
APPUNTI SISTEMI PER L'IOT

INSEGNAMENTO DI SISTEMI PER L'IOT TENUTO DAL PROF. LUCA ROMANELLI

STILATO DA
STEFANO ZIZZI, LUIS FRASHERI

*Università degli Studi di Urbino
Informatica Applicata*



MATRICOLA: 312793, 312972
DICEMBRE 2023

Indice

I	Introduzione	7
1	Fondamenti dei Sistemi per l'Internet delle Cose	9
1.1	Definizione di Sistemi per l'Internet delle Cose (IoT)	9
1.1.1	Edge Devices - Dispositivi Perimetrali	9
1.1.2	Modello Hub-and-Spoke (Concentratore e Raggi)	9
1.1.3	Modello Hub-and-Spoke - esempi in altri ambiti	10
1.1.4	Edge Computer	10
1.2	Storia Termine IoT e Alcune delle Tecnologie Correlate	11
1.3	Casi di Utilizzo	11
1.3.1	Premessa	11
1.3.2	Ambito Industriale e Manifatturiero	12
1.3.3	Ambito Consumer	12
1.3.4	Ambito Assistenza Sanitaria	12
1.3.5	Ambito Trasporto e Logistico	13
1.3.6	Ambito Agricoltura e Ambiente	13
1.3.7	Energia	14
1.3.8	Smart City	14
2	Architettura Generale di un Sistema IoT	15
2.1	Architettura generale di un sistema IoT	15
2.1.1	Livello di "campo"	15
2.1.2	Livello di rete locale	16
2.1.3	Livello di rete locale	16
2.1.4	Livello di rete geografica	16
2.1.5	Livello cloud o applicazione	17
2.2	IoT, M2M e SCADA	17
2.2.1	IoT vs M2M e SCADA	18
2.3	Il Valore Atteso di un Sistema IoT	18
2.3.1	Le Leggi di Metcalfe	18
2.3.2	La Legge di Beckstrom	19
II	Elementi di Base	21
3	Sensori e Attuatori	23
3.1	Sensori Passivi	26
3.2	Sensori Attivi	31
3.3	Sensori ad Alte Prestazioni	36
3.4	Integrazione di Sensori	39
3.5	Attuatori	40
4	Sorgenti ed Accumulatori di Energia	41
4.1	Sorgenti di Energia	41
4.1.1	Unità di Misura	41
4.1.2	Gestione dell'Assorbimento	42
4.2	Approvvigionamento di Energia	42
4.2.1	Energia Solare	43

4.2.2	Energia da Dispositivi Piezo-Meccanici, Elettrostatici, Elettromagnetici	44
4.2.3	Energia da RF	44
4.2.4	Energia a Gradienti Termici	44
4.2.5	Energia da Radioattività	45
4.3	Accumulatori di Energia	46
4.3.1	Batterie Alcaline	47
4.3.2	Batterie agli Ioni di Litio	48
4.3.3	Supercondensatori	48
III	Tecnologie per la Comunicazione	51
5	Non Basate su Protocollo IP	53
5.1	Elementi di base dei sistemi WPAN	53
5.2	Il protocollo Bluetooth	54
5.2.1	Storia del Bluetooth	54
5.3	Il protocollo Zigbee	55
5.3.1	Storia di Zigbee	55
5.3.2	Zigbee-tipo di standard e aspetti tecnici peculiari	56
5.3.3	Elementi dell'ecosistema Zigbee	56
5.3.4	Come un dispositivo Zigbee entra a far parte di una rete (associare un dispositi- vo Zigbee)	57
5.3.5	Zigbee e sicurezza	57
5.4	Il protocollo Z-wave e i sistemi HVAC	58
5.4.1	Sistemi HVAC	58
5.4.2	I meccanismi su cui si basa	58
5.4.3	I vantaggi di HVAC per l'ambiente e il risparmio energetico	58
5.4.4	BMS	59
5.4.5	Sistemi di riscaldamento	60
5.4.6	Sistemi di raffreddamento	60
5.4.7	Z-Wave	60
5.4.7.1	Controller device:	61
5.4.7.2	Nodi o Dispositivi "Slave"	61
5.4.7.3	Indirizzamento in Z-Wave	61
5.4.7.4	Topologia e instradamento	62
5.5	Lo standard KNX	62
6	Basate su Protocollo IP	67
6.1	Enti, Organizzazioni e Definizioni Preliminari	67
6.2	Il modello ISO/OSI	68
6.3	Elementi del Protocollo IP	68
6.3.1	Strato di Internetworking INTERNET PROTOCOL (IP - RFC 791)	68
6.3.2	Comunicazione basata su protocollo IP	69
6.4	Sistemi Wireless: Diagrammi di Irradiazione Antenne	70
6.4.1	Definizioni	70
6.4.2	Antenna Isotropica (Diagramma di Radiazione)	71
6.4.3	Diagrammi di Irradiazione	72
6.4.4	Diagrammi di irradiazione	73
6.5	La Suite di Protocolli IEEE 802.11	74
6.5.1	I Protocolli in IEEE 802.11 e le WLAN	74
6.6	IEEE 802.11 e la Suite di Protocolli	74
6.6.1	Evoluzione di IEEE 802.11	75
6.6.2	Topologie di Base per Sistemi 802.11	76
6.7	Allocazione dello Spettro di Frequenze	76
6.8	Tecniche di Modulazione e Codifica IEEE 802.11	77
6.9	Mitigazione delle Interferenze	78
6.10	IEEE 802.11 MIMO	79
6.11	Beamforming	79
6.12	Protocolli 802.11 per Sistemi IoT: 802.11ac/802.11p/802.11ah	80

6.12.1	IEEE 802.11ac	80
6.12.2	IEEE 802.11p	81
6.12.3	IEEE 802.11ah	82
6.12.3.1	Conclusione	84
6.13	Il Progetto 6LoWPAN	85
6.13.1	WPAN con IP - 6LoWPAN	85
6.13.2	Topologia 6LoWPAN	85
6.13.3	Sicurezza in 6LoWPAN	86
6.13.4	WPAN con IP-Thread	87
6.13.4.1	WPAN con IP-Architettura e Topologia del Thread	87
7	Sistemi e Protocolli di Comunicazione a Lungo Raggio	89
7.1	Tecnologie cellulari	89
7.1.1	Connettività cellulare	89
7.1.2	Modelli e Standard di Governance	90
7.1.2.1	Modelli e Standard di Governance 5G	92
7.2	Tecnologia LoRaWAN	95
7.3	Tecnologia Sigfox	96
IV	Architettura	99
8	Architettura Edge e Cloud	101
8.1	Significato ad Esempi di "Edge Computing"	101
8.1.1	Casi d'Uso	103
8.1.2	Protezione dell'Hardware	103
8.1.3	Sistemi Operativi	104
8.1.3.1	Windows 10 IoT	106
8.1.3.2	Virtualizzazione	106
8.1.3.2.1	Container	107
8.1.4	Cloud Computing per i Sistemi IoT	108
8.2	Protocolli per lo Scambio Dati tra Edge e Cloud: il Protocollo MQTT	110
8.2.1	MQTT	112
8.2.1.1	MQTT-SN	114
8.3	Ambient Computing, Sensori Sintetici	114
8.3.1	Sensori Sintetici	115
9	Vulnerabilità e Protezione dei Sistemi IoT	119
9.1	Classificazione dei Tipi di Attacco Informatico	119
9.2	Tecniche e Meccanismi di Difesa	121
9.3	Esempi di Attacchi Informatici a Sistemi IoT	123
9.3.1	MIRAI	123
9.3.2	STUXNET	125
9.3.3	REAZIONE A CATENA	125
9.3.3.1	Conclusioni	127

Parte I

Introduzione

Capitolo 1

Fondamenti dei Sistemi per l'Internet delle Cose

1.1 Definizione di Sistemi per l'Internet delle Cose (IoT)

IoT si occupa di quei dispositivi che **non necessariamente debbano essere connessi tra di loro o alla rete** e che **in passato** non erano dotati di alcuna, o poca, intelligenza computazionale e capacità di connessione.

Implicitamente stiamo affermando che fanno parte del dominio dell'IoT quei dispositivi che in passato hanno avuto limiti di costo, di assorbimento di energia, di peso, di dimensioni, di riscaldamento, etc. che non consentivano di utilizzarli in maniera pervasiva, limiti oggi superabili (e di fatto superati in molti casi) grazie allo sviluppo tecnologico. Vediamo allora quali potrebbero essere i requisiti per affermare che un dispositivo fa parte dell'IoT e, dandone una definizione più stringente, emergerà che le proiezioni sulla crescita dei dispositivi per IoT, per quanto esplosiva, potrebbero essere comunque sovrastimate perché vanno a sommare anche dispositivi esterni all'insieme specifico IoT.

Requisiti minimi affinché un dispositivo possa dirsi parte di IoT:

- Deve avere capacità computazionali
- Deve avere a bordo lo stack minimo di protocolli per consentire una comunicazione efficace attraverso l'internet, **non necessariamente una connessione diretta ad internet**, che richiederebbe hardware e potenza sufficienti per utilizzare un protocollo di trasporto "complesso"
- Non deve essere una periferica dotata **storicamente** di connessione ad internet, come ad esempio: PC, laptop, smartphone, tablet, server, etc.

1.1.1 Edge Devices - Dispositivi Perimetrali

Vanno inclusi tra i dispositivi del dominio IoT anche i dispositivi "perimetrali". I dispositivi perimetrali sono apparecchi progettati e realizzati per operare anche se esposti ad agenti atmosferici non controllabili e/o in situazioni in cui la fornitura di energia non è costante. Generalmente sono "nodi" gestiti che devono essere posizionati vicino alla sorgente del dato (ad esempio una sonda di temperatura) oppure vicino alla zona di operatività (ad esempio un attuatore motorizzato).

Un nodo non necessariamente coincide con un unico dispositivo ma può essere costituito anche da edge computer connesso a periferiche in modalità Hub-and-spoke.

1.1.2 Modello Hub-and-Spoke (Concentratore e Raggi)

Hub-and-spoke è un modello che si basa su un hub centralizzato che "facilita" le operazioni dei sistemi periferici ed è un modello molto generale che non riguarda solo i sistemi elettronici.

1.1.3 Modello Hub-and-Spoke - esempi in altri ambiti

Le **linee aeree** operano utilizzando aeroporti centralizzati che usano aeroporti regionali o periferici come punti per mezzo dei quali offrire voli.

In ambito **logistica** le merci viaggiano dall'hub verso stabilimenti o magazzini più piccoli, chiamati spokes, per l'ulteriore elaborazione e distribuzione.

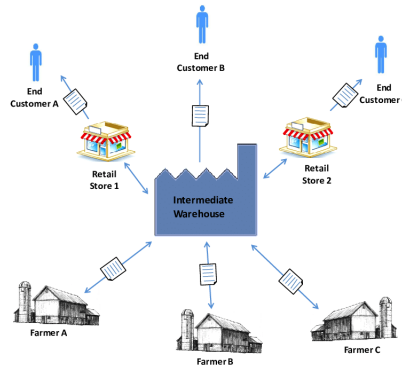


Figura 1.1: Esempio di modello Hub-and-Spoke per la logistica

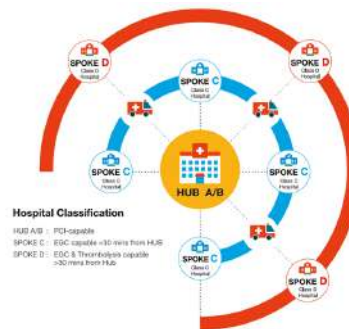


Figura 1.2: Esempio di modello Hub-and-Spoke per un sistema sanitario

In **ambito IoT** molti dispositivi non hanno capacità di connessione ma vengono utilizzati all'interno di un modello "hub-and-spoke" in cui l'hub è generalmente costituito da un piccolo elaboratore che è definito "edge computer".



Figura 1.3: Edge Computer

1.1.4 Edge Computer

Esiste un largo ventaglio di dispositivi connessi agli edge computer attraverso LPAN (Local Personal Area Networks), piccole reti basate su protocolli non-IP (ad esempio Bluetooth), protocolli industriali (ad esempio ModBus), protocolli "legacy" (ereditati) che sono derivati da altri ambiti (ad esempio RS232) o addirittura semplici segnali hardware.

1.2 Storia Termine IoT e Alcune delle Tecnologie Correlate

Il termine può essere probabilmente attribuito a Kevin Ashton che nel 1997, lavorando presso la Procter and Gamble utilizzò tag RFID per gestire una catena di approvvigionamento merci.

Nel 1999 passò al MIT creando insieme ad altri colleghi interessati al tema, il consorzio di ricerca “Auto-ID center” (centro per l’auto identificazione).

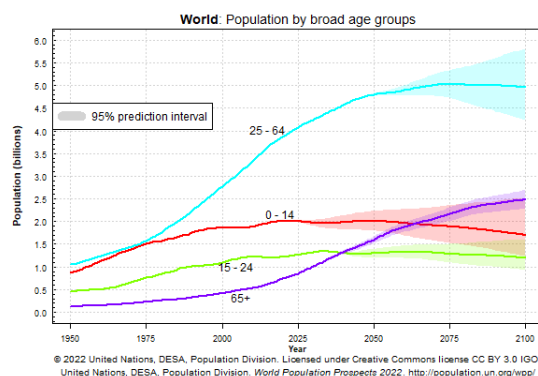
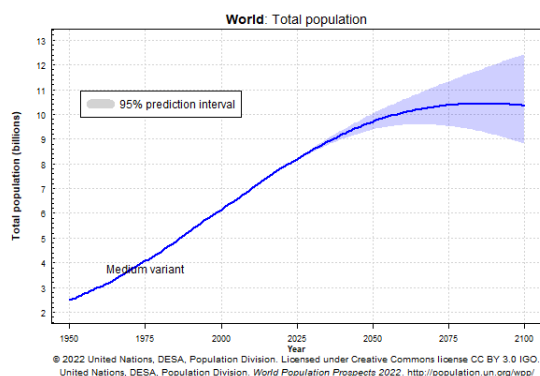
Da allora è nato un ecosistema che, partendo da semplici tag RFID, si stima avrà circa 1 trilione di dispositivi connessi entro il 2030.

Oggi IoT coinvolge segmenti dell’industria, delle imprese, della salute e anche dei prodotti di largo consumo.

La stima sul numero di dispositivi connessi per il 2020-2021 è di 33.4 miliardi e ARM ltd stima 1 trilione (1 miliardo di miliardi) di dispositivi connessi nel 2035.

ARM Holdings è una società di alta tecnologia con sede a Cambridge, Regno Unito. La società è nota principalmente per la sua linea di processori basata sull’architettura ARM (da cui il nome dell’azienda) sebbene sviluppi e venda anche system-on-a-chip, piattaforme hardware, infrastrutture e software sotto i marchi RealView e KEIL.

In generale, per il prossimo futuro, si stima una crescita del 20% all’anno, numero che potremmo comparare alla crescita della popolazione mondiale stimata tra 0.9 e 1.9 % anno.



Su quali settori impatterà questa crescita?

- Soluzioni per la green energy - nuove fonti di reddito
- Assistenza medica presso l’abitazione - riduzione dei costi
- Automazione della produzione – aumento della velocità di produzione
- Tracciamento delle risorse – miglioramento della logistica
- Settore dei beni deperibili - riduzione delle perdite di produzione e/o trasporto
- Sostituzione di dispositivi tradizionali (“cannibalismo”)

Teniamo sempre presente qual è il valore che una soluzione IoT porta in serbo, se si tratta solo di un nuovo gadget probabilmente avrà un mercato molto limitato ma se i benefici che introduce superano il costo industriale di produzione allora probabilmente sarà una soluzione di successo.

Per definire una soluzione di successo, una misura empirica potrebbe essere quella di verificare (o porsi come obiettivo) che ottenga un miglioramento di un fattore 5 rispetto alla tecnologia tradizionale.

1.3 Casi di Utilizzo

1.3.1 Premessa

Nel mondo IoT il segmento che cresce con maggiore velocità e con maggiore diffusione è l’**IIoT** cioè **Industrial IoT** che tradizionalmente viene identificato con **OT** (Operations Technology) la tecnologia

che riguarda il monitoraggio ed il controllo di hardware e software in **tempo reale** (real time).

Storicamente OT si è basata su computer e server in sede (on-premises) ed il sistema hardware e software è identificato dall'acronimo **SCADA** (Supervisory Control And Data Acquisition).

Vi sono differenze significative tra **OT** e **IT**:

- **OT**: misure in campo (sensoristica), dati in tempo reale, tempo di esercizio lungo, sicurezza del sistema.
- **IT**: servizi e distribuzione dei dati, utilizzo di applicazioni groupware, sicurezza (dei dati).

Il vincolo del "tempo reale" è una caratteristica peculiare dei sistemi industriali, sia nel caso dell'acquisizione dei dati che nell'attuazione dei comandi.

Altrettanto importanti sono il *tempo di esercizio* e la *sicurezza* e ciò comporta la necessità di ricorrere quasi ovunque alla **ridondanza**.

Inoltre i sistemi industriali spesso non seguono le "mode" della tecnologia **preferendo l'affidabilità alla velocità**.

Non è un caso che si trovino in funzione ancora sistemi con più di 30 anni di vita e la cui comunicazione è basata su interfacce seriali RS485 piuttosto che su tecnologie più recenti.

1.3.2 Ambito Industriale e Manifatturiero

Vedremo insieme questi casi di utilizzo:

- Utente finale (consumer)
- Assistenza sanitaria (healthcare)
- Trasporti e logistica
- Agricoltura e ambiente
- Energia
- Smart city

1.3.3 Ambito Consumer

L'ambito **consumer** è stato il primo segmento ad adottare dispositivi connessi ad internet e al giorno d'oggi milioni di case hanno termostati "NEST", lampadine "HUE", assistenti "ALEXA" o "GOOLE", le stesse persone sono connesse per mezzo di tecnologia indossabile, ad esempio il "fitbit".

Una delle difficoltà del mercato consumer è il proliferarsi degli standard, ad esempio abbiamo più protocolli per realizzare una WPAN (Wireless Personal Area Network): Bluetooth, Zigbee, Z-wave e purtroppo sono tutti non interoperabili (a meno dell'inserimento di appositi gateway).

Rientrano in ambito IoT per il mercato consumer tutti i sistemi "smart" per irrigazione, automazioni porte, illuminazione, sicurezza, riscaldamento/raffrescamento, etc.

1.3.4 Ambito Assistenza Sanitaria

Il segmento **healthcare** è stato il primo ad adottare dispositivi connessi ad internet e al giorno d'oggi coinvolge milioni di abitazioni.

L'assistenza sanitaria compete con l'industria manifatturiera e la logistica per il miglior ritorno sull'investimento (ROI).

IoT è la soluzione, ormai matura, per monitoraggio veloce e flessibile dei pazienti, ovunque si trovino.

I dati raccolti dalla sensoristica, in combinazione con l'analisi statistica e l'intelligenza artificiale (sistemi esperti) possono "osservare" il paziente e diagnosticare problemi, prescrivere cure se non addirittura salvare vite in situazioni di emergenza.

Ad oggi ci sono circa 500 milioni di dispositivi di monitoraggio indossabili con una crescita prevista in doppia cifra. Purtroppo per i sistemi afferenti al dominio della salute sono posti vincoli molto stringenti:

- Approvazione da parte degli enti regolatori (ad esempio HIPAA in USA)
- Gestione della sicurezza dei dati, che sono considerati (giustamente) “sensibili” dai regolamenti per la privacy (si veda il recente GDPR/2016 per l’Europa)
- I dispositivi devono essere realizzati con standard di qualità pari a quella dei sistemi ospedalieri
- Devono poter comunicare con i centri di monitoraggio 24/7, senza interruzioni (no downtime)
- In alcuni casi, ad esempio per un paziente in un veicolo di emergenza, devono potersi connettere alla rete ospedaliera

Alcuni degli scenari in cui potrebbero essere utilizzati sistemi IoT per la sanità, sono i seguenti:

- Monitoraggio del paziente presso l’abitazione (l’attuale pandemia ha fatto emergere drammaticamente il problema, si veda la situazione delle USCA in Italia)
- Monitoraggio di pazienti affetti da demenza o morbo di Alzheimer (sensori di caduta, etc.)
- Tracciamento dei farmaci e monitoraggio delle somministrazioni (tecnologie RFID)

1.3.5 Ambito Trasporto e Logistico

La **logistica** è uno degli ambiti in cui l’impatto dell’IoT potrebbe essere più significativo.

Partiamo dalla considerazione che attualmente, un veicolo è dotato, mediamente, di 100 sensori.

Questo numero potrebbe raddoppiare a breve per consentire la comunicazione veicolo-veicolo / veicolo-strada / guida automatica.

Questi aspetti si spingono ancora più avanti se si tratta di veicoli su rotaia o natanti.

Alcuni degli scenari in cui potrebbero essere utilizzati sistemi IoT per il trasporto e la logistica, sono i seguenti:

- Localizzazione di flotte
- Localizzazione della merce trasportata
- Pianificazione dei percorsi e monitoraggio di veicoli di servizio: trasporto pubblico, spazzaneve, nettezza urbana, etc.
- Garantire la catena del freddo nel trasporto di cibo congelato/surgelato
- Manutenzione preventiva dei veicoli circolanti

1.3.6 Ambito Agricoltura e Ambiente

L’IoT per **agricoltura e ambiente** si occupa della salute del bestiame, analisi del suolo, uso efficiente dell’acqua, predizione di condizioni climatiche o geologiche pericolose.

Alcuni degli scenari in cui potrebbero essere utilizzati sistemi IoT per l’agricoltura e l’ambiente, sono i seguenti:

- Sistemi intelligenti per l’irrigazione e la fertilizzazione
- Sistemi di illuminazione intelligente per l’allevamento
- Manutenzione preventiva remota dei mezzi e dispositivi per l’agricoltura
- Sorveglianza del territorio tramite droni
- Robotica per l’agricoltura
- Monitoraggio geologico, ad esempio per predire movimenti del terreno o eventi vulcanici

1.3.7 Energia

Attualmente uno degli utilizzi più diffusi è il monitoraggio della produzione di **energia** e del consumo attraverso sistemi di misura intelligenti e connessi che rivelano tipo e volume dei consumi in tempo reale. Si pensi ad esempio al fatto che molti siti di produzione dell'energia sono collocati in ambienti "ostili": deserto (matrici di pannelli solari), alta montagna o mare aperto (matrici di pale eoliche), stabilimenti protetti militarmente (centrali nucleari).

Alcuni dei scenari in cui potrebbero essere utilizzati sistemi IoT per l'energia, sono i seguenti:

- Monitoraggio remoto di campi di pannelli solari
- Monitoraggio real time dei consumi di elettricità, gas e acqua per dare risposta a richieste di picco e misurare consumi
- Movimento real time delle pale di sistemi eolici in funzione delle condizioni del vento

1.3.8 Smart City

Possiamo sintetizzare il termine **Smart City** come insieme di infrastrutture, cittadini e veicoli interconnessi in maniera intelligente.

Siccome comporta un ottimo rapporto costi/benefici è uno degli ambiti di maggior crescita tra quelli dell'IoT.

Ad esempio a Barcellona sono stati installati migliaia di sensori per il monitoraggio dei contenitori di raccolta dei rifiuti urbani per gestire in maniera efficiente la raccolta (si raccoglie solo nei punti in cui non è più disponibile capacità) e tenere traccia del tempo trascorso dall'ultimo prelievo.

Cominciamo ad intuire che sistemi così pervasivi vanno opportunamente dimensionati: se posizioniamo migliaia di telecamere di sorveglianza dovremo dimensionare opportunamente la rete di collegamento dati ma avremo problemi analoghi se posizioniamo qualche decina di migliaia di sensori che, singolarmente, comunicano dati con occupazione di banda minima.

Alcuni degli scenari in cui potrebbero essere utilizzati sistemi IoT per Smart City, sono i seguenti:

- Sensori per la raccolta dati relativi all'inquinamento
- Previsioni meteo locali
- Efficientamento del flusso di traffico veicolo attraverso il controllo intelligente di semafori e l'adozione di percorsi dinamici
- Riduzione dei consumi di energia elettrica implementando sistemi di illuminazione on demand
- Sistemi intelligenti di irrigazione dei parchi e spazi pubblici che tengano conto delle condizioni meteo e dell'utilizzo
- Sistemi intelligenti per il parcheggio che segnalino gli spazi disponibili on demand
- Monitoraggio delle infrastrutture (ponti, strade, etc.) per migliorare il servizio e garantire longevità

Capitolo 2

Architettura Generale di un Sistema IoT

2.1 Architettura generale di un sistema IoT

L'impianto IoT più semplice può partire da sensori remoti che trasducono effetti fisici analogici in segnali digitali. Tali segnali digitali poi vengono trasmessi, trasformati in più modi, fino a destinazione per giungere ad un data center. **La forza dei sistemi per IoT consiste nell'aggregare potenzialmente milioni di sensori, eventi, dispositivi.**

Nell'immagine è mostrato un modello a pila per mezzo del quale viene organizzata la connessione dai sensori al cloud attraverso comunicazione diretta o tramite edge gateways.

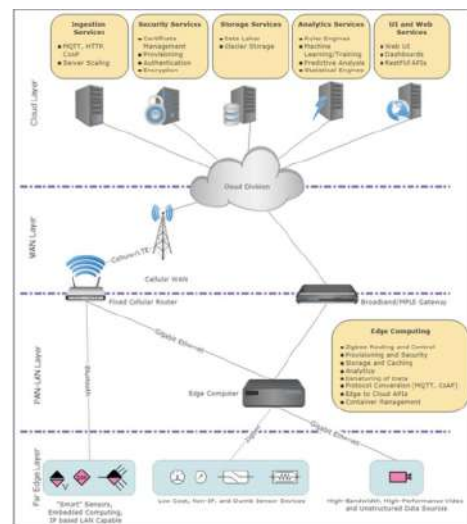


Figura 2.1: Modello Completo

2.1.1 Livello di "campo"

Sistemi embedded, sistemi operativi real time, sistemi di distribuzione dell'energia, sensori MEMs (Micro-Electro-Mechanical Systems)

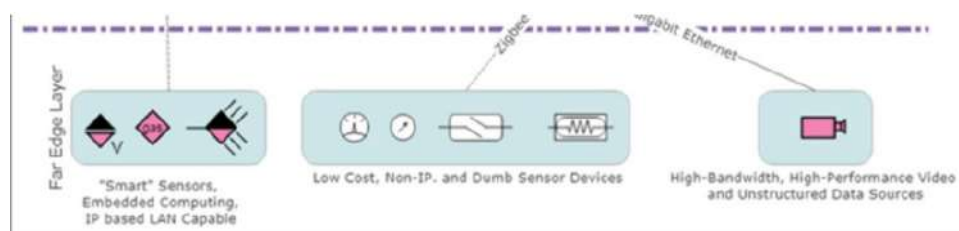


Figura 2.2: Modello di Campo

2.1.2 Livello di rete locale

Sistemi di comunicazione con i sensori (WPANs e LANs)

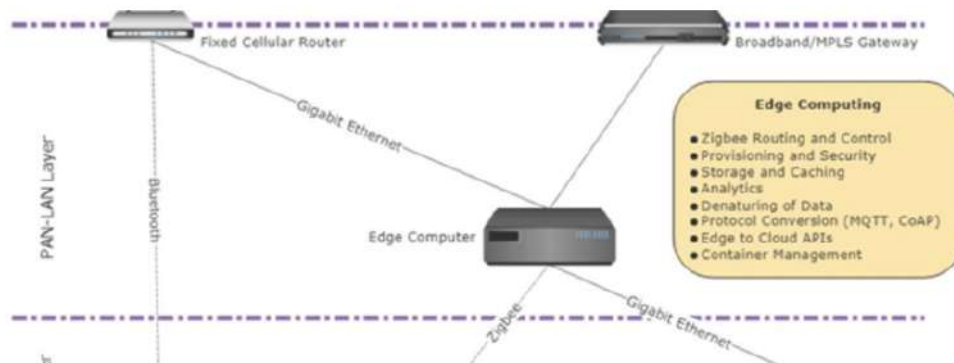


Figura 2.3: Livello di Rete Locale

Sistemi di comunicazione con i sensori (WPANs e LANs)

- **WPANs** : utilizzati per collegamenti a distanze fino a 100m, bassa velocità di trasmissione e bassi consumi, quasi sempre non basati su IP;
- **LANs** : utilizzati per collegamenti anche a grande distanza (fino a chilometri), spesso realizzati secondo le specifiche 802.11 (Ethernet WI-FI) e utilizzando protocollo IP per comunicazioni radio veloci, sia in configurazione peer-to-peer (ad esempio ponti radio) che a stella.

2.1.3 Livello di rete locale

Edge computing Si utilizzano computer più vicini alla sorgente del dato (edge computer), piuttosto che computer posizionati in data center per:

- risolvere problemi di latenza;
- migliorare i tempi di risposta in ottica real-time;
- gestire i “buchi” di connessione;
- fornire ridondanza.

Si deve quindi provvedere a sistemi dotati di CPU, RAM, spazio di archiviazione (storage). Vengono coinvolti produttori di moduli, componenti passivi, client “leggeri” (thin client), sistemi radio wireless o cellulari, dispositivi per la sicurezza perimetrale, sistemi di gestione dei certificati digitali, etc. etc.

2.1.4 Livello di rete geografica

Comunicazioni a distanza “geografica” che utilizzano, ad esempio, reti LTE, satelliti oppure reti estese ma a basso consumo (LPWAN = Low Power WAN) come Sigfox o LoRa.

Sfruttano il protocollo IP configurato appositamente per veicolare messaggi del mondo IoT come MQTT, CoAP ma anche HTTP.

Sono sistemi embedded, spesso anche economici:

- **Aggregatori** : consentono l’interconnessione di dispositivi con stesso protocollo a livello 2 della pila ISO/OSI;
- **Gateway** : dispositivi che mettono in comunicazione reti con protocolli diversi (spesso limitati al livello 2 della pila ISO/OSI);
- **Router** : dispositivi per instradamento di pacchetti IP.

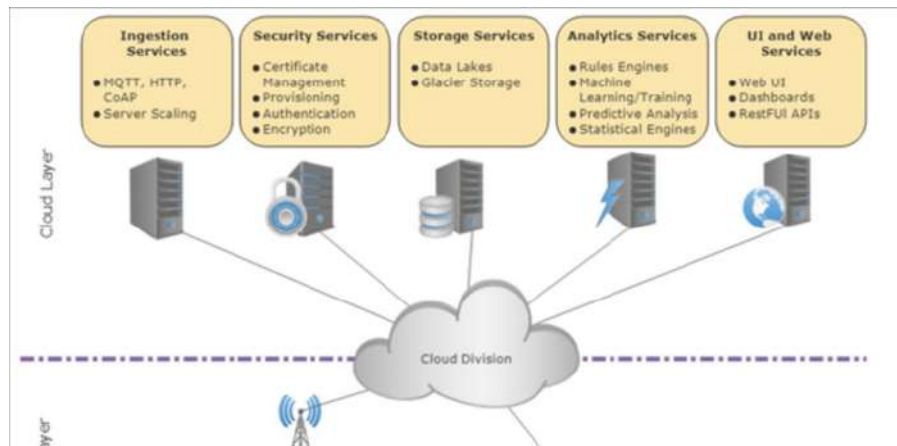


Figura 2.4: Caption

2.1.5 Livello cloud o applicazione

Qui troviamo fornitori di IaaS (Infrastructure as a Service), di SaaS (Software as a Service), di Database, di pacchetti per l'analisi dei dati, servizi di Machine Learning, etc.

- **Analisi dei dati (Data Analytics):** Lo scopo è dare valore alle informazioni raccolte in massa dalla periferia del sistema (edge) e per fare ciò vengono impiegate differenti tecniche di analisi: dalla statistica, ai generatori di regole fino alla più avanzata Machine Learning.
- **Sicurezza:** La sicurezza riguarda l'intera architettura in modalità end-to-end: dai componenti fisici (hardware), passando per i sistemi di comunicazione (siano essi radio o cablati), fino agli stessi protocolli di comunicazione. Ogni livello deve garantire almeno 3 cose:
 - sicurezza;
 - autenticità;
 - integrità.

Non possiamo consentire nessun anello debole nella catena perché IoT costituirà la più grande superficie di attacco del pianeta.

2.2 IoT, M2M e SCADA

Vediamo gli acronimi:

- **M2M** : Machine-To-Machine;
- **SCADA** : Supervisory Control And Data Acquisition.

I 2 termini, insieme al termine IoT, si riferiscono tutti a dispositivi interconnessi e possono utilizzare tecnologie simili ma non sono sovrapponibili al 100%.

M2M e SCADA nascono in ambito automazione industriale, il primo ad essere utilizzato è stato lo SCADA, ancora comunque diffusissimo.

SCADA Sistemi utilizzati nell'industria, nella gestione di automazioni di edificio (building automation) dagli anni '60. Solitamente coinvolgono dispositivi PLC (Programmable Logic Controller) che monitorano o controllano sensori e attuatori. Sono sistemi che solo recentemente, diciamo dai primi anni 2000 in poi, sono stati connessi ad internet diventando i protagonisti di industria 2.0.

Utilizzano protocolli di comunicazione quali ModBus, BACNET e Profibus.

M2M Un dispositivo comunica autonomamente con un altro dispositivo. Il termine “autonomo” si riferisce proprio alla capacità del dispositivo di instaurare e sostenere una comunicazione con un altro dispositivo, senza intervento umano.

Tipicamente non vengono utilizzati canali con protocollo IP ma protocolli che possono essere addirittura “proprietary”.

La differenza sostanziale dagli altri 2 sistemi è che IoT aggrega dati verso un edge computer o verso un gateway. Può incorporare nodi M2M, come ad esempio una rete di dispositivi che comunicano tra loro in bluetooth, ma i dati vengono poi convogliati o ricevuti verso/da endpoint collegati ad internet. Può basarsi anche su sensori e attuatori che hanno capacità computazionali e sono in grado di utilizzare lo stack di protocollo IP per comunicare.

La differenza sostanziale dagli altri 2 sistemi è che IoT aggrega dati verso un edge computer o verso un gateway.

Può incorporare nodi M2M, come ad esempio una rete di dispositivi che comunicano tra loro in bluetooth, ma i dati vengono poi convogliati o ricevuti verso/da endpoint collegati ad internet.

Può basarsi anche su sensori e attuatori che hanno capacità computazionali e sono in grado di utilizzare lo stack di protocollo IP per comunicare.

In definitiva, il fatto che abbia la possibilità di comunicare utilizzando internet è ciò che definisce IoT

2.2.1 IoT vs M2M e SCADA

Spostare i dati via internet consente anche al più semplice dei dispositivi (sensore, smart device, processore di campo, etc.) di utilizzare tutto l'arsenale di tecnologie disponibili in cloud.

Insieme alla maturazione delle tecnologie cloud, i sistemi radio (wireless) sono diventati pervasivi, le sorgenti di energia (solitamente batterie agli ioni di litio) si sono potute acquistare a basso costo e i modelli di machine learning si sono evoluti al punto da produrre valori tangibili.

Senza queste tecnologie, tutte insieme, saremo ancora in un mondo M2M perché mancherebbe lo scopo di trasferire dati verso una elaborazione, archiviazione più spinta, che produca valore.

2.3 Il Valore Atteso di un Sistema IoT

2.3.1 Le Leggi di Metcalfe

Nel 1980 *Robert Metcalfe* formulò il concetto che il valore di una rete è proporzionale al quadrato degli “utenti” connessi. $V \propto N^2$

$$V \propto N^2$$

[un esempio che valida tale legge è stato recentemente fornito utilizzando il valore delle blockchains e l'andamento delle crypto valute]

Nel caso di IoT gli “utenti” sono rappresentati da sensori o dispositivi di campo (edge devices) che abbiano qualche capacità di comunicazione. Se supponiamo che ogni sensore abbia un costo invariabile di 10 € avremo una retta crescente che rappresenta il costo totale del sistema IoT che incrocia in un punto la parabola della “legge” di Metcalfe, a destra di quel punto si troverà la zona che darà un ROI (Return On Investment) positivo.

Limiti Si faccia attenzione al fatto che la legge di Metcalfe non tiene conto di alcuni fattori che invece potrebbero influenzare anche pesantemente una rete “reale”. Non tiene conto della degradazione dei servizi al crescere del numero di utenti connessi in assenza di un corrispondente aumento di banda disponibile, della eventuale presenza di attori malevoli sulla rete (ad esempio attacchi DoS), della indisponibilità della rete in alcuni momenti (ad esempio per veicoli in movimento), etc. etc.

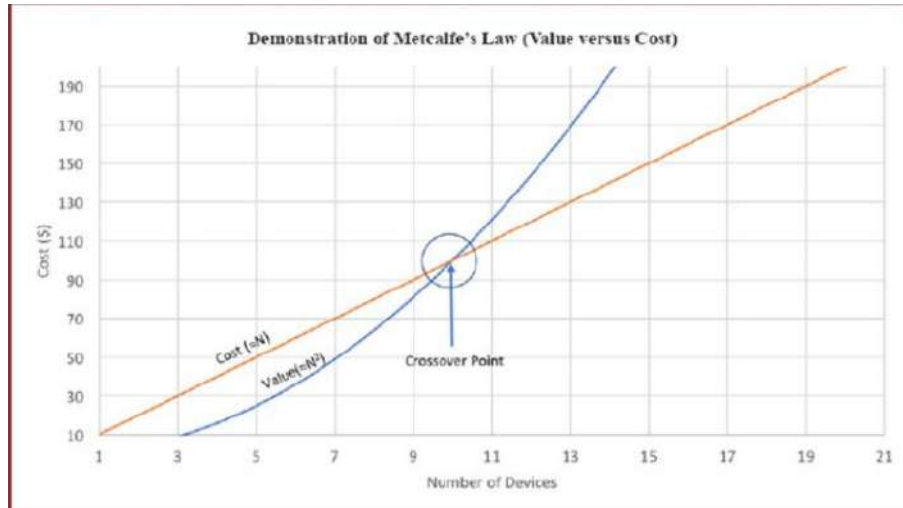


Figura 2.5: Relazione Costo-Dispositivi

2.3.2 La Legge di Beckstrom

$$\sum_{i=1}^n V_{i,j} = \sum_{i=1}^n \sum_{k=1}^m \frac{B_{i,j,k} - C_{i,j,k}}{(1 + r_k)^{t_k}}$$

- $V_{i,j}$: Rappresenta il valore presente del network per un dispositivo i sul network j ;
- i : Un utente individuale o dispositivo sul network;
- j : Il network;
- k : Una singola transazione;
- $B_{i,j,k}$: Il beneficio che il valore k porterà al dispositivo i sul network j ;
- $C_{i,j,k}$: Il costo di una transazione k ad un dispositivo i sul network j ;
- r_k : Il tasso di sconto di interesse al tempo della transazione k ;
- t_k : Il tempo passato (in anni) per transizionare k ;
- n : Il numero di individui;
- m : Il numero di transazioni.

Una legge più precisa è quella fornita da Beckstrom che sostanzialmente asserisce che per calcolare il valore di una rete (in questo contesto pensiamo sempre ad una soluzione IoT), dobbiamo calcolare il valore di ogni transazione generata da ogni dispositivo e sommare quei valori per ottenere il valore complessivo.

Ma in un modello IoT, come calcolare il beneficio (il valore) di una transazione ?

Partiamo da due casi pratici: Quanto potrebbe incidere 1 ora di assenza di connessione per un sensore di temperatura su una macchina? Probabilmente molto poco.

Se invece pensiamo ad un sensore di allagamento che, a causa di una batteria scarica, non comunica l'allagamento di un garage? Il valore potrebbe essere altissimo.

Il primo problema che si deve porre un progettista di sistemi IoT è *quale valore si sta realizzando perché potrebbe anche essere addirittura negativo* ma l'architetto di sistemi IoT ha anche la possibilità di scegliere tra milioni di combinazioni e tra queste quelle convenienti (così come quelle fallimentari) possono essere tante.

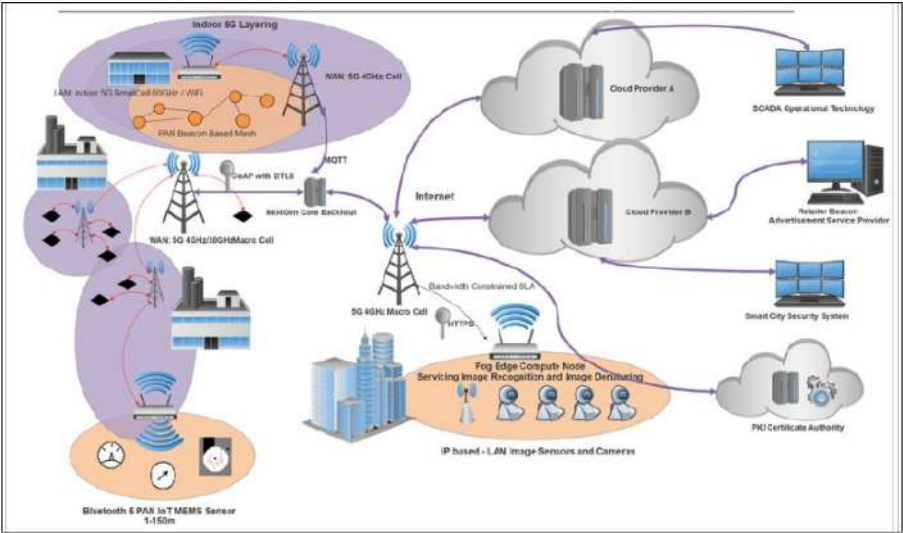


Figura 2.6: Network

Parte II

Elementi di Base

Capitolo 3

Sensori e Attuatori

Come abbiamo visto i sensori e gli edge computer sono elementi fondamentali di una architettura IoT. Ora approfondiremo il loro funzionamento dal punto di vista elettronico e dei principi che gli consentono di misurare le grandezze fisiche di interesse.

In poche parole *“quale tipo di sensore e/o dispositivo di campo dovrei prendere in considerazione per il problema che devo risolvere?”*

La categoria dei sensori è molto vasta e va dalle semplici termocopie fino ai più avanzati sistemi video.

Possiamo suddividerla in 3 grandi gruppi:

- Sensori passivi – rispondono a cambiamenti nell’ambiente;
- Sensori attivi – inviano segnali che servono a misurare tempo o spazio;
- Sensori ad alte prestazioni.

