare il web.

Lo standard IPV4 è quello che gestisce gli indirizzi IP formati da quattro byte, quello IPV6 gestisce gli indirizzi formati da sei byte, anche se è ancora poco diffuso. Dato che gli indirizzi IP sono dinamici, cioè cambiano ogni volta che ci si connette alla rete, è necessario che esista un distributore di indirizzi che li fornisca all'occorrenza. Il distributore si chiama provider. Esso compra un certo numero di indirizzi IP in una certa area geografica e poi li rivende ai singoli utenti. Quando ci connettiamo a internet il provider ci dà un indirizzo IP, quando disconnettiamo il nostro indirizzo viene assegnato a un altro utente. Se vogliamo che il nostro indirizzo IP sia sempre lo stesso, dobbiamo acquistarne uno statico. Ciò accade per quei siti che devono sempre essere univocamente rintracciabili, come i server web. In tal modo possiamo trovarli sia scrivendo i numeri che compongono il loro indirizzo IP sia inserendo, ad esempio, "www.google.it". Le lettere sono più usate poiché più vicine all'ambito umano.

La comunicazione tra reti locali e reti globali

Occupiamoci di come avviene il dialogo tra i due tipi di reti, locali e globali. In particolare parleremo, come esempio di rete globale, della rete Internet. Quando infatti ci colleghiamo a internet, la rete locale dalla quale ci colleghiamo deve dialogare con la rete globale.

Vediamo anzitutto una schematizzazione:



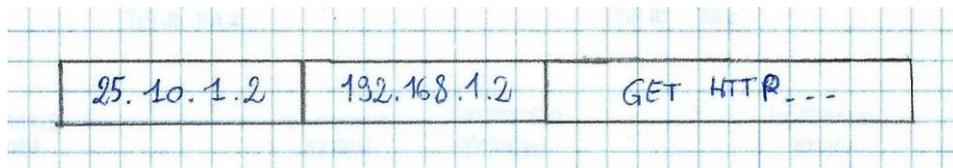
Genericamente, il client è la macchina che usufruisce del servizio fornito dal server. Ciò che permette di connettersi a internet è il modem ADSL. “Prima” del modem troviamo il computer, o, più in generale, la rete LAN da connettere. “Dopo” si trova la rete internet, quella a cui ci si connette. Il problema che sorge immediatamente è il fatto che la rete LAN funziona con indirizzi MAC, mentre la rete internet funziona con gli indirizzi IP. Gli indirizzi MAC hanno valori compresi tra 00:00:00:00:00:00 e FF:FF:FF:FF:FF:FF, quelli IP tra 0.0.0.0 e 255.255.255.255.

Ipotizziamo di voler aprire il motore di ricerca Google. Nella barra degli indirizzi di Internet Explorer o di Google Chrome andiamo a scrivere “www.google.it”. Premendo il tasto invio si apre la pagina richiesta. Ciò significa che il computer su cui stiamo lavorando si è collegato al server di Google. Come è avvenuto ciò? Il nostro computer, per quanto ne sappiamo, appartenendo a una rete LAN possiede solo un indirizzo MAC che non è valido in ambito globale. Nel paragrafo precedente abbiamo scoperto che connettendosi a internet il computer ottiene un indirizzo IP. Sarà dunque quest’ultimo a permettere il collegamento col server di Google, che a sua volta possiede un indirizzo IP. Tuttavia il nostro computer non viene connesso direttamente alla rete. Prima i dati che vuole inviare devono raggiungere il modem, cioè percorrere un tratto della rete LAN. In LAN vengono usati gli indirizzi MAC, non quelli IP. Ma procediamo con ordine.

