Riassunto

Fondamenti di Internet of Things

Prof. Sabrina Sicari – A.A. 2022/23

Federico Garegnani

25 marzo 2023

Indice

1 Progettazione di una WSN

La progettazione di una WSN è influenzata dai seguenti fattori

- Reliability (o fault tolerance)
- Scalabilità
- Costi di produzione
- Limitiazioni hardware
- Topologia
- Ambiente applicativo
- Mezzo di trasmissione
- Consumo di potenza (o lifetime)

Analizziamo di seguito ogni parametro nel dettaglio

1.1 Reliability

La *tolleranza ai guasti* (*fault tolerance*, *reliability* o *affidabilità*) è la capacità della rete di non interrompere il servizio anche in presenza di guasti. Un nodo può rompersi per perdita di potenza, danni fisici o interferenze ambientali ma questo non deve pregiudicare l'operatività di tutta la rete di sensori.

L'affidabilità di un nodo sensore K è modellata mediante una distribuzione di Poisson dalla seguente equazione

$$R_k(t) = e^{-\lambda_k t}$$

Essa fornisce la probabilità che sia verificata una rottura al tempo t, λ_k rappresenta il tasso di rottura (failure rate).

L'affidabilità di un broadcast range con N nodi sensori è ottenuta nel seguente modo

$$R(t) = 1 - \prod_{k=1}^{N} [1 - R_k(t)]$$

A seconda dell'affidabilità richiesta si possono variare i protocolli e gli algoritmi utilizzati.

1.2 Scalabilità

Il numero di sensori in una regione può variare considerevolmente, da pochi e diverse migliaia.

La densità dei sensori è il numero di nodi all'interno del *radio range* R

$$\mu(R) = \frac{N\pi R^2}{A}$$

dove N è il numero di sensori all'interno della regione di area A, mentre R è il range di trasmissione.

1.3 Costi di produzione

Il costo deve essere basso. L'obiettivo è di raggiungere un prezzo inferiore a 1 $/dispositivo, attualmente il costo varia tra $25 e $180.

1.4 Topologia

Nella gestione della topologia di una sensor network si distinguono tre fasi

1. **Pre-deployment e deployment phase**
   I sensori possono essere gettati da un aeroplano (*random deployment*) oppure possono essere disposti in modo organizzato (*regualar deployment*). Nel caso di sensori mobili questi possono muoversi autonomamente e cercare aree di interesse, compensare guasti nella rete oppure essere mossi da forze esterne come vento e acqua.

2. **Post-deployment phase**
   Variazioni nella topologia possono verificarsi a seguito di cambiamenti di posizione, problemi di raggiungibilità (a causa di jamming[1], rumore, ostacoli, etc.), energia insufficiente, malfunzionamenti.

3. **Re-deployment di nodi aggiuntivi**

1.5 Mezzo trasmissivo

I dispositivi possono comunicare tramite onde radio, dispositivi ottici (infrarossi) o acustici.

1.6 Consumo potenza

La sorgente di potenza (batteria) è molto limitata ed è il principale fattore che determina il tempo di vita del sensore; in molti scenari non è infatti possibile ricaricare le batterie, risulta quindi importante usare l'energia in modo parsimonioso. Spesso i dispositivi vengono dotati di due batterie, una batteria principale con il compito di alimentare il dispositivo ed una batteria tampone per l'invio dei dati raccolti prima del completo scaricamento.

Da questo punto di vista assume importanza anche la scelta dei protocolli utilizzati per la trasmissione dei dati, si parla di *Power Aware Communication Protocols* in quanto devono minimizzare la dimensione dei pacchetti e il numero di pacchetti scambiati. Alcuni studi hanno analizzato l'assorbimento energetico di ogni componente di un sensore e si è osservato che le fasi di trasmissione e ricezione risultano essere quelle più energivore (Figura 1).

2 Livello di rete

Per una serie di ragioni legate all'elevato numero di nodi, alla limitata potenza e all'approccio data centrico non è possibile utilizzare un identificativo unico per

[1] Il jamming è l'atto di disturbare volutamente le comunicazioni radio (wireless) facendo in modo che ne diminuisca il rapporto segnale/rumore, indice di chiarezza del segnale, tipicamente trasmettendo sulla stessa frequenza e con la stessa modulazione del segnale che si vuole disturbare.