Sensori **passivi** - rispondono a cambiamenti nell’ambiente.

- Sensori di temperatura;
- Sensori di corrente elettrica, luce, radiazione infrarossa.

Sensori **attivi** - inviano segnali che servono a misurare tempo o spazio

- LiDAR;
- MEMS:
 - Accelerometri e giroscopi;
 - Microfoni;
 - Sensori di pressione.

Sensori **attivi ad alte prestazione**

- Sistemi di visione

Teorema del Campionamento La maggior parte dei segnali che incontriamo nel mondo reale sono dei segnali a tempo continuo (la voce, la musica, le immagini). Gli algoritmi di elaborazione numerica del segnale vengono spesso usati per trattare i segnali a tempo continuo. Per farlo i segnali vengono dapprima campionati e codificati con un convertitore A/D (analogico/digitale), vengono elaborati per via numerica e poi riconvertiti in forma analogica con un convertitore D/A (digitale/analogico).

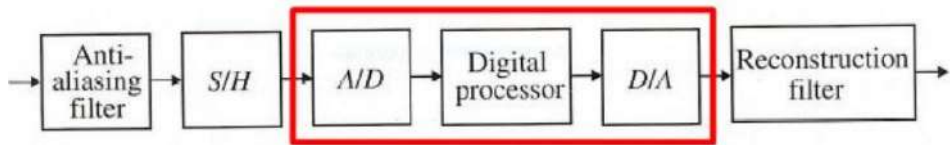


Figura 3.1: Elaborazione Numerica

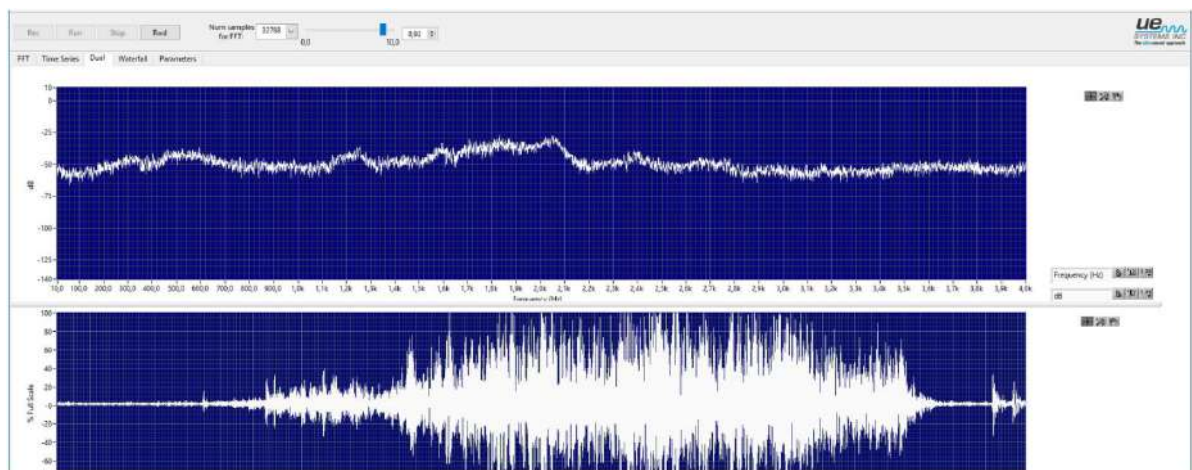


Figura 3.2: Esempi di forma d’onda di segnali analogici



Figura 3.3: Conversione analogico-digitale

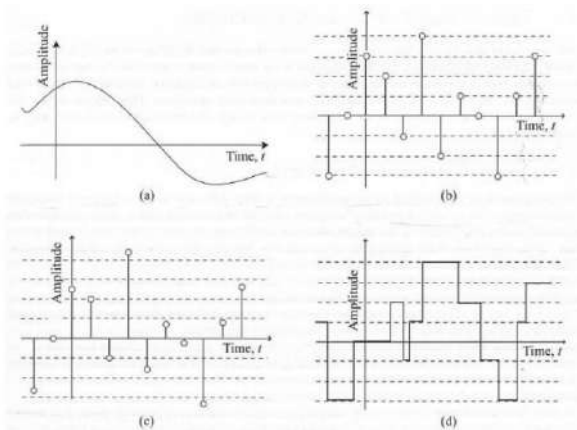


Figura 3.4: Le 3 operazione che possiamo effettuare su un segnale analogico

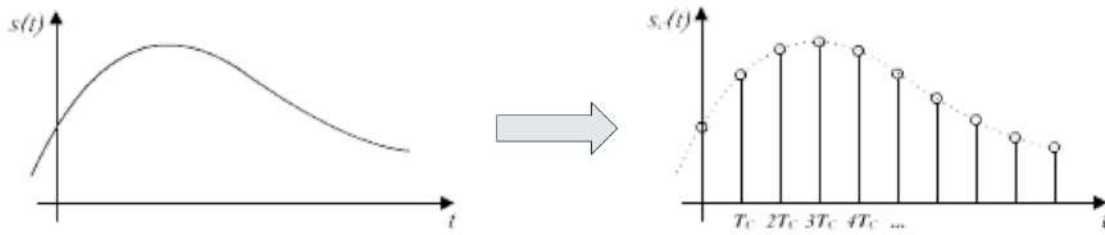


Figura 3.5: Campionamento di un segnale analogico

Campionamento di un segnale analogico Il campionamento è il primo passo del processo di conversione analogico-digitale di un segnale.

Consiste nel prelievo di campioni (samples) da un segnale analogico e continuo nel tempo ogni Delta tempo. Il valore Delta t è detto intervallo di campionamento, mentre $1/\Delta t$ è la frequenza di campionamento.

Il risultato è un segnale analogico in tempo discreto, che viene in seguito quantizzato, codificato e reso accessibile a qualsiasi elaboratore digitale.

Il teorema di Nyquist-Shannon (o teorema del campionamento dei segnali) stabilisce che, dato un segnale analogico $s(t)$ la cui banda di frequenze sia limitata dalla frequenza f_M e dato $n \in \mathbb{Z}$,

allora

il segnale $s(t)$ può essere univocamente ricostruito a partire dai suoi campioni $s(n\Delta t)$ presi a frequenza $f_S = 1/\Delta t$ se $f_S > 2f_M$

Teorema del campionamento Sia $g_a(t)$ un segnale a banda limitata, con $G_a(j\Omega) = 0$ per $|\Omega| > \Omega_m$. $g_a(t)$ è univocamente determinato dai suoi campioni $g_a(n, T)$, (ovvero, può essere fedelmente ricostruito dai suoi campioni $g_a(n, T)$) $-\infty < n < +\infty$, se la frequenza angolare di campionamento

$$\Omega_T \geq 2\Omega_m$$

In altre parole, se vogliamo che il segnale a banda limitata $g_a(t)$ sia ricostruibile dai suoi campioni, dobbiamo campionarlo con una frequenza di campionamento almeno pari a due volte la larghezza di banda.

Spettro di un segnale: trasformata di Fourier a tempo continuo (CTFT) La rappresentazione nel dominio della frequenza di un segnale a tempo continuo $x_a(t)$ è data dalla CTFT definita da

$$X_a(j\Omega) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} x_a(t) e^{-j\Omega t} dt$$

La CTFT viene anche chiamata *spettro di Fourier*, o semplicemente *spettro*, di un segnale continuo.

Il segnale a tempo continuo $x_a(t)$ può essere ricostruito dalla sua CTFT mediante la trasformata di Fourier a tempo continuo inversa (ICTFT - Inverse CTFT), definita da

$$x_a(t) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} X_a(j\Omega) e^{j\Omega t} d\Omega$$

Si noti che esiste una corrispondenza biunivoca tra il segnale $x_a(t)$ e la sua trasformata:

$$x_a(t) \xleftrightarrow{CTFT} X_a(j\Omega)$$

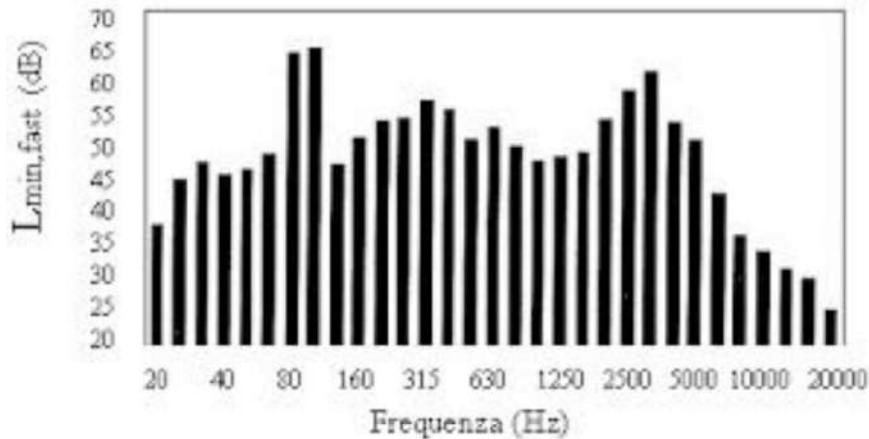


Figura 3.6: Spettro di un segnale

3.1 Sensori Passivi

I sensori di temperatura sono la categoria più diffusa, sono elementi dei termostati ma sono presenti anche nella logistica per garantire la catena del freddo, fino alle macchine industriali.

Termocoppie Sono dispositivi che non hanno bisogno di un segnale di eccitazione per funzionare ma, di contro, producono un segnale molto debole, dell'ordine dei microvolts.

Le termocoppie si basano sull'effetto elettromotrice Seebeck.

Effetto Seebeck Due conduttori metallici, di materiale diverso tra loro, vengono uniti nel punto in cui si vuole effettuare la temperatura. All'altro capo dei due conduttori si sviluppa una differenza di potenziale che varia al variare della temperatura e può essere misurata.

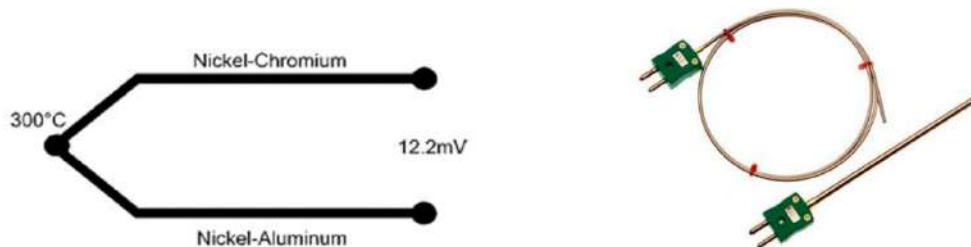


Figura 3.7: Effetto Seebeck

Sostanzialmente si tratta di sensori formati da una coppia di conduttori di diverso materiale uniti in un punto, nel quale è applicata la temperatura da misurare (giunto caldo), mentre le estremità dei conduttori sono lasciate libere (giunto freddo).

L'entità della differenza di potenziale, al capo opposto rispetto al punto di giunzione, dipende dai due materiali utilizzati.

Occorre tenere presente che la relazione tra la differenza di potenziale dei due materiali e la temperatura è non lineare e per questo motivo, il circuito elettronico che campiona la differenza di potenziale, è dotato di software che fa riferimento ad una tabella che esprime la temperatura in funzione della relazione non lineare fra i metalli scelti.

Sono utilizzate in applicazioni semplici perché sono generalmente poco precise, ottenere termocoppie di precisione elevata ha un costo che spesso non ne giustifica l'utilizzo.

Inoltre la precisione della misura degrada facilmente con l'andare del tempo.

Sono utilizzate in ambito industriale e ambienti ad alta temperatura.

Denominati RTD (Resistance Temperature Detectors), rispetto alle termocoppie operano in un range di temperature molto più stretto ma sono molto più accurati.

Vengono raramente utilizzati per applicazioni oltre i 600 °C e ciò ne limita l'uso in ambito industriale. Sono generalmente realizzati avvolgendo un sottile filo di platino su un cilindro di materiale ceramico o vetro e questo produce una relazione resistenza-temperatura.



Figura 3.8: Sensori Passivi RTD

Poiché si basano su una misura di resistenza, è richiesta una corrente di eccitazione, seppure debole, dell'ordine del milliampere (mA).

La resistenza di un RTD segue una retta con pendenza caratteristica. Ad esempio un 200 PT100 RTD ha una pendenza di $0.00200 \text{ } \Omega/^{\circ}\text{C}$ tra 0 e 100° C.

Un RTD può essere 2, 3 o 4 fili dove i 4 fili si utilizzano in applicazioni di alta precisione.

Sono spesso utilizzati in circuiti a ponte con applicazioni che poi linearizzano i risultati.

Termistori Sono essenzialmente delle resistenze che variano la propria impedenza al variare della temperatura.

Rispetto agli RTD producono una più grande variazione di resistenza a parità di variazione di temperatura e questa relazione è fortemente non lineare.

Sono utilizzati quando è necessaria una grande accuratezza in un range di temperature molto "stretto".



Figura 3.9: Sensori Passivi Termistori

Ci sono 2 tipi di termistori:

- NTC - la resistenza decresce all'aumentare della temperatura
- PTC - la resistenza cresce all'aumentare della temperatura

Sono utilizzati nella realizzazione di strumenti per la medicina, dispositivi scientifici, incubatori, conservazione di alimenti e termostati per abitazioni, uffici, etc.

DEFINIZIONI PRELIMINARI Il **campo elettrico (E)** è la proprietà elettrica dello spazio dovuta alla presenza di cariche elettriche: ciò provoca una perturbazione dell'ambiente circostante in conseguenza del quale altre cariche elettriche vengono attratte o respinte.

L'intensità del campo elettrico si misura in Volt per metro (V/m).

Il campo magnetico (H) è la proprietà magnetica dello spazio dovuta alla presenza di cariche elettriche in movimento (corrente) o di magneti (calamita, ago della bussola).

L'intensità del campo magnetico si misura in Ampere per metro (A/m)

Tuttavia si parla spesso di Tesla, che è l'unità di misura dell'induzione magnetica (B) (o densità del flusso magnetico) che tiene conto del mezzo (aria) in cui il campo magnetico si genera.

Il campo elettrico (E) deriva dai conduttori inseriti nelle prese quindi è sempre presente anche se un apparecchio elettrico è spento, purché sia collegato alla rete elettrica.

È facile schermare il campo elettrico tramite oggetti o pareti.

Il campo magnetico (H) viene prodotto quando un apparecchio elettrico viene messo in funzione e quindi in esso circola corrente elettrica.

Non è facile schermare il campo magnetico.

Il campo elettromagnetico è dato dal campo elettrico che varia nel tempo genera perpendicolarmente ad esso il campo magnetico.

Le onde elettromagnetiche, quindi, sono formate da campi elettrici e magnetici che si propagano nello spazio alla velocità della luce.

La due onde oscillano perpendicolarmente e in fase tra loro, cioè le rispettive ampiezze aumentano e diminuiscono simultaneamente.

Alle bassissime frequenze non si utilizza il termine campo elettromagnetico, ma si parla di campo elettrico e campo magnetico: questo perché le variazioni dei campi nel tempo sono così lente che i due campi si comportano come agenti fisici separati e i loro effetti vanno analizzati separatamente.

Le radiazioni elettromagnetiche si distinguono in:

- radiazioni ionizzanti (IR = Ionizing Radiations) come ad esempio i "raggi X" o i "raggi gamma"
- radiazioni non ionizzanti (NIR = Non Ionizing Radiations)

La differenza fra questi due tipi di radiazioni sta nella capacità o meno della loro energia in movimento di ionizzare gli atomi e quindi di modificare a livello atomico la materia.

Sensori per la misurazione di corrente elettrica (basati su effetto Hall) Consistono in una striscia di metallo che viene attraversata da una corrente, realizzando un dispositivo che sfrutta l'effetto Hall per misurare correnti alternate (AC) o continue (DC).



Figura 3.10: Sensori passivi di corrente elettrica

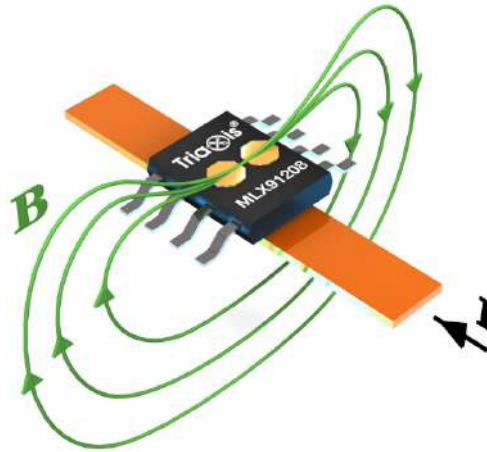


Figura 3.11: Sensori passivi di corrente elettrica-2

Vengono utilizzati per realizzare sensori di posizione, magnetometri, interruttori ad alta velocità di commutazione, rilevatori di livello liquidi.

In ambito industriale misurano velocità di rotazione di macchine o motori.

Sono molto economici e resistenti in ambienti con condizioni ambientali anche difficili.

Sensori fotoelettrici La sensibilità alla presenza di luce o all'intensità della luce viene utilizzata in molteplici sensori, ad esempio: sistemi di sicurezza, interruttori intelligenti, illuminazione pubblica smart (smart cities).

Distinguiamo **Fotoresistori** da **Fotodiodi**, come si può dedurre dai nomi, il primo varia la propria resistenza in funzione della intensità della luce mentre il secondo converte la luce in corrente elettrica.

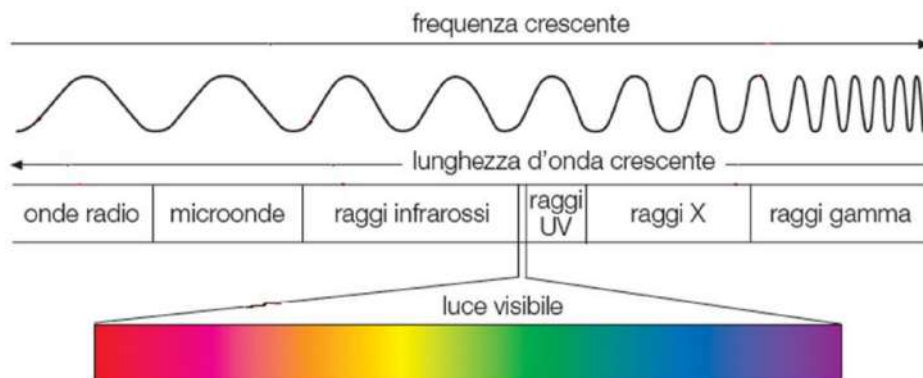


Figura 3.12: I range della radiazione elettromagnetica

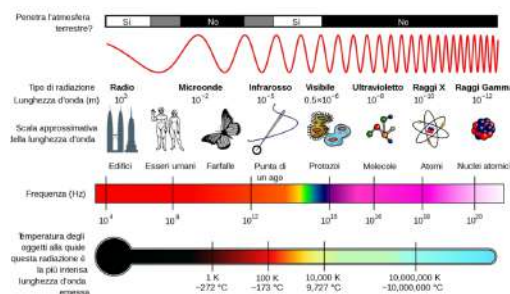


Figura 3.13: I range della radiazione elettromagnetica-2

Sensori fotoelettrici - fotoresistori I fotoresistori sono costruiti con un semiconduttore ad alta impedenza: la resistenza decresce in funzione della luce assorbita dal semiconduttore, in condizioni di buio assoluto la resistenza può essere misurata in megaohm (di fatto un circuito “aperto”). Sono sensibili alla lunghezza d’onda della luce incidente.



Figura 3.14: Sensori fotoelettrici - fotoresistori

Sensori fotoelettrici - fotodiodi I fotodiodi sono realizzati con una giunzione p-n e rispondono alla luce creando una coppia elettrone-lacuna: le lacune muovono verso l’anodo mentre gli elettroni muovono verso il catodo e si realizza una corrente elettrica. Le celle solari operano proprio in questo modo (Effetto fotovoltaico).

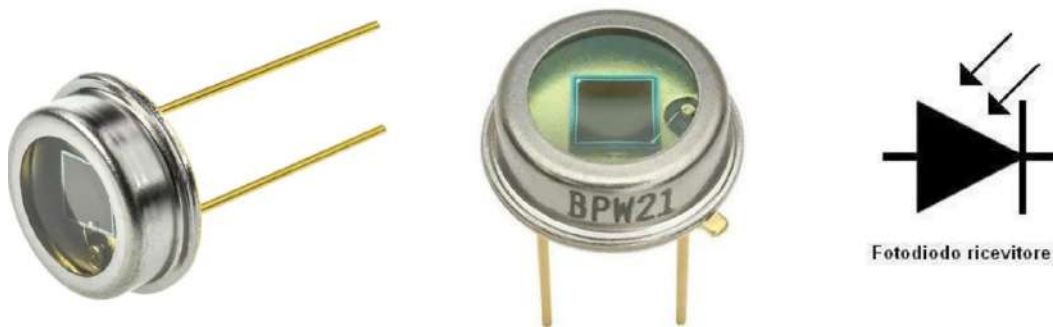


Figura 3.15: Sensori fotoelettrici - fotodiodi

Sensori PIR I sensori PIR (Pyroelectric InfraRed) basano il loro funzionamento su materiali che sono sensibili alla radiazione infrarossa e al calore.

I sensori PIR operano solitamente nell’intervallo di lunghezze d’onda 8-14 micrometri che sono tipiche del corpo umano.

Il materiale sensibile è posto dietro una lente di Fresnel e quando un corpo caldo entra in uno degli archi della lente o lo lascia, genera un segnale che viene rilevato da un transistor ad effetto di campo (FET) e passa poi ad un amplificatore.

Per fare rilevazioni su ampie superfici, ad esempio una intera stanza, si utilizzano array di PIR con molteplici lenti di Fresnel e di solito il dispositivo completo consente di regolare ampiezza, sensibilità e tempo di “hold” del segnale che specifica il tempo per il quale il segnale di uscita viene mantenuto “alto” dopo aver rilevato un movimento lungo il percorso del PIR, chiaramente minore è il tempo di “hold” maggiore è il numero di eventi che si rilevano.



Figura 3.16: Sensori di radiazione infrarossa

Lente di Fresnel La lente di Fresnel è stata inventata dal fisico francese Augustin-Jean Fresnel che la utilizzò per la prima volta nel 1822, per realizzare, in vetro, la lente di un faro per segnalazioni marittime. Le lenti di Fresnel sono utilizzate per raccogliere la luce o ingrandire immagini con un ingombro minore rispetto alle lenti realizzate con geometria tradizionale.

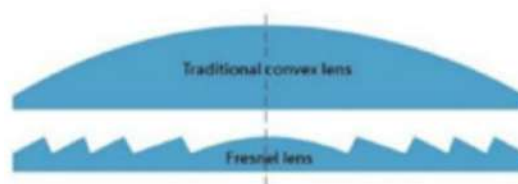


Figura 3.17: Lente di Fresnel

A differenza delle tipiche lenti ottiche sferiche o asferiche, le lenti di Fresnel sono composte da una serie di scanalature concentriche incise su un lato di un foglio di plastica. Una lente di Fresnel è una lente ottica sottile e piatta che consiste in una serie di piccole scanalature concentriche strette sulla superficie di un foglio di plastica leggero al fine di ridurre lo spessore, il peso e il costo. Ogni scanalatura ha un'angolazione leggermente diversa dalla successiva e con la stessa lunghezza focale per focalizzare la luce verso un punto focale centrale.

Ogni scanalatura può essere considerata come una piccola lente individuale per piegare le onde luminose di Fresnel parallele e focalizzare la luce. L'obiettivo elimina effettivamente alcune aberrazioni sferiche.

La lente di Fresnel è una lente ottica speciale. Ora può essere realizzato in plastica come lenti di Fresnel acriliche, PMMA, polivinilcloruro (PVC), policarbonato (PC) e HDPE. Una lente convessa in vetro tradizionale sarebbe spessa, pesante e molto costosa, ma una lente di Fresnel in plastica è un'alternativa sottile, piatta, leggera e a basso costo.

Un'elevata densità di scanalature ha un'immagine di proiezione di qualità migliore. L'obiettivo di Fresnel è una buona soluzione per immagini di qualità ed efficienza a un costo notevolmente inferiore.

3.2 Sensori Attivi

Sono sensori che pur utilizzando tecnologie differenti, operano basandosi sullo stesso principio di funzionamento:

generano un segnale (ad esempio un impulso laser) che ritorna ad un rilevatore (che è esso stesso un sensore) per costruire un'immagine o indicare che si è verificato un certo evento.

Sono molto più complessi dei sensori passivi e richiedono più potenza, costi e spazio di ingombro.

Sensori LiDAR (Light Detection And Ranging) Questi sensori misurano la distanza di un obiettivo basandosi sul "tempo di volo" di un impulso laser riflesso dal target ma sono anche in grado di

analizzare qualunque cosa attraversi il cammino dell'impulso o raggio laser, ad esempio gas, composizione delle nuvole, particelle in sospensione nell'aria, velocità di oggetti in movimento, etc.



Figura 3.18: Sensori attivi - LiDAR

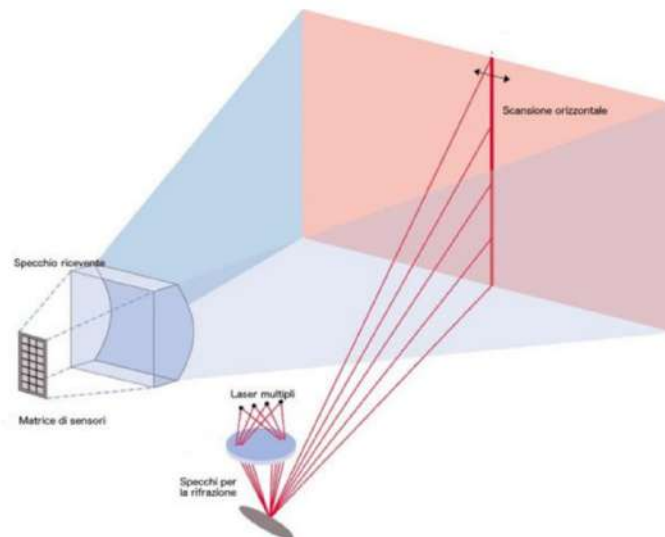


Figura 3.19: Schema di un sistema di visione basato su LiDAR per robot

Vengono utilizzati in agricoltura, veicoli a guida autonoma, robotica, sistemi di sorveglianza e studi ambientali.

I laser utilizzati lavorano tipicamente nel range di lunghezze d'onda da 600 a 1000 nanometri e sono relativamente economici, naturalmente la potenza è molto limitata per evitare problemi agli occhi.

A volte vengono utilizzati laser che operano a lunghezze d'onda maggiori, 1550 nanometri, perché questa lunghezza d'onda non danneggia gli occhi anche ad energie più alte.

Possono operare a distanze anche considerevoli, ad esempio da un satellite.

L'impulso laser emesso viene riflesso dall'oggetto target ed il segnale di ritorno viene rilevato da una batteria di fotodiodi.

Il calcolo della distanza è molto semplice:

$$d = \frac{C \times T}{2}$$

dove C = velocità della luce, T= tempo di volo.

Possono consentire di comporre anche figure 3D dell'ambiente utilizzando uno specchio in rotazione.

Sensori MEMS (MicroElectroMechanical System) Appartengono a questa categoria sensori che vengono utilizzati nell'industria dagli anni '80 anche se il loro antenato, un sensore di pressione basato su effetto piezoelettrico, fu realizzato nel 1960.

Essenzialmente incorporano strutture meccaniche miniaturizzate che interagiscono con componenti elettronici e, diversamente da tutti gli altri sensori, possono ruotare, allungarsi, muoversi o alterare la propria

forma generando un segnale elettrico.

Li possiamo trovare non solo come elementi di sensoristica ma anche a bordo di stampanti a getto di inchiostro, di proiettori DLP (Digital Light Processor).

La possibilità di realizzare sensori MEMS grandi come la testa di uno spillo li rende di fatto già protagonisti di una enorme diffusione in ambito IoT.

Appartengono a questa categoria:

- Accelerometri e giroscopi
- Microfoni
- Sensori di pressione

Accelerometri e giroscopi Un accelerometro risponde ad un cambiamento di velocità lineare mentre un giroscopio segnala movimenti di rotazione.

Oggigiorno sono utilizzati in maniera diffusa ad esempio in dispositivi per il fitness o per il tracciamento della posizione.

Il funzionamento si basa sul fatto che un dispositivo piezoelettrico produce una differenza di potenziale in risposta ad un movimento.

Accelerometri Un accelerometro ha una piccola massa centrale tenuta da molle. Il movimento produce uno spostamento della massa che fa variare la capacità di un circuito e questo fenomeno è misurabile.

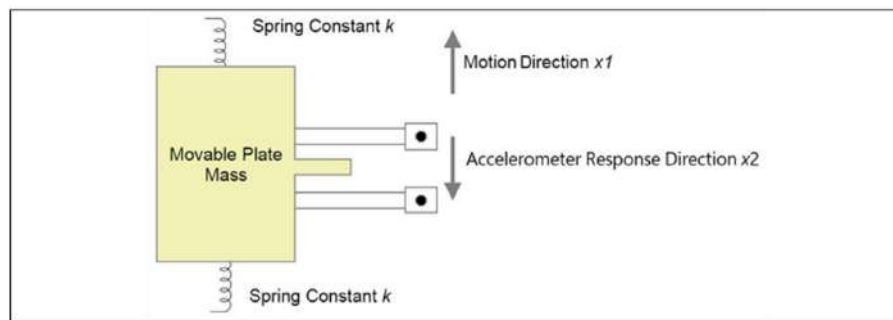


Figura 3.20: Accelerometro

Giroscopi Un giroscopio opera in maniera molto simile ma basandosi sulla misura di forze di Coriolis. La forza di Coriolis è una forza *apparente*, a cui risulta soggetto un corpo quando si osserva il suo moto da un sistema di riferimento che sia in moto rotatorio rispetto a un sistema di riferimento inerziale.

$$a = -2\omega \times v$$

a = accelerazione di Coriolis

v = velocità lineare dell'oggetto

ω = velocità angolare di rotazione

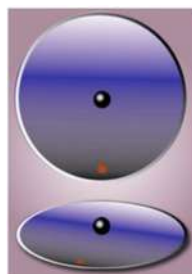


Figura 3.21: Giroscopi

Forza di Coriolis - esempio La forza di Coriolis si manifesta sui corpi che si spostano in direzione nord-sud sulla superficie della Terra.

Tutti i corpi ruotano infatti alla velocità della superficie terrestre, ma la superficie non ruota sempre alla stessa velocità: all'equatore si sposta più velocemente (un punto all'equatore percorre una circonferenza di circa 40 mila km in 24 ore) e decresce andando verso i poli: all'altezza di Milano un punto del terreno percorre nelle stesse 24 ore solo 28 mila km.

Un oggetto non vincolato alla superficie che si sposti dall'equatore verso Milano tenderà a mantenere la stessa velocità che aveva all'origine, quindi a spostarsi verso est dal punto di vista di chi è vincolato alla superficie, al contrario un oggetto che si muova verso l'equatore tenderà a ovest.

Il giroscopio MEMS viene realizzato utilizzando dischi di silicio che messi in rotazione fanno variare una capacità elettrica se sottoposti ad accelerazione di Coriolis.

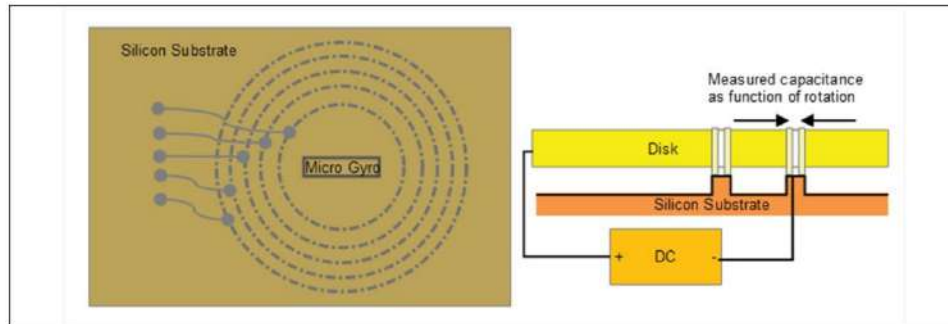


Figura 3.22: Giroscopi-2

Sensori di pressione Usati tipicamente per misurare la pressione di liquidi e gas, nei sistemi IoT si utilizzano ad esempio anche per il monitoraggio di infrastrutture. Il cuore del circuito è un circuito piezoelettrico il quale, cambiando forma, produce differenza di potenziale misurabile.

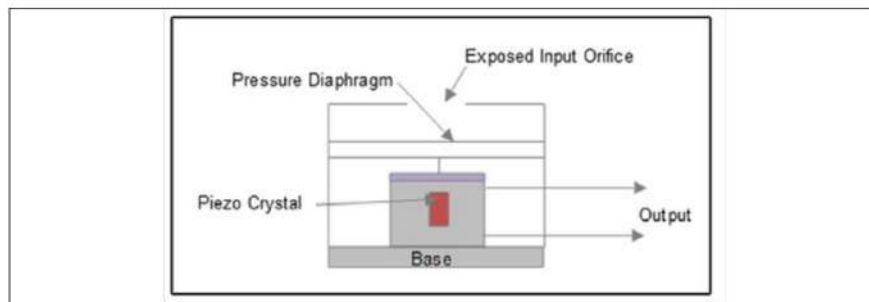
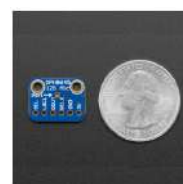
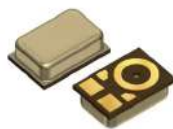


Figura 3.23: Sensore di pressione

Microfoni Sono atti a rilevare suoni o vibrazioni e sono molto utilizzati in industria per prevenire malfunzionamenti o incidenti. Ad esempio in un sistema di mescolamento, anche ad alta velocità di rotazione, la rilevazione di vibrazioni anomale potrebbe essere il segnale di problemi relativi al materiale trattato o difetti nel macchinario.

Un microfono MEMS può essere analogico o digitale.



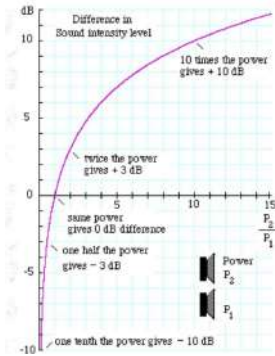
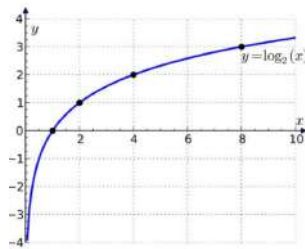
Il Decibel (dB) Il decibel (dieci volte un Bel) è un'attività di misura che esprime il rapporto tra due grandezze.

Non esprime un valore lineare ma logaritmico.

Il decibel è quindi un valore logaritmico adimensionale

Il dB è utilizzato per esprimere diverse grandezze: acustiche, elettriche, etc.

La rappresentazione esponenziale ha dei vantaggi nella rappresentazione di ordini di grandezza elevati.



Alcuni valori notevoli:
se si pone $1 \text{ dB} = 10 \text{ Log } N/D$

$$0 \text{ dB} \longrightarrow N = D$$

$$3 \text{ dB} \longrightarrow N = 2 \times D$$

$$10 \text{ dB} \longrightarrow N = 10 \times D$$

Per ottenere il valore in bel la formula è la seguente:

$$\text{Bell} = \log\left(\frac{p}{p_1}\right)$$

dove:

p = valore da convertire

p_1 = valore di riferimento

Supponiamo che il valore di riferimento sia 10 e che vogliamo convertire in decibel i valori 100, 500, 1000, otteniamo:

$$\log\left(\frac{100}{10}\right) = 1 \text{ Bell}$$

$$\log\left(\frac{500}{10}\right) = 1.69897 \text{ Bell}$$

$$\log\left(\frac{1000}{10}\right) = 2 \text{ Bell}$$

Il decibel non è altro che il Bell moltiplicato per 10, cioè:

$$\text{Decibel} = 10 \text{ Bell} = 10 \log\left(\frac{p}{p_1}\right)$$

Riproponendo gli esempi di prima ma questa volta in decibel:

$$10 \log\left(\frac{100}{10}\right) = 10 \text{ dB}$$

$$10 \log\left(\frac{500}{10}\right) = 16.98 \text{ dB}$$

$$10 \log\left(\frac{1000}{10}\right) = 20 \text{ dB}$$

Gli ambiti di utilizzo dei Decibel sono essenzialmente 3:

I decibel in acustica esprimono i livelli di pressione, intensità e potenza del suono. In questo ambito i più usati sono sicuramente i dBspl che si riferiscono alla pressione sonora.

I decibel in elettronica sono di vario tipo. Ad esempio in ambito audio engineering fanno solitamente riferimento alla rappresentazione elettrica di un segnale audio. Vedremo i dBi, unità di misura utilizzata in ambito radiotrasmissioni.

I decibel in ambito digitale sono i dBfs, che vengono usati per rappresentare il segnale audio all'interno di un sistema digitale.

3.3 Sensori ad Alte Prestazioni

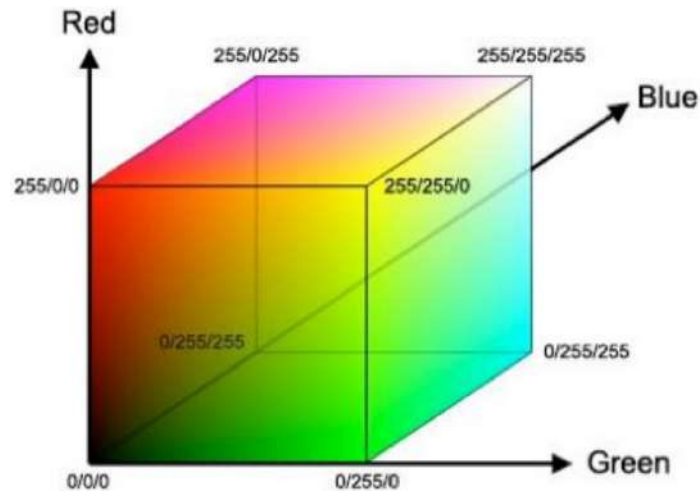


Figura 3.26: RGB-CMY

Spazio colore RGB-CMY

- Spazio di colore definito dalle quantità di luce rossa, verde e blu;
- Lo spazio dei colori è contenuto in un cubo.

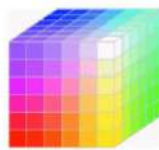
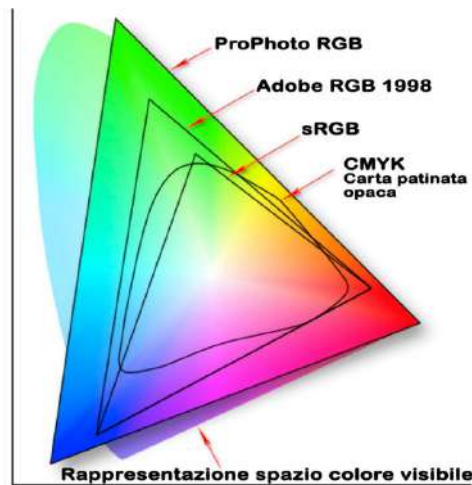


Figura 3.27: Cubo RGB-CMY

È una descrizione matematica della quantità di colore rappresentabile da una periferica, una sorta di carta d'identità.



Sono sensori che hanno elevate capacità di calcolo, parliamo in particolare di visione.

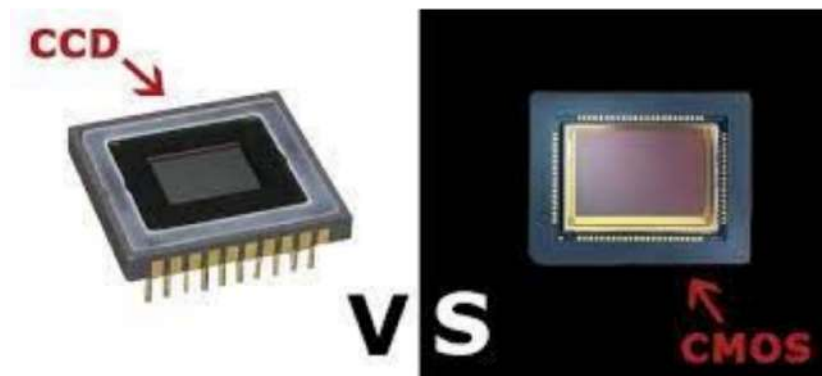


Figura 3.28: Sensori ad alta prestazione

Un sistema di visione è composto da almeno 2 elementi: sistema ottico (lenti) e circuiti sensibili alla luce.

I circuiti sensibili alla luce si distinguono in 2 categorie:

- **CCD** - Charge Coupled Devices;
- **CMOS** - Complementary Metal Oxide Semiconductor.

CCD - generano immagini ad alta risoluzione e con basso rumore, consumano però circa 100 volte più energia di un CMOS, rispetto ai quali richiedono anche processi di fabbricazione più complicati.

CMOS - ogni pixel, costituito da un transistor miniaturizzato, viene "letto" indipendentemente. E' molto più sensibile al rumore rispetto ad un CCD ma consuma molto meno.

CMOS Vengono solitamente organizzati in matrici bidimensionali, con transistor sensibili ai 3 colori fondamentali (RGB) su cui microlenti fanno convergere l'immagine da riprendere.

Necessitano di una pipeline di filtri digitali che effettuano vari tipi di elaborazioni (correzioni), la pipeline solitamente è racchiusa in una ISP = Image Signal Processor.

Le correzioni più comuni sono le seguenti: bilanciamento del bianco, riduzione del rumore, incremento della nitidezza (sharpening), correzioni Gamma (corregge la risposta non lineare dei CMOS), codifica JPEG, etc.

Sono molto più diffusi i sensori CMOS perché meno costosi (processo di fabbricazione più semplice dei CCD).