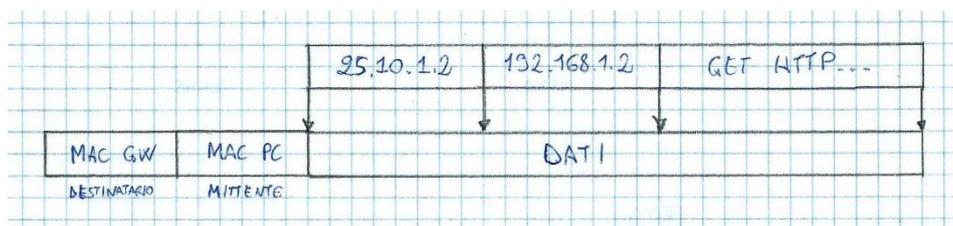
Immaginiamo che il server di Google abbia indirizzo statico 25.10.1.2. Per poter dialogare con esso il nostro computer deve possedere un indirizzo IP. Esso è però di tipo dinamico, non statico. Quando si connette a internet il computer chiede a tutti i dispositivi che appartengono alla rete LAN alla quale è connesso se c’è qualcuno in grado di fornirgli un indirizzo IP con cui navigare. L’azione descritta ha la caratteristica di essere rivolta a tutti i dispositivi. Tale caratteristica è definita “broadcast”. Il computer riceve risposta dal modem, che effettivamente è l’unico in grado di fornire indirizzi IP. Esso ne chiede uno dinamico al provider e lo assegna al computer che lo ha richiesto. Per assegnarlo il modem segue il protocollo DHCP (Dynamic Host Configuration Protocol) della LAN cui appartiene. Per poter interagire col provider che si trova in rete anche il modem ha bisogno di un indirizzo che lo identifichi in rete: ipotizziamo che esso sia 72.10.1.5. Per poter accedere a internet, le informazioni devono inoltre “passare” attraverso un Gateway. Pertanto il computer deve possedere l’indirizzo dello stesso. Lo richiede di nuovo broadcast. Infine il computer ha bisogno di conoscere l’indirizzo IP al quale deve inviare la propria richiesta. Nel nostro caso volevamo metterci in contatto con Google e avevamo supposto che il suo indirizzo IP fosse 25.10.1.2. Il DNS (Domain Name System) trasforma il nome di Google che abbiamo digitato nel dominio 25.10.1.2.

Tuttavia non abbiamo ancora risolto il nostro problema: far sì che le informazioni possano attraversare quel breve tratto di rete LAN per raggiungere il modem. Il pacchetto che circola nella

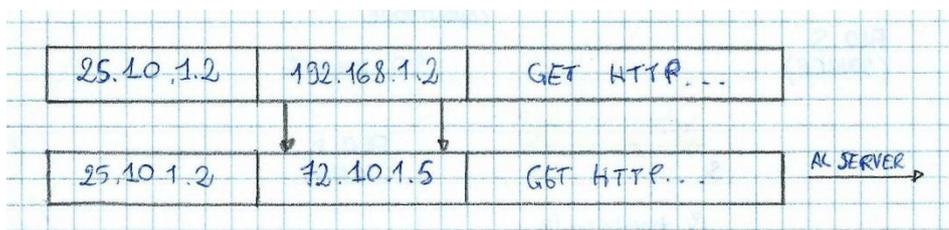
rete internet contiene indirizzo del destinatario (il server di Google), indirizzo del mittente (il nostro computer) e informazioni:



Affinché tale pacchetto possa attraversare la LAN viene inserito in un pacchetto ethernet. Esso diviene l'informazione del pacchetto ethernet, mentre l'indirizzo del destinatario è quello MAC del Gateway e l'indirizzo del mittente è quello MAC del computer:



Il modem recupera l'informazione (che è il pacchetto che era stato preparato dal computer per navigare in rete). Tuttavia essa non viene inviata direttamente. Prima il modem sostituisce il proprio indirizzo IP (nel nostro esempio, 72.10.1.5) a quello del computer (192.168.1.2) e poi immette l'informazione in rete inviandola finalmente al server. Graficamente:



Il server riceve la richiesta e risponde. Per la risposta si ha il processo inverso: l'informazione viaggia tra indirizzi IP nella rete e raggiunge il modem. Il modem la inserisce nella LAN per farla giungere al computer che ha inviato la richiesta. Si pone a questo punto un problema: come fa il modem a ricordare quale fosse il computer che aveva inviato la richiesta? Ancora, se su quel computer sono aperte più finestre del browser, come riesce il modem a inviare la risposta alla giusta finestra? La risposta è semplice: all'indirizzo IP vengono associati due byte che identificano le porte logiche attraverso cui l'informazione deve passare per giungere alla destinazione. Possiamo identificare allora 2^{16} informazioni diverse. L'indirizzo IP unito ai byte identificativi forma il cosiddetto SOCKET. Esso identifica l'applicativo che accede a internet e la porta logica ad esso associata. La porta 80 è quella predefinita per il server web.

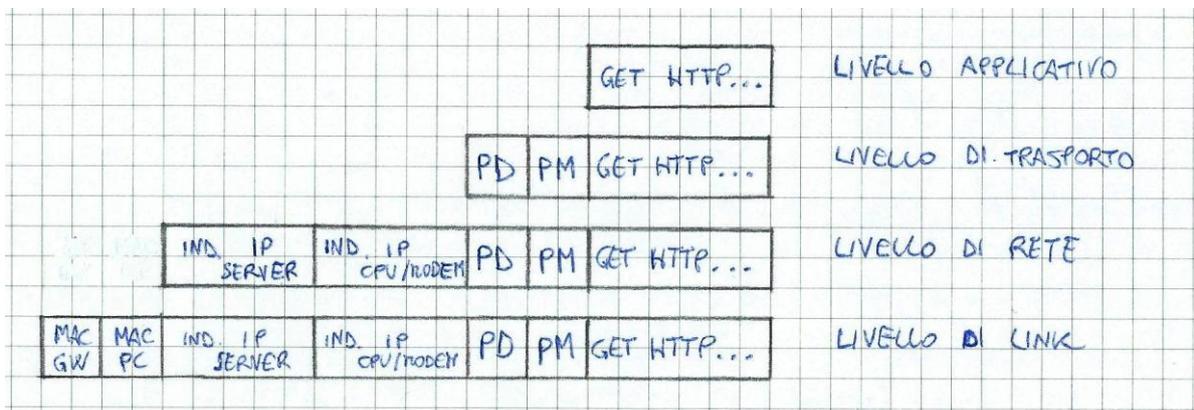
Nel corso della descrizione del processo di trasmissione dati (e nelle immagini ad esso relative) abbiamo genericamente inserito "GET HTTP..." a simboleggiare l'informazione richiesta. In realtà il fatto che debbano essere identificate le porte logiche porta alla necessità di inserire le stesse

subito prima dell'informazione. Troviamo quindi, in ordine da sinistra a destra: indirizzo IP del server, indirizzo di CPU o del Modem, porta destinatario della risposta, porta mittente della risposta e informazione:



Livelli di comunicazione

Il formato http che abbiamo menzionato definisce un protocollo che effettua gli scambi tra PC e server. In tal modo si possono caricare o scaricare dati. Il formato html invece esegue sul computer quanto deriva dal formato http. Esistono poi formati di trasporto dei dati, di rete e di riconoscimento dei link. A seconda dell'ambito cui sono adibiti i formati si definiscono diversi "livelli":

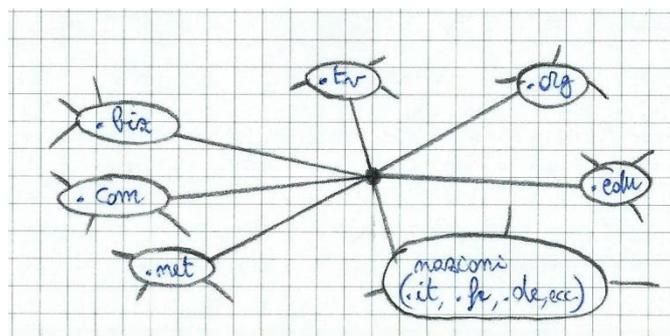


I domini

Ogni server viene raggiunto grazie al dominio che possiede. La rete internet è dunque "suddivisa" in gruppi cui appartengono i diversi siti che visitiamo. Inizialmente i siti vennero suddivisi in base all'ambito di appartenenza: i domini erano .org, .edn, .com, .net, ecc. In seguito i siti divennero talmente tanti da essere più conveniente una suddivisione in base alla Nazione di appartenenza. Nacquero così .it, .fr, .de, ecc.