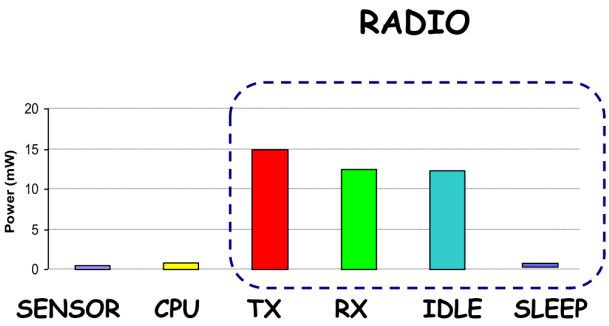

Figura 1: Consumo di energia per ogni componente del dispositivo

ogni dispositivo all'interno di una sensor network. Un indirizzamento dei nodi è tuttavia necessario per la gestione dei nodi, il *querying*, il *service discovery* e il routing.

Risulta necessario utilizzare nuovi algoritmi di routing che tengano conto della limitata potenza, capacità computazionale e memoria disponibili, dei frequenti cambiamenti nella topologia e dell'assenza di un identificativo globale dei dispositivi.

Tassonomia dei protocolli di routing Si distinguono tre categorie di protocolli di routing

- Data Centric Protocols
  Flooding, Gossiping, SPIN, SAR, Directed Diffusion, Rumor Routing, Constrained Anisotropic Diffused Routing, COUGAR, ACQUIRE
- Hierarchical Protocols
  LEACH, TEEN, APTEEN, PEGASIS, Energy Aware Scheme
- Location Based Protocols
  MECN, SMECN, GAF, GEAR

2.1 Flooding e Gossiping

Il *flooding* è l'approccio convenzionale e consiste nell'invio broadcast dei dati a tutti i nodi vicini; il *gossiping* prevede invece l'invio dei dati a un solo nodo vicino scelto casualmente. Sebbene semplici e reattive queste tecniche implicano una serie di svantaggi:

- Implosion
- Overlap
- Resource Blindness
- Power inefficient

Il gossiping è migliore rispetto al flooding in quanto invia i dati a un solo nodo per risparmiare energia ed evita implosion.

Protocollo di routing ideale Le caratteristiche che il protocollo idea dovrebbe avere sono:

- selezione del percorso più breve per l'invio dei dati
- evitare overlap
- minimo consumo di energia
- conoscenza globale della topologia

2.2 SPIN: Sensor Protocol for Information via Negotiation

Alla base del protocollo SPIN ci sono due idee: i sensori si scambiano informazioni su dati che sono già in loro possesso o che desiderano avere, risparmiano così energia e lavorano in modo efficiente; i sensori devono monitorare ed adattarsi alle proprie risorse energetiche.

Usa tre tipi di messaggi: ADV, REQ, DATA nel seguente modo:

- quando un sensore ha qualcosa da trasmettere invia in broadcast un *advertisement packet* (ADV)
- i nodi interessati invio un *request packet* (REQ)
- i dati sono inviati ai nodi che li richiedono
- la procedura viene ripetuta finché tutti i nodi non hanno una copia

SPIN è basato su *data centric routing*, ossia i nodi inviano in broadcast l'advertisement nel caso ci siano dati disponibili e aspettano le richieste delle sink interessate, risulta quindi un ottimo protocollo per disseminare le informazioni tra tutti i nodi.

SPIN-1 Si tratta di un protocollo di handshake a tre fasi, ha il vantaggio di essere semplice e di evitare l'implosione ma al contempo implica un elevato consumo di potenza a causa dell'overhead introdotto.

SPIN-2 Si tratta di una variante di SPIN-1 che modifica il comportamento del sensore sulla base dell'energia disponibile. Quando la disponibilità di energia è massima il protocollo si comporta come SPIN-1, quando invece l'energia residua scende al di sotto di una certa soglia il nodo riduce la propria partecipazione, attivandosi solo se dispone di un'autonomia sufficiente a fargli completare le fasi rimanenti (stato di *dormant*).

Conclusioni

SPIN-1 Previene l'invio di messaggi ridondanti e permette un risparmio di energia del 70% rispetto al flooding

SPIN-2 risparmio energetico del 10% rispetto a SPIN-1 e del 60% rispetto a flooding

2.3 Directed Diffusion Routing Algorithm

Si tratta di un algoritmo *data-centric* in quanto un sensore non necessita di un'identità univoca ma ogni dato trasmesso è identificato mediante i suoi attributi.