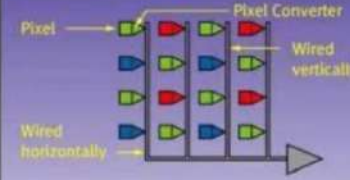
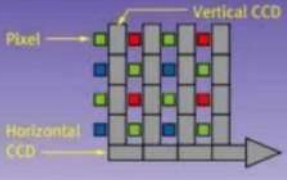
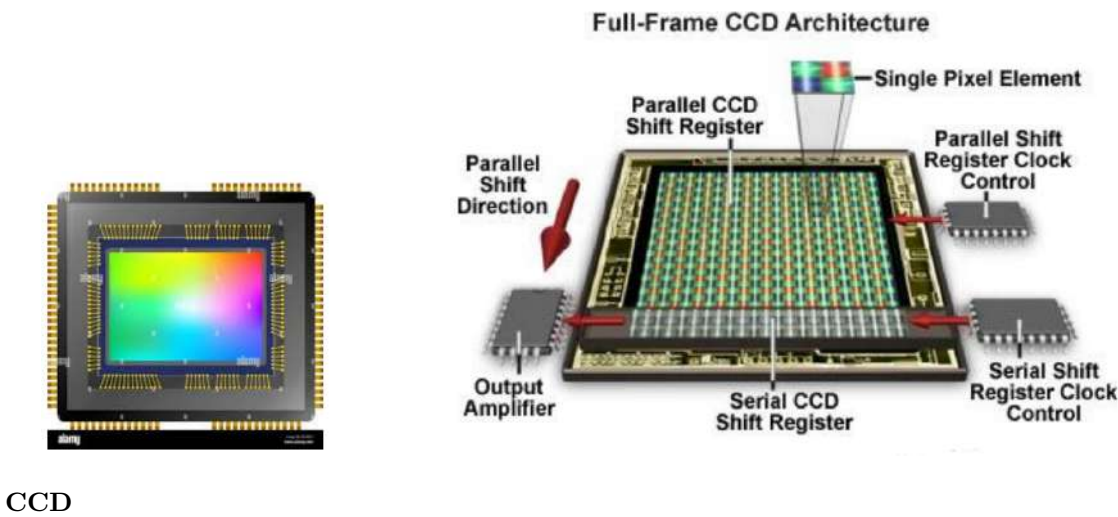
	CMOS Sensor	CCD (Charge Coupled Device)
Modello e lettura del segnale	 <p>Il segnale è letto su ogni pixel, quindi letto.</p>	 <p>Il segnale è letto sul basket di pixel, l'amplificazione è fatta alla fine sul basket</p>
Vantaggi	Basso consumo energetico, molto veloce, circuiteria semplice	Tecnologia affidabile e largamente usata <ul style="list-style-type: none">• Sensore semplificato• Basso rumore, alto rapporto S/N
Svantaggi	Più Rumore <ul style="list-style-type: none">• Pixel irregolari• Rumore random	Maggior consumo energetico, maggiore difficoltà a velocità alte, circuiteria più complessa

Figura 3.29: CMOS vs CCD



CCD

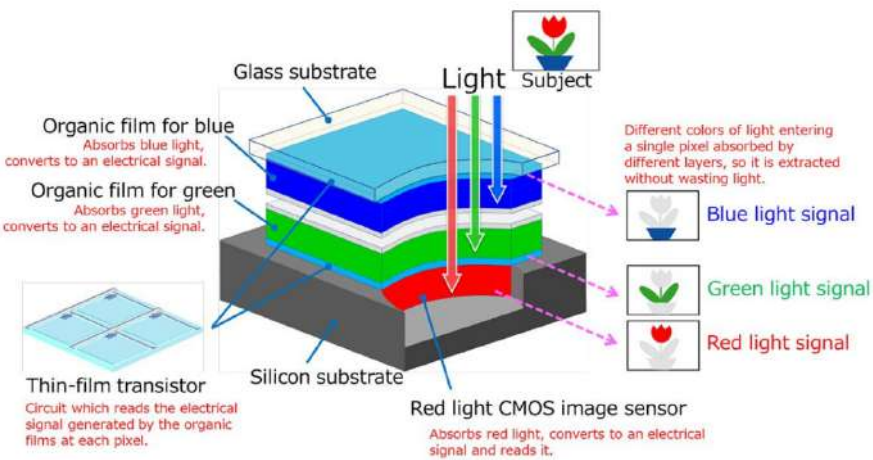


Figura 3.31: Three-layer color image sensor and operating principles

CMOS

Sistemi di visione I sistemi di visione sono largamente utilizzati per controlli di qualità nell'industria e più recentemente, con l'avvento della IA, intervengono anche attivamente nel processo ad esempio verificando l'orientamento di piccoli dettagli o la qualità di un tessuto.



Figura 3.32: Sistemi di visione

3.4 Integrazione di Sensori

E' l'operazione che consente di utilizzare più tipi di sensori diversi per rivelare informazioni più complesse sul contesto.

Ad esempio, in un sistema IoT, un sensore di temperatura da solo non potrebbe darci informazioni riguardo ad un improvviso aumento di temperatura. Se combinato con un sensore PIR per la rilevazione di movimento e un sensore di luce, potrebbe informarci che molte persone si stanno riunendo in un luogo con molto soleggiamento e quindi il sistema potrebbe decidere di incrementare il volume di aria fresca prodotto dal sistema di condizionamento o anche solo di aumentare la circolazione dell'aria o di aprire finestre automatiche o una combinazione di queste cose.

Correlazione dei dati In sostanza se abbiamo più dati correlati tra di loro temporalmente, possiamo probabilmente prendere decisioni migliori, noi o i sistemi di IA.

Questo è uno dei motivi principali per cui ha senso far convergere grandi quantità di dati verso il cloud, cosa che causa anche grande utilizzo di storage digitale.

Ci sono 2 modi di fare integrazione di sensori:

- **Centralizzato:** i dati vengono fatti convergere in un unico punto (cloud) e lì elaborati;
- **Decentralizzato:** i dati sono correlati direttamente a bordo dei sensori o in edge computer.

Dispositivo per prototipazione TI SensorTag CC2650 è un dispositivo della Texas Instruments ed è un buon esempio di sensore modulare per prototipazione e sviluppo di sistemi IoT.



Figura 3.33: TI SensorTag CC2650

In unico chip risiedono:

INPUT

- sensore di luce;
- sensore di temperatura (ambientale e infrarosso);
- sensore di umidità;
- accelerometro;
- giroscopio;
- magnetometro;
- altimetro (sensore di pressione);
- microfono (MEMS);
- 2 bottoni.

OUTPUT

- Speaker (buzzer);
- 2 LEDs.

COMUNICAZIONE

- Bluetooth a bassa energia;
- Zigbee;
- 6LoWPAN.

3.5 Attuatori

I sistemi di output in un ecosistema IoT possono essere una moltitudine di dispositivi, da un semplice LED fino ad un completo sistema video. Includono attuatori, motori passo-passo, altoparlanti, valvole, etc. Il controllo ed il processamento in uscita (output) può essere più o meno centralizzato, ad esempio un sistema video potrebbe richiedere uno streaming che parta da cloud server e venga poi processato dai dispositivi di uscita (schermi) alla frontiera del sistema (edge).

Capitolo 4

Sorgenti ed Accumulatori di Energia

4.1 Sorgenti di Energia

4.1.1 Unità di Misura

Unità di misura di potenza (kW) e energia (kWh).

Dalla definizione di Potenza:

$$P = \frac{E}{t}$$

deriva

$$E = P \times t$$

Ne discende che le unità di misura utilizzate sono: per la potenza il W e suoi multipli, in particolare il chilowatt ($\text{kW} = 10^3 \text{ W}$) per l'energia il Wh, anche in questo caso con il più diffuso multiplo kWh.

ATTENZIONE: per quanto sopra esposto è completamente errata la notazione kW/h.

Abbiamo detto di migliaia di sensori e dispositivi di campo (edge devices) che possono essere utilizzati anche in condizioni estreme e lontano da comode fonti di energia, gestire le sorgenti diventa quindi una sfida ineludibile.

La gestione dell'energia coinvolge sia hardware che software ed è determinante per il successo di un progetto IoT, per cui dobbiamo tenere conto di questi elementi:

- energia assorbita da sensori attivi e passivi;
- energia assorbita dagli attuatori;
- energia assorbita da eventuali sistemi di comunicazione wireless;
- energia assorbita dai microcontrollori (funzione della frequenza dei core);
- frequenza di raccolta dei dati (comunicazione);
- perdite dovute al trasporto e all'inefficienza delle batterie o accumulatori.

Per dimensionare bene la fonte di alimentazione dobbiamo considerare la somma di tutti questi fattori e tenere anche ben presente che, nel caso delle batterie, non abbiamo un comportamento lineare nel tempo per cui il voltaggio decresce man mano che le batterie si avvicinano all'esaurimento.

Inoltre i dispositivi alimentati possono avere caratteristiche particolari per cui, ad esempio, se il voltaggio di alimentazione di un sistema wireless cala, il microprocessore o la parte radio potrebbero non raggiungere più la soglia di operatività e non essere più in grado di funzionare affatto.

Esempio: Il sensore TI SensorTag CC2650
Caratteristiche di alimentazione:

- Standby mode: 0.24 mA;
- Running with all sensors disabled (only powering LEDs): 0.33 mA;
- Running with all sensors on at 100 ms/sample data rate and broadcasting: 12.08 mA;
 - BLE: 5.5 mA;
 - Temperature sensor: 0.84 mA;
 - Light sensor: 0.56 mA;
 - Accelerometer and gyros: 4.68 mA;
 - Barometric sensor: 0.5 mA.

Se viene alimentato con una batteria CR2032 che eroga 240 mAh, la durata massima attesa di funzionamento sarà di circa 20 ore (240 mAh/12.08 mA) e questo senza tenere conto della non linearità della scarica che ridurrebbe ulteriormente la durata di funzionamento.



Figura 4.1: Sensore TI SensorTag CC2650

4.1.2 Gestione dell'Assorbimento

Esistono diverse pratiche per ottimizzare il consumo di energia, ad esempio il *clock gating*, che riduce il clock dei componenti che non vengono utilizzati ma anche lo *sleep mode*, il *dynamic voltage*, il *frequency scaling*, tecniche spesso derivate dal mondo dei computer dove sono ampiamente adottate e la cui efficacia è stata validata da tempo mentre, tra le nuove tecniche citiamo *approximate computing* e *probabilistic design*.

4.2 Approvvigionamento di Energia

Bisogna fare attenzione affinché il sistema non sia mai sotto-alimentato o non alimentato affatto.

I sistemi stessi possono in qualche modo produrre energia ad esempio quando cambiano stato, cambia temperatura, cambia la condizione di illuminazione o vengono “colpiti” da un’onda elettromagnetica.



Figura 4.2: Sistema piezoelettrico

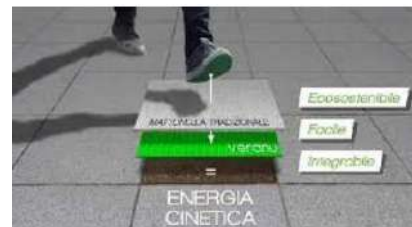


Figura 4.3: Sistema piezoelettrico

Alcuni dispositivi di fatto utilizzano questi cambiamenti come loro unica fonte di energia, altri utilizzano metodi ibridi. Se, ad esempio, un sistema piezoelettrico viene posto al di sotto di un marciapiede per produrre energia dal movimento delle persone, bisogna comunque prevedere un'alimentazione supplementare nei casi in cui non ci sia sufficiente movimento per raggiungere la soglia di alimentazione desiderata (o necessaria).

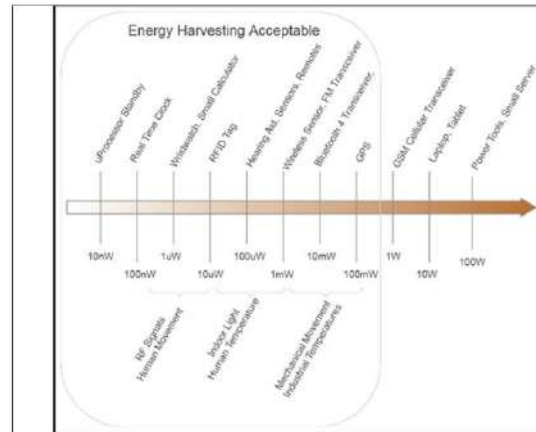


Figura 4.4: Assorbimenti di energia tipici per vari tipi di dispositivi

4.2.1 Energia Solare

L'energia della luce, che sia essa naturale o artificiale, può essere catturata e utilizzata come fonte di energia.

Gli stessi fotodiodi che abbiamo visto utilizzabili come sensori possono essere aggregati in matrici che ne contengano grandi quantità per costruire un pannello solare, chiaramente l'energia prodotta è limitata all'irraggiamento disponibile.

I pannelli così realizzati sono classificati in base al numero massimo di WATT che possono generare.

Nella mappa creata dal Laboratorio nazionale per le energie rinnovabili degli Stati Uniti, vengono graficati i kWh/m² di radiazione solare media giornaliera nell'arco di un anno. Aldilà delle considerazioni ovvie sul fatto che un pannello solare posizionato in Alaska produrrebbe meno energia di un pannello posizionato ai confini con il Messico, dalla colorazione si evidenzia immediatamente che anche un pannello installato a sud ovest degli Stati Uniti potrebbe produrre molta più energia di un pannello posizionato sulla costa opposta pur trovandosi magari alla stessa latitudine.

I pannelli fotovoltaici non sono in genere molto efficienti, hanno prestazioni che variano tra 8% e 20% con un 12% tipico., inoltre questa efficienza viene verificata solo se la luce incidente è perpendicolare, ad esempio se avessimo una efficienza dichiarata del 12%, scenderebbe al 9,6% quando il sole è posizionato a 30° rispetto alla verticale.

L'efficienza diminuisce ulteriormente con il passare del tempo.



Figura 4.5: Mappa Energia Rinnovabile

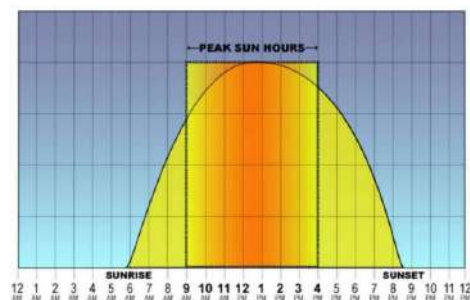


Figura 4.6: Prestazioni Pannello Solare

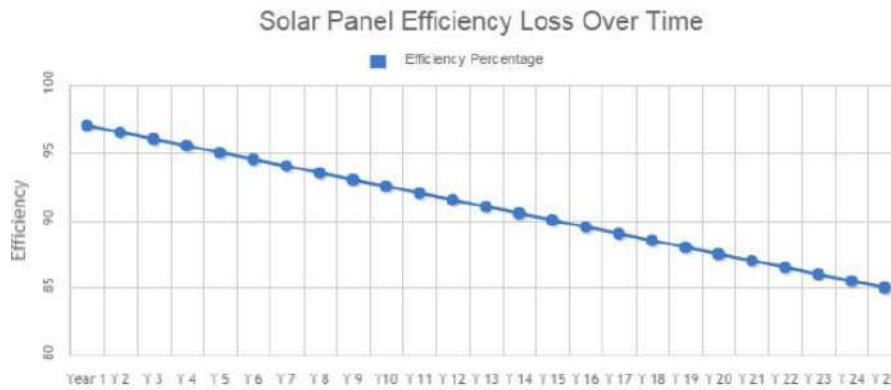


Figura 4.7: Efficienza Pannello Solare

4.2.2 Energia da Dispositivi Piezo-Meccanici, Elettrostatici, Elettromagnetici

Anche in questo caso, come per i fotodiodi, un dispositivo che abbiamo visto poter essere utilizzato come sensore, può essere utilizzato anche come generatore di energia. Possiamo produrre energia da movimento, vibrazioni e persino suono. L'energia prodotta è dell'ordine dei mWatt che è adatta per dispositivi molto piccoli che abbiamo anche una qualche capacità di accumulo di energia. Possono essere dispositivi piezo-meccanici ma anche sistemi elettrostatici o sistemi elettromagnetici

4.2.3 Energia da RF

Viene utilizzata da anni nei dispositivi RFID i quali, trattandosi di oggetti che operano in prossimità del transceiver ("antenna" che deve leggere il dato contenuto in RFID), beneficiano di questa vicinanza per poter essere alimentati per mezzo dell'onda elettromagnetica incidente emessa dal rilevatore.

Per applicazioni in cui invece non vi sia vicinanza (far-field) viene raccolta energia dalle emissioni broadcast (segnali televisivi, telefonici, radio). Le energie in gioco sono tra le minori nel panorama completo della produzione di energia, perché questo tipo di segnali a radiofrequenza hanno una bassissima densità di energia.

Tipicamente si raccolgono segnali nella banda 531 KHz – 1611 KHz (nella banda radio della modulazione di ampiezza, AM).

4.2.4 Energia a Gradienti Termici

É possibile produrre energia elettrica ovunque ci sia un gradiente di temperatura, in 2 modi:

- Sfruttando fenomeni termoelettrici (l'effetto Seebeck);
- Sfruttando fenomeni termoionici (noto anche come thermotunneling).

Gli effetti termoelettrici li abbiamo già visti sfruttandoli per costruire sensori di temperatura, in questo contesto piuttosto che parlare di termocoppie si parla di TEG = ThermoElectric Generator.

Possono generare una debole corrente semplicemente sfruttando il delta temperatura tra il corpo umano e la temperatura ambiente. L'energia prodotta è proporzionale al quadrato della differenza di potenziale (volt) e proporzionale alla differenza di temperatura tra i 2 elettrodi. Quanta energia possiamo generare? Una differenza di 5°C può generare circa 40 uW a 3V. Arriviamo quindi ad alimentare al massimo un RFID tag.

Un TEG con una differenza di 5°C può generare circa 40 uW a 3V e potrà alimentare al massimo un dispositivo della classe degli RFID.

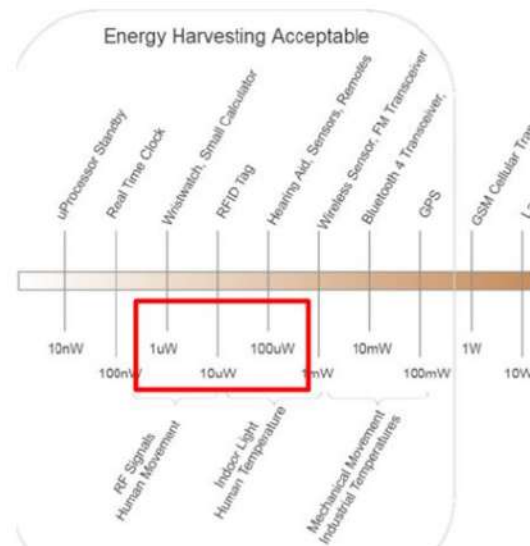


Figura 4.8: Accumulo Energia Accettabile

Siccome una termocoppia produce differenze di potenziale molto piccole, vengono solitamente messe in serie diverse termocoppie (si sommano le differenze di potenziale) a formare una Termopila. Il maggior problema di questi dispositivi è la bassa efficienza, mediamente minore del 10% ma sono molto piccoli, poco costosi e duraturi nel tempo (fino a 100.000 ore) inoltre, per poter generare energia con una certa costanza, occorre posizionarli in luoghi dove la differenza di temperatura sia anch'essa piuttosto costante.

I dispositivi termoionici si basano sul fenomeno fisico per cui gli elettroni vengono espulsi da un elettrodo caldo verso un elettrodo freddo attraverso una barriera di potenziale e sono più efficienti delle termocoppie ma l'energia richiesta per superare la barriera di potenziale, che è direttamente proporzionale alla differenza di temperatura tra i due elettrodi, è in genere così elevata che li rende non adatti ad operare in ambito IoT. Sono allo studio anche dispositivi basati su effetto di Tunneling quantico con esiti ancora tutti da verificare.

4.2.5 Energia da Radioattività

Una sorgente di energia radioattiva ad elevate densità di energia (105kJ/cm^3) può generare energia termica dall'energia cinetica delle particelle emesse nel range da Watt a KWatt. Sorgenti come il cesio-137 hanno un'emivita di 30 anni con una capacità di 0.015 W/gm .

[emivita = negli elementi chimici radioattivi è il tempo in cui decade metà della massa iniziale dell'elemento stesso.]

Altro problema per l'utilizzo effettivo è che il decadimento radioattivo ha un profilo di densità di potenza molto debole a causa dei lunghi tempi ma si può pensare di utilizzarli per caricare dei supercondensatori che provvedano poi ad erogare energia quando necessario. Ultimo problema è che dovendo avere schermature in piombo, aumenta molto il peso ed anche il costo del dispositivo: il cesio-137 richiede 80mm/W di schermatura.

Sono fonti di energia che vengono utilizzate spesso, e da tempo, nei veicoli spaziali come nel caso del rover marziano Curiosity e della navicella New Horizon



Figura 4.9: Rover Curiosity

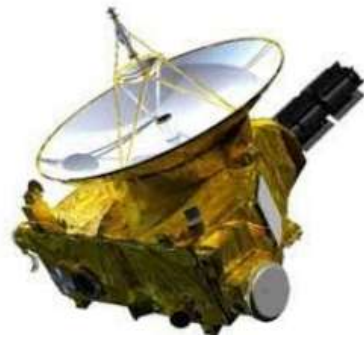


Figura 4.10: Navicella New Horizon

4.3 Accumulatori di Energia

La capacità di una batteria si misura in Ampere all'ora (A/h). Le stime sulla durata di una batteria si basano tutte sull'effetto Peukert: la capacità di una batteria decresce a velocità diverse quando si incrementa la scarica cioè scaricando velocemente una batteria viene rimossa maggiore capacità rispetto alla scarica lenta della stessa batteria.

Vediamo un esempio:

Se si impiegano 20 ore per scaricare una batteria da 100 Ah, ad esempio collegandola ad un carico resistivo che utilizza 5 A/h, se colleghiamo un carico resistivo che utilizza 10 A/h la durata sarà inferiore alle 10 ore attese (calcolate algebricamente) così come scaricandola più lentamente, carico resistivo da 2,5 A/h la durata sarà maggiore di 40 ore.

A seconda del materiale che compone la batteria, si possono disegnare differenti curve “di scarico”.

- **Batterie alcaline** – hanno un andamento lineare per un grande range di scarica;
- **Batterie agli ioni di Litio** – hanno un andamento lineare ma presentano uno “scalino” in prossimità della scarica completa che rende molto difficile predire il punto in cui non saranno più in grado di alimentare il dispositivo;
- **Batterie al Nichel Cadmio** – hanno meno differenza di potenziale ma un andamento di scarica curvilineo che è più facilmente prevedibile.

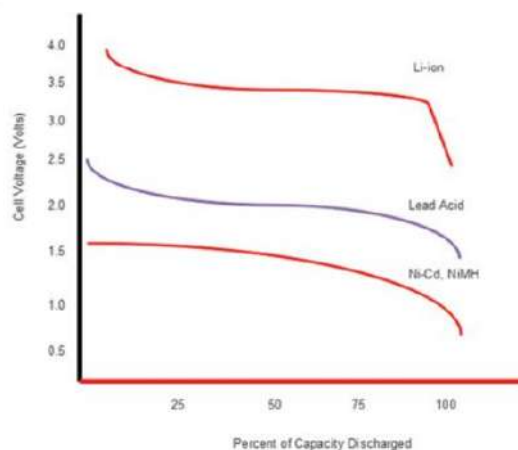


Figura 4.11: Percentuale Scaricata della Capacità

Diagramma di Ragone E' disegnato in scala logaritmica dove la densità di energia (Wh/kg) di una sorgente è confrontata con la densità di potenza (W/kg) e mostra anche il tempo di scarica per vari dispositivi accumulatori. Il diagramma evidenzia i dispositivi che durano più a lungo (batterie) rispetto a dispositivi che immagazzinano più energia (supercondensatori). Ad esempio le batterie agli ioni di litio

hanno più alta densità di energia e maggior velocità di scarica delle nichel-cadmio o nichel-ibrido.

Un diagramma di Ragone è un grafico utilizzato per comparare la densità di energia di vari dispositivi di accumulo.

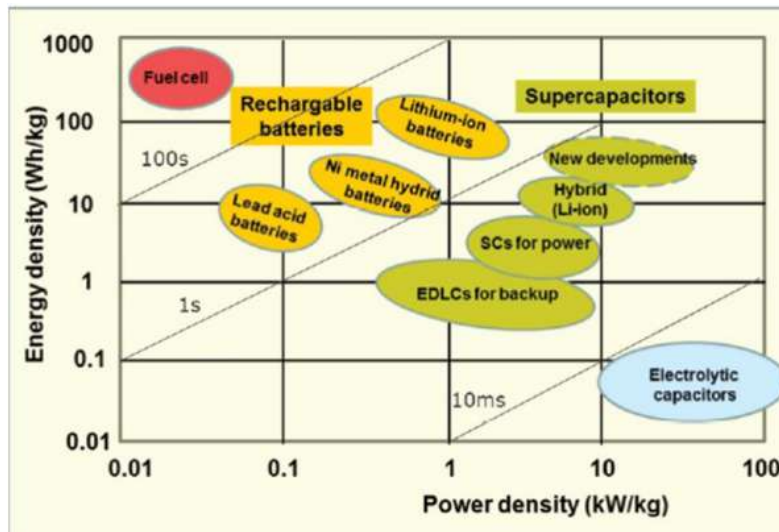


Figura 4.12: Esempio Diagramma di Ragone

4.3.1 Batterie Alcaline

Sappiamo che quando utilizziamo una batteria, questa inizia a scaricarsi. Quanto velocemente si scarica, dipende dal carico che applichiamo alla batteria e quindi dall'intensità della corrente di scarica, ovvero dalla potenza assorbita dall'utilizzatore. Nel caso di una batteria alcalina, se la utilizzo per applicazioni che assorbono molta corrente, diciamo 2A, mi fornisce solo $1.5V \times 2.5Ah = 3.75Wh$, ovvero la metà dell'energia fornita dalla batteria di un cellulare (che è agli ioni di Litio). La pila alcalina è eccellente (a parte il costo elevato) per applicazioni a basso assorbimento di corrente, come radio, giocattoli, piccole luci etc..

Guardando le specifiche di una cella D alcalina, per esempio, notiamo che la capacità in mAh diminuisce drasticamente quando aumenta la corrente assorbita. Passando dai 18000 mAh, quando praticamente non assorbo corrente a 2500mAh di quando spillo 2A dalla batteria.

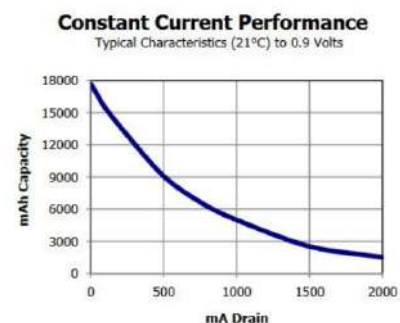


Figura 4.13: Performance Corrente Costante

4.3.2 Batterie agli Ioni di Litio

Una batteria agli ioni di Litio è la scelta tipica per alimentare dispositivi mobili grazie alla sua densità di energia. In questo tipo di batteria gli ioni di litio si muovono fisicamente dall'elettrodo negativo al positivo durante la fase di scarica mentre si muovono in verso opposto durante la carica (movimento ionico).

Dopo una serie di cicli di carica/scarica, le batterie sviluppano un effetto memoria che comporta perdita di capacità anche del 30% dopo 1.000 cicli. Questa perdita di capacità è anche collegata alla temperatura ambiente e quindi diventa importante considerare le condizioni climatiche del luogo in cui andremo a posizionare i dispositivi alimentati da questo tipo di batterie.

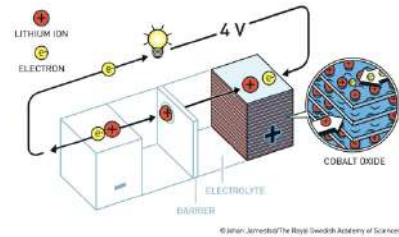


Figura 4.14: Caption



Figura 4.15: Caption

Un ultimo fattore che incide in maniera negativa sulla capacità della batteria è costituito dalle reazioni chimiche indesiderate perché comportano perdita di energia. La velocità della perdita dipende dalla composizione chimica e, ancora una volta, dalla temperatura. Una batteria agli ioni di litio ha perdite chimiche che incidono circa il 2% al mese il che comporta una durata complessiva di 10 anni mentre una batteria alcalina ha perdite del 15% circa al mese e quindi durerà anni di meno.

4.3.3 Supercondensatori

I supercondensatori immagazzinano energia in volumi molto più grandi dei condensatori. Un tipico condensatore ha densità di energia vicina a 0.01 Wh/kg mentre un supercondensatore ha densità di energia da 1 a 10 Wh/kg il che li posiziona nella zona di densità di energia tipica delle batterie (200 Wh/kg).



Figura 4.16: Supercondensatore



Figura 4.17: Supercondensatore

L'energia è immagazzinata elettrostaticamente su una piastra e non è coinvolto nessun processo chimico.

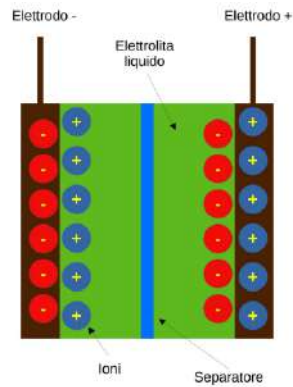


Figura 4.18: Processo Chimico

Sono dispositivi costruiti con materiali “esotici”, come ad esempio il grafene, il che impatta ovviamente sul costo, hanno però il vantaggio di caricarsi al limite anche in pochi secondi dove batterie agli ioni di litio impiegano invece minuti per raggiungere 80% della loro capacità.

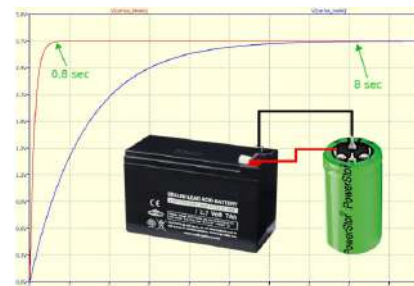


Figura 4.19: Carica Supercondensatore

Nelle condizioni reali il supercondensatore si carica in circa 8 secondi, un tempo sempre veloce e di tutto rispetto. Le curve del grafico evidenziano l'andamento della tensione di ricarica sul condensatore. Nel caso ideale, la resistenza che incontra la corrente è solamente quella del condensatore (0,00023 Ohm). Nel caso reale, invece, la resistenza è quella del condensatore (0,00023 Ohm) più quella interna della batteria (0,0025 Ohm).

Di contro i supercondensatori non possono essere sovraccaricati mentre le batterie agli ioni di litio, se sovraccaricate, possono risultare anche molto pericolose. I supercondensatori si presentano in 2 forme:

- **Condensatori a doppio strato (Electric double-layer capacitors = EDLC):** utilizzano un elettrodo al carbone attivo e immagazzinano energia solo elettrostaticamente;
- **Pseudocondensatori:** utilizzano un metallo e trasferiscono carica in maniera elettrochimica.

I supercondensatori hanno anche un altro notevole vantaggio rispetto alle batterie: l'energia rimanente (e quindi il tempo di utilizzo rimanente) può essere dedotto dal cambiamento di differenza di potenziale dei terminali. Come abbiamo visto, le batterie agli ioni di litio hanno un profilo di scarica piatto e questo rende difficile la stima del tempo (energia) rimanente. Invece l'andamento della differenza di potenziale (voltaggio) di un supercondensatore cambia continuamente dandoci informazioni sullo stato di scarica (energia disponibile rimanente).

Naturalmente, poiché il circuito alimentato dal supercondensatore richiede quasi sempre un voltaggio costante, occorrerà inserire un DC-DC converter per compensare le ampie variazioni di potenziale. Per concludere, i problemi principali connessi ai supercondensatori sono le basse correnti ed il costo ma hanno comunque un loro ruolo all'interno di un sistema IoT, ad esempio in soluzioni ibride dove occorra una corrente di spunto alta (veicolo elettrico in accelerazione) e poi possano intervenire batterie tradizionali per sostenere la potenza più a lungo.

Category	Li-Ion Battery	Supercap
Energy density	200 Wh/kg	8-10 Wh/kg
Charge-discharge cycles	Capacity drops after 100 to 1,000 cycles	Nearly infinite
Charge-discharge time	1 to 10 hours	Milliseconds to seconds
Operational temperature	-20°C to +65°C	-40°C to +85°C
operational voltage	1.2V to 4.2V	1V to 3V
Power delivery	Constant voltage over time	Linear or exponential decay
charge rate	(Very slow) 40 C/x	(Very fast) 1,500 C/x
operational life	0.5 to 5 years	5 to 20 years
Form factor	Very small	Large
Cost (\$/kWh)	Low (\$250 to \$1,000)	High (\$10,000)

Figura 4.20: Tabella Comparativa Batterie e Supercondensatori

Parte III

Tecnologie per la Comunicazione

Capitolo 5

Non Basate su Protocollo IP

5.1 Elementi di base dei sistemi WPAN

Sensori ed attuatori necessitano di un metodo per trasmettere e ricevere informazioni. Questo è il dominio della PAN (Personal Area Network) e comunicazioni a corto raggio.

In un ecosistema IoT la comunicazione può avvenire attraverso cavi in rame oppure via radio (senza fili = wireless).

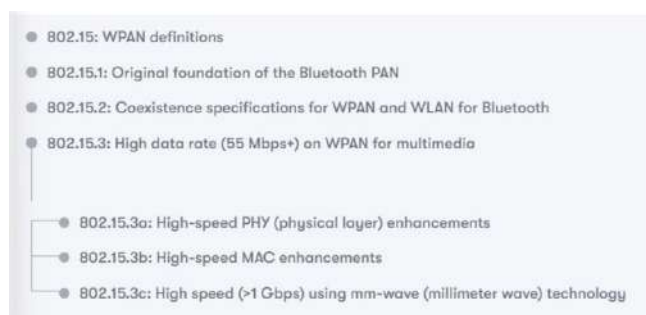
La WPAN, dove la W sta per **WIRELESS**, è il metodo prevalente mentre la comunicazione via cavo viene utilizzata prevalentemente in aree dove la radio frequenza potrebbe essere disturbata.

Andando dall'endpoint (solitamente il sensore o l'attuatore) verso internet, si possono attraversare diversi metodi di collegamento da quelli costruiti sul classico stack IP (ad esempio 6LoWPAN) a quelli che non utilizzano protocollo IP per minimizzare il consumo di energia (ad esempio BLE = Bluetooth Low Energy).

Gli standard 802.15 Molti dei protocolli e modelli di rete che andremo a descrivere sono basati su lavori dei gruppi partecipanti alla stesura degli standard IEEE 802.15.

802.15 nasce per studiare dispositivi indossabili e conia l'acronimo PAN ma ormai il progetto si è notevolmente espanso e ora si concentra su protocolli per l'alta velocità di trasmissione dati, distanze dal metro a chilometri, etc.

Attualmente vengono prodotte circa 1 milione di dispositivi al giorno che hanno a bordo qualche protocollo appartenente al progetto 802.15, quella che segue è una lista dei vari sottoprogetti di 802.15 che IEEE aggiorna e sviluppa.



- 802.15.5: Mesh networking
- 802.15.6: Body area networking for medical and entertainment
- 802.15.7: Visible light communications using structured lighting
- 802.15.7a: Extends range to UV and near-IR, changed name to optical wireless
- 802.15.8: Peer Aware Communications (PAC) infrastructure-less peer to peer at 10 Kbps to 55 Mbps
- 802.15.9: Key Management Protocol (KMP), management standard for key security
- 802.15.10: Layer 2 mesh routing, recommend mesh routing for 802.15.4, multi PAN
- 802.15.12: Upper layer interface, attempts to make 802.15.4 easier to use than 802.11 or 802.3

5.2 Il protocollo Bluetooth

Bluetooth è una tecnologia di collegamento wireless e a basso consumo usata pervasivamente in telefoni cellulari, sensori, periferiche (tastiere, mouse, stampanti, etc.), consolle per videogiochi, etc.

Il nome discende dal re Harald Blatand vissuto intorno al 958 d.c. in una zona che ora è identificabile con Norvegia e Svezia. Dente blu (bluetooth) sembra derivi dal fatto che amava cibarsi di mirtili o dei suoi nemici congelati ma aldilà di questo è stato scelto perché seppe ricomporre le tribù che al tempo erano in lotta tra di loro esattamente come è avvenuto con la iniziale formazione del Bluetooth SIG.

Bluetooth SIG Bluetooth Special Interest Group (BSIG) è un'organizzazione che sovrintende lo sviluppo degli standard Bluetooth e la concessione in licenza delle tecnologie e dei marchi Bluetooth ai produttori.

La SIG è una società senza fini di lucro, senza azioni, fondata il 20 maggio 1999.

La SIG non produce o vende prodotti abilitati Bluetooth.



5.2.1 Storia del Bluetooth

Nasce in Ericsson nel 1994 con l'intento di sostituire i cavi necessari a connettere le periferiche ad un PC.

Intel e Nokia si uniscono ad Ericsson espandendo lo scopo al collegamento di telefoni cellulari ai PC.

Nel 1996, nella sede di Ericsson in Svezia, le 3 aziende fondano il SIG. Nel 1998 i membri del Bluetooth SIG diventano 5 perché si aggiungono IBM e Toshiba e rilasciano la versione 1.0 delle specifiche Bluetooth.

Nel 2005 il SIG conta 4.000 membri e viene ratificata la versione 2.0 e successivamente, nel 2007 viene sviluppato l'Ultra Low Power Bluetooth che ora è conosciuto come BLE (Bluetooth Low Energy) e questo apre interi nuovi segmenti di mercato perché ora è possibile alimentare dispositivi con una piccola batteria a bottone (CR2032).

Nel 2010 sarà poi rilasciata la versione 4.0 che include ufficialmente il BLE.

Attualmente ci sono 2,5 miliardi di prodotti con a bordo Bluetooth e il SIG conta 30.000 membri.

Bluetooth è utilizzatissimo nello sviluppo di sistemi IoT, soprattutto in profilo LE, per sistemi di marketing (beacon), sensori, sistemi di tracciamento, controlli remoti, dispositivi per il monitoraggio medico, sistemi di allarme.

Il successo è da attribuire principalmente alla lungimiranza del progetto iniziale, alla pervasività nei dispositivi mobili e al tipo di licenza perché in tutti i passaggi dello sviluppo è rimasto sotto licenza GPL

e quindi essenzialmente in open source.

Version	Features	Release Date
Bluetooth 1.0 and 1.0B	Basic rate Bluetooth (1 Mbps) initial version released	1998
Bluetooth 1.1	IEEE 802.15.1-2002 standardized 1.0B specification defects resolved Non-encrypted channel support Received signal strength indicator (RSSI)	2002
Bluetooth 2.0	IEEE 802.15.1-2005 Rapid connection and discovery Adaptive frequency hopping spread spectrum (AFH) Host controller interface (three-wire UART) Flow control and retransmission modes	2003
Bluetooth 2.0 +EDR (optional)	Enhanced Data Rate Mode (EDR): 3 Mbps	2004
Bluetooth 2.1 +EDR (optional)	Secure Simple Pairing (SSP) using public key cryptography with four unique authentication methods Extended Inquiry Response (EIR) allows for better filtering and reduced power	2007

Bluetooth 2.0 +EDR (optional) +HS (optional)	LEAP enhanced retransmission mode (ERM) for reliable and scalable connection states Alternate MAC/PHY (AMP) 28 kbps using IEEE 802.11 PHY for unicast connectionless data for low latency Enhanced power control	2009
Bluetooth 4.0 (LE optional) +HS (optional) +LL (optional)	AKB BluetoothSmart Introduced Low Energy mode (LE) Introduced ATT and GATT protocols and profiles Dual mode: BR/EDR and LE mode Security manager with AES encryption	2010
Bluetooth 4.1	Mobile wireless service (MWS) coexistence Train nudging (coexistence feature) Interleaved scanning (coexistence feature) Devices support multiple simultaneous roles	2013
Bluetooth 4.2	LE secure connections Link layer privacy IPsec support profile	2014
Bluetooth 5.0	Slur availability mode (SAM): 2 Mbps PHY and LE LE long-range mode LE extended advertising modes Mesh networking	2016

Bluetooth 5.1	Direction finding GATT caching Randomized Advertising Channel Indexing Periodic advertising sync transfer	2019
---------------	--	------

5.3 Il protocollo Zigbee

Zigbee è un protocollo per WPAN basato su IEEE 802.15.4, nato per rispondere ai vincoli di costo, ingombro e alimentazione, tipici dei sistemi IoT, e destinato all'ambito commerciale.

Il nome è onomatopeico per un insetto volante, perché come un insetto vola di fiore in fiore portando polline, così il protocollo Zigbee trasporta pacchetti in una rete mesh, da dispositivo a dispositivo.

Opera nelle frequenze radio assegnate per scopi industriali, scientifici e medici (ISM, 868 MHz in Europa, 915 MHz negli Stati Uniti e 2,4 GHz nella maggior parte del resto del mondo), ma di fatto oggi le uniche vere implementazioni disponibili sul mercato sono quelle a 2,4 GHz.

E' una tecnologia nata con lo scopo di essere più semplice e più economica di Bluetooth prima che Bluetooth diventasse LE (Low Energy).

5.3.1 Storia di Zigbee

La Zigbee Alliance si forma nel 2002 con lo scopo di realizzare reti wireless a basso consumo, un concetto nato negli anni '90.

Viene ratificato nel 2004 con il progetto IEEE 802.15.4 e reso pubblico il 13 Giugno 2005, la road map è stata questa:

- 2005: rilasciato Zigbee 2004;
- 2006: rilasciato Zigbee 2006;
- 2007: rilasciato Zigbee 2007, conosciuto anche come Zigbee pro.

La relazione tra la Zigbee Alliance e il gruppo IEEE 802.15.4 è simile alla relazione esistente tra la Wi-Fi Alliance ed il gruppo IEEE 802.11:

Il gruppo 802.15.4 si occupa dei livelli 1 e 2 della pila ISO/OSI (livelli fisico e data link) mentre la Zigbee Alliance di tutto il resto (mantiente e pubblica gli standard, organizza la ricerca e gestisce la lista di applicazioni).

5.3.2 Zigbee-tipo di standard e aspetti tecnici peculiari

Zigbee è uno standard proprietario e chiuso, quindi richiede il pagamento di una licenza e una approvazione da parte della Zigbee Alliance ma, in cambio, offre la possibilità di utilizzare il logo e garantisce la compatibilità, cioè la interoperabilità con dispositivi a protocollo Zigbee prodotti da altri.

Zigbee ha da sempre avuto come obiettivo la semplicità e questo, nei calcoli della alleanza si traduce in un risparmio del 50% supporto software.