Se anziché scrivere www.google.it scriviamo www.google.it., mettendo cioè un punto dopo it, la pagina iniziale di google viene ugualmente caricata. Il disegno seguente consente di comprendere meglio il senso della suddivisione in domini. Il punto rappresenta la totalità della rete internet e da questo i diramano i vari domini come .com, .net, ecc.

A loro volta nascono le diramazioni da ogni dominio:



A questo punto possiamo chiederci: che differenza c'è tra indirizzi IP e domini? Sono due diverse sfaccettature del medesimo concetto. Per l'utente che usa il computer è comodo ricordare un sito con un nome espresso in lettere. A quel nome in lettere è associato un indirizzo espresso con cifre. Le cifre vengono utilizzate dal mondo dell'elettronica. Il nome in lettere del sito è detto DNS, "Domain Name Server".

Esiste a livello mondiale un gruppo di 13 server che rimangono costantemente accesi e conservano tutti i DNS esistenti con gli indirizzi IP ad essi associati. L'azienda che possiede i computer si chiama IANA. In ogni Stato esistono poi due computer che raccolgono tutti i DNS con dominio di quello Stato. In tal modo se uno dei due computer non funziona l'altro rimane in servizio.

Da ogni dominio possiamo diramare dei sottodomini. Da .it discendono tutti i siti con dominio .it. Uno di questi è il sito di unige: www.unige.it. Da unige.it discendono i siti dei vari Dipartimenti: dist.unige.it, dicca.unige.it, ecc. Dal dominio di ciascun Dipartimento discendono i quelli dei Professori che vi appartengono: un esempio è giordano.dist.unige.it.

Quando vogliamo aprire un sito dobbiamo quindi affittare un dominio. Inserito il nome che vogliamo assegnare scegliamo il dominio: .it, .com, ecc. Poi possiamo creare sottodomini associati al nostro DNS. Ad esempio a giorda.net sono sottoposte la telecamera larotonda.giorda.net e l'indirizzo e-mail al@giorda.net.

Chi ci affitta il dominio si chiama provider. Ci viene fornito anche un pannello di controllo che ci permette di controllare il nostro dominio e tutti i sottodomini ad esso associati.

Ora ci chiediamo: un sito come www.giorda.net possiede indirizzo IP statico o dinamico? Ha un indirizzo dinamico. Come si fa a ritrovare il sito quando l'indirizzo cambia? Il provider a cui il sito appartiene, dyndns, possiede un servizio che invia un messaggio di cambiamento dell'IP e lo aggiorna.

Pertanto alcuni server hanno IP statico poiché sono l'impalcatura della rete internet. Altri hanno IP dinamico. Un server può ospitare più siti: si parla in tal caso di multi-hosting.

La ricerca degli errori di trasmissione

In una rete, locale o globale, vengono trasmessi dati. È possibile quindi che si abbiano degli errori nella trasmissione. In alcuni casi essi non minano totalmente la qualità dei dati trasmessi. Può accadere ad esempio che in una voce registrata non si capisca una parola o in un filmato un'immagine in una lunga sequenza non si veda bene. Ancora, può succedere che in un testo un carattere non si legga.

Altri errori invece devono essere corretti poiché la qualità del dato trasmesso è vitale perché lo stesso possa essere utilizzato in seguito. Pensiamo ad esempio a un numero di telefono: se una delle cifre è sbagliata tutto il numero è inutilizzabile.

Si può allora correggere l'errore manualmente o richiedere un nuovo invio del documento al mittente. Ciò è però possibile solo se non lavoro in tempo reale. Negli altri casi esistono algoritmi per il riconoscimento e l'automatica correzione degli errori.