Naming scheme I dati generati dai sensori sono identificati dalla coppia attributo-valore; per generare una query occorre usare una coppia attributo-valore come nome dell'oggetto.

Directed Diffusion vs SPIN

Directed Diffusion	SPIN
Sink interroga i sensori per verificare la disponibilità di un dato inviando tramite flooding una regola.	I sensori dichiarano la disponibilità di dati, lasciando alla sink la possibilità di richiederli.

Vantaggi e svantaggi

Vantaggi	Svantaggi
• DD è data centric, non necessita quindi di uno schema di indirizzamento • ogni nodo può fare aggregazione, caching e sensing • DD è efficiente a livello energetico in quanto è *on demand* e non necessita di conoscere la topologia globale della rete	• Non è sempre utilizzabile in quanto si tratta di un *query driven data delivery model* • non è una buona scelta per applicazioni che necessita di un continuo invio di dati • i *naming schemes* sono dipendenti dall'applicazione • il processo di abbinamento tra dato e query causa un leggero overhead

2.4 Low Energy Adaptive Clustering Hierarchy (LEACH)

LEACH è un protocollo basato su cluster con il fine di minimizzare la dissipazione di energia nelle reti di sensori. L'idea alla base è quella di selezionare casualmente alcuni sensori come cluster head e generare i cluster sulla base dell'intensità del segnale ricevuto.
Si distinguono due fasi: *set-up phase* e *steady phase*.

Fase di setup Nella fase di set-up ogni sensore genera un numero casuale tra 0 e 1, se il numero è minore di una certa soglia il nodo diventa cluster head; dopo che i cluster head sono stati selezionati i cluster head stessi comunicano agli altri nodi che sono i nuovi cluster head; terminata la fase di set-up ogni nodo accede alla rete attraverso il cluster head che richiede la minore energia

per essere raggiunto.
Quando i nodi ricevono l'advertisement determinano il cluster di appartenenza sulla base dell'intensità del segnale ricevuto, informano quindi il cluster head di voler partecipare al cluster; a questo punto i cluster head assegnano ai nodi l'intervallo di tempo in cui possono trasmettere.

Fase steady I sensori rilevano e trasmettono dati ai cluster head, i quali aggregano i dati. Dopo un certo periodo trascorso nella fase steady, la rete torna nella fase di setup e ricomincia con la selezione di nuovi cluster head.

E' difficile determinare il numero ottimale di cluster: pochi cluster significa che i nodi si trovano lontani dal proprio cluster head, al contrario molti cluster implica che molti nodi inviano dati alla sink.

Vantaggi LEACH presenta numerosi vantaggi:

- risparmio energetico rispetto a comunicazioni dirette
- clustering dinamico che permette di sopperire agevolmente alla perdita di alcuni nodi
- l'essere completamente distribuito e non richiedere la conoscenza globale della rete
- adozione di un single hop routing, ossia ogni nodo trasmette direttamente al cluster head

Tuttavia presenta lo svantaggio di non essere applicabile per regioni molto ampie.

3 Protocollo SETA

Il protocollo SETA (SEcure sharing of TAsks in clustered wireless sensors networks) nasce al fine di garantire sicurezza e privacy all'interno delle rete di sensori, senza sacrificare il risparmio energetico. La principale caratteristica che caratterizza e distingue SETA è l'adozione di un'architettura ibrida (wireless sensors network, wireless mesh network), i nodi sensori sono organizzati in cluster e comunicano con i cluster heads, i quali sono i router in grado di comunicare con la sink (Figura 2). SETA mira a fornire integrità dei dati, anonimato, risparmio energetico e aggregazione dei dati sicura di tipo end-to-end.

Aspetti chiave Condivisione dei compiti sulla base delle diverse disponibilità energetiche e capacità computazionali tra sensori e cluster head. I nodi sensori svolgono il ruolo di soggetto e processore, si occupano solamente di registrare il dato e criptarlo; i cluster head svolgono il ruolo di soggetto nell'aggregare i dati, di processore nella criptazione dei dati e di controller nella verifica dell'integrità dei dati ricevuti.