Non ha un servizio di trasporto dati o un ambiente per l'esecuzione delle applicazioni ma è in grado di formare una rete, individuare altri dispositivi, fornire un primo livello di sicurezza e gestire la rete stessa.

5.3.3 Elementi dell'ecosistema Zigbee

In una rete Zigbee ci sono 3 elementi costituenti:

Zigbee controller (ZC): è un dispositivo con buone capacità computazionali che è utilizzato per formare la rete iniziale. Ogni rete Zigbee può avere un solo ZC e, una volta che la rete è stata instaurata, potrà poi comportarsi da Zigbee Router (ZR). Assegna indirizzi di rete logici e consente agli altri dispositivi di unirsi alla rete o abbandonarla.

Zigbee router (ZR): è un componente opzionale ma, comportandosi da instradatore e coordinatore, alleggerisce in parte il carico della rete mesh. Partecipa all'instradamento multi-hop dei messaggi e, come fa lo ZC, consente agli altri dispositivi di unirsi alla rete o abbandonarla.

Zigbee end device (ZED): è normalmente un dispositivo endpoint, ad esempio una lampada o un termostato. Ha solo la capacità di comunicare con il coordinatore (ZC) ma non ha logiche di instradamento (routing) e non può gestire autonomamente l'associazione ad una rete Zigbee. Ogni messaggio ricevuto da uno ZED che non è destinato a lui viene semplicemente eliminato.

I dispositivi Zigbee comunicano utilizzando 3 differenti modalità **Dati ciclici (o periodici):** il dato è inviato o ricevuto ad intervalli di tempo regolari definiti dall'applicazione (ad esempio un sensore che trasmette periodicamente la sua misura).

Dati ciclici a bassa latenza: vengono allocate delle finestre temporali (time slot) per la trasmissione che possono avere latenze anche molto basse, ad esempio gli eventi prodotti da un mouse o una tastiera di un PC.

Dati intermittenti: vengono inviati quando una applicazione o uno stimolo esterno creano un evento ad un tempo qualunque (random), ad esempio un interruttore.

Zigbee può operare con 3 topologie di base (che possono essere combinate tra loro)

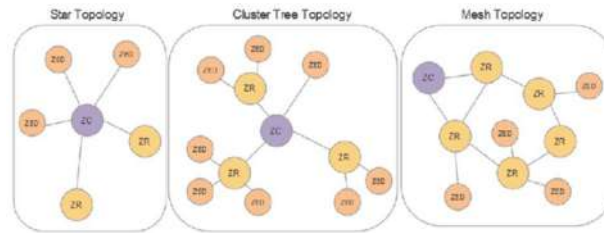


Figura 5.1: Topologie di Zigbee

- **Rete a stella:** Un solo ZC con ZED periferici, quindi un solo salto (hop) e questo limita la distanza massima tra ZC e ZED inoltre lo ZC rappresenta un unico punto di fallimento.
- **Rete ad albero:** Una rete multi-hop che impiega anche segnalazioni (beaconing) utilizzata per superare i limiti di estensione della topologia a stella. ZC e ZR possono avere “figli” e gli ZED rimangono endpoints che possono comunicare solo con i loro “genitori” mentre questi ultimi possono comunicare a valle verso i “figli” e a monte verso altri “genitori”. Anche in questo caso però abbiamo il problema del singolo punto di fallimento (SPOF) al centro.
- **Rete a maglia (mesh network):** I percorsi di collegamento sono stabiliti in maniera dinamica e l'instradamento può avere luogo da qualunque sorgente verso qualunque destinazione utilizzando algoritmi di instradamento basati su tabelle. Per eseguire questi compiti ZC e ZR possono essere attivati in qualunque momento e questo può comportare assorbimenti di energia importanti, riducendo la durata delle batterie. Inoltre, in una rete mesh, non è semplice calcolare la latenza perché diventa non deterministica. I vantaggi principali di una rete mesh sono 2: può crescere praticamente in maniera indefinita ed è meno sensibile all'interruzione di connessioni perché esistono più percorsi verso un qualunque dispositivo (ridondanza).

5.3.4 Come un dispositivo Zigbee entra a far parte di una rete (associare un dispositivo Zigbee)

Uno ZED non è in grado di fare instradamento ma comunica con un “genitore” che è anche un router. Se lo ZED vuole entrare a far parte di una rete Zigbee invia una richiesta broadcast a tutti i dispositivi autorizzati a far entrare ZED nella rete, chiedendo di inviare a loro volta un segnale di conferma. Inizialmente solo lo ZC è autorizzato a rispondere ma, con il crescere della rete, possono intervenire altri ZC o ZR.

Quando uno ZC consente ad un nuovo dispositivo di entrare a fare parte della rete attiva la procedura di “associazione”.

Se uno ZED perde contatto con il “genitore” può riassociarsi attraverso un processo detto “orfano”.

5.3.5 Zigbee e sicurezza

Zigbee realizza, per la sicurezza, quanto indicato nella IEEE 802.15.4 implementando 3 meccanismi diversi: ACLs, cifratura 128-bit AES e marcatori di durata del messaggio (message freshness timers).

Il modello di sicurezza è distribuito su quasi tutti i livelli della pila ISO/OSI:

Livello applicazione: creazione delle chiavi e servizi di trasporto verso ZDO.

Livello rete: viene utilizzata una chiave per il collegamento definita dall'instradamento, se disponibile, altrimenti viene utilizzata una chiave di rete.

Livello data link: gestita tramite una API e controllata dai livelli superiori.

Zigbee utilizza anche diversi tipi di chiavi:

Master key: può essere preinstallata dal costruttore o inserita con processo manuale dall'utente. **Network key:** procura una protezione a livello di rete. **Link key:** procura un'associazione sicura tra 2 dispositivi. Se i dispositivi hanno la possibilità di scegliere tra Link Key o Network Key scelgono sempre la Link Key che garantisce migliore sicurezza.

5.4 Il protocollo Z-wave e i sistemi HVAC

Z-Wave Z-Wave è una tecnologia WPAN che opera alla frequenza di 900MHz utilizzata principalmente per l'automazione domestica.

Nel 2021 si contavano circa 2.100 prodotti diversi che la utilizzavano, soprattutto in sistemi di illuminazione e controlli HVAC (Heating, Ventilation e Air Conditioning) ma di fatto ha una presenza sul mercato nettamente inferiore rispetto ai sistemi Bluetooth o Zigbee.

5.4.1 Sistemi HVAC

HVAC (Heading Ventilation Air Conditioning) sono sistemi per il riscaldamento, la ventilazione e la climatizzazione dell'aria.

Nell'ambito della building automation e del condizionamento industriale vengono progettati e utilizzati per il controllo e il monitoraggio dei sistemi di combustione, riscaldamento, ventilazione, e condizionamento dell'aria, da cui dipendono il comfort ambientale e i livelli di sicurezza.



Figura 5.2: HVAC

5.4.2 I meccanismi su cui si basa

Tutti i sistemi HVAC si basano sulla progettazione, lo sviluppo e la fornitura di componenti hardware e software con un elevato livello di connettività e digitalizzazione, da utilizzare per lo spostamento dell'aria tra le aree interne ed esterne, ma anche per il riscaldamento e raffreddamento di edifici sia residenziali che commerciali.

L'obiettivo di ogni sistema HVAC è quello di raggiungere il comfort termico per gli occupanti, garantendo il risparmio per quanto riguarda i costi di installazione, operazione e manutenzione, monitorando i consumi di energia negli ambienti.

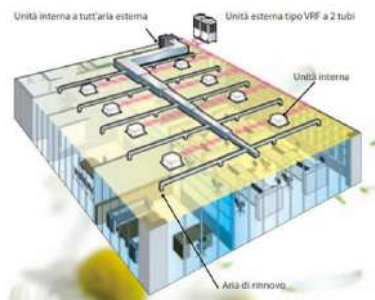


Figura 5.3: Esempio di sistema per ricambio aria in ambito terziario

5.4.3 I vantaggi di HVAC per l'ambiente e il risparmio energetico

A patto di effettuare una corretta manutenzione del sistema, gli impianti HVAC garantiscono un sensibile risparmio nei livelli di consumo energetico.

Altri vantaggi legati all'applicazione di questo tipo di impianto di condizionamento sono:

- controllo ottimale della temperatura;
- possibilità di utilizzo del free-cooling nelle mezze stagioni (sfruttando l'aria esterna quando presenta una temperatura leggermente più bassa di quella interna);

- minore portata totale di aria trattata rispetto ad altre soluzioni;
- possibilità di regolazione locale da parte dell'utente;
- manutenzione dell'impianto in zone non occupate.

L'Internet of Things sta rivoluzionando il modo in cui gli utenti si avvicinano alla tecnologia in tutti i settori.

Grazie all'IoT, oggi i sistemi HVAC possono fare molto di più che riscaldare e raffrescare gli ambienti su comando.

L'integrazione dell'IoT nelle soluzioni HVAC migliora nettamente i livelli di efficienza, comfort e user experience, garantendo un livello di connessione e comunicazione tra utente e tecnologia impensabile solo fino a poco tempo fa, trasformando definitivamente il modo in cui interagiamo con i sistemi di climatizzazione.

La connessione degli impianti alla rete consente agli utenti di restare sempre aggiornati sullo stato dell'aria di edifici e abitazioni, tramite app installate sui dispositivi smartphone che inviano avvisi di testo e aggiornamenti in tempo reale grazie alla presenza di sensori smart in grado di rilevare e monitorare in maniera costante i livelli di occupazione, temperatura, umidità ed emissioni di CO2 presenti nell'aria.

I sensori di temperatura trasmettono le informazioni tramite l'IoT ai sistemi HVAC, per un controllo della temperatura più preciso e ambienti più confortevoli. Naturalmente i sensori di temperatura sono importanti in un sistema HVAC, ma anche il rilevamento dell'umidità gioca un ruolo importante nel funzionamento di un sistema HVAC. I sensori di umidità consentono a un sistema HVAC di rilevare l'umidità all'interno e all'esterno di un edificio, per regolare con precisione il carico di lavoro e operare in modo più efficiente. Ma ci sono anche altri fattori ambientali che contribuiscono al comfort e alla qualità dell'aria negli spazi in cui viviamo. Sensori di rilevamento della CO2 accurati sono in grado di trasmettere i livelli di CO2 al sistema HVAC, per aiutare con la ventilazione e una migliore qualità dell'aria. L'IoT consente a tutti questi sensori di comunicare e lavorare in combinazione tra loro, per inviare al sistema di controllo centrale i dati per un funzionamento dell'impianto più efficiente sulla base dei rilevamenti effettuati, per ambienti sempre più confortevoli.

5.4.4 BMS

L'Internet of Things si lega strettamente anche all'ascesa delle soluzioni di controllo dei sistemi HVAC e **Building Management System**, ossia quei sistemi di controllo computerizzato che controllano e monitorano tutte le apparecchiature meccaniche ed elettriche di un edificio, come la ventilazione, l'illuminazione, l'alimentazione, i sistemi antincendio e i sistemi di sicurezza.

Grazie ai sistemi BMS, dunque, tutte le funzioni tecnologiche di un edificio possono oggi essere gestite in maniera integrata: sorveglianza, controllo accessi, rilevazione incendi, consumi energetici e sistemi HVAC, e sono proprio le soluzioni di controllo di questi ultimi a rappresentare, solitamente, l'elemento più complesso all'interno di un BMS, soprattutto in presenza di impianti su larga scala che prevedono la presenza di tante unità. La raccolta e l'analisi dei dati relativi alla climatizzazione tramite IoT, consente agli amministratori di sistema di comprendere meglio il funzionamento dell'edificio ed effettuare le opportune modifiche gestionali in modo tale da ottimizzare il flusso di operazioni e la programmazione e ridurre i consumi energetici, migliorando la gestione dell'edificio a livello generale e complessivo. L'integrazione nel cloud permette la comunicazione tra i diversi sistemi agevolando il monitoraggio e il controllo, e la connettività continua consente ai sistemi di apprendere i modelli di impiego e consumo energetico per evolversi e regolare automaticamente le proprie azioni.

La connessione dei dispositivi ai sistemi di controllo centrale consente inoltre di diagnosticare eventuali guasti e malfunzionamenti a distanza, ma anche fare manutenzione preventiva. Invece di effettuare arbitrariamente la manutenzione dei prodotti a periodi fissi, l'IoT permette di identificare i problemi sul nascere, risparmiando sulla riparazione e allungando la durata di vita dei prodotti.

5.4.5 Sistemi di riscaldamento

I sistemi di riscaldamento possono assumere forme diverse. Alcuni sono forni che bruciano materiale per fornire aria riscaldata attraverso le condotte; ci sono poi le caldaie che riscaldano l'acqua per i radiatori a vapore o i sistemi ad acqua forzata con radiatori a battiscopa, energia elettrica e pompe di calore. Le caldaie funzionano generalmente a gas naturale o propano.

Un'altra opzione è un pavimento radiante, noto anche come sistema di riscaldamento idronico. Questi utilizzano tubazioni sottopavimento e sono costituiti da tubi flessibili che vengono riempiti con acqua o una soluzione glicolica. Possono riscaldare qualsiasi tipo di pavimento, compreso il cemento, e sono un metodo efficace per fornire calore in una casa. Possono anche essere installati a posteriori su pavimenti in legno, anche se devono essere installati con cura nella guaina per pavimenti in legno.



Figura 5.4: Sistema di riscaldamento

5.4.6 Sistemi di raffreddamento

I condizionatori d'aria sono disponibili in molte forme, dagli enormi box progettati per raffreddare un'intera casa a un modello portatile montato sulla finestra che può essere utilizzato nelle zone a clima più freddo durante le estati brevi.

Molti condizionatori d'aria possono anche essere installati da sé, con i sistemi mini-split senza condotto. L'installazione è ancora un progetto importante, poiché gli elementi interni ed esterni del sistema devono essere correttamente collegati, ma sono relativamente economici da acquistare e gestire.

Per i climi più secchi si possono utilizzare i raffreddatori evaporativi. Questi attirano l'aria esterna nel sistema, facendola passare attraverso cuscinetti saturi d'acqua, che raffreddano e inumidiscono l'aria prima di spingerla nello spazio abitativo e spostare l'aria calda.

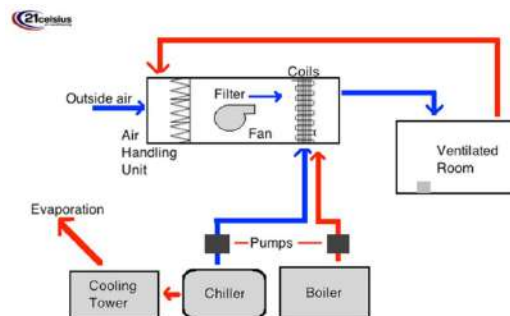


Figura 5.5: Sistema di raffreddamento

5.4.7 Z-Wave

Nasce nel 2001 alla Zensys, azienda danese che sviluppa sistemi di controllo luci.

Nel 2005 si forma la Z-Wave Alliance promossa da Zensys, Leviton Manufacturing, Danfoss e Ingersoll-Rand.

Nel 2008 entra a far parte dell'alleanza anche la Sigma Designs che ora è l'unico produttore di moduli hardware Z-Wave.

L'alleanza Z-Wave oggi ha tra i suoi membri anche SmartThings, Honeywell, Belkin, Bosch, Carrier, ADT e LG.

E' un protocollo proprietario che tutt'ora è in larga parte non divulgato con il vincolo di usare poca banda per comunicare con sensori e interruttori e si basa su specifiche ITU (International Telecommunications Union) per comunicazioni radio a raggio corto con banda stretta.

In termini di bit, può operare a 3 diverse velocità di trasmissione rispettivamente a 3 diverse frequenze radio:

- 100 Kbps: 916.0 MHz con estensione di 400 KHz;
- 40 Kbps: 916.0 MHz con estensione di 300 KHz;
- 9.6 Kbps: 908.4 MHz con estensione di 300 KHz.

Una rete Z-Wave è composta di nodi con funzioni specifiche:

5.4.7.1 Controller device:

è il dispositivo di più alto livello sulla rete mesh e ne controlla la tabella di routing. Ci sono 2 tipi di controller:

- **Controllore primario:** è il master e per ogni rete ne può esistere uno solo. Mantiene la topologia e gerarchia di tutti gli altri dispositivi. Può includere o escludere nodi e ha il compito di allocare gli identificativi dei vari nodi (node IDs);
- **Controllore secondario:** assiste il controllore primario nella gestione degli instradamenti.

I controllori possono essere statici o portatili:

- **Un controllore portatile** è in sostanza un comando remoto che può quindi cambiare posizione nell'ambiente e, una volta mosso, ricalcolerà gli instradamenti più veloci sulla rete;
- **Un controllore statico** non cambia posizione, ad esempio un gateway inserito in una presa di corrente.

5.4.7.2 Nodi o Dispositivi "Slave"

Dispositivi che eseguono azioni in base ai comandi ricevuti (sono degli attuatori) e non possono comunicare con altri nodi senza che questo gli venga richiesto tramite un comando. Possono memorizzare informazioni di instradamento ma non mantengono o calcolano tabelle, possono invece operare come semplici ripetitori.

5.4.7.3 Indirizzamento in Z-Wave

Il meccanismo di indirizzamento è molto più semplice degli analoghi in Bluetooth e Zigbee ed è così perché tutto è progettato in funzione della minimizzazione del traffico e del consumo.

Ci sono solo 2 identificatori:

- **Home ID:** è un indirizzo di 32 bit pre-programmato nei dispositivi di controllo e identifica in maniera univoca una rete Z-Wave rispetto ad un'altra. All'avvio della rete tutti i nodi (i dispositivi "slave") hanno una Home ID uguale a zero, il controllore li popola tutti con il suo home ID;
- **Node ID:** è un indirizzo di 8 bit che è assegnato ad ogni slave dal controllore (256 nodi al massimo).

Quando un nuovo dispositivo deve entrare a far parte di una rete esistente, occorre avviare una procedura di accoppiamento (pairing) e aggiunta che parte solitamente con la pressione meccanica di un tasto da parte dell'utente che invoca il controller primario che provvederà ad assegnare un nuovo Home ID + Node ID.

5.4.7.4 Topologia e instradamento

Un singolo controllore primario gestisce la rete e stabilisce gli instradamenti, la tabella è molto semplice perché specifica solo quale vicino è connesso ad determinato nodo guardando solo 1 passo avanti (1 solo hop). La tabella è costruita dal controllore primario chiedendo ad ogni nodo quali dispositivi sono raggiungibili dalla sua posizione. Quando un dispositivo riceve un messaggio lo inoltra al nodo successivo nella catena.

Instradamento Ad esempio nella tabella di routine seguente il percorso più breve da "Bridge" a "Slave 5" segue questo percorso: Bridge — Slave 3 — Slave 2 — Slave 5

	Slave 1	Slave 2	Slave 3	Slave 4	Enhanced Slave 5	Primary Controller	Secondary S/S	Bridge	Portable Controller
Slave 1	0	1	1	0	0	1	0	0	0
Slave 2	1	0	1	0	1	0	0	0	1
Slave 3	1	1	0	0	0	0	0	1	0
Slave 4	0	0	0	0	1	0	0	0	0
Enhanced Slave 5	0	1	0	1	0	0	1	0	0
Primary Controller	0	0	0	0	0	0	0	0	0
Secondary S/S	0	0	0	0	1	0	0	0	0
Bridge	1	0	1	0	0	0	0	0	0
Portable Controller	0	1	0	0	0	0	0	0	0

Figura 5.6: Tabella di routing

Topologia Esempio di topologia Z-Wave: un unico controllore primario, 4 slave, 1 slave avanzato, 1 bridge controller (agisce da gateway verso una rete Wi-Fi), un controllore portatile e un controllore secondario (inserito sulla rete mesh solo per assistere il controllore primario).

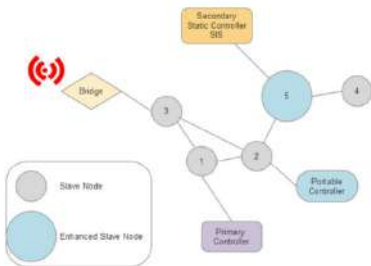


Figura 5.7: Esempio di Topologia Z-Wave

5.5 Lo standard KNX

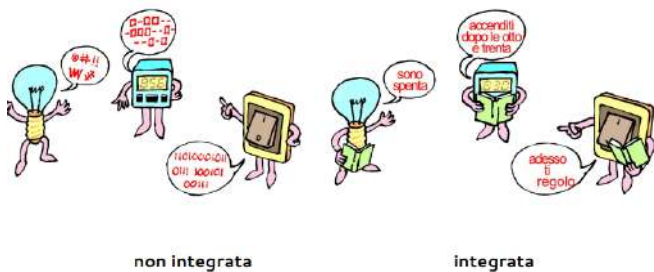


Figura 5.8



Figura 5.9: Connettere ingressi e uscite (in IoT sono sensori e attuatori)

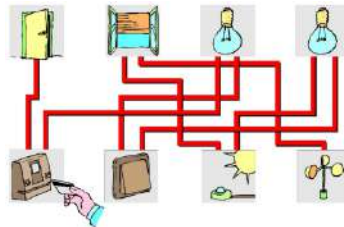


Figura 5.10: Connettere ingressi e uscite con un cablaggio tradizionale

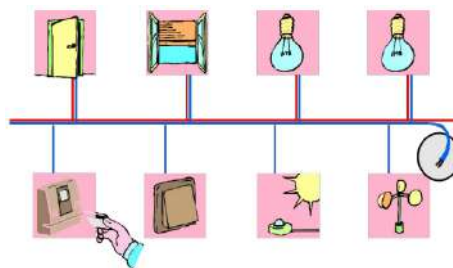


Figura 5.11: Connettere ingressi e uscite con un sistema bus

Installazione tradizionale	Installazione "intelligente"
• cablaggio punto punto	• cablaggio libero su linea dedicata (bus)
• maggiore quantità di cavi	• minore quantità di cavi
• presenza di un quadro o centralina di controllo	• assenza di centraline di controllo (faccitive) (intelligenza distribuita)
• dispositivi periferici tendenzialmente privi di "intelligenza"	• dispositivi con "intelligenza"
• dispositivi dedicati ad una sola applicazione	• dispositivi con applicazione configurabile
• rischio di contatto diretto nei dispositivi di comando	• possibilità di operare sui dispositivi di comando sotto tensione (SELV)
• maggior rischio di incendio	• riduzione del rischio di incendio
• inter-operatività stabilita dal cablaggio	• inter-operatività flessibile e configurabile

Figura 5.12: Sistema cablato VS Sistema bus

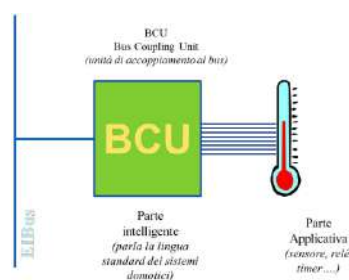


Figura 5.13: Dispositivo per sistema bus

KNX - sintesi cronostoria dispositivi

- 1992 - prima generazione
- 1998 - seconda generazione
- 2004 - terza generazione

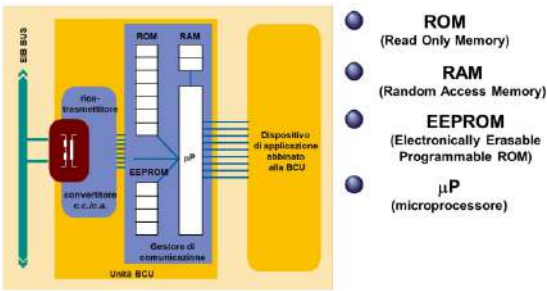


Figura 5.14: Dispositivo per sistema bus-dettaglio



Figura 5.15

- 2009 - quarta generazione
- 2016 - quinta generazione

KONNEX diventa standard ISO/IEC 14543-3 KNX è l'unico STANDARD INTERNAZIONALE per la building automation e la domotica approvato in tutto il mondo. Risponde ai requisiti di entrambi gli standard europei, quali CENELEC EN50090 e CEN EN 13321-1. Da novembre 2006 approvato come standard internazionale (ISO/IEC 14543-3), a conferma dell'importanza globale dello standard KNX.

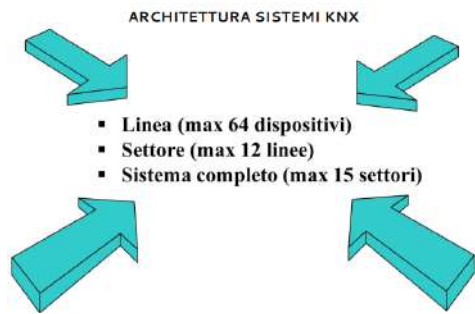


Figura 5.16

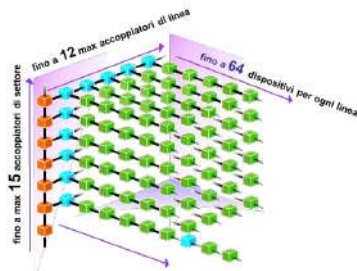


Figura 5.17: Potenzialità architettura sistema KNX

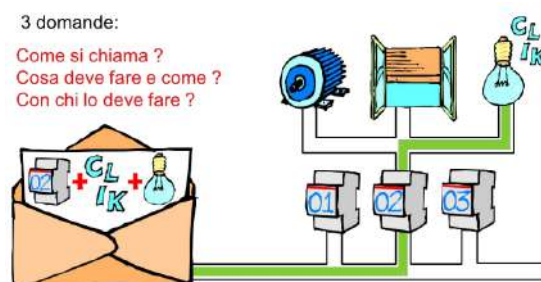


Figura 5.18

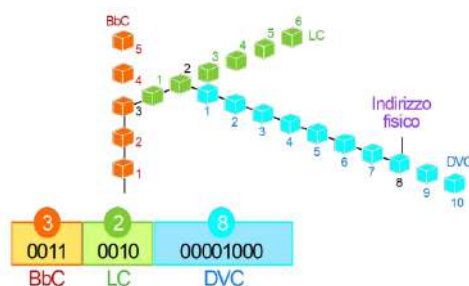


Figura 5.19: Come si chiama? - indirizzo fisico

Esempio (ingresso 4)

1. Rilevazione di allarme con riporto nei 3 (o N) sinottici di piano;
2. Scenario luminoso (un pulsante fa accendere solo alcune luci);
3. Accesso in camera via badge con sblocco elettroserratura, accensione luce cortesia, riporto info su PC Hall.

Esempio (uscita 6)

1. Illuminazione esterna con comandi manuale, crepuscolare, temporizzato, da PC;
2. Irrigazione con comandi manuale, temporizzato, automatico da sensore umidità, da PC;
3. Comando luci da 3 (o N) punti in corridoio molto lungo

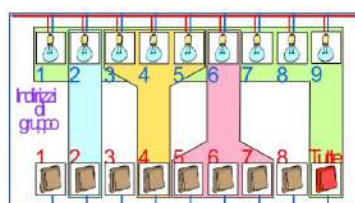


Figura 5.20: Cosa deve fare? - applicazione

Capitolo 6

Basate su Protocollo IP

6.1 Enti, Organizzazioni e Definizioni Preliminari

IETF (Internet Engineering Task Force) E' un organismo senza scopo di lucro che promuove standard di rete e Internet. Questi standard sono pubblicati come protocolli "aperti". Ciò significa che sono liberi per chiunque voglia leggerli.

Chiunque può presentare nuovi protocolli per la IETF per la sponsorizzazione. Qualsiasi protocollo in fase di sviluppo è assegnato a un gruppo di lavoro. Un gruppo di lavoro si estende su una superficie di specializzazione nel networking.

Tutti i membri del comitato lavorano gratis, sulla base del fatto che i loro datori di lavoro potranno destinare il tempo di lavoro IETF. I collaboratori e membri del comitato sono tutti identificati, insieme con la loro società datrice di lavoro.

<https://www.ietf.org/>

ISO (International Standard Organization) E' il principale ente internazionale di standardizzazione che si occupa anche di reti di calcolatori.

<https://www.iso.org/home.html>

IEEE (Institute of Electrical and Electronics Engineers) E' l'organizzazione mondiale degli ingegneri elettrici ed elettronici con gruppi che si occupano di reti.

<https://www.ieee.org/>

ICANN (Internet Corporation for Assigned Names and Numbers) È l'associazione, costituita nel 1998 per gestire gli indirizzi degli Internet provider (i fornitori dell'accesso alla rete) e controllare la registrazione dei nuovi domini internazionali (le "regioni" in cui si suddivide Internet). È un ente sotto il diretto controllo del Ministero del commercio degli Stati Uniti.

<https://www.icann.org/>

IANA (Internet Assigned Numbers Authority) E' un organismo dell'ICANN che ha la responsabilità nell'assegnazione degli indirizzi IP.

<https://www.iana.org/>

RFC (Request for Comments) Indica un documento che può rappresentare simultaneamente la definizione di un nuovo standard per Internet e/o, come indica appunto il nome, la richiesta di commenti e proposte di miglioramento ad un particolare protocollo.

Accade molto spesso che una Request for Comments tratti gli stessi argomenti di una RFC precedente: in questo caso si può dire che la RFC precedente è stata aggiornata (updated) o addirittura che è stata resa obsoleta (obsoleted).

I documenti RFC non vengono mai modificati: in caso di aggiornamento di uno standard o di altra tipologia di documento, viene rilasciato un nuovo RFC con un numero progressivo che indica l'avvenuto

aggiornamento di una RFC precedente.

Al momento gli RFC sono circa 3000 ma viene pubblicato, periodicamente, un RFC che funge da indice e da descrittore dello stato degli RFC.

6.2 Il modello ISO/OSI

Il modello OSI è un progetto della fine degli anni '70. Scopo del modello: fungere da riferimento per le reti di calcolatori. Ha un approccio a 7 livelli (layers) che eseguono ciascuno una specifica funzione. Ha raggiunto lo scopo di fungere da elemento coordinatore tra tutte le attività di standardizzazione. I livelli 1 e 2 sono oggi standardizzati, ma è stato meno efficace ai livelli 3/4/5/6/7, soprattutto per l'impatto sui dispositivi di instradamento.



Figura 6.1: Livelli OSI

Indipendentemente dal protocollo utilizzato a livello di sensori, il dato dovrà giungere ad un cloud pubblico o privato (o ibrido) per essere analizzato, controllato o monitorato. Al di fuori di una WPAN, il mondo è basato su TCP/IP.

OSI (Open Source Interconnection) Model			
Layers	Purpose / Function	Protocol Used	Fundamental Data Type
7. Application	User Application Layer: browser, FTP, app, etc. (handle file access, resource sharing, LDAP, SMTP)	SMTP, FTP	Data
6. Presentation	Syntax Layer: encrypt, compress (optional) (data representation, codes, translation)	PGP, ASCII, ROT13	Data
5. Session Layer	Synchronization and Logical Port Routing: session establishment, start & terminate, security, logging, name recognition	RPC, NFS, NetBIOS	Data
4. Transport Layer	TCP: host to host & flow control (end-to-end connection & reliability, message segmentation, acknowledgment, session multiplexing)	TCP / UDP	TCP Segments, UDP Datagrams
3. Network Layer	Packets: IP address (path determination, logical addressing, routing, multi-cast, flow segmentation, packet management)	IP, SPX, ICMP	Packets
2. Data Link Layer	Data Frames: MAC address, packet (physical addressing, Media Access Control, LLC, frame error checking, sequencing and monitoring)	PPP/SLIP	Frames
1. Physical Layer	Physical Device: cables, fibers, RF spectrum (data encoding, media attachment, synchronization, signaling, binary transmission)	coax, Modem, wireless	Bits / Signals

Figura 6.2: Funzione Livelli

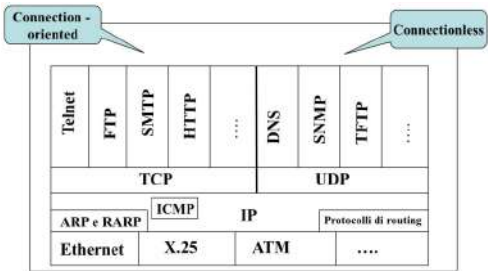


Figura 6.3: Protocolli

6.3 Elementi del Protocollo IP

6.3.1 Strato di Internetworking INTERNET PROTOCOL (IP - RFC 791)

Inserito nel modello OSI al livello 3. Il protocollo definisce un meccanismo di consegna dati di tipo “non affidabile” e non orientato alla connessione (connectionless). Ogni pacchetto IP può essere suddiviso in Header IP e DATA. Header IP contiene, tra gli altri, questi campi 2 campi fondamentali:

- Indirizzo mittente;
- Indirizzo destinatario.

0	4	8	15	19	24	31
Version	HLEN	Service type	Total length			
Identification			Flags	Fragment Offset (13 bit)		
Time to Live		Protocol	Header checksum			
Source IP Address						
Destination IP address						
IP Option					Padding	
Data						

Figura 6.4: Header Pacchetto IP

In una rete locale il pacchetto IP viene “incapsulato” in una trama Ethernet ma potrebbe anche essere trasportato su reti con livelli 1 e 2 completamente diversi.

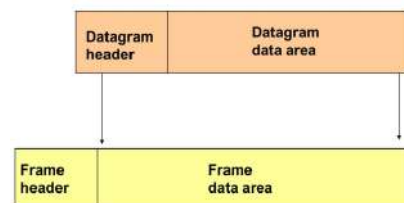


Figura 6.5: Incapsulamento

Il protocollo IP prevede la consegna a 3 tipi diversi di destinatari:

- **Unicast delivery** (il pacchetto è inviato a un singolo computer);
- **Broadcast delivery** (il pacchetto viene inviato a tutti i computer di una data rete);
- **Multicast delivery** (il pacchetto viene inviato ad uno specifico sottoinsieme di host).

6.3.2 Comunicazione basata su protocollo IP

Dal punto di vista dell'IoT, avvicinare l'IP all'origine dei dati consente di collegare tra loro due mondi: IT e OT. La tecnologia dell'informazione (IT) gestisce l'infrastruttura, la sicurezza e il provisioning delle reti mentre la tecnologia operativa (OT) gestisce la salute e il throughput del sistema che funziona per produrre qualcosa. Questi due ruoli sono stati tradizionalmente separati, poiché elementi come sensori, contatori e controller programmabili in passato non erano collegati direttamente.

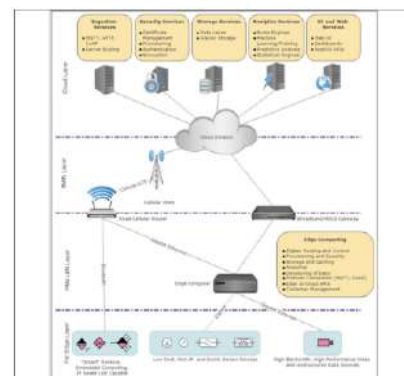


Figura 6.6: Comunicazione

Dobbiamo chiederci forse ,come mai IP sia lo standard per la comunicazione globale?

Onnipresente - quasi tutti i sistemi operativi sono dotati di stack IP ed il protocollo di comunicazione IP è in grado di funzionare su sistemi WPAN, cellulari, connessioni in rame, fibra ottica, bus PCI

Express e sistemi satellitari.

Perché IP è lo standard “de facto” per la comunicazione globale?

Longevo - TCP è stato istituito nel 1974 e lo standard IPv4 ancora in uso oggi è stato progettato nel 1978. Ha resistito alla prova del tempo per 40 anni. La longevità è fondamentale per molte soluzioni IoT industriali (ma non solo) che devono supportare dispositivi e sistemi per decenni. Vari altri protocolli proprietari sono stati progettati, come AppleTalk, SNA, DECnet e Novell IPX, ma nessuno ha guadagnato la diffusione sul mercato che ha IP.

Basato su standard - TCP/IP è governato dall’Internet Engineering Task Force (IETF). L’IETF mantiene una serie di standard aperti incentrati su IP.

Scalabile - IP ha dimostrato scalabilità e adattamento, utilizzato da miliardi di utenti e ancor più numerosi dispositivi. IPv6 potrebbe fornire un indirizzo IP univoco a ogni atomo della Terra e supportare altri 100 mondi in più.

Gestibilità - esistono vari strumenti per gestire reti e dispositivi IP: applicativi per la modellazione, sniffer di rete, applicativi per la diagnostica ed anche dispositivi per facilitare la creazione, il dimensionamento e la manutenzione delle reti.

Affidabilità - IP è un protocollo affidabile per la trasmissione dei dati. Pur essendo un protocollo “connectionless”, cioè ogni pacchetto viene trattato indipendentemente dagli altri, la consegna viene garantita utilizzando il meccanismo “best-effort” per mezzo del quale vengono effettuati tutti i tentativi di trasmettere un pacchetto attraverso vari percorsi.

La forza di questo modello ci consente di sostituire i livelli uno e due dello stack senza modifiche ai pacchetti (ad esempio Wi-Fi con rete 4G) avendo comunque la certezza che la consegna avverrà con esito positivo.

Se poi allo strato superiore, quello di “rete” nel modello ISO/OSI (il quarto), viene utilizzato il protocollo TCP ne viene anche controllato il flusso (cosa che non avviene con UDP che delega tale controllo all’applicazione).

Più nel dettaglio, mentre l’IP risponde alla necessità di avere un livello di rete ben supportato e robusto, TCP e UDP (Universal Datagram Protocol) sono responsabili della comunicazione end-to-end. TCP viene utilizzato per le trasmissioni orientate alla connessione, mentre UDP viene utilizzato per le trasmissioni senza connessione. UDP è naturalmente molto più semplice da implementare rispetto a TCP, ma non altrettanto resiliente. Entrambi i servizi forniscono il riordino dei segmenti poiché non è garantito che i pacchetti vengano consegnati in ordine utilizzando un protocollo IP. TCP fornisce il controllo del flusso utilizzando finestre scorrevoli e algoritmi di prevenzione della congestione. UDP fornisce un metodo leggero e ad alta velocità per trasmettere dati a vari dispositivi che possono essere o meno presenti o affidabili.

6.4 Sistemi Wireless: Diagrammi di Irradiazione Antenne

6.4.1 Definizioni

dBi = guadagno in potenza di un’antenna isotropica.

dBi indica il valore in decibel del guadagno di un’antenna rispetto a un’ipotetica antenna isotropica, supponendo che ad entrambe le antenne sia fornita la stessa potenza.

È solo un valore teorico, in quanto l’antenna isotropica non esiste nella realtà e non può essere progettata né costruita.

Isotropico viene dalle parole greche “isos” (uguale, identico) e “trópos” (rotazione). In ambito scientifico si applica a quegli oggetti che dimostrano proprietà identiche e uniformi in tutte le direzioni.

6.4.2 Antenna Isotropica (Diagramma di Radiazione)

L'antenna isotropica teorica rappresenta un punto infinitamente piccolo nel vuoto, radiante idealmente uniformemente in qualsiasi direzione dello spazio, senza riflessioni e perdite (il suo diagramma di radiazione è una sfera).

Per calcolare il guadagno in potenza di un'antenna espresso in dBi , si usa la seguente formula

$$G(dBi) = 10 \cdot \log_{10}(G) = 10 \cdot \log_{10}\left(\frac{P}{P_i}\right)$$

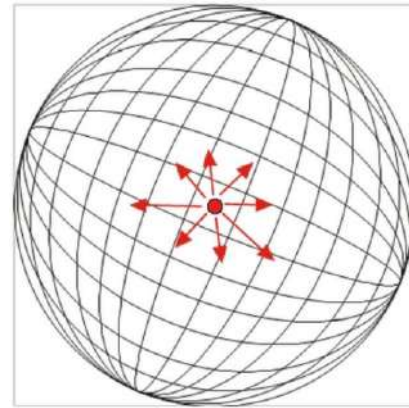


Figura 6.7: Raggio Antenna Isotropica

Dove:

P = potenza della antenna in esame

P_i = potenza della antenna isotropica

$G(dBi)$ = guadagno di un'antenna rispetto a un'isotropica, espresso in decibel

L'antenna isotropica ha un guadagno di 0 dBi. Convertendo in scala lineare si ottiene:

$$G = 10^{\frac{G(dBi)}{10}}$$

Esempio:

Calcoliamo di quanto sia superiore un'antenna con guadagno di 17 dBi rispetto a un'antenna isotropica nella ricezione/trasmissione del segnale.

$$G = 10^{\frac{G(dBi)}{10}} = 10^{\frac{17}{10}} = 10^{1.7} = 50.11$$

Dunque, un'antenna con guadagno di 17 dBi riceve/trasmette il segnale 50.11 volte più fortemente da un'antenna isotropica.

L'unità dBi e il termine "antenna isotropica" vengono utilizzati per calcolare *E.I.R.P.* (Effective Isotropic Radiated Power).

EIRP rappresenta la potenza che dovrebbe essere irradiata da un'antenna isotropica ipotetica per ottenere lo stesso livello di segnale nella direzione di radiazione massima di un'antenna. Viene utilizzato nella progettazione e nei calcoli dei parametri delle reti Wi-Fi, collegamenti satellitari, ecc.

La normativa vigente dell'Unione Europea stabilisce la potenza massima a cui è consentita la trasmissione in una gamma di frequenza Wi-Fi:

- 2400,0 – 2483,5 MHz (banda a 2,4 GHz) - la potenza non può superare i 100 mW *E.I.R.P.* (20 dBm);
- 5150 – 5350 MHz (banda a 5 GHz) - la potenza non può superare i 200 mW *E.I.R.P.* (23 dBm);
- 5725 – 5875 MHz (banda a 5 GHz) - la potenza non può superare i 1000 mW *E.I.R.P.* (30 dBm).

Al fine di non superare i valori limite di *E.I.R.P.* è necessario considerare:

- Potenza di uscita del trasmettitore (ad esempio, di una scheda di rete, punto di accesso);

- Tipo di cavo, la sua lunghezza e attenuazione alla frequenza di lavoro e attenuazione dei connettori;
- Guadagno in potenza dell'antenna.

Ricordiamo che i produttori dei punti di accesso (Access Points) spesso indicano la potenza del trasmettitore in *E.I.R.P.*. Ciò significa che il dispositivo è conforme alle normative solo ed esclusivamente con antenna fornita in dotazione o antenna integrata.

Se si decide di costruire un dispositivo Wi-Fi, è necessario fare semplici calcoli per verificare se ci si trova all'interno della gamma di potenza consentita dalla legge.