Partiamo dal caso di trasmissione di dati su cui non lavorare in tempo reale (come ad esempio una e-mail). La trasmissione, ossia il livello di trasporto, è gestito, come abbiamo visto, dai protocolli TCP e UDP. Il TCP si accerta che il destinatario riceva. La trasmissione si svolge in tre fasi: si

chiede se il destinatario è pronto a ricevere, il destinatario risponde, il mittente manda. Finita la trasmissione del primo “pezzo” di dati, il destinatario dà automaticamente l’ok e il mittente invia un altro pezzo. Se c’è un errore si può inviare nuovamente il “pezzo” danneggiato.

Durante una telefonata? I dati di una telefonata vengono utilizzati in tempo reale. Non è possibile inviare nuovamente un suono che si è sentito male. Si sfrutta allora la ridondanza e si correggono in modo probabilistico gli errori con il meccanismo FEC. In una serie di bit inviati, supponiamone sette, si sceglie di considerare solo il primo, il quarto e il settimo. I bit adiacenti a questi contengono le stesse informazioni: il secondo la stessa del primo, il terzo e il quinto la stessa del quarto, il sesto la stessa del settimo. Se arriva un bit in un’altra posizione, lo si sostituisce a quella più vicina che doveva essere considerata. Inoltre, come ulteriore controllo, per un certo numero di dati trasmessi viene aggiunto un bit di controllo. Si rende pari il numero di bit inviati. In tal modo se si riceve un numero dispari di bit inviati, in particolare se ne manca uno, sono sicuro che ci sia un errore. Se invece mancano due bit non riesco ad accorgermi dell’errore. In alternativa si può fare la somma dei bit e vedere se ha lo stesso valore per mittente e destinatario.

Gli indirizzi IP e la maschera di rete

Esistono vari tipi di indirizzi IP. La maggior parte di quelli oggi esistenti sono di tipo IPV4. Si fa uso di 32 bit per gestire 2^{32} indirizzi. Perciò sono disponibili 7 indirizzi per km^2 .

Ci sono poi gli indirizzi IPV6. Si fa uso di 128 bit per avere 2^{128} indirizzi. Si hanno $6,7 \times 10^{17}$ indirizzi per mm^2 . Nasce dunque il concetto di Internet delle cose, cioè la possibilità di connettere a Internet tutti gli oggetti della realtà per ottenere un’enorme rete di identificazione e localizzazione.

Quando ci troviamo in una rete locale è necessario che non ci sia confusione nell’invio delle informazioni: ogni computer deve avere un indirizzo univoco e si deve sapere se il computer cui si inviano i dati appartiene alla LAN o meno. Serve una maschera di rete, cioè un indirizzo che consenta di riconoscere l’appartenenza del destinatario alla LAN. Dobbiamo inoltre assegnare una serie di indirizzi.

Il primo è quello della rete che non può essere utilizzato da nessuna macchina. Immaginiamo sia 192.168.1.0. Poi ci sono l’indirizzo del router ADSL, 192.168.1.1, del nostro PC, 192.168.1.2, della stampante, 192.168.1.10, ecc. Diciamo ancora che l’indirizzo della maschera di rete è 255.255.255.0.

Se ho 255.255.255.0 posso connettere 255 computer alla LAN (più l’indirizzo di rete). Se ho 255.255.254.0 posso connetterne 511, se ho 255.255.0.0 posso connetterne 2^{16} compreso l’indirizzo di rete.

Facendo la somma logica tra indirizzo della maschera e quello del PC destinatario veniamo a sapere se esso appartiene o meno alla rete:

255.	255.	255.	0	
				AND
192.	168.	1.	X	

Cioè:

11111111.	11111111.	11111111.	00000000	
192.	168.	1.	X	AND
<hr/>				
192.	168.	1.	0	

Il risultato è proprio l'indirizzo di rete: il PC appartiene alla rete.