Procedura di trasmissione

1. I nodi sensori acquisiscono i dati dall'ambiente, li criptano e li inviano al CH

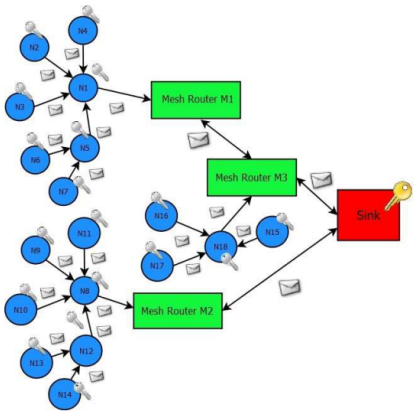

Figura 2: Architettura di rete del protocollo SETA

2. Ogni CH attraversato verifica l'integrità dei dati

3. La sink decifra i dati e genera le statistiche

La verifica di integrità è svolta dal CH nel seguente modo: in caso di violazione della privacy viene inviata alla sink una notifica di errore, altrimenti il dato può essere aggregato o meno a seconda del livello di congestione della rete e inviato alla sink.
L'assegnamento del ruolo di controllore di sicurezza ai CH riduce il traffico sulla rete e di conseguenza l'overhead, con un sensibile risparmio energetico. Gli aspetti negativi della congestione di rete sono lo spreco energetico e la compromissione dell'accuratezza delle stime. Ogni qual volta il numero di messaggi del buffer in trasmissione supera una certa soglia, il CH aggrega i dati al fine di evitare l'overflow del buffer; l'aggregazione dei dati è possibile senza che i messaggi vengano decriptati grazie a un algoritmo di *homomorphic enryption*.

Valuzione delle prestazioni In simulazione SETA è stato confrontato con DyDAP: in condizioni ideali DyDAP presenta prestazione migliori, situazione che però si inverte in presenza di nodi malevoli, dove SETA è in grado di generare un minor carico sulla rete; SETA è in grado di garantire una maggior accuratezza dei dati in seguito ad aggregazione; DyDAP presenta un leggere vantaggio nei tempi di consegna dei messaggi alla sink; SETA garantisce un consumo energetico sensibilmente inferiore.

4 Wireless Multimedia Sensors Networks (WMSN)

La disponibilità di videocamere e microfoni miniaturizzati ha permesso la nascita delle reti di sensori multimediali, in cui vengono trasmessi segnali audio-video dell'ambiente che si desidera monitorare. L'alto volume di flussi multimediali insieme con la ricchezza di informazione generano problemi di congestione, privacy e sicurezza all'interno della rete. La prima sfida da affrontare è quella

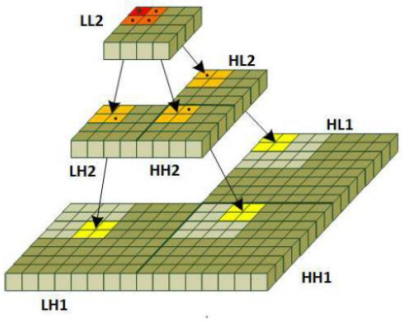

Figura 3: Trasmissione video su una WMSN con differenti bitstreams (codec FS-SPIHT)

di limitare gli effetti della perdita di pacchetti, un pacchetto perso infatti non solo peggiora la qualità del segnale audio-video ricostruito, ma aumenta anche il consumo energetico a causa della ritrasmissione.

I parametri da monitorare per evitare la congestione della rete sono: carico del canale, tempo di arrivo tra i pacchetti, occupazione del buffer locale.

4.1 Secure Selective Dropping Congestion Control (S^{2}DCC)

Il protocollo S^{2}DCC mira a risolvere i problemi sopra elencati relativamente alle reti wireless di sensori. La prima innovazione è l'azione di tecniche di codifica scalabili, il contenuto multimediale viene infatti trasmesso con differenti flussi di bit che possono essere sommati tra di loro o ignorati per modificare la risoluzione a seconda della banda disponibile (Figura 3). Il protocollo è poi in grado di garantire sicurezza e supporta sia WMSN gerarchiche che ibride, permettendo la distribuzione dei compiti nella rete sulla base delle capacità di ogni nodo.

5 Radio Frequency Identification (RFID)

L'idea di RFID si è sviluppata come evoluzione elettronica dei codici a barre, per un'identificazione più rapida dei beni.