Per calcolare *E.I.R.P.* di un sistema costituito da un trasmettitore (ad esempio, router wireless), cavo e antenna, si utilizza la seguente formula:

$$E.I.R.P. = P - l \cdot Tk + Gi$$

Dove:

P = potenza del trasmettitore espressa in dBm

l = lunghezza del cavo in metri

Tk = attenuazione di 1 metro di cavo alla frequenza di lavoro del trasmettitore

Gi = guadagno in potenza di un'antenna isotropica espresso in decibel

Esempio:

Si costruisce una rete Wi-Fi nella banda a 2,4 GHz e si ha a disposizione:

- $P = 16$ dBm
- 8 metri di cavo *TRI - LAN* - 240 (l'attenuazione a 2,4 GHz è di 0,4 dB/metro), ossia $8 \cdot 0,4$ dB = 3,2 dB,
- Due connettori - cioè attenuazione $+ 2 \cdot 0,5$ dB = 1 dB.
- $Gi = 8$ dBi

$$E.I.R.P. = P - l \cdot Tk + Gi$$

$$E.I.R.P. = 16 \text{ dBm} - 3,2 \text{ dB} - 1 \text{ dB} + 8 \text{ dBi} = 19,8 \text{ dBm}$$

Ci si trova all'interno della gamma consentita dalla legge (potenza ≤ 20 dBm)

Se sostituissimo l'antenna, mettendone una con guadagno superiore agli 8 dBi della precedente, ad esempio 13 dBi:

$$E.I.R.P. = 16 \text{ dBm} - 3,2 \text{ dB} - 1 \text{ dB} + 13 \text{ dBi} = 24,8 \text{ dBm}$$

Superiamo la gamma consentita dalla legge (potenza ≤ 20 dBm)

Attenzione che non tutti i punti di accesso sono in grado di ridurre la potenza di uscita e quindi il guadagno di antenna diventa il solo parametro modificabile. In generale è consigliabile utilizzare un'antenna con alto guadagno e un trasmettitore con meno potenza rispetto a un'antenna con basso guadagno e un trasmettitore con più potenza perché i dispositivi non funzionano solo in trasmissione ma anche in ricezione e quindi la sensibilità del ricevitore è importante.

6.4.3 Diagrammi di Irradiazione

L'irradiazione è il termine usato per rappresentare l'emissione o la ricezione del fronte d'onda dell'antenna, specificandone l'intensità. In ogni progetto RF si può notare sempre un particolare grafico che rappresenta la radiazione dell'antenna. Esso è il diagramma di radiazione e permette facilmente di risalire alle caratteristiche e alla direttività di un'antenna.

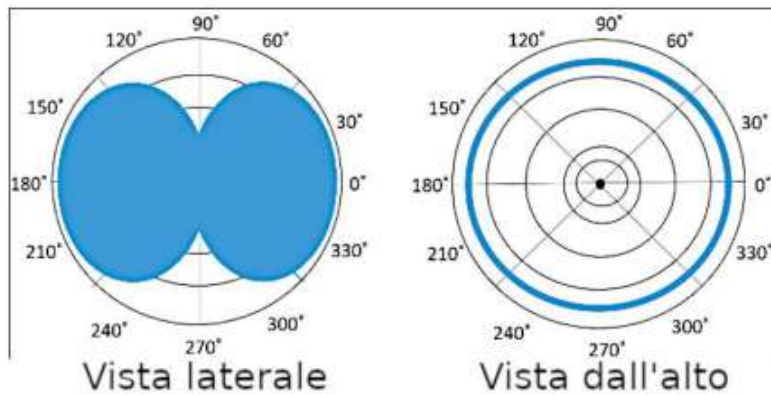


Figura 1: un esempio di diagramma nei piani orizzontale e verticale

Figura 6.8: Diagramma Irradiazione

6.4.4 Diagrammi di irradiazione

Un diagramma può contenere i seguenti lobi:

- **Lobo di radiazione principale:** è il picco più intenso della radiazione. Questa è la parte in cui esiste la massima energia irradiata. La direzione di questo lobo indica la direttività dell'antenna;
- **Lobo minore:** qualsiasi lobo di radiazione diverso dal lobo principale;
- **Lobo laterale:** è il lobo di radiazione in qualsiasi altra direzione diversa da quella prevista;
- **Lobo posteriore:** è il lobo di radiazione opposto a quello principale. Qui si spreca una notevole quantità di energia.

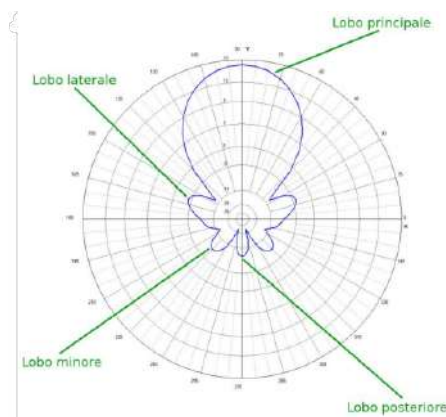


Figura 6.9: Diagramma di Irradiazione

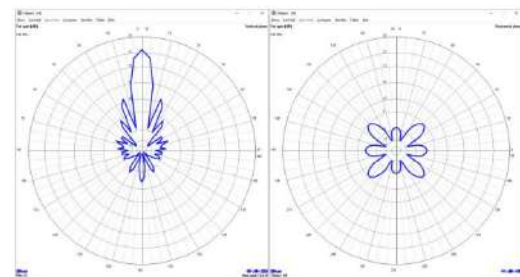


Figura 6.10: Irradiazione Antenna Direttiva

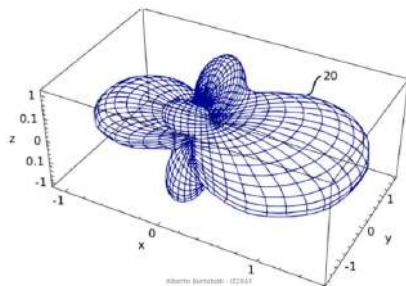


Figura 6.11: Diagramma 3D Irradiazione Antenna Omni-Direzionale

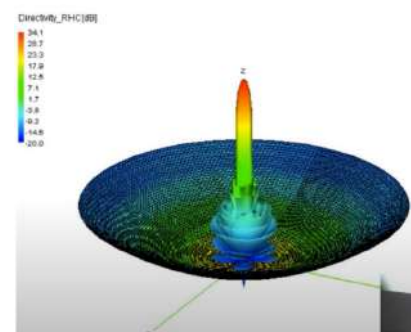


Figura 6.12: Diagramma 3D Irradiazione Antenna Parabolica

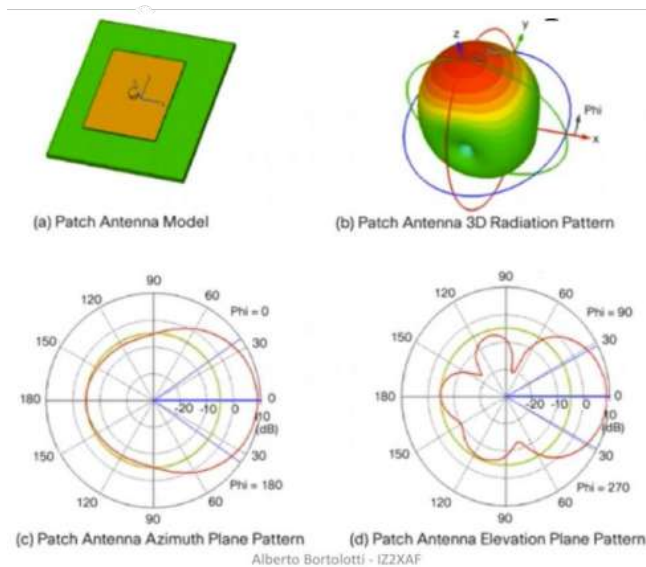


Figura 6.13: Diagrammi Irradiazione Antenna a Pannello

6.5 La Suite di Protocolli IEEE 802.11

6.5.1 I Protocolli in IEEE 802.11 e le WLAN

Il progetto inserito nel modello OSI ai livelli 1 e 2 per reti locali (LAN) e metropolitane (MAN) ha prodotto una voluminosa serie di standard noti con sigle del tipo 802.x ed ha contrastato la tendenza a creare nuovi tipi di reti locali per ragioni commerciali. Oggi è suddiviso in molteplici gruppi, tra i più conosciuti abbiamo:

- IEEE 802.3 CSMA/CD Access Method [Ethernet]
- IEEE 802.11 Wireless local area network
- IEEE 802.16 WiMAX - Broadband wireless access

[si veda <https://www.ieee802.org/>]

6.6 IEEE 802.11 e la Suite di Protocolli

IEEE 802.11 è una suite di protocolli con una ricca storia e diversi casi d'uso. 802.11 è la specifica che definisce il Media Access Control (MAC) e il livello fisico (PHY) di uno stack di rete. La definizione e le specifiche sono regolate dall'IEEE LAN/MAN Standards Committee. Wi-Fi è la definizione di WLAN basata sugli standard IEEE802.11 ma mantenuta e regolata dalla Wi-Fi Alliance senza scopo di lucro.

IEEE 802.11 deve la sua creazione alla NCR nel 1991, che per prima ha sviluppato il protocollo wireless come mezzo per collegare in rete i suoi registratori di cassa. Nel 1999, con la costituzione della Wi-Fi Alliance, la tecnologia è diventata onnipresente e pervasiva nel mercato dei PC e dei notebook. Il protocollo originale era molto diverso dai moderni protocolli 802.11 b/g/n/ac, supportando solo una velocità dati di 2 Mbps con correzione degli errori.

Il successo di IEEE 802.11 può essere attribuito all'approccio stack a strati del modello OSI, che ha consentito di utilizzare facilmente l'infrastruttura TCP/IP esistente.

Oggi, quasi tutti i dispositivi mobili, notebook, tablet, sistemi embedded, giocattoli e videogiochi incorporano un sistema radio IEEE 802.11 di qualche tipo.

Il modello di sicurezza 802.11 originale era basato sul meccanismo di sicurezza UC Berkeley Wired

Equivalent Privacy (WEP), che in seguito si è dimostrato inaffidabile e facilmente compromettibile. Diversi exploit degni di nota, tra cui la violazione dei dati di TJ Maxx attraverso 802.11 WEP nel 2007, hanno portato al furto di 45 milioni di carte di credito.

Oggi, Wi-Fi Protected Access (WPA) e WPA2 che utilizzano chiavi precondivise AES a 256 bit hanno sicuramente rafforzato la sicurezza e il WEP viene utilizzato raramente.



Figura 6.14: Trend Attacchi Cibernetici

Tra tutti i protocolli della suite IEEE 802.11, esamineremo più dettagliatamente quelli di particolare importanza per il sistemista IoT:

- **IEEE 802.11ac;**
- **IEEE 802.11p** (per la comunicazione V2V);
- **IEEE 802.11ah** (HaLow).

6.6.1 Evoluzione di IEEE 802.11

L'IEEE LAN/MAN Standards Committee mantiene e governa le specifiche IEEE 802. Dall'802.11 originale si è evoluto a 802.11ac nel 2013, esaminando varianti specifiche per casi d'uso come l'interconnessione IoT a bassa potenza/bassa larghezza di banda (802.11ah), la comunicazione da veicolo a veicolo (802.11p), il riutilizzo dello spazio RF analogico televisivo (802.11af) e della banda vicina al metro per audio/video (802.11ad), ed infine il successore dello standard 802.11ac: 802.11ax.

Il protocollo 802.11 rappresenta una famiglia di comunicazioni radio wireless basate su diverse tecniche di modulazione nelle bande ISM a 2,4 GHz e 5 GHz dello spettro senza licenza. 802.11b e 802.11g risiedono nella banda a 2,4 GHz, mentre 802.11n e 802.11ac aprono la banda a 5 GHz. Il Wi-Fi è suscettibile allo stesso rumore e interferenza di Bluetooth e Zigbee e implementa una varietà di tecniche per garantire robustezza e resilienza.

IEEE 802.11 Protocol	Use Case	Release Date	Frequency (GHz)	Bandwidth (MHz)	Streaming Data Rate per Channel min-max (Mbps)	Allowable MIMO Streams	Modulation	Indoor Range (m)	Outdoor Range (m)	Typical Dissipated Power per Chip (mW)
802.11	First 802.11 design	Jun-97	2.4	22	1 to 2	1	DSSS, F-SS	20	20	50
a	Release simultaneously with 802.11b. Less prone to interference than 802.11b.	Sep-99	5	20	6 to 54	1	Q-PSK (SSA)	30	120	50
b	Release simultaneously with 802.11a. Significant speed increase over 802.11a at improved range.	Sep-99	2.4	22	1 to 11	1	DSSS (S-SS)	50	150	7 to 50
g	Speed increase over 802.11b.	Jun-02	2.4	20	6 to 54	1	Q-PSK, BPSK (SSC)	28	140	50
n	Multiple antenna technology for improved speed, and range.	Oct-05	2.4 / 5	20	7.2 to 72.2	4	OFDM (MIMO)	70	250	40
ac	Better performance and coverage over 802.11n. Wider channel and improved modulation. Allows multiple users using MU-MIMO. Introduced beam forming.	Dec-13	5	30	7.2 to 16.8	8	Q-PSK (MU-MIMO)	35	35	40
ah	"WiFi Hetero"	Dec-16	2.4 / 5	1 to 26	34.7	4	Q-PSK	3000	1000	100 but goal is low power
p	"Vehicle Access or Vehicular Environments" "Intelligent Transport Systems" Dedicated Short Range Communication Transport uses cases: toll collection, safety and collision avoidance, vehicular networking.	Jun-08	5.8	10	27	1	Q-PSK	NA	400 to 1000	40
af	"White WiFi" or "Super WiFi" Deploy around spectrum in TV bands, to provide fast and connectivity in India, Korea, Singapore, the US, and others.	Nov-13	0.470 to 1.710	6 to 8	568	4	Q-PSK	NA	6000-100,000	100
ad	80 GHz. Wireless for HD video and producer's Audio and video transport and cable replacement.	Dec-12	60	2160	4056	>32	16-QAM (MU-MIMO)	10	30	100
ax	"High Efficiency Wireless (HEW)" Next gen 802.11. An increase in capacity over 802.11ac. Average increase of 4x speed per user over 802.11ac. Backwards-compatible to 802.11a/2g/n/ac. License deployment scenario.	2019	2.4 / 5	20	40 to 3000	8	OFDMA (MU-MIMO)	25	35	100

Figura 6.15: Protocolli

6.6.2 Topologie di Base per Sistemi 802.11

I sistemi 802.11 supportano tre topologie:

- **Infrastruttura:** un dispositivo endpoint 802.11 (o STA = station), ad esempio uno smartphone, comunica con un punto di accesso centrale (AP). Un AP può essere un gateway per altre reti (WAN), un router o un vero punto di accesso in una rete più ampia. Questa topologia è a stella ed è anche nota come BSS (Infrastructure Basic Set Service).
- **Ad hoc:** i nodi 802.11 possono formare quello che viene chiamato un servizio di set di base indipendente (IBSS) in cui ciascuna stazione comunica e gestisce l'interfaccia con altre stazioni. In questa configurazione non viene utilizzato nessun punto di accesso ed è una topologia peer-to-peer.
- **Sistema distribuito (DS):** il DS combina due o più reti BSS indipendenti tramite interconnessioni di punti di accesso (AP) cioè di interconnessioni dei centro stella.

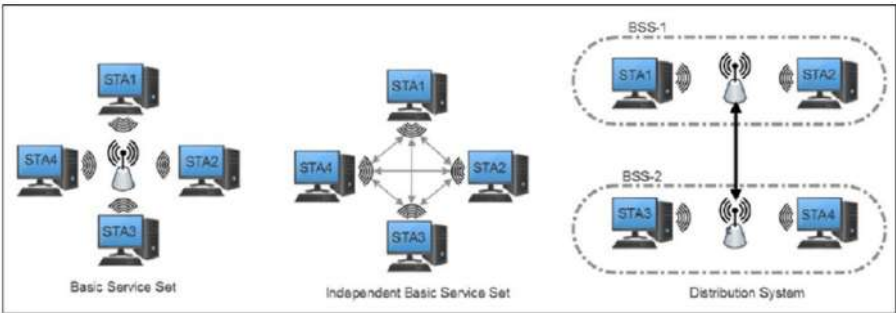


Figura 6.16: Topologie di Base Sistemi 802.11

Il protocollo 802.11 consente di associare fino a 2.007 STA ad un unico punto di accesso.

6.7 Allocazione dello Spettro di Frequenze

Il primo protocollo 802.11 utilizzava uno spettro nella regione ISM a 2 GHz e 5 GHz e canali distanziati uniformemente a circa 20 MHz l'uno dall'altro. La larghezza di banda del canale era di 20 MHz, ma gli emendamenti successivi di IEEE consentivano il funzionamento anche di 5 MHz e 10 MHz. Negli Stati Uniti, 802.11b e g consentono 11 canali (altri paesi possono supportarne fino a 14)

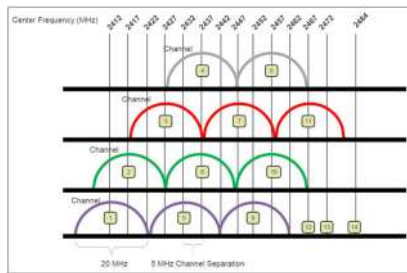


Figura 6.17: Separazione dei Canali

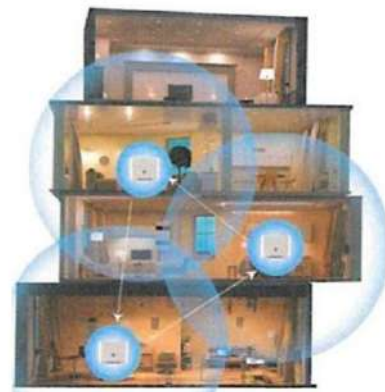


Figura 6.18: Copertura Spettro di Frequenze

6.8 Tecniche di Modulazione e Codifica IEEE 802.11

Questa sezione descrive le tecniche di modulazione e codifica nel protocollo IEEE 802.11.

Queste tecniche non sono esclusive di 802.11; si applicano anche ai protocolli 802.15 e, come vedremo, anche ai protocolli cellulari.

I metodi di frequency hopping, modulazione e phase-shift keying sono metodi fondamentali poiché insieme bilanciano gamma, interferenza e throughput.

I dati digitali trasmessi da un segnale RF devono essere trasformati in analogici. Ciò si verifica al PHY indipendentemente dal segnale RF descritto (Bluetooth, Zigbee, 802.11 e così via). Un segnale portante analogico sarà modulato da un segnale digitale discreto.

Questo forma quello che viene chiamato un simbolo o un alfabeto di modulazione.

Un modo semplice per pensare alla modulazione dei simboli è un pianoforte con quattro tasti. Ciascuna chiave rappresenta due bit (00, 01, 10, 11). Se riesci a suonare 100 tasti al secondo, ciò significa che puoi trasmettere 100 simboli al secondo. Se ogni simbolo (tono del pianoforte) rappresenta due bit, equivale a un modulatore a 200 bps.

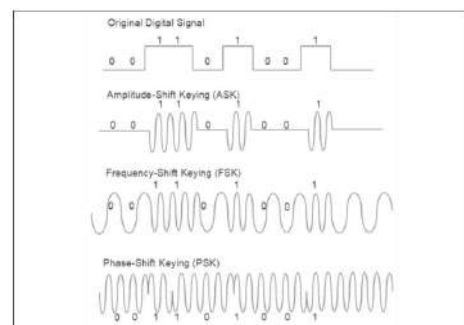


Figura 6.19: Tecniche di Modulazione e Codifica

Sebbene ci siano molte forme di codifica dei simboli da studiare, le tre forme di base includono:

- **Amplitude-shift keying (ASK):** questa è una delle modulazioni di ampiezza possibili. Il binario 0 è rappresentato da una modulazione di ampiezza diversa dalla modulazione di ampiezza che rappresenta 1. In figura è mostrato un modulo semplice ma moduli più avanzati possono rappresentare i dati in gruppi utilizzando livelli di ampiezza aggiuntivi;
- **FSK (Frequency-Shift Keying):** viene modulata una frequenza portante per rappresentare 0 o 1. La forma più semplice (quella rappresentata in figura) è la Binary Frequency-Shift Keying (BFSK), che è la forma utilizzata in 802.11. Nell'ultimo capitolo abbiamo parlato di Bluetooth e Z-Wave. Questi protocolli utilizzano una forma di FSK chiamata Gaussian frequency-shift keying (GFSK), che filtra i dati attraverso un filtro gaussiano, che attenua l'impulso digitale (-1 o +1) e lo modella per limitare la larghezza spettrale;

- **Phase-shift keying (PSK)**: modula la fase di un segnale di riferimento (segnale portante). Utilizzato principalmente nei tag 802.11b, Bluetooth e RFID, PSK utilizza un numero finito di simboli rappresentati come diversi cambiamenti di fase. Ogni fase codifica un numero uguale di bit. Un modello di bit formerà un simbolo. Il ricevitore avrà bisogno di un segnale di riferimento contrastante e calcolerà la differenza per estrarre i simboli e quindi demodulare i dati. Un metodo alternativo non richiede alcun segnale di riferimento per il ricevitore. Il ricevitore ispezionerà il segnale e determinerà se c'è un cambiamento di fase senza fare riferimento a un segnale secondario. Questo è chiamato DPSK (differenziale phase-shift keying) e viene utilizzato in 802.11b.

6.9 Mitigazione delle Interferenze

Gli standards 802.11 utilizzano diverse tecniche di mitigazione dell'interferenza:

- **Frequency-hopping spread spectrum (FHSS)**: diffonde un segnale su 79 canali non sovrapposti larghi 1 MHz nella banda ISM a 2,4 GHz. Utilizza un generatore di numeri pseudocasuali per avviare il processo di salto. Il tempo di sosta si riferisce al tempo minimo di utilizzo di un canale prima del salto (400 ms);
- **Direct-sequence spread spectrum (DSSS)**: utilizzato per la prima volta nei protocolli 802.11b, dispone di canali di 22 MHz di larghezza. Ogni bit è rappresentato da più bit nel segnale trasmesso. I dati trasmessi vengono moltiplicati per un generatore di rumore. Ciò diffonderà efficacemente il segnale sull'intero spettro in modo uniforme utilizzando una sequenza numerica pseudocasuale, chiamata codice pseudo-rumore (PN);
- **OFDM**: utilizzato in IEEE 802.11a e protocolli più recenti. Questa tecnica divide un singolo canale da 20 MHz in 52 sottocanali (48 per i dati e quattro per la sincronizzazione e il monitoraggio) per codificare i dati utilizzando QAM e PSM. Una veloce trasformata di Fourier (FFT) viene utilizzata per generare ciascun simbolo OFDM. Un insieme di dati ridondanti circonda ogni sottocanale.

Questa banda ridondante di dati è chiamata intervallo di guardia (GI) e viene utilizzata per prevenire l'interferenza intersimbolica (ISI) tra le sottoportanti vicine.

Si noti che le sottoportanti sono molto strette e non hanno bande di guardia per la protezione del segnale. Ciò era intenzionale perché ogni sottoportante è distanziata equamente rispetto al reciproco del tempo del simbolo.

Cioè, tutte le sottoportanti trasportano un numero completo di cicli sinusoidali che, una volta demodulati, si sommeranno a zero. Per questo motivo, il design è semplice e non richiede il costo aggiuntivo dei filtri passa-banda. IEEE 802.11a utilizza 250.000 simboli al secondo. OFDM è generalmente più efficiente e più denso (quindi più larghezza di banda) di DSSS e viene utilizzato nei protocolli più recenti.



Figura 6.20: IEEE 802.11 MIMO

6.10 IEEE 802.11 MIMO

MIMO è l'acronimo di multiple-input multiple-output. MIMO sfrutta un fenomeno RF chiamato multipath (multipercorso) che implica che i segnali si riflettano su pareti, porte, finestre e altri ostacoli. Un ricevitore vedrà molti segnali, tutti in arrivo in momenti diversi attraverso percorsi diversi.

Il multipath tende a distorcere i segnali e causare interferenze, che alla fine degradano la qualità del segnale. (questo effetto è chiamato dissolvenza multipath.)

Con l'aggiunta di più antenne, un sistema MIMO può aumentare linearmente la capacità di un dato canale semplicemente aggiungendo più antenne.

Esistono due forme di MIMO:

- **Diversità spaziale:** si riferisce alla diversità di trasmissione e ricezione. Un singolo flusso di dati viene trasmesso su più antenne contemporaneamente utilizzando la codifica spazio-temporale. Questi forniscono miglioramenti nel rapporto segnale-rumore e sono caratterizzati dal miglioramento dell'affidabilità del collegamento e della copertura del sistema.

Il salto di frequenza è un metodo per superare il problema della dissolvenza del multipath modificando costantemente gli angoli del multipath.

Ciò ha l'effetto di distorcere la dimensione del segnale RF. I sistemi Bluetooth in genere hanno un'antenna, rendendo così difficile il MIMO. Per quanto riguarda il Wi-Fi, solo lo standard 802.11 originale supportava una forma di FHSS. I sistemi OFDM mantengono un blocco del canale e quindi possono essere soggetti al problema del fading multipath.

- **Multiplexing spaziale:** viene utilizzato per fornire ulteriore capacità di dati utilizzando i percorsi multipli per trasportare traffico aggiuntivo, ovvero aumentando la capacità di throughput dei dati. In sostanza, un singolo flusso di dati ad alta velocità verrà suddiviso in più trasmissioni separate su diverse antenne. Una WLAN dividerà i dati in più flussi chiamati flussi spaziali. Ciascun flusso spaziale trasmesso utilizzerà un'antenna diversa sul trasmettitore.

IEEE 802.11n consente quattro antenne e quattro flussi spaziali. Utilizzando più flussi inviati separatamente per antenne distanziate l'una dall'altra, la diversità spaziale in 802.11n dà una certa sicurezza che almeno un segnale sarà abbastanza forte da raggiungere il ricevitore. Sono necessarie almeno due antenne per supportare la funzionalità MIMO. Un processore di segnale digitale sul trasmettitore e sul ricevitore regolerà gli effetti multipath e ritarderà la trasmissione in visibilità ottica (percorso diretto) di un tempo sufficiente per allinearla perfettamente con i percorsi non in visibilità ottica. Ciò farà sì che i segnali si rafforzino.

Il protocollo IEEE 802.11 identifica i flussi MIMO con la notazione $M \times N : Z$, dove M è il numero massimo di antenne di trasmissione, N è il numero massimo di antenne di ricezione e Z è il numero massimo di flussi di dati che possono essere utilizzati contemporaneamente. Quindi, un MIMO di 3×2 : 2 implica che ci sono tre antenne di flusso di trasmissione e due antenne di flusso di ricezione, ma possono inviare o ricevere solo due flussi simultanei.

6.11 Beamforming

802.11n ha anche introdotto la funzione opzionale del beamforming (percorso della linea di comunicazione radio). 802.11n definisce due tipi di metodi di beamforming: feedback implicito e feedback esplicito.

Il disegno della slide precedente illustra gli effetti del beamforming in una situazione senza visibilità ottica tra trasmettitore e ricevitore. Nel peggiore dei casi, i segnali arrivano sfasati di 180 gradi e si annullano a vicenda. Con il beamforming, i segnali possono essere regolati in fasi per rafforzarsi a vicenda sul ricevitore.

Il beamforming si basa su più antenne distanziate per focalizzare un segnale in una posizione particolare. I segnali possono essere regolati in fasi e grandezze per arrivare alla stessa posizione e rafforzarsi

a vicenda, fornendo una migliore potenza e portata del segnale. Purtroppo 802.11n non ha standardizzato un singolo metodo per il beamforming e lo ha lasciato all'implementatore. Diversi produttori utilizzano tecniche differenti e quindi possono garantire solo che funzioni con hardware identico, ciò ha comportato che il beamforming non venisse adottato in modo diffuso

Vedremo che le tecnologie MIMO sono state utilizzate in molte altre aree, come 802.11ac e nella comunicazione a lungo raggio utilizzando radio 4G LTE cellulari.

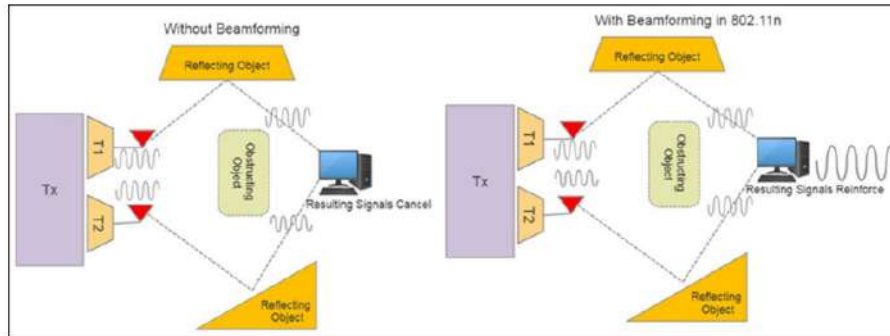


Figura 6.21: Beamforming

6.12 Protocolli 802.11 per Sistemi IoT: 802.11ac/802.11p/802.11ah

6.12.1 IEEE 802.11ac

IEEE802.11ac è la WLAN di nuova generazione e il seguito della famiglia di standard 802.11.

IEEE802.11ac è stato approvato come standard nel dicembre 2013 dopo cinque anni di lavoro. L'obiettivo è fornire un throughput multistazione di almeno 1 Gbps e un throughput di collegamento singolo di 500 Mbps. La tecnologia raggiunge questo obiettivo attraverso una larghezza di banda del canale più ampia (160 MHz), più flussi spaziali MIMO e modulazione a densità estrema (256-QAM).

802.11ac esiste solo nella banda 5 GHz, ma coesisterà con gli standard precedenti (IEEE802.11a/n).

Le specifiche e le differenze tra IEEE802.11ac e IEEE802.11n sono:

- Larghezza minima del canale 80 MHz con larghezza massima del canale 160 MHz;
- Otto flussi spaziali MIMO:
 - Introdotto da MU-MIMO downlink con un massimo di quattro client downlink;
 - Più STA con più antenne possono ora trasmettere e ricevere in modo indipendente su più flussi.
- Modulazione opzionale 256-QAM con la possibilità di utilizzare il beamforming standardizzato 1024-WAM.

MIMO multiutente merita ulteriori dettagli.

802.11ac estende 802.11n da quattro flussi spaziali a otto. Uno dei maggiori fattori che contribuiscono alla velocità 802.11ac è il multiplexing a divisione spaziale (SDM), a cui si è fatto riferimento in precedenza. Se combinata con aspetti multiutente o client multipli di 802.11ac, questa tecnica prende il nome di accesso multiplo a divisione di spazio (SDMA). In sostanza, MU-MIMO in 802.11ac è un analogo wireless di uno switch di rete.

802.11ac estende anche la costellazione di modulazione da 64-QAM a 256-WAM. Ciò implica che ci sono 16 livelli di ampiezza e 16 angoli di fase, che richiedono hardware molto preciso per essere implementati. 802.11n rappresentava sei bit per simbolo, mentre 802.11ac rappresenta otto bit per simbolo.

I metodi di beamforming sono stati formalmente standardizzati dal comitato IEEE. Ad esempio, il comitato ha convenuto che il feedback esplicito è l'approccio standard all'associazione beamforming. Ciò consentirà al beamforming e ai vantaggi in termini di prestazioni di essere disponibili da più fornitori.

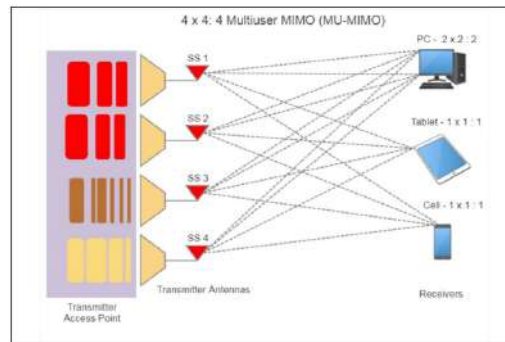


Figura 6.22: Sistema MU-MIMO 802.11ac 4 x 4: 4 con tre client

La maggiore larghezza di banda per canale (fino a 80 MHz con l'opzione di 160 MHz o due blocchi da 80 MHz) fornisce un aumento sostanziale del throughput nello spazio a 5 GHz.

Teoricamente, utilizzando un dispositivo 8×8 : 8, canali larghi 160 MHz e modulazione 256-QAM, si potrebbe sostenere un throughput aggregato di 6,933 GBps.

6.12.2 IEEE 802.11p

Le reti veicolari (a volte chiamate reti veicolari ad hoc o VANET) sono spontanee e non strutturate, funzionano come un'auto che si muove in una città mentre interagisce con altri veicoli e infrastrutture. Questo modello di rete utilizza i modelli da veicolo a veicolo (V2V) e da veicolo a infrastruttura (V2I). Nel 2004, il gruppo di lavoro 802.11p ha formato e sviluppato la prima bozza entro aprile 2010. 802.11p è considerato un canale dedicato di comunicazione a corto raggio (DSRC) all'interno del Dipartimento dei trasporti degli Stati Uniti. L'obiettivo di questa rete è fornire un sistema V2V e V2I standard e sicuro utilizzato per la sicurezza dei veicoli, la riscossione dei pedaggi, lo stato del traffico/avvisi, l'assistenza stradale e l'e-commerce all'interno di un veicolo.

La topologia e il caso d'uso generale per una rete IEEE 802.11p sono illustrati nella figura seguente. Esistono due tipi di nodi nella rete. Il primo è l'unità lato strada (RSU), che è un dispositivo di postazione fissa molto simile a un punto di accesso. Serve per collegare veicoli e dispositivi in movimento a Internet per l'utilizzo del servizio applicativo e l'accesso alle autorità di fiducia. L'altro tipo di nodo è l'unità di bordo (OBU), che risiede nel veicolo. È in grado di comunicare con altre OBU e le RSU fisse quando necessario.

Le OBU possono comunicare con le RSU e tra loro per trasmettere dati sul veicolo e sulla sicurezza. Le RSU vengono utilizzate per collegare i servizi dell'applicazione e le autorità di fiducia per l'autenticazione.

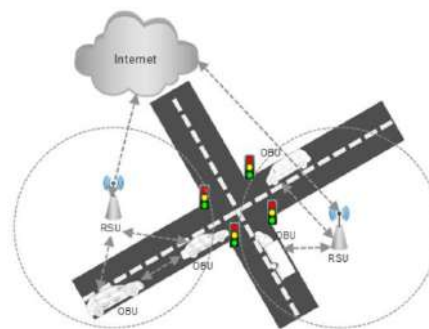


Figura 6.23: Esempio di utilizzo e topologia 802.11p: OBU all'interno dei veicoli e RSU con infrastruttura fissa

Per i sistemi di trasporto, esistono diverse sfide nella comunicazione wireless. C'è un elevato livello di sicurezza necessario nella comunicazione e nel controllo veicolare. Effetti fisici come spostamenti Doppler, effetti di latenza e solide reti ad hoc sono alcuni problemi da considerare.

Molte delle differenze rispetto agli standard 802.11 consistono nel garantire qualità e autonomia rispetto alla velocità di trasmissione. Altri fattori sono le modifiche per ridurre la latenza nell'avvio di una connessione. Di seguito è riportato un riepilogo delle caratteristiche di IEEE 802.11p e delle differenze rispetto agli standard IEEE 802.11a:

- La larghezza del canale è 10 MHz invece dei 20 MHz utilizzati in 802.11a;
- IEEE 802.11p opera nella larghezza di banda di 75 MHz dello spazio di 5,9 GHz. Ciò implica che sono disponibili un totale di sette canali (uno di controllo, due critici e quattro di servizio);
- Supporta la metà dei bit rate rispetto a 802.11a, ovvero 3/4/5/6/9/12/18/24/27 Mbps;
- Ha gli stessi schemi di modulazione, come BPSK/QPSK/16QAM/64QAM e 52 sottoportanti;
- La durata del simbolo è diventata doppia rispetto a 802.11a: IEEE 802.11p supporta 8 µs, mentre 11a supporta 4 µs;
- L'intervallo di tempo di guardia è 1,6 µs in 802.11p, mentre 11a ha 0,8 µs
- Tecniche come MIMO e beamforming non sono necessarie o fanno parte delle specifiche.

Il modello di utilizzo fondamentale dell'802.11p consiste nel creare e associare rapidamente reti ad hoc. Questo collegamento andrà e verrà man mano che i veicoli si allontanano l'uno dall'altro e gli altri veicoli alla fine ricostituiscono il tessuto. Nel modello 802.11 standard, un BSS sarebbe la topologia di rete da utilizzare, che richiede la sincronizzazione, l'associazione e l'autenticazione per la formazione di una rete wireless. 802.11p fornisce un BSSID jolly nell'intestazione di tutti i frame scambiati e consente ai collegamenti di iniziare a scambiare frame di dati immediatamente dopo l'arrivo su un canale di comunicazione.

Lo stack del protocollo IEEE 802.11p è derivato da 802.11a ma apporta importanti modifiche per affrontare la sicurezza e la protezione dei veicoli. La figura seguente illustra lo stack del protocollo. Una differenza significativa rispetto ad altri stack IEEE 802.11 è l'uso degli standard IEEE 1609.x per affrontare i modelli di applicazioni e sicurezza.

Lo stack completo è chiamato accesso wireless in ambienti veicolari (WAVE) e combina livelli PHY e MAC 802.11p con livelli IEEE 1609.x.

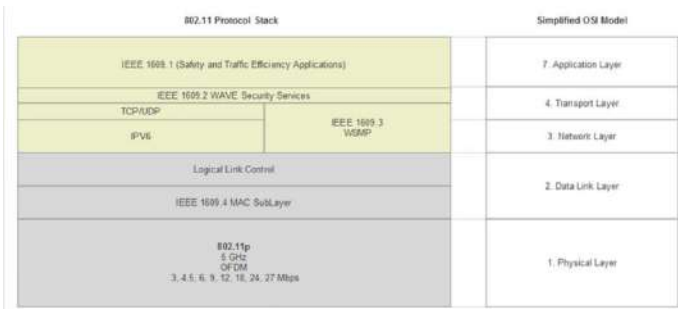


Figura 6.24: Stack del protocollo automobilistico 802.11p

6.12.3 IEEE 802.11ah

Basato su architettura 802.11ac e PHY, 802.11ah è una variante dei protocolli wireless destinati all'IoT. Il progetto tenta di ottimizzare i dispositivi con sensori vincolati che richiedono una lunga durata della batteria e possono ottimizzare la portata e la larghezza di banda. 802.11ah è anche indicato come HaLow, che è essenzialmente un gioco di parole con "ha" che si riferisce ad "ah" all'indietro e "basso" che implica bassa potenza e frequenza più bassa. Messa insieme, forma un derivato di "hello".

L'intento del gruppo di lavoro IEEE 802.11ah era quello di creare un protocollo con una portata estesa per le comunicazioni rurali e l'offload del traffico cellulare. Lo scopo secondario era quello di utilizzare il protocollo per comunicazioni wireless a bassa velocità nella gamma sub-gigahertz. La specifica è stata pubblicata il 31 dicembre 2016.

L'architettura è molto diversa da altre forme di standard 802.11 nei seguenti modi in particolare:

- Funziona nello spettro di 900 MHz. Ciò consente una buona propagazione e penetrazione dei materiali e delle condizioni atmosferiche;
- La larghezza del canale varia e può essere impostata su 2, 4, 8 o 16 MHz di larghezza. I metodi di modulazione disponibili sono diversi e includono le tecniche di modulazione BPSK, QPSK, 16-QAM, 64-WAM e 256-QAM;
- Modulazione basata sullo standard 802.11ac con modifiche specifiche. Un totale di 56 sottoportanti OFDM con 52 dedicate ai dati e 4 dedicate ai toni pilota. La durata totale del simbolo è di 36 o 40 microsecondi;
- Supporta il beamforming SU-MIMO;
- Associazione rapida per reti di migliaia di STA utilizzando due diversi metodi di autenticazione per limitare la contesa;
- Fornisce connettività a migliaia di dispositivi in un unico punto di accesso. Include la possibilità di inoltrare per ridurre la potenza sulle STA e consentire una forma grezza di rete mesh utilizzando un metodo di portata one-hop;
- Consente una gestione avanzata dell'alimentazione su ogni nodo 802.11ah;
- Consente la comunicazione della topologia non a stella tramite l'uso di finestre ad accesso limitato (RAW);
- Consente la settorizzazione, che consente di raggruppare le antenne per coprire diverse regioni di un BSS (denominate settori). Ciò si ottiene utilizzando il beamforming adottato da altri protocolli 802.11.

Il throughput minimo sarà di 150 Kbps, basato sulla modulazione BPSK su un singolo flusso MIMO con una larghezza di banda del canale di 1 MHz. Il throughput teorico massimo sarà 347 Mbps basato su una modulazione 256-WAM utilizzando 4 flussi MIMO e canali 16 MHz.

La specifica IEEE 802.11ah richiede che le STA supportino larghezze di banda del canale da 1 MHz e 2 GHz. I punti di accesso devono supportare canali a 1, 2 e 4 MHz. I canali da 8 MHz e 16 MHz sono opzionali. Più stretta è la larghezza di banda del canale, maggiore è l'intervallo ma più lento è il throughput.

Più ampia è la larghezza di banda del canale, più breve è la gamma ma più veloce è il throughput.

La larghezza del canale varia a seconda della regione in cui è distribuito 802.11ah. Alcune combinazioni non funzioneranno a causa di normative in regioni specifiche.

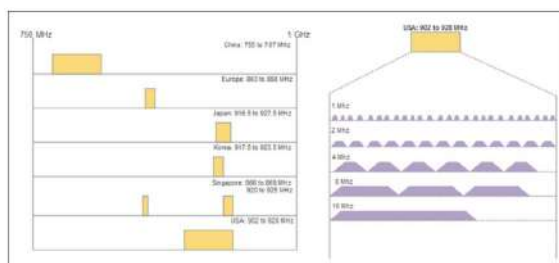


Figura 6.25

L'obiettivo di connettere diverse migliaia di dispositivi a un singolo AP viene raggiunto anche utilizzando un'assegnazione di un identificatore di associazione univoco (AID) di 13 bit. Ciò consente il raggruppamento di STA in base a criteri (luci di corridoio, interruttore della luce e così via). Ciò consente

a un AP di connettersi a oltre 8191 STA. (802.11 potrebbe supportare solo STA 2007.) Che molti nodi, tuttavia, hanno il potenziale per indurre un numero enorme di collisioni di canale. Anche se il numero di STA collegate è aumentato, l'obiettivo era ridurre la quantità di dati in transito per indirizzare queste stazioni.

