

MSA - 17.03.2021

Come dicevamo ieri, chiudiamo il capitolo sulla presentazione del protocollo 802.11. Abbiamo acquistato una visione abbastanza completa dell'architettura complessiva del protocollo, dal punto di vista statico che dal punto di vista dinamico della gestione dell'accesso. Ultimo capitolo quella che è una delle evoluzioni più recenti del protocollo. Si è introdotta pochissimi anni fa un'innovazione sostanziale, e la spinta è stata l'emergere di scenari applicativi per i quali l'architettura che abbiamo finito di descrivere risultava non più adeguata. Si rifanno tutti e due in un certo senso al tema dell'IoT, emerso come scenario non più futuristico ma che del già presente. Un primo scenario è quello in cui ho una grande quantità di sensori distribuiti in un'area anche vasta. Ogni singolo sensore, collettore di dati che potrebbe essere un oggetto in grado di eseguire azioni, trasmette pochi dati, ogni entità in gioco ha tassi di trasmissione nell'ordine dei kb/s. Si tratta quasi sempre di entità stazionarie e pochi mobili, presenti in grande numero, nell'ordine delle migliaia o decine di migliaia, rispetto ad un singolo AP che colleziona le informazioni trasmesse da questi sensori. Molto spesso le entità in gioco sono alimentate da batteria. Un altro scenario, un raffinamento del precedente, si prefigura che anche per ragioni di risparmio energetico, i sensori dispersi nell'ambiente comunicano usando un protocollo di più basso impatto energetico come 802.15 (bluetooth) con una portata e banda di trasmissione minore. Se i nodi ultimi utilizzano un protocollo di questo genere per comunicare, i dati che trasmettono non vanno lontano e bisogna pensare a dei gateway che collezionano queste informazioni e le inoltrano, usando un protocollo tipo 802.11, verso un AP centrale che fa da punto di collezione. Il numero di stazioni in gioco che raccolgono i dati può essere un ordine di grandezza inferiore, siamo in ambienti esterni.

Cosa possiamo dire di questi scenari nella prospettiva dell'utilizzo del protocollo 802.11? Sono scenari, questi due più un altro paio, che hanno delle caratteristiche in comune. Un grande numero di nodi interessati a trasmettere, il fatto che i nodi possono essere dispersi su uno spazio geografico anche ampio, il fatto che ogni singolo pacchetto trasmesso è mediamente di piccole dimensioni, e molto spesso si tratta di stazioni difficilmente mantenibili. In uno scenario caratterizzato da questi parametri, come si trova un protocollo come quello che abbiamo descritto, che lavora su una banda dai 2.4 ai 5 ghz se non oltre (60 ghz). Ci sono delle caratteristiche cozzano con i punti caratterizzanti dei scenari descritti. Sono bande che hanno problemi di propagazione nel senso che il segnale si attenua rapidamente o penetrazione di ostacoli fisici. Il protocollo richiede un header di dimensioni non trascurabili. La configurazione attuale non consente la presenza di un numero superiore a 2007 di stazioni che si connettono allo stesso AP. Il meccanismo di risparmio energetico non è ottimale in certe situazioni. Questo perché il protocollo descritto è stato pensato per uno scenario con un alto throughput richiesto, il numero delle stazioni che si connettono ad uno stesso AP è limitato, le distanze da percorrere sono tipicamente brevi pensando a scenari indoor. Se lo scenario di riferimento non è più questo, un protocollo pensato per questo scenario non va bene per una situazione differente.

Il gruppo di lavoro 802.11, visti i nuovi scenari, si è posto l'obiettivo di introdurre una nuova versione del protocollo, nota con il suffisso h, che utilizza bande trasmissive sotto al Ghz, andando a pescare bande non coperte da licenza, ed oltre questo altri obiettivi, dal punto di vista logico si vuole consentire un maggiore numero di nodi connessi oltre la soglia magica di 2007, la possibilità di utilizzare header e segnali di beacon più piccoli per risparmiare banda ed energia, più altri obiettivi. Nelle slide che seguono ho fatto una selezione dei punti su cui il comitato di gestione del protocollo è intervenuto. La scelta è arbitraria.