La rete RFID si basa sulla presenza di due dispositivi di base: un lettore e un tag. Un lettore è composto da modulo a radiofrequenza (antenna), memoria, CPU, Batteria; un tag è composto da antenna, circuito di recupero dell'energia (trasmessa dal lettore), memoria non volatile per memorizzare l'ID, e in alcuni casi anche di sensori e una piccola CPU. Si distinguono 3 tipi di tag:

Passivi Funzionano solo grazie al circuito di recupero dell'energia che cattura quella trasmessa dal lettore

Semi-passivi Dotati di batteria

Attivi Dotati di batteria e trasmettitore

Electronic Product Code (EPC) Ogni tag è dotato di un EPC, si tratta di uno standard che definisce un codice univoco del prodotto su 96 bit organizzati nel seguente modo:

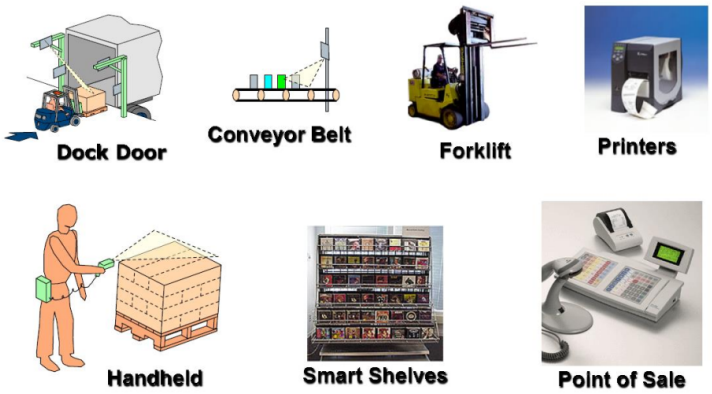

Figura 4: Applicazioni reali che sfruttano RFID

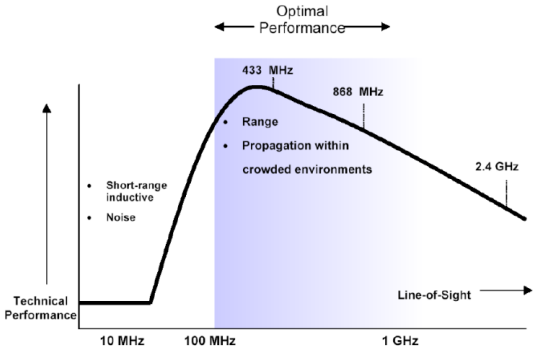

Figura 5: Performance di trasmissione al variare della frequenza per RFID

- *Header*. 8 bit, numero di versione del tag
- *EPC Manager*. 28 bit, ID del produttore
- *Object class*. 24 bit, ID del prodotto
- *Serial number*. ID dell'unità

Le sfide nella realizzazione di un tag RFID sono la costruzione di un circuito di recupero dell'energia efficiente, la miniaturizzazione e il contenimento del costo.

Le sfide nella realizzazione di un lettore RFID riguardano invece la realizzazione di antenne in grado di fornire sufficiente energia al tag, grande differenza nella magnitudo del segnale inviato e ricevuto e l'integrazione con i sistemi industriali più diffusi.

RFID sfrutta le frequenze tra 100 MHz e 1 GHz in quanto risultano essere quelle ottimali per la propagazione del segnale in ambienti affollati (Figura 5).

6 Near Field Communication (NFC)

Connessione tra due dispositivi

RFID	Difference	Barcode
$0.5/1 unit.	Price	$0.005/1 unit.
More than 2^{16} . High capacity	Storage capacity	2^{7} · Low capacity(about 20 letter)
Radio frequency.	How to use	Visible light(infrared ray)
Possible.(by using satellite or mobile radio communication network)	Information understanding in real time	Impossible.
Long distance.	Distance with reader	Close distance.
Read + write.	Read/write capacity	Only read.
Virtually none. Once up and running, the system is completely automated.	Human Capital	Large requirements. Laborers must scan each tags.
High.	Durability	Low.
High.	Security	Low.