Dal punto di vista della topologia, ci sono tre tipi di stazioni in una rete 802.11ah:

- **Punto di accesso radice (root access point):** la radice principale. In genere, funge da gateway per altre reti (WAN);
- **STA:** la tipica stazione 802.11 o client endpoint;
- **Nodo di inoltro (relay node):** un nodo speciale che combina un'interfaccia AP con STA che risiedono su un BSS inferiore e un'interfaccia STA con altri nodi di inoltro o un AP radice sul BSS superiore.

Questa architettura differisce sostanzialmente da altri protocolli 802.11 nell'uso di nodi di inoltro a hop singolo che agiscono per creare un BSS identificabile. La gerarchia dei relè forma una rete più ampia. Ogni staffetta funge da AP e STA.

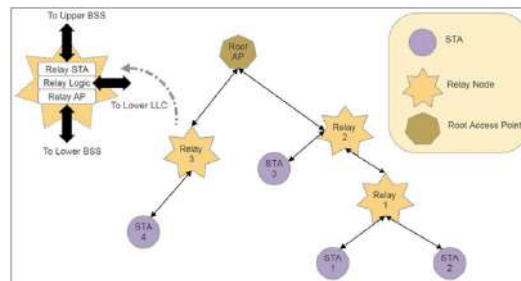


Figura 6.26: Topologia IEEE802.11ah

Oltre ai tipi di nodo di base, ci sono tre stati di risparmio energetico in cui una STA può risiedere:

- **Mappa di indicazione del traffico (TIM):** ascolta l'AP per il trasferimento dei dati. I nodi riceveranno periodicamente informazioni sui dati memorizzati per loro dal proprio punto di accesso. Il messaggio inviato prende il nome di elemento informativo TIM;
- **Stazioni non TIM:** negozia con AP direttamente durante l'associazione per ottenere il tempo di trasmissione su Periodic Restricted Access Windows (PRAW);
- **Stazioni non programmate:** non ascolta i beacon e utilizza il polling per accedere ai canali.

6.12.3.1 Conclusione

L'alimentazione è fondamentale nei sensori IoT e nei dispositivi edge basati su batterie a bottone o sull'energy harvesting. I protocolli 802.11 sono noti per le richieste di potenza elevata. Per rimediare alla potenza di questo protocollo wireless, 802.11ah utilizza un valore Max Idle Period, che fa parte delle normali specifiche 802.11. In una rete 802.11 generale, il periodo di inattività massimo è di circa 16 ore in base a un tempo di risoluzione a 16 bit. In 802.11ah, i primi due bit del timer a 16 bit sono un fattore di scala che consente alla durata del sonno di superare i cinque anni.

La potenza aggiuntiva viene mitigata attraverso modifiche al beacon. Come spiegato in precedenza, i beacon trasmettono informazioni sulla disponibilità di frame bufferizzati. I beacon porteranno una bitmap TIM, che ne gonfia le dimensioni poiché 8191 STA faranno crescere notevolmente la bitmap. 802.11ah utilizza un concetto chiamato segmentazione TIM in cui alcuni beacon trasportano porzioni della bitmap complessiva. Ogni STA calcola quando arriverà il rispettivo beacon con informazioni bitmap e consente al dispositivo di entrare in una modalità di risparmio energetico fino al momento in cui deve riattivarsi e ricevere informazioni beacon.

Un'altra funzionalità di risparmio energetico è denominata Target Wake Time (TWT), destinata agli

STA che raramente trasmettono o ricevono dati. Questo è molto comune nelle implementazioni IoT come i dati dei sensori di temperatura. Una STA e il suo AP associato negozieranno per arrivare a un TWT concordato e la STA entrerà in uno stato di sospensione fino a quando non verrà segnalato quel timer.

6.13 Il Progetto 6LoWPAN

6.13.1 WPAN con IP - 6LoWPAN

Nel tentativo di portare l'indirizzamento IP a dispositivi più piccoli e con risorse limitate, nel 2005 è stato formato il concetto di 6LoWPAN. Un gruppo di lavoro ha formalizzato il progetto nell'IETF secondo la specifica RFC 4944 (RFC = Request For Comment) e successivamente aggiornato con RFC 6282 per la compressione dell'intestazione e RFC 6775 per il rilevamento dei dispositivi vicini. Il consorzio ora è terminato ma lo standard può essere implementato ed utilizzato da chiunque.

6LoWPAN è un acronimo che sta per IPV6 su reti personali wireless a bassa potenza. L'intento è quello di consentire di realizzare reti IP su sistemi di comunicazione RF a bassa potenza per dispositivi che sono limitati in termini di alimentazione e spazio e non necessitano di servizi di rete a larghezza di banda elevata. Il protocollo può essere utilizzato con altre comunicazioni WPAN come 802.15.4, nonché Bluetooth, protocolli RF sub-1 GHz e sistemi di trasmissione su Power Line (PLC = Power Line Controller). Il principale vantaggio di 6LoWPAN è che il più semplice dei sensori può essere dotato di indirizzamento IP e agire come soggetto di rete su router 3G/4G/LTE/Wi-Fi/Ethernet. Un effetto secondario è che IPV6 ha uno spazio di indirizzamento che genera fino a 2^{128} o $3,4 \times 10^{38}$ indirizzi univoci. Ciò coprirebbe sufficientemente circa 50 miliardi di dispositivi connessi a Internet e continuerebbe a coprire quei dispositivi ben oltre. Pertanto, 6LoWPAN è adatto per la crescita dell'IoT.

6.13.2 Topologia 6LoWPAN

Le reti 6LoWPAN sono reti mesh che risiedono alla periferia di reti più grandi. Le topologie sono flessibili, consentendo reti ad hoc e disgiunte senza alcun legame con Internet o altri sistemi, oppure possono essere collegate al backbone o a Internet utilizzando router edge. Le reti 6LoWPAN possono essere unite a più router edge; questo è chiamato multihoming. Inoltre, è possibile creare reti ad hoc senza richiedere la connettività Internet di un router edge.

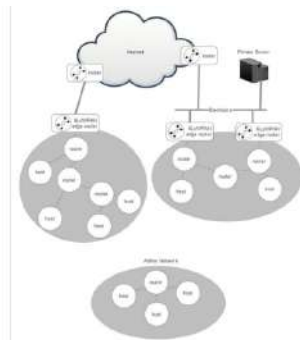


Figura 6.27: Le topologie 6LoWPAN

Un router edge (noto anche come router di confine) è necessario per un'architettura 6LoWPAN in quanto ha quattro funzioni:

- Gestisce la comunicazione con i dispositivi 6LoWPAN e trasmette i dati a internet;
- Esegue la compressione delle intestazioni IPv6 riducendo un'intestazione IPv6 a 40 byte e intestazioni UDP a 8 byte per l'efficienza in una rete di sensori. Una tipica intestazione IPv6 a 40 byte può comprimere da 2 a 20 byte a seconda dell'utilizzo;
- Avvia la rete 6LoWPAN;

- Scambia i dati tra i dispositivi sulla rete 6LoWPAN;

I router perimetrali formano reti mesh 6LoWPAN su perimetri di rete tradizionali più ampi. Possono anche mediare scambi tra IPv6 e IPv4, se necessario. I datagrammi vengono gestiti in modo simile a una rete IP, che presenta alcuni vantaggi rispetto ai protocolli proprietari. Tutti i nodi all'interno di una rete 6LoWPAN condividono lo stesso prefisso IPv6 stabilito dal router perimetrale. I nodi si registreranno con i router perimetrali come parte della fase di Network Discovery (ND).

ND controlla come gli host e i router nella mesh 6LoWPAN locale interagiranno tra loro. Il multihoming consente a più router edge 6LoWPAN di gestire una rete, ad esempio quando è necessario disporre di più supporti (4G e Wi-Fi) per il failover o la tolleranza agli errori.

Esistono tre tipi di nodi all'interno della mesh 6LoWPAN:

- **Nodi router:** questi nodi effettuano il marshalling dei dati da un nodo mesh 6LoWPAN a un altro. I router possono anche comunicare verso l'esterno con la WAN e Internet;
- **Nodi host:** gli host nella rete mesh non possono instradare i dati nella mesh e sono semplicemente endpoint che consumano o producono dati. Gli host possono essere in stato di sospensione, svegliandosi occasionalmente per produrre dati o ricevere dati memorizzati nella cache dai router principali;
- **Router edge:** come affermato, questi sono i gateway e i controller mesh di solito su un edge WAN. Una mesh 6LoWPAN verrebbe amministrata sotto l'edge router.

I nodi sono liberi di muoversi e riorganizzarsi/riassemblarsi in una mesh. Del resto, un nodo può spostarsi e associarsi a un router edge diverso in uno scenario multihome o persino spostarsi tra diverse mesh 6LoWPAN. Queste modifiche alla topologia possono essere causate da vari motivi, come cambiamenti nella potenza del segnale o movimento fisico dei nodi. Quando si verifica una modifica della topologia, anche l'indirizzo IPv6 dei nodi associati cambierà naturalmente.

In una mesh ad hoc senza un edge router, un nodo router 6LoWPAN potrebbe gestire una mesh 6LoWPAN. Questo sarebbe il caso quando la connettività WAN a Internet non è necessaria. In genere, questo è raramente visto poiché l'indirizzabilità IPv6 per una piccola rete ad hoc non è necessaria.

Il nodo router sarebbe configurato per supportare due funzioni obbligatorie:

- Generazione di indirizzi unicast locali univoci;
- Esecuzione della registrazione ND di neighbor discovery.

Il prefisso IPv6 mesh ad hoc sarebbe un prefisso locale anziché il prefisso IPv6 WAN globale più ampio.

6.13.3 Sicurezza in 6LoWPAN

Poiché, in un sistema WPAN, è facile ascoltare (sniffing) la comunicazione, 6LoWPAN fornisce sicurezza a più livelli. Al livello 802.15.4 del protocollo, 6LoWPAN si basa sulla crittografia dei dati AES-128. Inoltre, 802.15.4 fornisce un contatore con modalità CBC-MAC (CCM) per fornire la crittografia e un controllo di integrità. La maggior parte dei chipset che forniscono un blocco di rete 802.15.4 include anche un motore di crittografia hardware per il miglioramento delle prestazioni.

Al livello tre (il livello di rete) del protocollo, 6LoWPAN ha la possibilità di utilizzare la sicurezza standard IPsec (RFC4301). Ciò comprende:

- **Gestore di autenticazione (AH):** come definito in RFC4302 per la protezione dell'integrità e l'autenticazione;
- **Encapsulating Security Payload (ESP):** in RFC4303, aggiunge la crittografia per garantire la riservatezza dei pacchetti.

ESP è di gran lunga il formato di pacchetto sicuro di livello tre più comune. Inoltre, una modalità di ESP definisce il riutilizzo di AES/CCM utilizzato nell'hardware di livello due anche per la crittografia di livello tre (RFC4309). Ciò rende la sicurezza di livello tre adatta per nodi 6LoWPAN vincolati.

Oltre alla sicurezza a livello di collegamento, 6LoWPAN utilizza anche Transport Layer Security (TLS) per il traffico TCP e Datagram Transport Layer Security (DTLS) per il traffico UDP.

6.13.4 WPAN con IP-Thread

Thread è un protocollo di rete relativamente nuovo per IoT e si basa su IPV6 (6LoWPAN). Il suo obiettivo principale è la connettività domestica e la domotica. Thread è stato lanciato nel luglio del 2014 con la formazione della Thread Group Alliance, che comprende aziende come Alphabet (la holding di Google), Qualcomm, Samsung, ARM, Silicon Labs, Yale (locks) e Tyco.

Basato sul protocollo IEEE 802.15.4 e 6LoWPAN, ha in comune con Zigbee e altre varianti 802.15.4, ma con una differenza significativa essendo Thread è indirizzabile IP. Questo protocollo IP si basa sui dati e sui livelli fisici forniti da 802.15.4 e funzionalità come sicurezza e routing da 6LoWPAN. Thread è anche basato su mesh, il che lo rende interessante per i sistemi di illuminazione domestica con un massimo di 250 dispositivi in una singola mesh.

La filosofia con Thread è che abilitando l'indirizzabilità IP nei più piccoli sensori e sistemi di automazione domestica, è possibile ridurre il consumo energetico e i costi perché il sensore abilitato per Thread non ha bisogno di mantenere lo stato dell'applicazione nello storage. Il thread si basa su datagrammi a livello di rete che, per sua stessa natura, elimina la necessità di elaborare le informazioni a livello di applicazione, risparmiando energia al sistema.

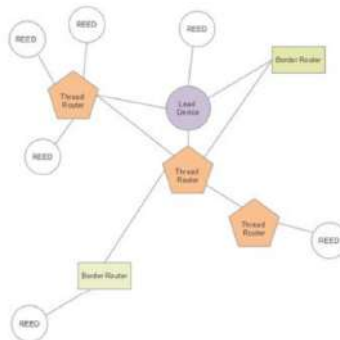
Infine, la conformità a IPV6 fornisce opzioni di sicurezza attraverso la crittografia utilizzando l'Advanced Encryption Standard (AES). Possono esistere fino a 250 nodi su una mesh Thread, tutti con trasporto e autenticazione completamente crittografati. Un aggiornamento software consente a un dispositivo 802.15.4 preesistente di essere compatibile con Thread.

6.13.4.1 WPAN con IP-Architettura e Topologia del Thread

Basato sullo standard IEEE 802.15.4-2006, Thread utilizza la specifica per definire i livelli MAC (Medium Access Control) e fisici (PHY). Funziona a 250 Kbps nella banda GHz.

Dal punto di vista della topologia, Thread stabilisce comunicazioni con altri dispositivi attraverso un router di frontiera (di solito un segnale Wi-Fi in una famiglia). Il resto della comunicazione si basa su 802.15.4 e forma una rete autoriparante.

Esempio di topologia di rete di thread contenente router di confine, router di thread, dispositivi lead e dispositivi IoT idonei che possono essere combinati nella mesh. Le interconnessioni sono variabili e autorigeneranti.



Di seguito sono riportati i ruoli di vari dispositivi in un'architettura Thread:

Router di Frontiera (border router): un router di frontiera è essenzialmente un gateway. Nella rete domestica, questo sarebbe un crossover di comunicazione da Wi-Fi a Thread e formerebbe il punto di ingresso a Internet da una rete Thread che corre sotto un router di confine. Più router di frontiera sono consentiti in base alla specifica Thread.

Dispositivo principale (lead device): il dispositivo principale gestisce un registro di ID router assegnati. Il lead controlla anche le richieste di dispositivi finali idonei al router (REED) da promuovere ai router. Un leader può anche fungere da router e avere figli all'estremità del dispositivo. Il protocollo per l'assegnazione degli indirizzi del router è il Constrained Application Protocol (CoAP). Le informazioni sullo stato gestite da un dispositivo principale possono essere archiviate anche negli altri router di thread. Ciò consente l'auto-guarigione e il failover se il leader perde la connettività.

Thread router: i thread router gestiscono i servizi di routing della mesh. I router di thread non entrano mai in uno stato di sospensione, ma le specifiche consentono di eseguire il downgrade per diventare un REED.

REED: un dispositivo host che è un REED può diventare un router o un leader. I REED non sono responsabili del routing nella mesh a meno che non siano promossi a router o leader. Inoltre, i REED non possono inoltrare messaggi o unire dispositivi alla mesh. I REED sono essenzialmente endpoint o nodi foglia nella rete.

Dispositivi finali (end devices): alcuni endpoint non possono diventare router. Questi tipi di REED hanno altre due categorie a cui possono abbonarsi: dispositivi finali completi (FED) e dispositivi finali minimi (MED).

Dispositivi finali inattivi (sleepy end devices): i dispositivi host che sono entrati in uno stato di sospensione comunicano solo con il router thread associato e non possono inoltrare messaggi.

Capitolo 7

Sistemi e Protocolli di Comunicazione a Lungo Raggio

Finora abbiamo visto reti personali wireless (WPAN) e reti locali wireless (WLAN).

Questi tipi di comunicazione collegano i sensori a una rete locale ma non necessariamente a Internet o ad altri sistemi.

Dobbiamo ricordare che l'ecosfera IoT includerà sensori, attuatori, fotocamere, dispositivi smart embedded, veicoli e robot nei luoghi più remoti. Per i collegamenti a lungo raggio si parla di WAN (Wide Area Network). Vedremo essenzialmente tecnologie cellulari (standard 4G LTE e 5G) e daremo cenni di LoRa e Sigfox.

Tratteremo dei sistemi di comunicazione cellulare a lungo raggio dal punto di vista dei dati, tralasciando le parti analogiche e vocali dei dispositivi mobili.

La comunicazione a lungo raggio è solitamente un servizio che viene fornito in abbonamento da un operatore che gestisce o noleggia l'infrastruttura ed è quindi differente dalle precedenti architetture WPAN e WLAN poiché di solito sono realizzate con dispositivi che il cliente acquista.

Il sistemista IoT deve essere consapevole che la scelta di uno specifico abbonamento o di un contratto di servizio (SLA) ha effetti sull'architettura dei sistemi e lo vincola.

7.1 Tecnologie cellulari

7.1.1 Connettività cellulare

La forma di comunicazione più diffusa è la radio cellulare e in particolare i dati cellulari. Sebbene i dispositivi di comunicazione mobile esistessero da molti anni prima della tecnologia cellulare, avevano una copertura limitata, uno spazio di frequenza condiviso ed erano essenzialmente radio a due vie.

Bell Labs ha costruito alcune tecnologie di telefonia mobile di prova negli anni '40 (servizio di telefonia mobile) e negli anni '50 (servizio di telefonia mobile migliorato) ma il successo è stato molto limitato.

All'epoca non esistevano standard per la telefonia mobile.

Fu solo quando Douglas H. Ring e Rae Young proposero il concetto di "cellulare" nel 1947, poi realizzato da Richard H. Frenkiel, Joel S. Engel e Philip T. Porter presso i Bell Labs negli anni '60, che il cellulare cominciò ad avere possibilità di diffusione reali.

Il trasferimento tra le celle è stato concepito e costruito da Amos E. Joel Jr., anche lui di Bell Labs.

Tutte queste tecnologie si sono combinate per formare il primo sistema telefonico cellulare.

Il primo telefono cellulare fu utilizzato per effettuare la prima chiamata cellulare il 3 aprile 1979 da Martin Cooper di Motorola.

Di seguito è riportato un modello di infrastruttura cellulare ideale in cui le cellule sono rappresentate come aree esagonali di posizionamento ottimale.

Le tecnologie e i progetti di prova alla fine portarono alle prime implementazioni commerciali e all'accettazione pubblica dei sistemi di telefonia mobile nel 1979 da parte di NTT in Giappone e poi in

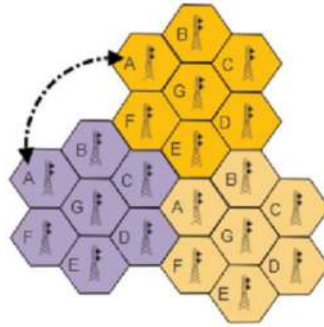


Figura 7.1: Il modello esagonale garantisce la separazione delle frequenze tra celle adiacenti. Non ci sono due frequenze simili entro uno spazio esadecimale l'una dall'altra, come mostrato nel caso della frequenza A in due regione diverse. Ciò consente il riutilizzo della frequenza.

Danimarca, Finlandia, Norvegia e Svezia nel 1981.

Le Americhe non avevano un sistema cellulare fino al 1983. Queste prime tecnologie sono conosciute come 1G o la prima generazione di tecnologia cellulare.

7.1.2 Modelli e Standard di Governance

L'International Telecommunication Union (ITU) è un'agenzia specialistica fondata nel 1865; prese il nome attuale nel 1932, prima di diventare un'agenzia specialistica dell'ONU.

Svolge un ruolo significativo a livello mondiale negli standard di comunicazione wireless, navigazione, dispositivi mobili, Internet, dati, voce e reti di nuova generazione.

Comprende 193 nazioni membri e 700 organizzazioni pubbliche e private. Ha anche una serie di gruppi di lavoro chiamati settori.

ITU-R è il settore delle radiocomunicazioni ed quello più rilevante per gli standard cellulari.

L'ITU-R è l'organismo che definisce gli standard e gli obiettivi internazionali per le varie generazioni di comunicazioni radio e cellulari. Questi includono obiettivi di affidabilità e velocità dati minime.



Figura 7.2

L'ITU-R ha prodotto due specifiche fondamentali che hanno governato la comunicazione cellulare nell'ultimo decennio. Il primo è stato l'International Mobile Telecommunications-2000 (IMT-2000), che specifica i requisiti per un dispositivo da commercializzare come 3G.

Più recentemente, l'ITU-R ha prodotto una specifica dei requisiti denominata International Mobile Telecommunications-Advanced (IMT-Advanced).

Il sistema IMT-Advanced si basa su un sistema wireless a banda larga mobile all-IP.

L'IMT-Advanced definisce ciò che può essere commercializzato come 4G in tutto il mondo.

L'ITU è stato il gruppo che, nell'ottobre del 2010, ha approvato la tecnologia LTE (Long-Term Evolution) nella tabella di marcia 3GPP (Third Generation Partnership Project) per supportare gli obiettivi della comunicazione cellulare 4G.

L'ITU-R continua a guidare i nuovi requisiti per 5G.

Esempi della serie di requisiti ITU-Advanced per un sistema cellulare da etichettare 4G includono:

- Deve essere una rete all-IP a commutazione di pacchetto interoperabile con il wireless esistente;
- Una velocità dati nominale di 100 Mbps quando il client è in movimento e 1 Gbps quando il client è fisso;

- Condivide ed utilizza dinamicamente le risorse di rete per supportare più di un utente per cella;
- Larghezza di banda del canale scalabile da 5 a 20 MHz;
- Connettività senza interruzioni e roaming globale su più reti.

Il problema è che spesso non vengono raggiunti tutti gli obiettivi posti dall'ITU e questo genera confusione di denominazione e marchio:

Feature	1G	2/2.5G	3G	4G	5G
First availability	1979	1999	2002	2010	2020
ITU-R specification	NA	NA	IMT-2000	IMT-Advanced	IMT-2020
ITU-R frequency specification	NA	NA	400 MHz to 3 GHz	450 MHz to 3.6 GHz	600 MHz to 6 GHz 24-86 GHz (mmWave)
ITU-R bandwidth specification	NA	NA	Stationary: 2 Mbps Moving: 384 Kbps	Stationary: 1 Gbps Moving: 100 Mbps	Min down: 20 Gbps Min up: 10 Gbps

Figura 7.3

Typical bandwidth	2 Kbps	14.4-64 Kbps	500 to 700 Kbps	100 to 300 Mbps (peak)	1 Gbps
Usage/features	Mobile telephony only	Digital voice, SMS text, caller-ID, one-way data	Superior audio, video, and data Enhanced roaming	Unified IP and seamless LAN/WAN/WLAN	IoT, ultra density, low latency
Standards and multiplexing	AMPS	2G: TDMA, CDMA, GSM 2.5G: GPRS, EDGE, 1xRTT	FDMA, TDMA WCDMA, CDMA-2000, TD-SCDMA	CDMA	CDMA
Handoff	Horizontal	Horizontal	Horizontal	Horizontal and vertical	Horizontal and vertical
Core network	PSTN	PSTN	Packet Switch	Internet	Internet

Figura 7.4

Switching	Circuit	Circuit for access network and air network	Packet-based except for air interface	Packet-based	Packet-based
Technology	Analog cellular	Digital cellular	Broad bandwidth CDMA, WiMAX, IP-based	LTE Advanced Pro-based	LTE Advanced Pro-based, mmWave

Figura 7.5

Il Third Generation Partnership Project (3GPP) è l'altro organismo standard nel mondo cellulare. È un gruppo di sette organizzazioni di telecomunicazioni (note anche come partner organizzativi) che gestiscono e governano la tecnologia cellulare.

Il gruppo si è formato nel 1998 con la partnership di Nortel Networks e AT&T Wireless e ha rilasciato il primo standard nel 2000.

Partner organizzativi e rappresentanti di mercato contribuiscono al 3GPP da Giappone, Stati Uniti, Cina, Europa, India e Corea.

L'obiettivo generale del gruppo è riconoscere gli standard e le specifiche per il Global System for Mobile Communications (GSM) nella creazione delle specifiche 3G per la comunicazione cellulare.

Il lavoro di 3GPP è svolto da tre gruppi di specifiche tecniche (STG) e sei gruppi di lavoro (WG). I gruppi si incontrano più volte all'anno in diverse regioni.

Il fine principale delle versioni 3GPP è rendere il sistema compatibile con le versioni precedenti e successive (per quanto possibile).

C'è un certo grado di confusione nel settore per quanto riguarda le differenze tra le definizioni ITU, 3GPP e LTE.

Il modo più semplice per concettualizzare la relazione è il seguente percorso in 3 step:

1. ITU definisce gli obiettivi e propone gli standard mondiali per un dispositivo da etichettare 4G o 5G;
2. 3GPP risponde agli obiettivi con tecnologie migliorative per LTE;
3. ITU conferma che tali progressi LTE soddisfano i requisiti per essere etichettati 4G o 5G.

LTE sta per Long-Term Evolution ed è il percorso seguito per raggiungere velocità e requisiti ITU-R (che inizialmente erano piuttosto aggressivi).

I fornitori di telefoni cellulari rilasciano nuovi smartphone utilizzando la tecnologia back-end legacy come il 3G.

Gli operatori commercializzano la connettività 4G LTE se hanno dimostrato un sostanziale miglioramento della velocità e delle funzionalità rispetto alle loro reti 3G legacy.

Tra la metà e la fine degli anni 2000, molti vettori non soddisfacevano le specifiche ITU-R 4G pur avvicinandosi abbastanza.

I gestori hanno utilizzato tecnologie legacy e in molti casi si sono rinominati come 4G.

LTE-Advanced è un altro miglioramento che si avvicina ancora di più agli obiettivi ITU-R.

In sintesi, la terminologia può essere fonte di confusione e fuorviante e un sistemista IoT deve saper leggere oltre le etichette del marchio per comprendere la tecnologia.

7.1.2.1 Modelli e Standard di Governance 5G

5G (o 5G-NR per New Radio) è il nuovo standard di comunicazione basato su IP.

Utilizza alcune tecnologie del 4G LTE ma presenta alcune differenze sostanziali e nuove funzionalità.

Il 5G promette di fornire funzionalità fondamentali per casi d'uso IoT:

- Migliora la larghezza di banda;
- Riduce la latenza;
- Aumenta la densità;
- Abbassa il costo.

Piuttosto che creare servizi cellulari diversi e categorie per ogni caso d'uso, il 5G tenta di essere un unico standard ombrello per servirli tutti.

Nel frattempo, 4G LTE continuerà a essere la tecnologia predominante per la copertura cellulare e continuerà ad evolversi.

Il 5G non è una evoluzione del 4G; deriva dal 4G ma è un nuovo insieme di tecnologie.

Il 5G è stato lanciato per i clienti finali nel 2020; tuttavia, la distribuzione e l'adozione di massa potrebbero seguire per la metà degli anni '20.

Gli obiettivi e l'architettura del 5G sono ancora in evoluzione e lo sono dal 2012.

Ci sono tre obiettivi distinti per il 5G (<http://www.gsmhistory.com/5g/>):

- Convergenza tra infrastruttura in fibra e infrastruttura cellulare;
- Cellulari ultra veloci che utilizzano piccole celle;
- Riduzione del costo dei dispositivi cellulari.

Anche in questo caso, l'ITU-R ha approvato le specifiche e gli standard internazionali, mentre il 3GPP sta seguendo una serie di standard che corrispondono alla sequenza temporale dell'ITU-R.

La RAN 3GPP ha già iniziato ad analizzare gli elementi di studio a partire dalla Release 14. L'intento è quello di produrre una versione in due fasi delle tecnologie 5G.

Il grafico presente nella prossima slide illustra la tabella di marcia ITU e 3GPP per il 5G. E' evidente il percorso di lancio graduale per il 5G.

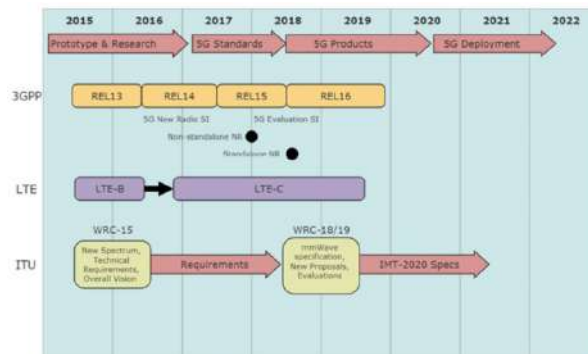


Figura 7.6: 3GPP ha accelerato il 5G utilizzando un concetto definito non autonomo(NSA). NSA riutilizza l'infrastruttura principale LTE. Al contrario, standalone (SA) si baserà esclusivamente sull'infrastruttura 5G Next Generation Core

Con la tecnologia 5G si vogliono raggiungere 3 macro obiettivi:

1. Banda larga mobile avanzata (eMBB):

- Connessioni da 1 a 10 GBps a UE/endpoint sul campo (non teoriche)
- Copertura del 100% in tutto il mondo (o percezione di)
- Da 10 a 100 volte il numero di dispositivi collegati su 4G LTE
- Connettività a una velocità di 500 km/h

2. Comunicazioni ultra affidabili e a bassa latenza (URLLC):

- < 1 ms di latenza di andata e ritorno end-to-end
- 99,999% di disponibilità (o percezione di)

3. Comunicazioni massive tra macchine (mMTC):

- 1000 volte la larghezza di banda per unità di area; implica circa 1 milione di nodi in 1 kmq
- Durata della batteria dei nodi di collegamento degli endpoint IoT fino a 10 anni.
- Riduzione del 90% del consumo di energia della rete

Vengono presi in considerazione due tipi principali di implementazioni 5G:

- La prima implementazione è il wireless mobile tradizionale che utilizza portante inferiore a 2 GHz ed ottimizza portata (raggio di azione), mobilità e consumi energetici. Si basa su macrocelle e sull'attuale compatibilità LTE. Inoltre deve essere in grado di penetrare segnali contro pioggia, nebbia e altri ostacoli. La larghezza di banda di picco è di appena 100 MHz
- La seconda implementazione è wireless fisso che utilizza frequenze al di sopra della gamma di 6 GHz. Si basa fortemente sulla nuova infrastruttura a piccole celle ed è destinato a casi a bassa mobilità o geografici fissi. La portata sarà limitata, così come la capacità di penetrazione. La larghezza di banda, tuttavia, sarà notevole a 400 MHz.

Occorre tenere presente che la tecnologia mmWave ha problemi di attenuazione e di penetrazione.

7.2 Tecnologia LoRaWAN

LPWAN include anche tecnologie proprietarie e non sponsorizzate da 3GPP.

LoRa è un livello fisico per un protocollo IoT a lungo raggio e a bassa potenza mentre LoRaWAN rappresenta il livello MAC.

Queste tecnologie e vettori proprietari LPWAN hanno il vantaggio di utilizzare lo spettro senza licenza e questo, semplicemente, abbatte il costo dell'utilizzo dei dati.



Figura 7.8: Tecnologia LoRaWAN

LoRa/LoRaWAN può essere costruito, personalizzato e gestito da chiunque. Non vi è alcuna necessità di contratto di servizio con operatori.

In genere, tecnologie come LoRaWAN (e anche Sigfox che vedremo tra poco) hanno una velocità di trasmissione dati da 5 a 10 volte inferiore rispetto alle tradizionali connessioni 3G o LTE per implementazioni di grandi volumi (100.000 unità).

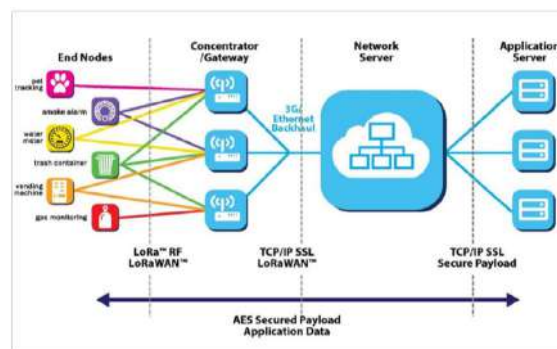


Figura 7.9: AES



Figura 7.10: Contatore d'acqua elettronico LoRaWAN

Scheda prodotto:

- Diametro nominale tubazione: 15mm
- Pressione operativa acqua: fino a 1MPa
- Temperatura operativa acqua: +5...+90 C
- Temperatura operativa ambiente: +5...+50 C
- LORAWAN: Classe A; ADR-Adaptive; Conferma pacchetti Si/No; 16 canali.

7.3 Tecnologia Sigfox

Sigfox è un protocollo LPWAN a banda stretta sviluppato nel 2009 a Tolosa, in Francia. La società fondatrice ha lo stesso nome. Questa è un'altra tecnologia LPWAN che utilizza le bande ISM senza licenza per un protocollo proprietario. Sigfox ha alcuni tratti che ne restringono significativamente l'utilità. Sebbene ci siano rigide limitazioni di Sigfox in termini di velocità effettiva e utilizzo, è destinato ai sistemi che inviano treni di dati piccoli e poco frequenti.



Figura 7.11

I dispositivi IoT come sistemi di allarme, semplici misuratori di potenza e sensori ambientali sarebbero candidati. I dati per vari sensori possono in genere rientrare nei vincoli, come i dati di temperatura/umidità rappresentati in 2 byte con una precisione di 0,004 gradi. È necessario prestare attenzione al grado di precisione fornito dal sensore e alla quantità di dati che possono essere trasmessi.

Stazione Meteo MeteoHelix® IoT Pro - Sigfox Micro-Weather Station Stazione micro-meteo wireless all-in-one professionale con schermo solare.

Autoricaricabile con pannello solare integrato e batteria interna che può durare fino a 6 mesi senza ricarica.

Disponibile in due versioni wireless: Sigfox e LoRaWAN.

Funzionalità:

- Misura temperatura dell'aria con precisione WMO
- Misura umidità dell'aria con precisione WMO con uscita del punto di rugiada e gelo
- Misura pressione atmosferica
- Misura irradiazione solare (piranometro)



Figura 7.12



Figura 7.13

Conclusioni Sebbene vi sia una certa comunanza tra i tipi di tecnologie di comunicazione a lungo raggio, ognuno si rivolge a diversi casi d'uso ed il sistemista IoT dovrebbe scegliere saggiamente quale adottare perché, come avviene per altri componenti del sistema IoT, LPWAN è difficile da modificare una volta implementato.

La scelta della LPWAN corretta dovrebbe quindi considerare una serie di aspetti, ne elenchiamo i più salienti:

- Quale velocità dati deve utilizzare la distribuzione IoT?
- La soluzione può scalare con la stessa LPWAN in tutte le regioni? C'è una copertura adeguata o deve essere costruita?
- Quale raggio di trasmissione è appropriato?
- C'è qualche requisito di latenza per questa soluzione IoT? La soluzione può funzionare con latenze molto elevate (più secondi)?
- Gli endpoint IoT sono alimentati a batteria e qual è il costo della loro manutenzione? Quali sono i vincoli di costo degli endpoint?

Parte IV

Architettura

Capitolo 8

Architettura Edge e Cloud

8.1 Significato ed Esempi di "Edge Computing"

L'IoT riceve molta attenzione dall'industria e dall'economia a causa del numero di dispositivi che verranno implementati e della quantità di dati che tali dispositivi produrranno. Esistono due metodi per quanto riguarda il modo in cui i dispositivi perimetrali e i sensori funzioneranno e comunicheranno con Internet:

- I sensori e i dispositivi di livello edge avranno un percorso diretto verso il cloud. Ciò implica che questi nodi e sensori edge-level disporranno di risorse, hardware, software e accordi sul livello di servizio sufficienti per trasmettere i dati direttamente attraverso la WAN;
- I sensori edge-level formeranno aggregazioni e cluster attorno a gateway e router per fornire aree di staging, conversioni di protocollo e capacità di elaborazione edge/fog, gestendo anche la sicurezza e l'autenticazione tra i sensori e la WAN.

Le piattaforme di elaborazione che funzionano al di fuori degli ambienti aziendali e dei data center gestiti dall'IT non sono nuove.

I sistemi edge sono essenzialmente sistemi di elaborazione remota e prendono in prestito elementi da domini ingegneristici consolidati quali (ma non solo) i sistemi embedded, la sicurezza informatica e le telecomunicazioni.

Una delle prime forme di elaborazione gestita in remoto esisteva molto prima dei sistemi cloud e dell'informatica generale.

Negli anni '30 e '40, negli Stati Uniti, la produzione di energia elettrica stava crescendo rapidamente utilizzando progetti idroelettrici ed estendendo la rete elettrica.

Per gestire l'ampia distesa di centrali elettriche è stato realizzato un sistema di telecontrollo tramite fili pilota. Possiamo pensare ai cavi pilota come a segnali di controllo della banda laterale al di fuori delle linee ad alta potenza e collegati a un circuito di emergenza per controllare il flusso di energia e affrontare in remoto i problemi di capacità e rilevare i guasti.

Gli operatori remoti potevano controllare la rete nazionale da centri operativi piuttosto che presidiare ciascuna centrale elettrica e sottostazione. Successivamente, questi dispositivi sono avanzati e sono diventati quelli che vengono chiamati controllori logici programmabili (PLC).

Ai fini della definizione, faremo riferimento ai componenti near-edge e far-edge come segue:

- **Componenti Near-Edge:** parte dell'infrastruttura tra il Far Edge e gli strati cloud.

I sistemi near-edge possono coesistere con l'infrastruttura del vettore WAN, come la co-localizzazione dell'hardware nei ripetitori e nelle stazioni di commutazione cellulare. Questo livello può ospitare servizi computazionalmente complessi;

- **Componenti Far-edge:** sono costituiti da dispositivi di elaborazione in grado di comunicare, gestire e scambiare dati con dispositivi cloud e/o near-edge. Questo livello è il più lontano dal

livello globale del cloud, ma mantiene comunque una relazione con il cloud ed è il più vicino agli utenti finali o ai sistemi di sensori.

Ha requisiti quali design in tempo reale, bassa latenza e può realizzare grandi reti PAN utilizzando poi dei gateway.

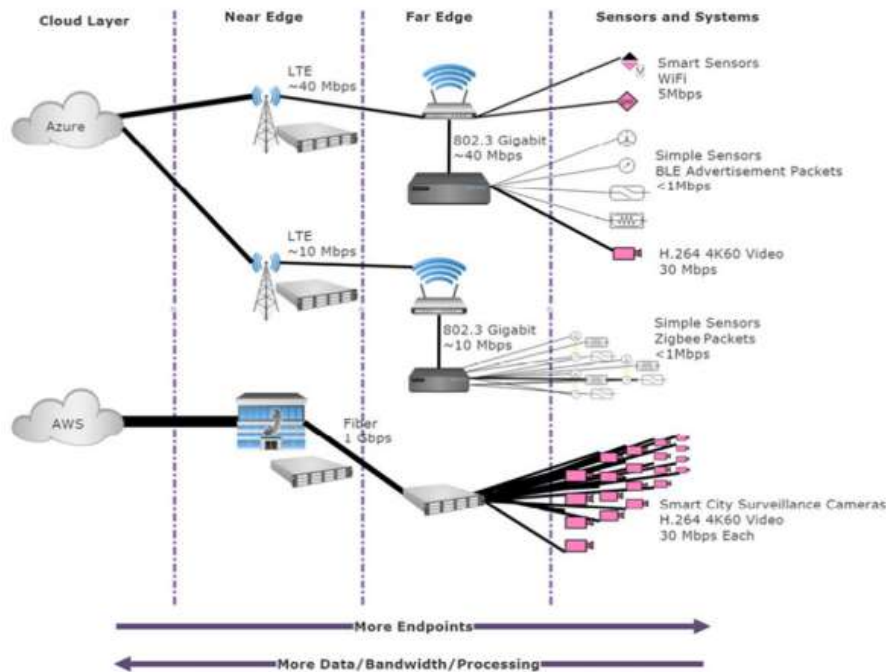


Figura 8.1: Differenza di prossimità e risorse tra near-edge computing e far-edge computing

Dovrebbe essere chiaro che il semplice collegamento di un sensore a un dispositivo informatico in una posizione remota non è edge computing, proprio come una persona che utilizza un PC non connesso per giocare a un videogioco non è un giocatore online.

Tuttavia, non appena quel sensore fornisce dati a un servizio cloud o l'essere umano gioca a un gioco di realtà virtuale interattivo, sta lavorando con dispositivi di edge computing che devono agire di concerto con servizi cloud centralizzati

In generale, più sono lontani dai livelli cloud, più le risorse informatiche saranno limitate. Tuttavia, ci saranno più endpoint man mano che ci allontaniamo dagli strati cloud sotto forma di sistemi non IP come sensori Bluetooth e sistemi SCADA. Chiamiamo questa espansione **fan-out**.

Oltre all'ampia definizione di edge computing esiste anche un gergo relativo all'approccio alla progettazione di edge computing:

- **Fog computing:** il fog computing si riferisce ad un'architettura di servizi cloud che vanno dai sistemi cloud dei data center centrali ai dispositivi near-edge e far-edge.

La nebbia rappresenta una singola astrazione di un insieme geograficamente disperso di cloud e computer periferici che si comportano e agiscono come un'unica entità;

- **Multi-access edge computing (MEC):** precedentemente chiamato mobile edge computing. MEC consente l'esistenza e la distribuzione di applicazioni a bassa latenza, larghezza di banda elevata e in tempo reale ai margini di reti più grandi. MEC è definito dall'European Telecommunications Standards Institute (ETSI) e prevede la possibilità per gli sviluppatori di eseguire applicazioni all'interno della rete di accesso radio (RAN, cioè sistemi radio e protocolli di comunicazione a lungo raggio, in analogia con l'acronimo WAN).

Tipicamente, la RAN esiste fisicamente con il controller di rete radio su una stazione base cellulare. MEC potrebbe consentire lo streaming video a bassa latenza o il gioco basato su cloud;

- **Cloudlet:** un cloudlet è un data center cloud su piccola scala, potremmo immaginarlo come un "cloud-in-a-box". In altre parole, è un dispositivo per supportare casi d'uso ad alta intensità di risorse nei tipi di applicazioni client-server. E' simile al concetto MEC per facilitare tempi di risposta migliori e una latenza inferiore ma non è necessariamente associato a un'infrastruttura di telecomunicazione o vettore.