Intanto, primo punto, l'aumento del numero di stazioni connettabili ad un certo AP. Questo numero nella versione descritto finora, cosa che non vi ho mai detto, visto che questo 2007 non emergeva. Si annida questo limite in quello che è stato scelto come formato di un oggetto, la mappa TIM che viene utilizzata quando viene attuata la politica di risparmio energetico, utilizzata per indicare quali stazioni devono rimanere accese. Tra le varie decisioni, c'è stata quella sul formato che è quello riportato. Un certo numero di byte, dove i primi 5 hanno il ruolo indicato. Il secondo serve a specificare la lunghezza di ciò che segue. La dimensione della mappa dipende dal numero di stazioni da informare se ci sono dati per loro. La mappa può avere dimensione da 1 fino a 251 byte. Il campo lunghezza ha dimensione 1 byte, che serve a dire da quanti byte è formato ciò che viene dopo questo campo, incluso lui stesso. Contando lui, abbiamo 4 byte, ne rimangono 251 al massimo. L'uso pratico della bitmap è questo: ad ogni stazione viene assegnata una posizione all'interno della bitmap: se quel bit sta a 0 o 1, indica se ci sono o no dati in attesa di essere trasmessi. Ci possono essere massimo 251 byte. 8×251 fa 2008. Il bit 0 non viene utilizzato, quindi rimangono 2007 posizioni occupabili. Il rimedio è altrettanto banale. Si è cambiato formato. Il nuovo formato della TIM ha aumentato la dimensione del campo lunghezza, da uno diventato due byte, ed ecco che la dimensione della bitmap è cresciuta fino a 1018 byte, riuscendo a codificare poco meno di 10000 stazioni. Questo in un uso banale, usando una strutturazione piatta della bitmap. In schemi più sofisticati è possibile estendere il numero di stazioni rappresentabili. Questo ha richiesto anche la modifica sul modo di interpretare nel frame che circola nelle comunicazioni gestite dal protocollo, l'interpretazione dei secondi due byte. I due byte hanno tipicamente il ruolo di campo durata, i byte che informano se devono attivare il meccanismo NAV di cui abbiamo parlato ieri. C'è però uno scenario particolare, quello della richiesta di associazione di un nodo al BSS gestito da un particolare AP. In questo caso questi due byte vengono utilizzati per assegnare alla stazione il suo identificativo all'interno del BSS. Nella versione prima della versione h, l'utilizzo di questi due byte è riportato in questa tabella. Se il bit 15 ha valore 0 va interpretato come un campo durata. Se il bit numero 15 vale 1, l'interpretazione è diversa: in particolare se i due bit sono entrambi uno ed i bit da 0 a 13 non sono tutti 0, qui la scelta fatta è conseguenza della scelta illustrata nella slide precedente. I valori da 1 a 2007 servono ad assegnare un identificativo ad una particolare stazione. Gli altri valori sono riservati per usi di natura diversa, non specificati. Questo modo di utilizzare questi bit non è in linea con l'innovazione della TIM. Il modo di leggere questi due byte deve essere aggiornato. Si estende la gamma di valori possibili che possono essere utilizzati per associare un identificativo alle stazioni che fanno parte di un BSS. E' una cosa banale, sicuramente tutto ciò appare molto banale, ma richiede di definire una macchina a stati differente da realizzare ed implementare nella scheda di rete del protocollo.

Altra modifica introdotta riguarda la riduzione dei dati da fare circolare sulla rete. Se devo trasmettere 100 o meno byte, e devo agganciare al carico utile un header di 28 bytes nel caso di tre campi indirizzo (34 se devo usare anche il quarto), l'overhead diventa significativo. L'idea è quella di consentire l'uso di header di dimensioni più piccole. Questo nuovo header è incompatibile con il precedente, non è sovrapponibile. A causa di questo, nei primi due bit del Protocol version, rimasti prima sempre a 0, con l'ingresso della variante h i due bit assumono la configurazione 01 per segnalare che ciò che segue va interpretato in modo diverso. Quello che cambia è che si possono utilizzare due soli campi indirizzo, i campi tre e 4 possono anche occupare 0 bit. Il campo tre nelle versioni precedenti è necessario in quanto tra chi invia e riceve ci possono essere quattro ruoli. Tipicamente negli scenari descritti, di questi ruoli due coincidono. E' necessario comunque avere tre campi indirizzo. In questo protocollo invece i campi indirizzo possono essere ridotti a 2 in quanto essendo in una situazione statica in cui la configurazione non cambia, c'è un preambolo iniziale in cui ci si conosce reciprocamente, si assegnano i ruoli del mittente e destinatario logico e fisico, ed una volta incamerata questa informazione cessa la necessità di ripetere l'informazione ridondante. Risparmio un campo indirizzo, ma uno dei due indirizzi, quello che individua il partner nella coppia di chi trasmette e chi riceve, questo nodo non AP può essere individuato con un campo indirizzi di dimensione ridotta, non 6 byte ma soltanto 2. Mettendo assieme tutte le possibili riduzioni la dimensione totale dell'header si riduce notevolmente. Altra cosa da rimarcare, può anche essere omissa il campo Durata. Questo perché il protocollo si dota di un meccanismo diverso, un po' più grossolano, per fare il carrier sensing virtuale.

Altro obiettivo del protocollo creato dal grande numero di stazioni, quindi grande probabilità di collisioni. Allora, questa variante del protocollo introduce due meccanismi per ridurre la probabilità di collisione ed aumentare il coefficiente di utilizzazione. Prima modalità è questa: l'idea è quella di partizionare tutte le stazioni in gruppi distinti, ed in parallelo anche il canale da un punto di vista temporale, viene suddiviso in slot consecutivi, con una corrispondenza 1-1 tra gruppi e slot. Ogni stazione che fa parte di un gruppo è autorizzata a provare ad accedere al canale solo nello slot di pertinenza per il gruppo, in cui l'accesso avviene in modalità DCF a contesa. L'informazione su come è la ripartizione in gruppi e la conseguente organizzazione dell'asse dei tempi viene trasmessa in broadcast tramite il segnale di beacon. In questo modo riduco il numero di stazioni che possono contemporaneamente provare ad accedere al canale riducendo la contesa. Caso limite, se suddivido le stazioni in gruppi quanto le stazioni, questo schema di accesso diventa un TDMA. Questo schema per essere utilizzato richiede una sincronizzazione precisa tra tutte le stazioni. Ogni stazione deve avere una nozione precisa di quando scatta e quando scade lo slot in cui è autorizzato a trasmettere.

Uno schema ancora più semplice è questo, pensato soprattutto per ovviare a problemi generati dalla presenza del fenomeno del terminale nascosto: l'idea è quella di ripartire lo spazio attorno all'AP in settori, a questo punto il tempo viene organizzato in slot successivi e la stazione, l'AP, concede il diritto di trasmettere in maniera circolare solo alle stazioni appartenenti ad un determinato settore. Le stazioni che fanno parte dello stesso settore sono in grado di ascoltarsi l'uno con l'altra, riduco la probabilità di collisioni, ed aumento throughput e risparmio energia.