Figura 6: Comparazione tra RFID e codici a barre

1. La bobina nel primo dispositivo è attraversata da una corrente che genera un campo magnetico percepito dal secondo dispositivo
2. il secondo dispositivo identifica il campo magnetico ricevuto come un segnale valido e offre la connessione
3. il primo dispositivo accetta la connessione e inizia la transazione

Grazie ai dispositivi dotati di NFC gli utenti possono eseguire pagamenti o usare coupon tramite il dispositivo, senza estrarre la carta di credito o debito; trasferire file; scaricare informazioni circa oggetti, servizi o luoghi; mostrare documenti digitale come le carte di imbarco.

I rischi legati all'utilizzo di questa tecnologia sono essenzialmente legati alla privacy (quali dati vengono trasmessi, processati e memorizzati?), alla sicurezza (cosa succede se smarrisco lo smarthphone?) e al *sentinel hacking*, in cui un tag NFC potrebbe essere nascosto in un punto e registrare le informazioni degli smartphone con cui viene in contatto.

I principali vantaggi della tecnologia NFC restano comunque la comunicazione bi-direzionale, l'alto livello di sicurezza grazie alle tecniche di cifratura, l'alta velocità di riconoscimento con una minima probabilità di errore.

Alternative Un'alternativa a NFC sono i codici a barre o i codici QR. Ognuna di queste tecnologie presenta comunque alcuni punti di forza, motivo per cui NFC non mira a sostituire i lettori ottici. In Figura 6 è riportato un confronto.

7 Nanotecnologie

La nanotecnologia è un ramo della scienza applicata e della tecnologia che si occupa del controllo della materia su scala dimensionale nell'ordine del nanometro e della progettazione e realizzazione di dispositivi in tale scala.

La nascita di questa nuova branca della tecnologia ci pone subito di fronte ad alcuni rischi, quali effetti avversi sulla salute umana o sull'ambiente in

seguito ad un'esposizione volontaria o accidentale e la potenziale proprietà esplosiva delle nanostrutture. Risulta molto difficile fare una valutazione dei rischi associati a questa tecnologia per diversi fattori, in primo luogo c'è bisogno di personale specializzato e apparecchiature sofisticate; risulta poi difficile prevedere come determinate particelle si comporteranno una volta ingerite o disperse nell'ambiente; infine bisogna valutare la potenziale presenza di sostanze tossiche e la loro persistenza all'interno del corso o in natura.

Al momento sono attive due grosse aree di ricerca: *Electromagnetic Nano Communication* e *comunicazione molecolare*. Nella comunicazione molecolare (*molecular communication*) l'informazione è codificata all'interno di DNA, proteine, peptidi, etc ed è trasmessa per diffusione o trasporto attivo[2].

Nella comunicazione elettromagnetica si viene invece a creare una BAN (Body Area Network), composta da molti sensori che comunicano con un micro-gateway (tramite frequenze nell'ordine del THz).

Information Centric Network (ICN) Con l'aumento smisurato della mole di dati prodotta e scambiata attraverso la rete non è più possibile sfruttare una connessione di tipo host-to-host come si è fatto fino ad ora con IP (*host centric network*), bisogna evolvere verso una *information centric network* in cui la rete è in grado di immagazzinare i dati prodotti e fornirli a chi ne faccia richiesta successivamente; si parla così di *in-network caching*. In una rete orientata al dato non collego più direttamente mittente e destinatario ma interrogo la rete e aspetto che qualcuno mi fornisca una risposta, sempre che questa non sia già stata inviata e sia quindi già presente. Chiunque ascolti la richiesta e sia in possesso di una valida copia del dato può rispondere; il dato restituito è firmato e opzionalmente anche criptato, così che la sua integrità ed associazione con il nome possano essere verificate. Si genera una *disallocazione spaziale e temporale* del dato, tutto ciò di cui ho bisogno è il nome del contenuto a cui sono interessato. Le Figure 7 e 8 riportano uno schema di funzionamento di una ICN e una panoramica dei vantaggi offerti.

Secondo lo schema della rete internet tradizionale le applicazioni erano orientate ai servizi offerti da alcuni server, una connessione di tipo host-to-host era quindi preferibile per l'interazione client-server. Al giorno d'oggi la rete internet risulta sensibilmente cresciuta rispetto a cinquant'anni fa, si parla di circa un miliardo di dispositivi attivi, di un trilione di pagine web indicizzate e annualmente vengono prodotti dati nell'ordine degli exabyte (10^{18} Byte). Di fronte a una tale mole di dati IP mostra alcune criticità: necessita di Content Distribution Network [3] (CDN), offre un limitato supporto alla mobilità, basa la

[2]Il trasporto attivo è il trasporto di molecole attraverso la membrana plasmatica mediato da una proteina transmembrana detta trasportatore di membrana. A differenza di quanto avviene nel trasporto passivo, nel trasporto attivo è richiesta una spesa energetica ed è sempre necessaria la mediazione di un trasportatore. In questa forma di trasporto le molecole si muovono contro un gradiente elettrico, chimico o elettrochimico.