8.1.1 Casi d'Uso

I sistemi perimetrali sono collocati, potremmo dire per definizione, vicino a dove vengono generati i dati o si trovano le persone.

Attualmente, circa il 20% dei dati nelle aziende viene raccolto al di fuori delle mura aziendali.

Gartner prevede che entro il 2023 fino al 75% dei dati aziendali sarà raccolto e gestito da sistemi al di fuori dei confini fisici dei data center.

L'edge computing serve quattro casi d'uso principali:

- **Latenza ridotta:** i sistemi perimetrali possono essere posizionati più vicino agli utenti finali e ai servizi. Questo naturalmente può evitare vari salti di rete e propagazione. Alcune applicazioni sensibili alla latenza, come i giochi basati su cloud e lo streaming video, hanno severi requisiti di latenza e prestazioni in tempo reale. Ciò include anche i dispositivi che richiedono un processo decisionale in tempo reale o l'esecuzione di motori di regole per i macchinari critici per la sicurezza;
- **Preservare la larghezza di banda:** alcuni ambienti hanno una larghezza di banda limitata da e/o verso il sistema perimetrale. In altri casi, i costi dei dati per il cloud o i data center possono crescere fino a diventare proibitivi su larga scala o con un volume considerevole. Molti vettori hanno limiti di dati o piani tariffari basati sull'utilizzo. L'edge computing può servire al meglio questo problema attraverso tecniche di filtraggio, memorizzazione nella cache e compressione dei dati per massimizzare in modo efficiente la larghezza di banda disponibile;
- **Computazione resiliente:** alcune situazioni non hanno una comunicazione affidabile. Un esempio potrebbe essere un'applicazione di trasporto o logistica che tiene traccia di una flotta di veicoli e merci in tempo reale, nonché dati critici sulla temperatura del carico. Un veicolo in movimento può, a volte, perdere la sua connessione con il vettore mentre viaggia attraverso tunnel, aree rurali e sottopassi. In questi casi d'uso è inaccettabile semplicemente "perdere" i dati. Pertanto, questi tipi di sistemi devono esaminare la memorizzazione nella cache locale perimetrale per archiviare i dati fino al ripristino della comunicazione. Questi sistemi possono anche disporre di tecniche di routing di failover o handover per passare a vettori diversi in caso di perdita di un vettore principale.
- **Sicurezza e privacy:** alcune situazioni devono proteggere o addirittura rimuovere determinati dati prima che viaggino ulteriormente nel cloud o in altri edge. Ciò è particolarmente vero per le situazioni che coinvolgono dati sanitari o immagini di sistemi di videosorveglianza. In molti casi, la sicurezza dei dati è definita dalle normative governative. Ad esempio, un sistema perimetrale utilizzato per la videosorveglianza potrebbe dover bonificare immagini contenenti bambini se le immagini vengono utilizzate in una trasmissione pubblica. Ciò può comportare enormi quantità di risorse informatiche all'edge.

Con questi modelli, alcuni casi d'uso comuni nel settore includono quanto segue:

8.1.2 Protezione dell'Hardware

Quando si distribuisce hardware che deve esistere all'esterno o all'interno di un'area in cui le condizioni ambientali non possono essere controllate, il sistemista IoT deve scegliere come proteggere l'elettronica dalla contaminazione e dall'umidità, nonché mitigare i problemi termici.

In genere l'elettronica sarà alloggiata in un contenitore.

L'industria elettronica utilizza una convenzione internazionale per qualificare la durezza delle custodie elettroniche.

Questo standard è chiamato "Ingress Protection mark" o IP in breve.

Categoria	Descrizione	Destinazione d'Uso
Automazione dei dispositivi	Interazione, controllo e monitoraggio di "cose", sensori, sistemi e ambienti edge-based. Questi richiedono un' integrazione basata su cloud, ma hanno esigenze di requisiti in tempo reale.	Sistemi di controllo industriale, veicoli autonomi.
Ambienti Immersivi	Tipi di interazione AR e VR, chirurgia remota, sistemi di comando vocale. (Alexa, Google Home, Siri, etc...).	Sistemi perimetrali utilizzati per ridurre la latenza aumentare la larghezza di banda per le applicazioni sensibili a tempi e sincronizzazione adeguati.
Monitoraggio del Paziente	Soluzione per l'assistenza sanitaria, l'assistenza domiciliare e il monitoraggio dei pazienti che devono essere robuste, infallibili e sicure.	I sistemi perimetrali possono coesistere con i pazienti per fornire comunicazioni sicure e resilienti con i sistemi degli operatori sanitari a monte.
Aggregazione PAN	Ambienti con sistemi non basati su IP e ambienti mesh che richiedono un bridge e una traduzione tra stack di protocolli .	I sistemi perimetrali che fungono da hub, bridge e gateway possono essere gestiti e partecipare come componenti protetti in una rete aziendale.
Gestione resiliente della comunicazione	Società di trasporto e logistica, flotte e autotrasporti che richiedono una comunicazione coerente con i sistemi cloud o data center.	I sistemi perimetrali possono monitorare e mantenere la resilienza in un ambiente con problemi di comunicazione tramite memorizzazione nella cache, tecniche di failover e metodi di commutazione del vettore.
Intrattenimento immersivo e reti di distribuzione dei clienti	Giochi basati su cloud, streaming video e intrattenimento mobile.	I sistemi perimetrali possono essere collocati in posizioni strategiche per aiutare a bilanciare la latenza e la capacità per applicazioni di streaming di giochi e video su larga scala generalmente ospitate in centri dati ampi e disparati.
Elaborazione IoT	Gestione di più sensori e ingressi. I dati devono essere filtrati, cancellati per rilevare eventuali anomalie, impacchettati e compressi. I dati possono anche essere passati attraverso motori di regole, motori di inferenza o hardware di elaborazione del segnale e agiti localmente.	I sistemi perimetrali offrono la possibilità di elaborare i dati in blocco localmente in tempo reale in modo efficiente senza il costo dello spostamento di tali dati nel cloud.
Gestione dei dispositivi	Dispositivi IoT e perimetrali che richiedono gestione del sistema, aggiornamenti del firmware e patch.	I sistemi perimetrali possono mantenere un "manifest" di patch e aggiornamenti qualificati autenticati dei dispositivi che gestisce. Possono organizzare e qualificare gli aggiornamenti del firmware senza l'intervento umano.

Il test IP verifica la capacità di un prodotto di proteggere da infiltrazioni di acqua, polvere e corpi estranei.

In generale, i prodotti che richiedono test IP includono computer, apparecchiature di laboratorio, molti dispositivi medici, lampade e prodotti che devono rimanere privi di polvere o resistenti all'umidità. Anche gli articoli che sono sigillati e che verranno probabilmente collocati in luoghi pericolosi necessitano della classificazione IP.

Esistono standard relativi ai test IP, in particolare MIL-STD-810 (militare), RTCA/DO-160 (Commissione tecnica radiofonica per l'aeronautica) e IEC 60529 (Commissione elettrotecnica internazionale). Sono definiti secondo la norma internazionale EN 60529.



Figura 8.2: Lettere Caratteristiche IP

Da sinistra										
Grado di protezione (protezione di oggetti solidi)										
Test	IP00	IP01	IP02	IP03	IP04	IP05	IP06	IP07	IP08	IP09
Test 1: Protezione contro i corpi solidi	IP00	IP01	IP02	IP03	IP04	IP05	IP06	IP07	IP08	IP09
Test 2: Protezione contro l'acqua	IP00	IP01	IP02	IP03	IP04	IP05	IP06	IP07	IP08	IP09
Test 3: Protezione contro la polvere	IP00	IP01	IP02	IP03	IP04	IP05	IP06	IP07	IP08	IP09
Test 4: Protezione contro i corpi solidi	IP00	IP01	IP02	IP03	IP04	IP05	IP06	IP07	IP08	IP09
Test 5: Protezione contro l'acqua	IP00	IP01	IP02	IP03	IP04	IP05	IP06	IP07	IP08	IP09
Test 6: Protezione contro la polvere	IP00	IP01	IP02	IP03	IP04	IP05	IP06	IP07	IP08	IP09
Test 7: Protezione contro i corpi solidi	IP00	IP01	IP02	IP03	IP04	IP05	IP06	IP07	IP08	IP09
Test 8: Protezione contro l'acqua	IP00	IP01	IP02	IP03	IP04	IP05	IP06	IP07	IP08	IP09
Test 9: Protezione contro la polvere	IP00	IP01	IP02	IP03	IP04	IP05	IP06	IP07	IP08	IP09
Test 10: Protezione contro i corpi solidi	IP00	IP01	IP02	IP03	IP04	IP05	IP06	IP07	IP08	IP09
Test 11: Protezione contro l'acqua	IP00	IP01	IP02	IP03	IP04	IP05	IP06	IP07	IP08	IP09
Test 12: Protezione contro la polvere	IP00	IP01	IP02	IP03	IP04	IP05	IP06	IP07	IP08	IP09
Test 13: Protezione contro i corpi solidi	IP00	IP01	IP02	IP03	IP04	IP05	IP06	IP07	IP08	IP09
Test 14: Protezione contro l'acqua	IP00	IP01	IP02	IP03	IP04	IP05	IP06	IP07	IP08	IP09
Test 15: Protezione contro la polvere	IP00	IP01	IP02	IP03	IP04	IP05	IP06	IP07	IP08	IP09
Test 16: Protezione contro i corpi solidi	IP00	IP01	IP02	IP03	IP04	IP05	IP06	IP07	IP08	IP09
Test 17: Protezione contro l'acqua	IP00	IP01	IP02	IP03	IP04	IP05	IP06	IP07	IP08	IP09
Test 18: Protezione contro la polvere	IP00	IP01	IP02	IP03	IP04	IP05	IP06	IP07	IP08	IP09
Test 19: Protezione contro i corpi solidi	IP00	IP01	IP02	IP03	IP04	IP05	IP06	IP07	IP08	IP09
Test 20: Protezione contro l'acqua	IP00	IP01	IP02	IP03	IP04	IP05	IP06	IP07	IP08	IP09
Test 21: Protezione contro la polvere	IP00	IP01	IP02	IP03	IP04	IP05	IP06	IP07	IP08	IP09
Test 22: Protezione contro i corpi solidi	IP00	IP01	IP02	IP03	IP04	IP05	IP06	IP07	IP08	IP09
Test 23: Protezione contro l'acqua	IP00	IP01	IP02	IP03	IP04	IP05	IP06	IP07	IP08	IP09
Test 24: Protezione contro la polvere	IP00	IP01	IP02	IP03	IP04	IP05	IP06	IP07	IP08	IP09
Test 25: Protezione contro i corpi solidi	IP00	IP01	IP02	IP03	IP04	IP05	IP06	IP07	IP08	IP09
Test 26: Protezione contro l'acqua	IP00	IP01	IP02	IP03	IP04	IP05	IP06	IP07	IP08	IP09
Test 27: Protezione contro la polvere	IP00	IP01	IP02	IP03	IP04	IP05	IP06	IP07	IP08	IP09
Test 28: Protezione contro i corpi solidi	IP00	IP01	IP02	IP03	IP04	IP05	IP06	IP07	IP08	IP09
Test 29: Protezione contro l'acqua	IP00	IP01	IP02	IP03	IP04	IP05	IP06	IP07	IP08	IP09
Test 30: Protezione contro la polvere	IP00	IP01	IP02	IP03	IP04	IP05	IP06	IP07	IP08	IP09
Test 31: Protezione contro i corpi solidi	IP00	IP01	IP02	IP03	IP04	IP05	IP06	IP07	IP08	IP09
Test 32: Protezione contro l'acqua	IP00	IP01	IP02	IP03	IP04	IP05	IP06	IP07	IP08	IP09
Test 33: Protezione contro la polvere	IP00	IP01	IP02	IP03	IP04	IP05	IP06	IP07	IP08	IP09
Test 34: Protezione contro i corpi solidi	IP00	IP01	IP02	IP03	IP04	IP05	IP06	IP07	IP08	IP09
Test 35: Protezione contro l'acqua	IP00	IP01	IP02	IP03	IP04	IP05	IP06	IP07	IP08	IP09
Test 36: Protezione contro la polvere	IP00	IP01	IP02	IP03	IP04	IP05	IP06	IP07	IP08	IP09
Test 37: Protezione contro i corpi solidi	IP00	IP01	IP02	IP03	IP04	IP05	IP06	IP07	IP08	IP09
Test 38: Protezione contro l'acqua	IP00	IP01	IP02	IP03	IP04	IP05	IP06	IP07	IP08	IP09
Test 39: Protezione contro la polvere	IP00	IP01	IP02	IP03	IP04	IP05	IP06	IP07	IP08	IP09
Test 40: Protezione contro i corpi solidi	IP00	IP01	IP02	IP03	IP04	IP05	IP06	IP07	IP08	IP09
Test 41: Protezione contro l'acqua	IP00	IP01	IP02	IP03	IP04	IP05	IP06	IP07	IP08	IP09
Test 42: Protezione contro la polvere	IP00	IP01	IP02	IP03	IP04	IP05	IP06	IP07	IP08	IP09
Test 43: Protezione contro i corpi solidi	IP00	IP01	IP02	IP03	IP04	IP05	IP06	IP07	IP08	IP09
Test 44: Protezione contro l'acqua	IP00	IP01	IP02	IP03	IP04	IP05	IP06	IP07	IP08	IP09
Test 45: Protezione contro la polvere	IP00	IP01	IP02	IP03	IP04	IP05	IP06	IP07	IP08	IP09
Test 46: Protezione contro i corpi solidi	IP00	IP01	IP02	IP03	IP04	IP05	IP06	IP07	IP08	IP09
Test 47: Protezione contro l'acqua	IP00	IP01	IP02	IP03	IP04	IP05	IP06	IP07	IP08	IP09
Test 48: Protezione contro la polvere	IP00	IP01	IP02	IP03	IP04	IP05	IP06	IP07	IP08	IP09
Test 49: Protezione contro i corpi solidi	IP00	IP01	IP02	IP03	IP04	IP05	IP06	IP07	IP08	IP09
Test 50: Protezione contro l'acqua	IP00	IP01	IP02	IP03	IP04	IP05	IP06	IP07	IP08	IP09
Test 51: Protezione contro la polvere	IP00	IP01	IP02	IP03	IP04	IP05	IP06	IP07	IP08	IP09
Test 52: Protezione contro i corpi solidi	IP00	IP01	IP02	IP03	IP04	IP05	IP06	IP07	IP08	IP09
Test 53: Protezione contro l'acqua	IP00	IP01	IP02	IP03	IP04	IP05	IP06	IP07	IP08	IP09
Test 54: Protezione contro la polvere	IP00	IP01	IP02	IP03	IP04	IP05	IP06	IP07	IP08	IP09
Test 55: Protezione contro i corpi solidi	IP00	IP01	IP02	IP03	IP04	IP05	IP06	IP07	IP08	IP09
Test 56: Protezione contro l'acqua	IP00	IP01	IP02	IP03	IP04	IP05	IP06	IP07	IP08	IP09
Test 57: Protezione contro la polvere	IP00	IP01	IP02	IP03	IP04	IP05	IP06	IP07	IP08	IP09
Test 58: Protezione contro i corpi solidi	IP00	IP01	IP02	IP03	IP04	IP05	IP06	IP07	IP08	IP09
Test 59: Protezione contro l'acqua	IP00	IP01	IP02	IP03	IP04	IP05	IP06	IP07	IP08	IP09
Test 60: Protezione contro la polvere	IP00	IP01	IP02	IP03	IP04	IP05	IP06	IP07	IP08	IP09
Test 61: Protezione contro i corpi solidi	IP00	IP01	IP02	IP03	IP04	IP05	IP06	IP07	IP08	IP09
Test 62: Protezione contro l'acqua	IP00	IP01	IP02	IP03	IP04	IP05	IP06	IP07	IP08	IP09
Test 63: Protezione contro la polvere	IP00	IP01	IP02	IP03	IP04	IP05	IP06	IP07	IP08	IP09
Test 64: Protezione contro i corpi solidi	IP00	IP01	IP02	IP03	IP04	IP05	IP06	IP07	IP08	IP09
Test 65: Protezione contro l'acqua	IP00	IP01	IP02	IP03	IP04	IP05	IP06	IP07	IP08	IP09
Test 66: Protezione contro la polvere	IP00	IP01	IP02	IP03	IP04	IP05	IP06	IP07	IP08	IP09
Test 67: Protezione contro i corpi solidi	IP00	IP01	IP02	IP03	IP04	IP05	IP06	IP07	IP08	IP09
Test 68: Protezione contro l'acqua	IP00	IP01	IP02	IP03	IP04	IP05	IP06	IP07	IP08	IP09
Test 69: Protezione contro la polvere	IP00	IP01	IP02	IP03	IP04	IP05	IP06	IP07	IP08	IP09
Test 70: Protezione contro i corpi solidi	IP00	IP01	IP02	IP03	IP04	IP05	IP06	IP07	IP08	IP09
Test 71: Protezione contro l'acqua	IP00	IP01	IP02	IP03	IP04	IP05	IP06	IP07	IP08	IP09
Test 72: Protezione contro la polvere	IP00	IP01	IP02	IP03	IP04	IP05	IP06	IP07	IP08	IP09
Test 73: Protezione contro i corpi solidi	IP00	IP01	IP02	IP03	IP04	IP05	IP06	IP07	IP08	IP09
Test 74: Protezione contro l'acqua	IP00	IP01	IP02	IP03	IP04	IP05	IP06	IP07	IP08	IP09
Test 75: Protezione contro la polvere	IP00	IP01	IP02	IP03	IP04	IP05	IP06	IP07	IP08	IP09
Test 76: Protezione contro i corpi solidi	IP00	IP01	IP02	IP03	IP04	IP05	IP06	IP07	IP08	IP09
Test 77: Protezione contro l'acqua	IP00	IP01	IP02	IP03	IP04	IP05	IP06	IP07	IP08	IP09
Test 78: Protezione contro la polvere	IP00	IP01	IP02	IP03	IP04	IP05	IP06	IP07	IP08	IP09
Test 79: Protezione contro i corpi solidi	IP00	IP01	IP02	IP03	IP04	IP05	IP06	IP07	IP08	IP09
Test 80: Protezione contro l'acqua	IP00	IP01	IP02	IP03	IP04	IP05	IP06	IP07	IP08	IP09
Test 81: Protezione contro la polvere	IP00	IP01	IP02	IP03	IP04	IP05	IP06	IP07	IP08	IP09
Test 82: Protezione contro i corpi solidi	IP00	IP01	IP02	IP03	IP04	IP05	IP06	IP07	IP08	IP09
Test 83: Protezione contro l'acqua	IP00	IP01	IP02	IP03	IP04	IP05	IP06	IP07	IP08	IP09
Test 84: Protezione contro la polvere	IP00	IP01	IP02	IP03	IP04	IP05	IP06	IP07	IP08	IP09
Test 85: Protezione contro i corpi solidi	IP00	IP01	IP02	IP03	IP04	IP05	IP06	IP07	IP08	IP09
Test 86: Protezione contro l'acqua	IP00	IP01	IP02	IP03	IP04	IP05	IP06	IP07	IP08	IP09
Test 87: Protezione contro la polvere	IP00	IP01	IP02	IP03	IP04	IP05	IP06	IP07	IP08	IP09
Test 88: Protezione contro i corpi solidi	IP00	IP01	IP02	IP03	IP04	IP05	IP06	IP07	IP08	IP09
Test 89: Protezione contro l'acqua	IP00	IP01	IP02	IP03	IP04	IP05	IP06	IP07	IP08	IP09
Test 90: Protezione contro la polvere	IP00	IP01	IP02	IP03	IP04	IP05	IP06	IP07	IP08	IP09
Test 91: Protezione contro i corpi solidi	IP00	IP01	IP02	IP03	IP04	IP05	IP06	IP07	IP08	IP09
Test 92: Protezione contro l'acqua	IP00	IP01	IP02	IP03	IP04	IP05	IP06	IP07	IP08	IP09
Test 93: Protezione contro la polvere	IP00	IP01	IP02	IP03	IP04	IP05	IP06	IP07	IP08	IP09
Test 94: Protezione contro i corpi solidi	IP00	IP01	IP02	IP03	IP04	IP05	IP06	IP07	IP08	IP09
Test 95: Protezione contro l'acqua	IP00	IP01	IP02	IP03	IP04	IP05	IP06	IP07	IP08	IP09
Test 96: Protezione contro la polvere	IP00	IP01	IP02	IP03	IP04	IP05	IP06	IP07	IP08	IP09
Test 97: Protezione contro i corpi solidi	IP00	IP01	IP02	IP03	IP04	IP05	IP06	IP07	IP08	IP09
Test 98: Protezione contro l'acqua	IP00	IP01	IP02	IP03	IP04	IP05	IP06	IP07	IP08	IP09
Test 99: Protezione contro la polvere	IP00	IP01	IP02	IP03	IP04	IP05	IP06	IP07	IP08	IP09
Test 100: Protezione contro i corpi solidi	IP00	IP01	IP02	IP03	IP04	IP05	IP06	IP07	IP08	IP09

Figura 8.3: Tabella Grado di Protezione IP

8.1.3 Sistemi Operativi

Le scelte di sistema operativo per un sistema edge sono numerose e meritano un'attenzione significativa. Dal punto di vista di un architetto, la scelta deve ricevere un discreto grado di diligenza poiché sarà la base per eventuali generazioni di soluzioni software distribuite in quell'ambiente. Un sistema operativo esiste come astrazione software e livello di protezione tra hardware e applicazioni. Tuttavia, un sistema operativo fornisce anche un'interfaccia binaria dell'applicazione (ABI) su cui il software può esistere. Può fornire una risposta in tempo reale e una manutenzione garantita, forma processi software e protezioni a livello di thread e fornisce interfacce per condividere memoria e IO tra applicazioni software e gestisce la memoria e le risorse di sistema.

In molte situazioni, un OEM hardware fornirà o consiglierà un sistema operativo e un pacchetto di supporto della scheda (BSP) per l'hardware che ha progettato. Altre volte, la scelta del sistema operativo non sarà così netta, il che è vero per hardware appositamente costruito e personalizzato. Le seguenti sono domande che l'architetto dovrebbe considerare e soppesare nella scelta del sistema operativo:

- Costo della distribuzione del sistema operativo. È una licenza pubblica come Linux o una licenza commerciale come Windows?
- C'è un contratto di assistenza? C'è una necessità in tempo reale, il sistema operativo supporta RT o ci sono estensioni?
- Su quale architettura del processore è supportato il sistema operativo (ARM, x86)?
- Il sistema operativo supporta tutte le caratteristiche del processore necessarie: memoria virtuale, cache multilivello, estensioni SIMD, emulazione a virgola mobile?
- Come si ottengono pacchetti o estensioni? Linux: APT, yum, RPM, PACMAN.
- Quanti pacchetti o estensioni sono stati creati per questa distribuzione del sistema operativo?
- Come avvierai il dispositivo: Flash, rete (PXEboot)? Quale forma di servizi di sicurezza sono integrati nel sistema operativo e nel kernel?
- Per i sistemi profondamente embedded e lontani, quanto piccolo può essere ridotto il sistema operativo in RAM e dimensioni di archiviazione?
- Il tempo di avvio è un problema?
- Quali filesystem sono necessari?
- Se il supporto per software e driver viene fornito con l'hardware periferico, a quale sistema operativo si rivolge?
- Il sistema operativo supporta le forme di comunicazione, rete e stack di protocollo desiderate?
- C'è bisogno di una GUI?

Questo non è un elenco esaustivo, ma dovrebbe indicare che ci sono molti fattori nella scelta di un sistema operativo. Il sistema operativo esiste come livello fondamentale su cui si basa il resto del sistema. La modifica di un sistema operativo spesso richiede un significativo refactoring di software e driver. Più avanti esamineremo un approccio basato su container alla distribuzione del software che aiuta in qualche modo ad attrarre un sistema operativo, ma a un livello fondamentale, deve esistere un sistema operativo (o kernel).

Un sistema operativo come Linux consente la flessibilità di aggiungere volumi di framework, strumenti, utilità e pacchetti diversi all'immagine del firmware del sistema. Tuttavia, è necessario prestare attenzione quando si costruiscono sistemi edge-based affidabili e robusti. La filosofia che dovresti usare nella creazione dell'immagine di base di una macchina edge è quella di fornire l'insieme minimo di pacchetti e librerie necessari per eseguire l'attività data. Questa metodologia differisce dalle tradizionali installazioni di macchine virtuali basate su cloud fornite con immagini Linux preconfezionate costituite da gigabyte di pacchetti e software.

Ai margini, siamo preoccupati per la sicurezza del dispositivo, quanto sia robusto il dispositivo per il suo ambiente e quanto sia difficile da mantenere. Nella maggior parte dei casi, il computer perimetrale ha risorse limitate e i pacchetti non critici consumano spazio di archiviazione e RAM che potrebbero essere utilizzati per casi d'uso effettivi.

Il software e le applicazioni in esecuzione sull'edge ne definiscono lo scopo. Man mano che si ridimensionano i dispositivi perimetrali, la loro gestione in remoto diventa la sfida. Certamente esistono modelli di controllo e distribuzione personalizzati che vengono utilizzati in produzione. **Oggi disponiamo di framework commerciali per la gestione dell'edge pronto all'uso**, nonché di metodologie basate su container che alleggeriscono l'onere di distribuire il software, in modo sicuro e controllato, ai computer perimetrali remoti.

In qualunque caso vogliamo che il software e il sistema siano:

- **Robusti:** in grado di ricevere, creare nuovamente immagini e rieseguire il software mentre viene distribuito;
- **Controllabili:** avere un cloud o un servizio centrale che gestisce e monitora l’implementazione;
- **Reattivi:** segnalazione di informazioni sul successo o sul fallimento del reimaging del software.

8.1.3.1 Windows 10 IoT

Windows 10 IoT è un membro della famiglia Windows 10 che offre potenza, sicurezza e gestibilità di livello aziendale per Internet delle cose. Sfruttando l’incorporamento dell’esperienza, dell’ecosistema e della connettività cloud di Windows, consente alle organizzazioni di creare una propria soluzione Internet delle cose con dispositivi sicuri che potranno essere sottoposti rapidamente a provisioning ed essere gestiti e connessi a una strategia cloud globale con facilità. Esiste in 2 declinazioni:

- Windows 10 IoT Core è il membro di dimensioni più ridotte della famiglia del sistema operativo Windows 10. Pur eseguendo solo una singola app, offre comunque la gestibilità e la sicurezza che ci si aspetta da Windows 10. E’ una versione di Windows 10 ottimizzata per dispositivi di dimensioni ridotte con o senza display, eseguibile su dispositivi sia ARM che x86/x64.
- è invece una versione completa di Windows 10 con funzionalità specializzate per creare dispositivi dedicati bloccati per uno specifico set di applicazioni e periferiche.

Funzionalità/Edizione	Windows 10 IoT Core	Windows 10 IoT Enterprise
Esperienza utente	Un app UWP in primo piano alla volta (vedi la documentazione di IoT Shell per la gestione del backstack delle app con servizi e app in background di supporto).	Shell di Windows tradizionale con funzionalità avanzate di blocco
Supporto headless	Sì	Sì
Architettura app supportata	Solo interfaccia utente UWP	Supporto completo interfaccia utente Windows (ad esempio UWP, WinForms e così via)
Cortana	SDK per Cortana	Sì
Aggiunta a un dominio	Solo AAD	Domino tradizionale e AAD
Gestione	MDM	MDM

Figura 8.4: Differenze tra Windows 10 IoT Core e Windows 10 IoT Enterprise

Tecnologie di sicurezza dei dispositivi	TPM, Avvio protetto, BitLocker, Device Guard e Attestazione dell'integrità del dispositivo	TPM, Avvio protetto, BitLocker, Device Guard e Attestazione dell'integrità del dispositivo
Architettura CPU supportata	x86, x64 e ARM	x86 e x64
Gestione delle licenze	Contratto di licenza online e contratti OEM incorporati, a titolo gratuito	Contratti OEM incorporati, diretti e indiretti
Scenari di utilizzo	Segnaletica digitale ¹² , edifici intelligenti, gateway IoT, HMI, domotica, dispositivi indossabili	Tablet aziendali, POS (Point of Service) per vendita al dettaglio, chioschi multimediali, segnaletica digitale ¹² , sportelli bancomat, dispositivi medici, dispositivi di produzione, thin client

Figura 8.5: Differenze tra Windows 10 IoT Core e Windows 10 IoT Enterprise

Architettura ARM L’architettura ARM (precedentemente Advanced RISC Machine, prima ancora Acorn RISC Machine), in elettronica e informatica, indica una famiglia di microprocessori RISC a 32-bit e 64-bit sviluppata da ARM Holdings e utilizzata in una moltitudine di sistemi embedded. Grazie alle sue caratteristiche di basso consumo elettrico, rapportato alle prestazioni, l’architettura ARM domina il settore dei dispositivi mobili dove il risparmio energetico delle batterie è fondamentale.

8.1.3.2 Virtualizzazione

Possiamo distinguere vari tipi di virtualizzazione:

Virtualizzazione hardware: un’astrazione a livello hardware generalmente in grado di eseguire qualsiasi software eseguibile su bare metal. Utilizza un hypervisor per gestire una o più macchine virtuali sul processore e può supportare la replica virtuale dell’hardware su più sistemi operativi virtuali tramite la virtualizzazione dell’IO hardware. Queste tecniche richiedono il supporto del processore e dell’hardware per la virtualizzazione che di solito si trova su processori di fascia alta come le parti della serie ARM Cortex A.

Come sottocategoria, esistono due tipi di hypervisor: gli hypervisor di tipo 1 vengono eseguiti direttamente su bare metal e di tipo 2 hanno un sistema operativo sottostante ospitato. Un esempio di hypervisor di tipo 1 è Microsoft HyperV. Un esempio di hypervisor di tipo 2 è Microsoft Virtual PC.

Paravirtualizzazione: fornisce un livello di astrazione chiamato livello di astrazione hardware (HAL) e richiede driver speciali. Questi driver sono collegati tramite l'hypervisor di sottolineatura e accedono all'hardware tramite hypercall. Richiede modifiche al sistema operativo guest per abilitare questa forma di virtualizzazione e offre al sistema operativo guest prestazioni più elevate e la capacità di comunicare direttamente con l'hypervisor.

8.1.3.2.1 Container

Gestisce l'astrazione a livello di applicazione. Non esiste un hypervisor o un sistema operativo guest. Piuttosto, i container richiedono solo il sistema operativo di hosting per fornire servizi di base. I container mantengono la separazione l'uno dall'altro, fornendo un livello di protezione simile a una macchina virtuale. I gestori di container possono anche adattarsi alle mutevoli risorse della macchina. Ad esempio, possono assegnare dinamicamente più memoria a un contenitore in fase di esecuzione.

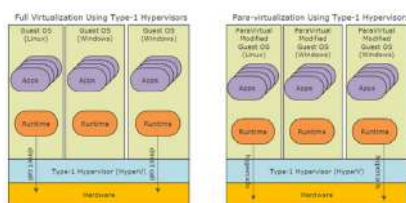


Figura 8.6: Full Virtualization vs Para-virtualization

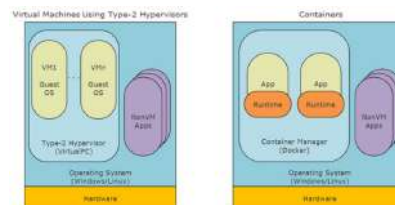


Figura 8.7: Type 2 VM vs Containers

Per alcune applicazioni di edge computing, le astrazioni basate su container sono particolarmente interessanti. I container richiedono risorse a livello di sistema che devono essere prese in considerazione, come le prestazioni di calcolo, la capacità di archiviazione e persino le funzionalità del processore. Offrono un metodo molto leggero e portatile per creare e distribuire applicazioni ai computer perimetrali. Poiché l'approccio containerizzato non utilizza sistemi operativi guest, è naturalmente più snello e più efficiente in termini di risorse rispetto alla virtualizzazione tradizionale. Questo è fondamentale per i dispositivi perimetrali con risorse limitate.

Inoltre, un'immagine qualificata e funzionante può essere containerizzata e modifiche e test possono essere eseguiti su tale immagine. Un container è anche molto portatile e può essere distribuito in qualsiasi ambiente e su quasi tutti i sistemi operativi host. Per questo motivo, ci concentreremo sui container come metodo per la distribuzione edge.

Sebbene un hypervisor di tipo 2 possa apparire simile al design di un contenitore nel diagramma precedente, non è analogo ai contenitori. Un hypervisor di tipo 2 ha ancora un elevato impatto sulle prestazioni e sui costi generali. In sostanza, stai eseguendo almeno un sistema operativo guest su un sistema operativo host completo. Tenendo conto dell'hypervisor e dei servizi di runtime, la quantità di memoria e di elaborazione richiesta è molto maggiore rispetto ai contenitori leggeri.

I contenitori sono un metodo per virtualizzare l'hardware e i servizi sottostanti come una macchina virtuale (VM). Mentre una macchina virtuale tradizionale richiede un hypervisor che si trova al di sopra dell'hardware e fornisce un livello di astrazione, un contenitore non richiede hypervisor semplicemente perché i suoi servizi risiedono al di sopra del livello del sistema operativo.

L'atto di creare un contenitore ed eseguire un'applicazione come processo in esso è chiamato containerizzazione.

Ci sono due definizioni fondamentali necessarie per comprendere gli elementi di base di un contenitore:

- **Container:** è una singola istanza di un'immagine contenitore. Possono esistere più istanze su un singolo host.
- **Immagine:** l'immagine del contenitore è un insieme di file che non contengono lo stato ma costituiscono il pacchetto (o lo snapshot) di un contenitore.

Docker Per comprendere l'architettura del contenitore possiamo fare riferimento a Docker.

È uno strumento per creare e gestire container e offre una versione gratuita chiamata DockerEE con servizi di base. Una distribuzione del contenitore consiste in un motore di gestione del contenitore dell'applicazione e un repository.

Per creare l'immagine di un contenitore di un'applicazione, iniziamo raccogliendo il codice dell'applicazione e le dipendenze richieste. Queste dipendenze sono librerie, file binari, middleware e componenti software associati che potrebbero essere necessari all'applicazione. Tutte le dipendenze devono essere incluse nell'immagine del contenitore per garantire che siano residenti anche per funzionalità che potrebbero essere eseguite raramente. *Questa aggregazione di applicazioni e dipendenze è chiamata contenitore.*

La creazione di un nuovo contenitore in Docker è semplice. Innanzitutto, scegliamo un'immagine di base a cui fare riferimento.

Docker ha molte immagini di base di vari sistemi operativi e ambienti, è possibile trovare un elenco di immagini di base minime su <https://hub.docker.com/search?q=&type=image>.

Successivamente, creiamo un file immagine Docker. Questo file descrive in dettaglio come verrà costruita l'immagine. Un esempio è il seguente:

```
1 FROM ubuntu
2 RUN apt-get update
3 RUN apt-get install iostat -y
4 CMD ["/usr/bin/iostat"]
```

In questo esempio, il dockerfile estrae dall'immagine di base di Ubuntu e quindi utilizza lo strumento di installazione apt-get per installare l'utilità iostat e quindi lo esegue.

Dopo aver costruito il dockerfile, è necessario creare la sua immagine Docker:

```
1 docker build -t <dockerID>/<image-name>
```

Il campo dockerID è necessario solo se si intende caricare l'immagine nel sistema globale Docker Hub, dove devi registrarti per un account. L'immagine risultante può quindi essere distribuita e istanziata ovunque dando il seguente comando sull'host:

```
1 docker run [options] [dockerID/image-name] [command]
```

Un dispositivo edge computer, così come qualsiasi altro sistema connesso che utilizza Docker, può estrarre questa nuova immagine ed eseguirla in modo simile. Ciò facilita notevolmente il compito di distribuzione e sviluppo.

Inoltre, consente ai modelli di sviluppo per i dispositivi edge di essere come processi su grandi soluzioni SaaS (software as a service) utilizzando tecniche come l'integrazione continua e la distribuzione continua (CI/CD).

Conclusioni L'edge computing è un dominio ampio che richiede competenze ingegneristiche ed informatiche (oltre che di networking, di gestione e di sicurezza) per creare una soluzione solida.

È fondamentale comprendere l'architettura hardware delle macchine edge e degli aspetti connessi alla sua protezione in un'area remota e non sorvegliata. Segue poi la progettazione del software e del sistema operativo. Costruire un sistema robusto e affidabile sull'edge significa che il software deve essere autogestito, snello ed esso stesso affidabile.

Abbiamo esplorato modi per ridurre la manutenzione, le dimensioni dell'immagine e le potenziali vulnerabilità arrivando alla conclusione che il sistemista IoT dovrebbe considerare i paradigmi di virtualizzazione basati su container e valutare come potrebbero funzionare su larga scala.

8.1.4 Cloud Computing per i Sistemi IoT

Microsoft Azure IoT Edge Una intera piattaforma di gestione, da utilizzare insieme ad altri strumenti IT, servizi di sicurezza e requisiti di gestione, può essere progettata da zero ma esistono diverse

piattaforme di gestione che forniscono molti dei servizi necessari per implementare l'edge computing come flotta. Microsoft Azure IoT Edge è un motore di distribuzione di container ed un servizio di gestione che viene eseguito su un computer o dispositivo perimetrale (edge) Windows o Linux.

Fornisce un runtime gratuito e open source, una piattaforma container (Docker), un processo di gestione e distribuzione dei container compatibile per i dispositivi perimetrali, un'API di interfacciamento cloud all'hub IoT di Azure e servizi di provisioning. Vengono fornite anche molte attività perimetrali di routine, quali:

- possibilità di operare offline o con connessioni intermittenti;
- memorizzare nella cache e memorizzare localmente i dati;
- sincronizzazione su cloud on demand;
- gestire servizi di sicurezza per la protezione dalle minacce end-to-end;
- filtrare/denaturare i dati prima di inviare i risultati al cloud per l'archiviazione o l'analisi.

I requisiti generali per Azure IoT Edge sono costituiti da:

- Un computer perimetrale che esegue x64, AMD64, ARM32v7 o ARM64
- Un sistema operativo edge: Linux o Windows;
- Varianti Linux testate: Ubuntu Server 16.04, Ubuntu Server 18.04. Altri non pienamente qualificati: CentOS, Debian 8, Debian 9, Debian 10, Raspian Buster, RHEL 7.5, Wind River 8, Yocto Linux;
- Varianti di Windows testate: Windows 10 IoT Core, Windows 10 IoT Enterprise, Windows Server 2019, Windows Server IoT 2019;
- Runtime del contenitore: un runtime del contenitore compatibile con OCI come il motore basato su Moby. Le immagini del contenitore Docker CE/EE sono compatibili con Moby;
- Risorse per carico di lavoro e caso d'uso: archiviazione per la memorizzazione nella cache offline di dati, RAM ed elaborazione per carichi di lavoro dei moduli,
- Un'interfaccia WAN TCP/IP upstream verso l'hub IoT di Azure tramite i protocolli MQTT o AMPQ.

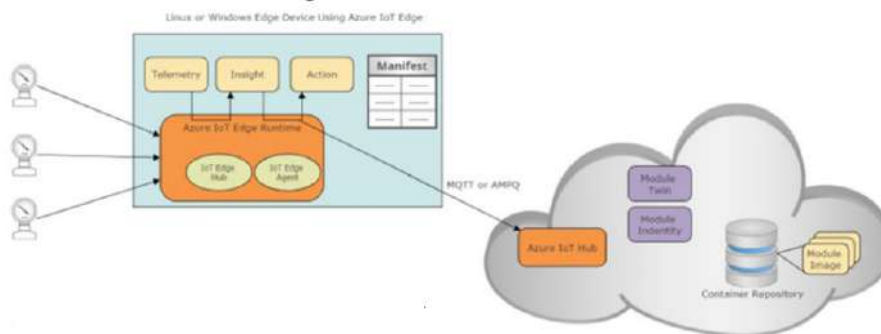


Figura 8.8: L'immagine mostrata il dispositivo perimetrale che ospita il runtime "IoT Edge". Più moduli vengono gestiti come distribuzioni di contenitori tramite il runtime insieme all'hub IoT di Azure. Un manifest controlla quali moduli sono qualificati per l'uso e il routing generale tra i moduli

L'architettura complessiva di Azure IoT Edge insieme all'hub IoT di Azure in esecuzione nel cloud è visibile nell'immagine precedente che mostra tre sensori collegati a un edge computer.

Il computer perimetrale ospita il servizio di runtime di Azure IoT Edge. Il runtime del servizio perimetrale è leggero e rappresenta il cuore del sistema. Il runtime perimetrale gestisce le installazioni di moduli/carichi di lavoro, la sicurezza, il monitoraggio dello stato e tutte le comunicazioni.

L'hub ha due ruoli:

- **IoT Edge Agent** - il servizio agente gestisce i moduli;
- **IoT Edge Hub** - il servizio hub gestisce la comunicazione e funge da proxy verso l'hub IoT di Microsoft Azure.

L'hub IoT in esecuzione nel cloud di Azure esegue un superset di funzioni ed è l'interfaccia principale usata in Azure per connettersi ai dispositivi IoT.

L'hub IoT Edge non è una versione completa dell'hub IoT, ma consente a un programmatore di interfacciarsi con l'edge come farebbe per interfacciarsi con l'hub IoT nel cloud tramite Azure IoT Device SDK.

L'hub IoT Edge gestisce anche un *manifest*. Questo manifest identifica i moduli qualificati e autenticati che possono essere eseguiti sul dispositivo perimetrale. Dichiara anche le regole di instradamento tra i diversi moduli in esecuzione sull'edge.

L'agente IoT Edge gestisce l'immagine del contenitore per ogni modulo in esecuzione sul dispositivo, le credenziali per accedere ai registri dei contenitori privati e le regole sulla creazione e gestione dei moduli.

Il **manifest** controlla il runtime e indica al dispositivo quali moduli installare e come configurarli perché possano interagire. Tutti i dispositivi IoT Edge devono essere configurati con un manifesto della distribuzione. Un runtime IoT Edge appena installato segnala un codice di errore finché non verrà configurato con un manifesto valido.

Uno dei concetti più potenti di Azure IoT Edge è che alcuni servizi e funzionalità progettati specificamente per l'esecuzione nel cloud di Azure possono ora essere eseguiti localmente nel dispositivo perimetrale se soddisfa i requisiti minimi del runtime.