[3]Content Delivery Network o Content Distribution Network (in sigla, CDN), ovvero *Rete per la consegna di contenuti*, descrive un sistema di computer collegati in rete attraverso Internet sotto forma di sistema distribuito per ripartire contenuti (specialmente contenuti multimediali di grandi dimensioni in termini di banda, come l'IPTV) agli utenti finali ed erogare servizi di streaming audio e video. Grazie alla CDN si riducono notevolmente i tempi di caricamento di una pagina perché quando un contenuto viene richiesto, a rispondere è il server più vicino geograficamente e ciò si ripercuote positivamente sulle prestazioni del sito.

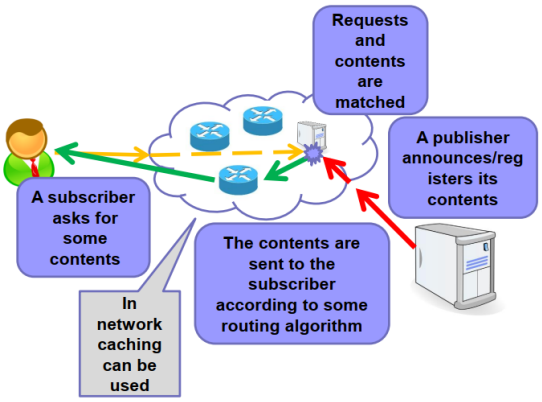

Figura 7: Schema di una ICN

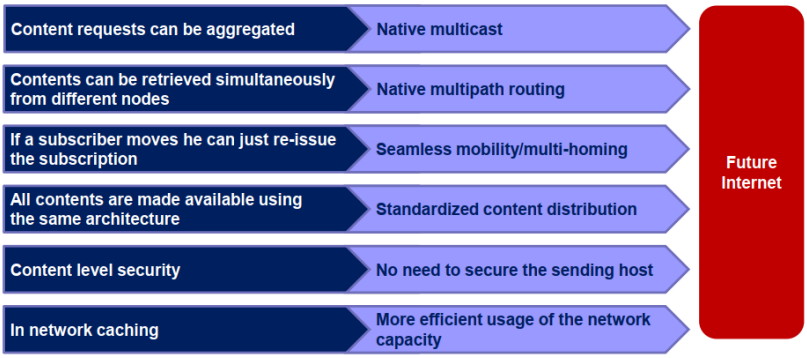

Figura 8: panoramica delle caratteristiche e dei vantaggi offerti dalle ICN

sicurezza sulla protezione degli host invece della protezione dei contenuti, non supporta il multicast su larga scala.

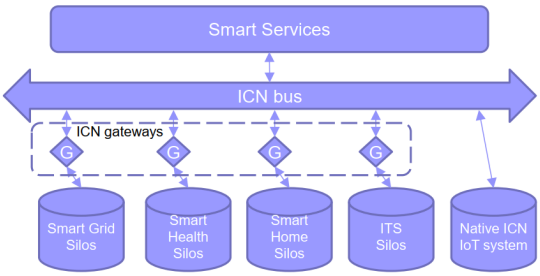

Figura 9: schema generale di un'architettura ICN-IoT

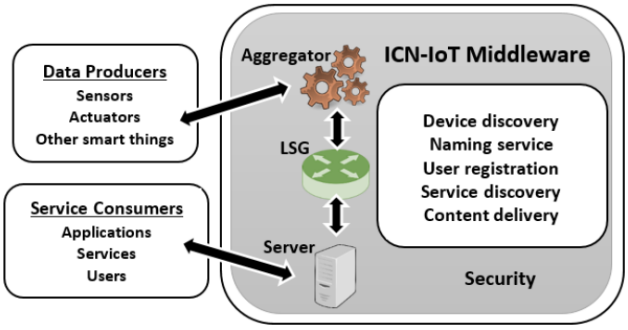

Figura 10: ICN-IoT middleware