Tali servizi e funzionalità includono:

- Distribuzione e uso di Funzioni di Azure in un modulo Azure IoT Edge;
- Utilizzo dei sistemi di analisi di flusso di Azure come modulo IoT Edge;
- Uso dei sottosistemi di apprendimento automatico di Azure all'interno di un modulo IoT Edge;
- Esecuzione della classificazione delle immagini con il servizio di visione personalizzato di Azure; come modulo IoT Edge
- Esecuzione di database SQL come contenitore IoT Edge.

Questa capacità di migrare i servizi di classe del data center cloud consente uno sviluppo rapido e facilità di esecuzione.

8.2 Protocolli per lo Scambio Dati tra Edge e Cloud: il Protocollo MQTT

Abbiamo visto che i dispositivi posti al limite della architettura di un sistema IoT generano una grande quantità di dati ed eventi.

Abbiamo compreso che esistono diversi mezzi e tecnologie di telecomunicazione per spostare i dati da WPAN, WLAN e WAN e che esistono molte complessità e sottigliezze nella creazione e nel bridging di queste connessioni di rete da reti PAN non basate su IP a reti WAN basate su IP. Verificheremo ora che esistono anche conversioni di protocollo che devono essere comprese.

I protocolli standard sono gli strumenti che legano e incapsulano i dati grezzi da un sensore e li trasformano in qualcosa di significativo e formattato per essere accettato dal cloud.

Una delle principali differenze tra un sistema IoT e un sistema Machine-to-Machine (M2M) è che M2M può comunicare su una WAN con un server o un sistema dedicato senza alcun protocollo incapsulato. Ad esempio, un sistema di automazione industriale SCADA può utilizzare BACNET o ModBus esclusivamente per la comunicazione dai macchinari ai vari computer di controllo.

L'IoT per definizione si basa invece sulla comunicazione tra endpoint e servizi con Internet come tessuto di rete comune. Descriveremo in dettaglio uno dei protocolli prevalenti in ambito IoT: il Message Queue Telemetry Transport (MQTT).

Protocolli, perchè non HTTP? Una domanda naturale è: perché esistono protocolli al di fuori di HTTP per trasportare i dati attraverso la WAN?

HTTP ha fornito servizi e capacità significative per Internet per oltre 20 anni, ma è stato progettato per l'elaborazione generica in modalità client/server. I dispositivi IoT possono essere molto limitati, remoti e con larghezza di banda limitata.

Pertanto, sono necessari protocolli più efficienti, sicuri e scalabili per gestire una pletora di dispositivi in varie topologie di rete come le reti mesh.

Detto questo, HTTP viene utilizzato e ha uno scopo nei sistemi IoT. Sebbene HTTP non sia efficiente in una rete, i protocolli HTTP2 e HTTP3 sono relativamente efficienti. Inoltre, la sicurezza tramite TLS è naturale e comune nelle sessioni HTTP. Infine, HTTP è ovunque e regolarmente utilizzato in un assortimento di comunicazioni e API RESTful.

Protocolli TCP e UDP Nel trasporto di dati su Internet, i protocolli TCP e UDP sono le scelte ovvie nella comunicazione di dati, con TCP che è significativamente più complesso nella sua implementazione rispetto a UDP.

UDP, tuttavia, non ha la stabilità e l'affidabilità del TCP, costringendo alcuni progetti a compensare aggiungendo resilienza nei livelli dell'applicazione sopra UDP. Vale anche la pena notare che UDP viene utilizzato per alcuni protocolli di comunicazione IoT come NB-IoT (comunicazione a lungo raggio).

Implementazioni MOM Molti dei protocolli elencati in questo capitolo sono implementazioni MOM (Message-Oriented Middleware). L'idea di base di una MOM è che la comunicazione tra due dispositivi avvenga utilizzando code di messaggi distribuite.

Una MOM consegna i messaggi da un'applicazione dello spazio utente a un'altra. Alcuni dispositivi producono dati da aggiungere a una coda mentre altri consumano i dati archiviati in una coda. Alcune implementazioni richiedono che un broker o intermediario sia il servizio centrale. In tal caso, produttori e consumatori hanno una relazione di tipo pubblicazione e sottoscrizione con il broker.

AMQP, MQTT e STOMP sono implementazioni MOM; altri includono CORBA e servizi di messaggistica Java. Un'implementazione MOM che utilizza le code può usarle per la resilienza nella progettazione: i dati possono persistere nelle code, anche se il server si guasta.

L'alternativa all'implementazione MOM è RESTful. In un modello RESTful, il server possiede lo stato di una risorsa ma lo stato non viene trasferito dal client al server in un messaggio. I progetti RESTful utilizzano metodi HTTP come GET, PUT, POST e DELETE per inserire richieste sull'URI (Universal Resource Identifier) di una risorsa.

Nessun broker o intermediario è necessario in questa architettura. Poiché si basano sullo stack HTTP, godono della maggior parte dei servizi offerti, come la sicurezza HTTPS. I design RESTful sono tipici delle architetture client-server. I client avviano l'accesso alle risorse tramite modelli di richiesta-risposta sincroni.

L'URI viene utilizzato come identificatore per il traffico dati basato sul Web. L'URI più importante è l'URL (Universal Resource Locator), come `http://www.iotforarchitects.net:8080/iot/?id="temperature"`. L'URI può essere suddiviso in parti componenti utilizzate da vari livelli dello stack di rete:

```

1 Schema: http://
2 Autorita': www.iotforarchitects.net
3 Porto: 8080
4 Percorso: /iot

```

5 Domanda: ?id="temperatura"

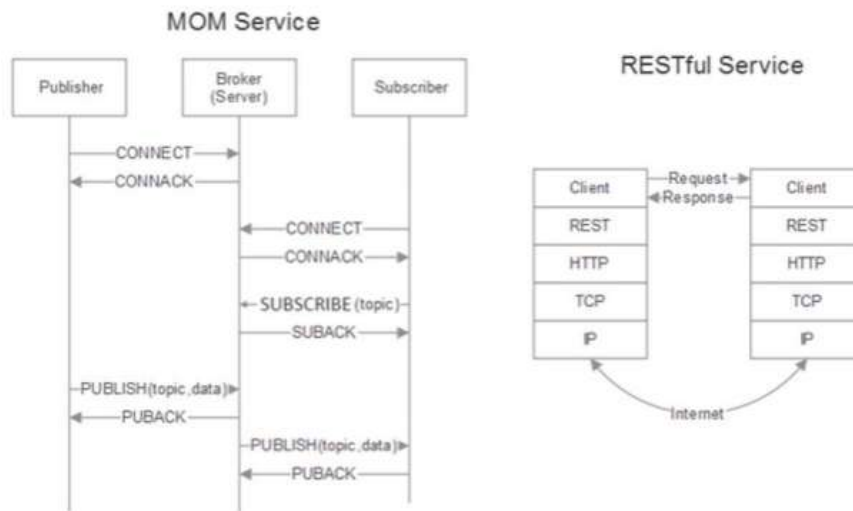


Figura 8.9: La figura illustra il processo di un servizio MOM rispetto a un servizio RESTful. Sulla sinistra c'è un servizio di messaggistica (basato su MQTT) che utilizza un server intermediario e editori e sottoscrittori a un evento. Molti client possono essere sia editori che sottoscrittori e le informazioni possono o meno essere archiviate in code per la resilienza. Sulla destra c'è un design RESTful in cui l'architettura è basata su HTTP e utilizza i paradigmi HTTP per comunicare dal client al server

8.2.1 MQTT

La tecnologia IBM WebSphere Message Queue è stata concepita per la prima volta nel 1993 per affrontare i problemi nei sistemi distribuiti indipendenti e non concorrenti e aiutarli a comunicare in modo sicuro.

Un derivato di WebSphere Message Queue è stato creato da Andy Stanford-Clark e Arlen Nipper presso IBM nel 1999 per affrontare i vincoli particolari del collegamento di oleodotti e gasdotti remoti su una connessione satellitare.

Quel protocollo divenne noto come MQTT.

Gli obiettivi di questo protocollo di trasporto basato su IP sono:

- Essere semplice da implementare;
- Fornire una qualche forma di gestione/controllo della qualità del servizio;
- Essere molto leggero ed efficiente in termini di larghezza di banda;
- Essere indipendente dai dati;
- Avere una consapevolezza della sessione continua;
- Affrontare i problemi di sicurezza.

MQTT prevede questi requisiti. L'organismo standard MQTT.org (mqtt.org) presenta un riassunto molto ben definito del protocollo:

MQTT sta per MQ Telemetry Transport. È un protocollo di messaggistica di pubblicazione/sottoscrizione, estremamente semplice e leggero, progettato per dispositivi vincolati e reti a bassa larghezza di banda, alta latenza o inaffidabili. I principi di progettazione consistono nel ridurre al minimo la larghezza di banda della rete e i requisiti di risorse del dispositivo, tentando anche di garantire affidabilità e un certo grado di garanzia di consegna. Questi principi si rivelano anche per rendere il protocollo ideale per l'emergente mondo "machine-to-machine" (M2M) o "Internet delle cose" dei dispositivi connessi e per le applicazioni mobili in cui la larghezza di banda e la carica della batteria sono un valore aggiunto.

MQTT è stato un protocollo interno e proprietario per IBM per molti anni fino al rilascio nella versione 3.1 nel 2010 come prodotto royalty-free. Nel 2013 MQTT è stato standardizzato e accettato nel consorzio OASIS. Nel 2014, OASIS lo ha rilasciato pubblicamente come versione MQTT 3.1.1. MQTT è anche uno standard ISO (ISO/IEC PRF 20922). Più recentemente, OASIS ha rilasciato la specifica MQTT 5.

MQTT Specification	Release Date	Features
Initial Release	1999	Initial creation and invention
MQTT 3.1	2010	Royalty-free release
HiveMQ	2013	Initial public release
MQTT3.1.1	2014	OASIS standard
MQTT 3.1.1	2016	ISO standard
HiveMQ 4	2018	MQTT compatibility release
HiveMQ MQTT Client	2019	Client public release
MQTT5	2019	OASIS standard
HiveMQ	2019	Open source edition

Figura 8.10: La tabella seguente illustra la versione delle funzionalità e la sequenza temporale dello standard OASIS MQTT, nonché le versioni di HiveMQ. HiveMQ è un fornitore leader di broker MQTT e software e soluzioni client.

MQTT 5 è stato rilasciato nel 2019 e risolve due problemi con il protocollo MQTT 3.1.1 ampiamente utilizzato:

- MQTT 3.1.1 presenta problemi di personalizzazione o aggiunta di metadati al protocollo, cosa comune nei dati HTTP.
- MQTT 3.1.1 ha difficoltà con l'interoperabilità durante la comunicazione tra piattaforme, librerie e percorsi dati diversi.

Per risolvere questi problemi, MQTT 5 ha introdotto le “proprietà utente”. Vedremo poi queste nuove proprietà e abilità.

MQTT, acronimo che contiene i termini “Message Queue” è in realtà improprio. Non ci sono code di messaggi inerenti al protocollo. Sebbene sia possibile accodare i messaggi, non è necessario e spesso non viene fatto. MQTT è basato su TCP e quindi include alcune garanzie che un pacchetto venga trasferito in modo affidabile.

É anche un **protocollo asimmetrico**. Supponiamo che il nodo A debba comunicare con il nodo B. Un protocollo asimmetrico tra A e B richiede che tutte le informazioni necessarie per il riassettaggio dei pacchetti devono essere contenute nell'header inviato da A.

I sistemi asimmetrici hanno un master e uno slave (l'FTP ne è un classico esempio). A o B possono assumere il ruolo di master o slave (Telnet è un esempio di questo caso). MQTT ha ruoli distinti che hanno senso nella topologia sensore/nuvola.

MQTT può **conservare un messaggio su un broker a tempo indeterminato**. Questa modalità di funzionamento è controllata da un flag. Un messaggio conservato su un broker viene inviato a qualsiasi client che si iscrive a quel ramo di MQTT, il messaggio viene trasmesso immediatamente a quel nuovo client. Ciò consente a un nuovo client di ricevere uno stato o un segnale da un ramo appena sottoscritto senza attendere. In genere, un client che si iscrive ad un ramo potrebbe dover attendere ore o addirittura giorni prima che un client pubblichi nuovi dati.

Il tempo massimo di keep alive è di 18 ore, 12 minuti e 15 secondi. L'impostazione di keep-alive internal

su 0 disabilerà la funzionalità keep-alive. Il timer è controllato dal cliente e può essere modificato dinamicamente per riflettere le modalità di sospensione o le variazioni della potenza del segnale.

8.2.1.1 MQTT-SN

Un derivato di MQTT è chiamato MQTT-SN (a volte chiamato MQTT-S) per le reti di sensori. Mantiene la stessa filosofia di MQTT come protocollo leggero per dispositivi edge ma è progettato specificamente per le sfumature di una WPAN tipica degli ambienti con sensori.

Queste caratteristiche includono il supporto per i collegamenti a larghezza di banda ridotta, l'errore di collegamento, la lunghezza del messaggio breve e l'hardware con risorse limitate. *MQTT-SN è, infatti, così leggero che può essere eseguito con successo su BLE e Zigbee.*

MQTT-SN non richiede uno stack TCP/IP. Può essere utilizzato su un collegamento seriale (modo preferito), dove può utilizzare un semplice protocollo di collegamento di dimensioni, in termini di dati, molto ridotte. In alternativa può essere utilizzato su UDP, che richiede sensibilmente meno risorse di TCP.

Ci sono quattro componenti fondamentali in una topologia MQTT-SN:

- **Gateway:** ha la responsabilità della conversione del protocollo da MQTT-SN a MQTT e viceversa (sebbene siano possibili altre traduzioni). I gateway possono anche essere aggregati o trasparenti;
- **Spedizionieri:** un percorso tra un sensore e un gateway MQTT-SN può prendere molti percorsi e passare attraverso diversi router lungo il percorso. I nodi tra il client di origine e il gateway MQTT-SN sono chiamati forwarder e semplicemente reincapsulano i frame MQTT-SN in frame MQTT-SN nuovi e invariati che vengono inviati alla destinazione fino a quando non arrivano al gateway MQTT-SN corretto per la conversione del protocollo;
- **Clienti:** i client si comportano allo stesso modo di MQTT e sono in grado di sottoscrivere e pubblicare dati;
- **Broker:** i broker si comportano allo stesso modo di MQTT.

8.3 Ambient Computing, Sensori Sintetici

Con il termine *ambient computing (elaborazione nell'ambiente)* sintetizziamo il concetto di elaborazione che si fonde in un ambiente, mentre il rilevamento sintetico (la sensoristica) offre la possibilità di aggregare sensori di livello perimetrale in un sistema più ampio che monitora l'ambiente. Dispositivi come sensori MEM, sensori termici e sensori di movimento PIR sono progettati per misurare tratti specifici dell'ambiente come pressione, temperatura o movimento.

Questi sensori da soli non hanno la capacità di comprendere azioni più complesse all'interno di un ambiente, ad esempio un fornello lasciato acceso o se il caffè ha completato l'erogazione.

Sebbene queste attività possano essere eseguite semplicemente aggiungendo sensori direttamente a ciascun dispositivo (ad esempio ad ogni singolo elettrodomestico), il rilevamento sintetico tenta di condensare tutti i sensori in un unico dispositivo hardware che analizza l'intero ambiente, sia esso una stanza, un ufficio, una sala conferenze, una sala ristorante, un cinema, un ipermercato, etc. etc.

L'ambient computing, a volte detto anche *pervasive computing*, è un nuovo paradigma di interazione uomo-macchina. Solitamente le interazioni umane con un computer coinvolgono un'interfaccia uomo-macchina realizzata con monitor, tastiera e mouse o touchscreen. Indipendentemente dal tipo di interfaccia, è chiara la distinzione tra noi (esseri umani) ed il dispositivo informatico con cui si deve interagire.

Nell'ambient computing, non c'è un computer con cui interagire: il computer è l'ambiente. L'interazione è con l'ambiente in cui ci si trova. Le cose stesse vengono usate, manipolate, consumate o indossate. Le cose stesse fanno parte del tessuto informatico. I dispositivi interagiscono con l'intelligenza artificiale per assistere e servire quando necessario, per poi passare in secondo piano quando non è necessario.

Funziona integrando ciò che abbiamo imparato finora su sensori, IoT e edge computing. In estrema sintesi si tratta di raccogliere e analizzare dati all'interno di un ambiente e agire su di esso. Dal punto di vista del computer periferico, l'obiettivo è farlo senza essere invadente.

I principi dell'ambient computing sono 4:

- **Invisibile:** i sistemi non dovrebbero attirare l'attenzione su se stessi. L'informatica dovrebbe essere onnipresente e ovunque ma non associata a un dispositivo. La tecnologia è a portata di mano quando necessario e nascosta quando non è necessaria;
- **Incorporato:** gli oggetti possono incorporare intelligenza sotto forma di sensori e capacità di elaborazione.
- **Fluidi:** l'ambient computing consiste nel rendere fluide (trasparenti) le più complesse interazioni con i computer. Mette gli esseri umani al centro piuttosto che un PC, un mouse e una tastiera. L'elaborazione avviene dove sei e conosce lo stato dell'ambiente in cui ti trovi;
- **Interconnesso:** un ambiente composto da cose e oggetti diversi che interagiscono comunicando e lavorando insieme. Questo diventa difficile con standard e protocolli concorrenti.

L'edge computing è il cuore dell'ambient computing.

I primi segnali di implementazione di questo nuovo paradigma sono dati da dispositivi come Amazon Alexa o Google Home, che tentano di posizionare un assistente vocale, connesso al cloud, all'interno di un ambiente.

Si potrebbe non essere nemmeno consapevoli che questo "edge computer" si trova nell'ambiente finché non lo si interpella tramite un messaggio vocale ("Alexa" o "ok Google"). Possono interfacciarsi con alcuni sensori ambientali e dispositivi IoT e interagire in modo invisibile.

La sfida per il progettista di edge computing è prendere le tecnologie descritte in precedenza in questo capitolo e costruire un'appliance che interagisca con gli esseri umani ma non sia vista dagli esseri umani. I sistemi perimetrali possono avere una grande quantità di potenza di calcolo, tanto quanto i blade dei data center ma, per quanto rendano evidente la loro funzionalità, rendere trasparente la loro struttura fisica è la sfida ingegneristica.

Ciò implica che dobbiamo considerare hardware, sistemi di comunicazione e infrastruttura secondo nuovi parametri di forma, spazio occupato, suono prodotto e visibilità: niente luci intermittenti, niente ventole ad alta velocità, nessun cablaggio invadente, etc. Un buon esempio di elaborazione ambientale che utilizza un dispositivo di edge computing è il rilevamento sintetico, che esamineremo nella sezione seguente.

8.3.1 Sensori Sintetici

Ad esempio, uno di questi computer con sensori sintetici può essere posizionato all'interno di una stanza e sapere quale fornello è acceso, se la lavastoviglie è in funzione, se il rubinetto è stato lasciato aperto. Nel diagramma che seguirà, sono mostrati cinque sensori che accumulano dati su un singolo edge computer. Questi segnali in tempo reale vengono elaborati e inviati tramite un motore di inferenza addestrato per rilevare gli eventi.

Ciò che manca a questo set di sensori è una vera fotocamera. Il rilevamento sintetico non può mai utilizzare dati video o fotografici. Il video rappresenta un "blob" di dati non strutturati che richiede una diversa forma di apprendimento automatico. Tutti gli altri insiemi di caratteristiche ambientali sono raccolti come un insieme di segnali *correlati nel tempo*.

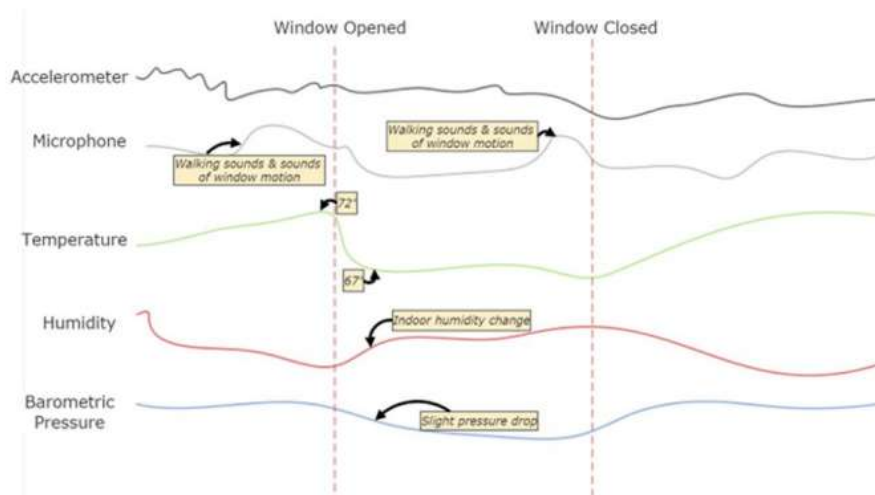


Figura 8.11: rilevamento sintetico che utilizza cinque sensori ambientali che accumulano segnali correlati al tempo, elaborati in tempo reale tramite un motore di inferenza edge. Gli eventi vengono tracciati in base alla firma di questi segnali: ad esempio basandosi sul modo in cui variano temperatura e umidità possono indicare l'apertura o la chiusura di una finestra all'interno di una stanza.

Quando il dispositivo periferico acquisisce i dati del sensore, tenta prima di rimuovere il rumore e le aberrazioni nei segnali. Successivamente tenterà di normalizzare le basi temporali e i timestamp dei dati raccolti. In altre parole, alcuni sensori possono essere campionati a una frequenza più alta e altri a frequenze più basse; il software regolerà correttamente i numerosi campioni su una base di tempo comune. Dopo aver corretto i segnali, il dispositivo periferico eseguirà un modello di inferenza addestrato sui dati, in tempo reale. Il prodotto finale sarà una classificazione dello stato in cui si trova l'ambiente.

Da un punto di vista hardware, un edge computer potrebbe utilizzare qualcosa come un circuito integrato Particle Photon che incorpora un microcontrollore ST Micro STM32F205. L'obiettivo è unire il maggior numero possibile di ingressi di sensori correlati al tempo all'interno di un singolo modulo.



Figura 8.12: Copertura Spettro di Frequenze

In questo caso, i sensori inclusi sono:

- Grayscale AMG833 per il rilevamento del livello a infrarossi;
- Sensore di colore e luminosità AMS TCS34725;
- Magnetometro a tre assi Xtrinsic MAG3110;
- Sensore di temperatura, barometro e umidità Bosch BM280;
- Sensore giroscopio e accelerometro a sei assi TDK InvenSense MPU6500;
- Rilevamento Wi-Fi a 2,4 GHz delle variazioni dell'intensità del campo RSSI;
- Sensore di movimento PIR Panasonic AMN2111;
- Sensore microfono MEMS analogico ADMP401;
- Sensore EMI induttore da 100 mH.

Tutti i dispositivi, il processore e la memoria, nonché l'interfaccia di rete USB e Wi-Fi, si adattano a un modulo da 45,4 mm per 48,6 mm. Ciò segue il paradigma dell'informatica ambientale creando un

dispositivo di rilevamento dello spazio che è cognitivo dei comportamenti incentrati sull'utente ma non richiede una pletora di sensori e dispositivi per essere collegato a ciascun oggetto. Il dispositivo può essere nascosto alla vista ma fornisce informazioni utili sullo stato di un ambiente al proprietario di casa o all'imprenditore.

Gierad Laput, Yang Zhang e Chris Harrison in "Sensori sintetici: verso il rilevamento per scopi generici" (Atti della conferenza CHI 2017 sui fattori umani nei sistemi informatici. pp 3986-3999. ACM, 2017). Video della conferenza, durata 20m:41s : <https://www.youtube.com/watch?v=DJMqNJAMNEs>

Capitolo 9

Vulnerabilità e Protezione dei Sistemi IoT

9.1 Classificazione dei Tipi di Attacco Informatico

Esamineremo il problema della sicurezza dell'IoT da un punto di vista olistico, dal sensore al cloud, facendo riferimento a specifici attacchi verso sistemi IoT realmente accaduti negli ultimi anni, nonché metodi per contrastare tali attacchi in futuro.

Abbiamo già illustrato quali siano le dimensioni ed il potenziale di crescita dell'Internet of Things (IoT). Attualmente ci sono miliardi di dispositivi e la crescita a due cifre del collegamento del mondo analogico a Internet costituisce anche la più grande superficie di attacco sulla Terra.

Diversi tipi di malware sono già stati sviluppati, implementati e diffusi a livello globale, interrompendo innumerevoli attività.

In qualità di sistemisti IoT, siamo responsabili della comprensione della pila di tecnologie IoT e della loro sicurezza.

Poiché posizioniamo dispositivi che non sono mai stati connessi ad Internet prima di questi anni, siamo responsabili della loro progettazione.

In molte realizzazioni IoT la sicurezza è stata spesso considerata per ultima, sbagliando.

Capita anche che i sistemi IoT realizzati siano ormai così vincolati che implementare il livello di sicurezza di cui godono i moderni sistemi Web e PC diventa difficile se non impossibile.

Vedremo alcuni attacchi particolarmente distruttivi incentrati sull'IoT come spunto di riflessione su quanto sia debole la sicurezza nell'IoT e quanti danni possano essere arrecati da una sottovalutazione del tema.

Esistono metodi per garantire la sicurezza a ogni livello dello stack: dai dispositivi fisici ai sistemi di comunicazione e reti fino alla protezione dei dati tramite perimetri software-defined e blockchain.

(Si veda lo United States Cybersecurity Improvement Act del 2017 e cosa potrebbe significare per i dispositivi IoT).

La cosa più importante in questo ambito è garantire la sicurezza a tutti i livelli, dal sensore, passando per gli edge computer, i sistemi di comunicazione, i router fino al cloud.

Linguaggio della Sicurezza Informatica: La sicurezza informatica ha una serie di definizioni che descrivono diversi tipi di attacchi, vediamo le principali.

Termini di Attacco e Minaccia: Di seguito sono riportati i termini e le definizioni di diversi attacchi o minacce informatiche malevole.

Attacco di Amplificazione: ingrandisce la larghezza di banda utilizzata dalla vittima.

Un utente malintenzionato potrebbe utilizzare un servizio legittimo, ad esempio NTP o DNS, per portare l'attacco. NTP può amplificare 556 volte e l'amplificazione DNS può aumentare la larghezza di banda di 179 volte.

Spoofing ARP: un tipo di attacco che invia un messaggio ARP falsificato che produce una falsa associazione tra l'indirizzo MAC dell'attaccante e l'IP di un sistema legittimo.

Scansioni Banner: tecnica utilizzata per fare l'inventario dei sistemi su una rete che può essere utilizzata anche da un utente malintenzionato per ottenere informazioni su un potenziale bersaglio, eseguendo richieste HTTP e ispezionando le informazioni restituite dal sistema operativo (SO) e/o dalle applicazioni o servizi (ad esempio, nc www.target.com 80) e/o dall'hardware.

Distributed Denial-of-Service (DDoS): un attacco che tenta di interrompere o rendere non disponibile un servizio online sovraccaricandolo da più fonti (distribuite).

Botnet: dispositivi connessi a Internet infettati e compromessi da malware che operano collettivamente mediante un controllo comune, utilizzati principalmente all'unisono per generare massicci attacchi DDoS da più client.

Forza Bruta: un metodo per tentativi ed errori per accedere a un sistema o aggirare la crittografia.

Overflow del Buffer: sfrutta un bug o un difetto nell'esecuzione del software che semplicemente sovraccarica un buffer o un blocco di memoria con più dati di quelli allocati. Questo sovraccarico può sovrascrivere altri dati in indirizzi di memoria adiacenti. Un utente malintenzionato può inserire codice dannoso in quell'area e forzare l'esecuzione del puntatore all'istruzione da lì. I linguaggi compilati come C e C++ sono particolarmente suscettibili agli attacchi di buffer overflow poiché mancano di protezione interna. La maggior parte dei bug di overflow sono il risultato di un software mal costruito che non controlla i limiti dei valori di input.

Slitte NOP: una sequenza di istruzioni di assemblaggio NOP iniettate utilizzate per "far scorrere" il puntatore di istruzioni di una CPU nell'area desiderata di codice dannoso. Di solito fa parte di un attacco di overflow del buffer.

Exploit RCE: esecuzione di codice in modalità remota che consente a un utente malintenzionato di eseguire codice arbitrario. Questo di solito si presenta sotto forma di un attacco di overflow del buffer su HTTP o altri protocolli di rete che inietta codice malware.

Attacco di Analisi della Potenza Correlata: Consente di scoprire le chiavi di crittografia segrete archiviate in un dispositivo attraverso quattro passaggi:

1. Esamina il consumo energetico dinamico di un target e lo registra per ogni fase del processo di crittografia.
2. Forza la destinazione a crittografare diversi oggetti di testo in chiaro e registrarne il consumo energetico.
3. Attacca piccole parti della chiave (sottochiavi) considerando ogni possibile combinazione e calcolando il coefficiente di correlazione di Pearson tra la potenza modellata e quella effettiva.
4. Mette insieme la migliore sottochiave per ottenere la chiave completa.

Attacco a Dizionario: metodo per accedere a un sistema di rete inserendo sistematicamente parole da un file dizionario, I dizionari vanno intesi, in questo contesto, non solo come raccolte ordinate di parole in lingua ma come grandi elenchi di testi, ad esempio titoli ed incipit di libri, titoli di film, titoli di opere liriche, etc. etc.

Fuzzing: un attacco che consiste nell'invio di dati non corretti o non standard a un dispositivo e nell'osservare come reagisce il dispositivo. Ad esempio, se un dispositivo funziona male o mostra effetti negativi, l'attacco fuzz potrebbe aver esposto una vulnerabilità.

Attacco Man-in-the-middle (MITM): una forma comune di attacco che pone un dispositivo nel mezzo di un flusso di comunicazione tra due parti ignare. Il dispositivo ascolta, filtra e si appropria delle informazioni dal trasmettitore e ritrasmette le informazioni selezionate al ricevitore. Un MITM può inserirsi direttamente nel collegamento, fungendo da ripetitore, o può essere in banda laterale per ascoltare la trasmissione.

Attacco di Riproduzione: un attacco di rete in cui i dati vengono ripetuti o riprodotti in modo dannoso dall'attaccante che intercetta i dati, li archivia e li trasmette a piacimento.

Attacco di Programmazione Orientata al Ritorno (ROP): un difficile exploit di sicurezza che un utente malintenzionato può utilizzare per sovvertire potenzialmente le protezioni con memoria non in esecuzione o esecuzione di codice dalla memoria di sola lettura. Se un utente malintenzionato ottiene il controllo di uno stack di processi attraverso un overflow del buffer o altri mezzi, può passare a sequenze di istruzioni legittime e invariate già presenti. L'attaccante cerca sequenze di istruzioni per chiamare gadget che possono essere messi insieme per formare un attacco malevolo.

Return-to-libc: un tipo di attacco che inizia con un buffer overflow in cui l'attaccante inietta codice per passare a libc o ad altre librerie comunemente utilizzate nello spazio di memoria dei processi nel tentativo di chiamare direttamente le routine di sistema. Bypassa la protezione offerta dalla memoria non eseguibile e dalle bande di guardia. Questa è una forma specifica di attacco ROP.

Rootkit: in genere, il software dannoso (sebbene spesso utilizzato per sbloccare gli smartphone) utilizzato per rendere non rilevabili i payload di altri software. I rootkit utilizzano diverse tecniche mirate come gli overflow del buffer per attaccare i servizi del kernel, gli hypervisor e i programmi in modalità utente.

Attacco side-channel: un attacco utilizzato per ottenere informazioni dal sistema di una vittima osservando gli effetti secondari del sistema fisico anziché trovare exploit di runtime o exploit zero-day. Esempi di attacchi del canale laterale includono l'analisi della potenza di correlazione, l'analisi acustica e la lettura dei dati residui dopo che sono stati eliminati dalla memoria.

Ingegneria Sociale: per quanto riguarda la sicurezza delle informazioni, una forma di attacco basata sulla manipolazione psicologica e sull'inganno personale per ottenere informazioni private.

Spoofing: la parte o il dispositivo dannoso impersona un altro dispositivo o utente su una rete.

SQL Injection: una forma di attacco con l'intento di distruggere il contenuto del database. Viene utilizzato su database che utilizzano istruzioni SQL (Structured Query Language).

SYN Flood: si verifica quando un host invia un pacchetto TCP:SYN che un agente canaglia falsificherà e falsificherà. Ciò farà sì che l'host crei connessioni semiaperte a molti indirizzi inesistenti causando l'esaurimento di tutte le risorse dell'host.

XSS: una vulnerabilità, chiamata anche cross-site scripting, nelle applicazioni Web in cui gli aggressori iniettano script lato client. È stata la forma più diffusa di attacchi informativi fino al 2007.

Exploit zero-day: difetti di sicurezza o bug nel software commerciale o di produzione sconosciuti al progettista o al produttore.

9.2 Tecniche e Meccanismi di Difesa

Concluso l'elenco dei più diffusi termini di attacco e minaccia, riportiamo ora i termini e le definizioni delle diverse tecniche e meccanismi di difesa informatica.

Address Space Layout Randomization (ASLR): un meccanismo di difesa che protegge la memoria e contrasta gli attacchi di overflow del buffer mediante la randomizzazione del punto in cui un eseguibile viene caricato in memoria. Un buffer overflow che inietta malware non può prevedere dove verrà caricato in memoria, quindi manipolare il puntatore di istruzioni diventa estremamente difficile. Protegge dagli attacchi di ritorno a libc.

Buco nero (sinkhole): difende dagli attacchi DDoS. Dopo aver rilevato un attacco DDoS, vengono stabiliti percorsi dal server DNS o dall'indirizzo IP interessato per forzare i dati non autorizzati a un buco nero o a un endpoint inesistente. Le buche eseguono ulteriori analisi per filtrare i dati validi.

Prevenzione esecuzione dati (DEP): contrassegna un'area come eseguibile o non eseguibile. Ciò impedisce a un utente malintenzionato di eseguire codice iniettato in modo dannoso in tale regione tramite un attacco di overflow del buffer. Il risultato è un errore di sistema o un'eccezione.

Deep Packet Inspection (DPI): un metodo per ispezionare ogni pacchetto (dati ed eventualmente informazioni di intestazione) in un flusso di dati per isolare intrusioni, virus, spam e altri criteri che vengono filtrati.

Firewall: un costrutto di sicurezza di rete che concede o rifiuta l'accesso di rete ai flussi di pacchetti tra una zona non attendibile e una zona attendibile. Il traffico può essere controllato e gestito tramite elenchi di controllo di accesso (ACL) sui router. I firewall possono eseguire filtri con stato e fornire regole basate sulle porte di destinazione e sullo stato del traffico.

Bande di protezione e memoria non eseguibile: protegge le aree di memoria scrivibili e non eseguibili. Protegge dalle slitte NOP. Intel: bit NX, bit ARM XN.

Honeypot: strumento di sicurezza per rilevare, deviare o decodificare gli attacchi dannosi. Gli Honeypot appaiono come siti Web legittimi o nodi accessibili in una rete, ma in realtà sono isolati e monitorati. I dati e le interazioni con il dispositivo vengono registrati.

Controllo dell'accesso alla memoria basato sulle istruzioni: una tecnica per separare la parte dei dati di uno stack dalla parte dell'indirizzo di ritorno. Questa tecnica aiuta a proteggere dagli attacchi ROP ed è particolarmente utile nei sistemi IoT vincolati.

Sistema di rilevamento delle intrusioni (IDS): un costrutto di rete per rilevare le minacce in una rete attraverso l'analisi fuori banda del flusso di pacchetti; pertanto, non in linea con la sorgente e la destinazione in modo da influenzare la risposta in tempo reale.

Sistema di prevenzione delle intrusioni (IPS): blocca le minacce a una rete tramite una vera analisi in linea e il rilevamento statistico o delle firme delle minacce.

Milkers: uno strumento difensivo che emula un dispositivo botnet infetto e si collega al suo host malevolo consentendo di comprendere e "mungere" i comandi del malware inviati alla botnet controllata.

Scansione delle porte: un metodo per trovare una porta aperta e accessibile su una rete locale.

Root of Trust (RoT): avvia l'esecuzione su un dispositivo di avvio a freddo da una fonte di memoria attendibile immutabile (come la ROM). Se il software di avvio/BIOS anticipato può essere modificato senza controllo, non esiste RoT. Il RoT è solitamente la prima fase di un avvio sicuro multifase.

Chiave pubblica: una chiave pubblica viene generata con una chiave privata ed è accessibile a entità esterne. Una chiave pubblica può essere utilizzata per decrittografare gli hash.

Infrastruttura a chiave pubblica (PKI): fornisce una definizione di gerarchie di verificatori per garantire l'origine di una chiave pubblica. Un certificato è firmato dalle CA.

Chiave privata: generata con una chiave pubblica, mai rilasciata esternamente e archiviata in modo sicuro. Viene utilizzato per crittografare gli hash.

Avvio sicuro: una serie di passaggi di avvio per un dispositivo che si avvia a un RoT e procede attraverso il caricamento del sistema operativo e dell'applicazione in cui ogni firma del componente viene verificata come autentica. La verifica viene eseguita tramite chiavi pubbliche caricate nelle precedenti fasi di avvio attendibile.

Stack canary: protegge lo spazio dello stack di elaborazione dai sovraccarichi dello stack e impedisce l'esecuzione di codice da uno stack.

Ambiente di esecuzione attendibile (TEE): un'area sicura di un processore che garantisce la protezione del codice e dei dati che risiedono all'interno di questa zona. Questo è solitamente un ambiente di esecuzione sul core del processore principale in cui il codice per l'avvio sicuro, i trasferimenti monetari o la gestione della chiave privata verrà eseguito con un livello di sicurezza più elevato rispetto alla maggior parte del codice.

9.3 Esempi di Attacchi Informatici a Sistemi IoT

L'area della sicurezza informatica è un argomento ampio e imponente che va oltre lo scopo di questo capitolo.

Vedremo però, a titolo di esempio, tre attacchi verso sistemi IoT perché comprendendo a fondo i comportamenti di queste minacce, il sistemista IoT può derivare tecnologie e processi preventivi per garantire che eventi simili non accadano o siano perlomeno mitigati.

- Mirai - l'attacco DoS più dannoso della storia, generato da dispositivi IoT non sicuri in aree remote
- Stuxnet - un'arma informatica di uno stato nazionale che prende di mira dispositivi IoT SCADA industriali che controllano le centrifughe di arricchimento dell'uranio e causano danni sostanziali e irreversibili al programma nucleare iraniano
- Reazione a catena - un metodo di ricerca per sfruttare il PAN utilizzando nient'altro che una lampadina, senza bisogno di Internet

9.3.1 MIRAI

Mirai è il nome del malware che ha infettato i dispositivi IoT Linux nell'agosto del 2016.

L'attacco è arrivato sotto forma di una botnet che ha generato una massiccia tempesta DDoS.

Gli obiettivi di alto profilo includevano Krebs on Security, un popolare blog sulla sicurezza di Internet; Dyn, un provider DNS molto popolare e ampiamente utilizzato per Internet e Lonestar cell, un grande operatore di telecomunicazioni in Liberia.

Gli obiettivi più piccoli includevano siti politici italiani, server Minecraft in Brasile e siti di aste russi.

Il DDoS su Dyn ha avuto effetti secondari su altri provider (anche di dimensioni rilevanti) che utilizzavano i loro servizi, come i server Sony Playstation, Amazon, GitHub, Netflix, PayPal, Reddit e Twitter. Sono stati infettati complessivamente 600.000 dispositivi IoT come parte del collettivo di botnet.

Il codice sorgente di Mirai è stato rilasciato su hackforums.net (un blog di hacker).

I ricercatori hanno scoperto come funzionava e si svolgeva l'attacco Mirai analizzando i sorgenti e attraverso tracce nei log:

- Scansione delle vittime: MIRAI ha prima eseguito una rapida scansione asincrona utilizzando i pacchetti TCP SYN per sondare indirizzi IPV4 casuali.
Ha cercato specificamente la porta SSH/Telnet TCP 23 e la porta 2323.
Se indirizzo IP e porta risultavano raggiungibili, venivano registrati per poi essere utilizzati dalla fase due.
Mirai includeva una lista nera codificata di indirizzi da evitare. La lista nera era composta da 3,4 milioni di indirizzi IP appartenenti al servizio postale degli Stati Uniti, Hewlett-Packard, GE e il Dipartimento della Difesa degli Stati Uniti.
Mirai era in grado di scansionare a velocità di circa 250 byte al secondo, relativamente bassa trattandosi di una botnet (attacchi come SQL Slammer hanno generato scansioni a 1,5 Mbps).

Il motivo principale di tale bassa velocità è che i dispositivi IoT in genere sono molto più limitati nella potenza di elaborazione rispetto ai dispositivi desktop e mobili.

- **Telnet a forza bruta:** Mirai ha tentato di stabilire una sessione Telnet con ogni vittima individuata in fase 1, inviando casualmente 10 coppie di nome utente e password utilizzando un attacco dizionario di 62 coppie.

Se l'accesso riusciva, registrava l'host su un server C2 centrale.

Le varianti successive di Mirai hanno evoluto il bot per eseguire exploit RCE.

- **Infect:** il server MIRAI provvedeva ad inviare un loader (programma di caricamento) alla potenziale vittima.

Il programma era responsabile dell'identificazione del sistema operativo e dell'installazione di malware specifico per quel dispositivo.

Cercava poi altri processi concorrenti utilizzando la porta 22 o 23 e ne interrompeva l'esecuzione (kill) insieme ad altri malware che potrebbero essere stati già presenti sul dispositivo.

Il file binario del loader veniva poi eliminato ed il nome del processo veniva offuscato per nascondere la presenza. Il bot rimaneva inattivo fino a quando non riceveva il comando di attacco.

I dispositivi presi di mira erano dispositivi IoT costituiti da telecamere IP, DVR, router consumer, telefoni VOIP, stampanti e set-top box, utilizzando file binari di malware progettati appositamente per processori ARM a 32 bit, MIPS a 32 bit e X86 a 32 bit e specifici per il dispositivo IoT oggetto di attacco.

La prima scansione è avvenuta il 1 agosto 2016 da un sito di hosting web statunitense. La scansione ha richiesto 120 minuti prima che trovasse un host con una porta aperta e una password nel dizionario. Dopo un altro minuto, altri 834 dispositivi sono stati infettati. In 20 ore sono stati infettati 64.500 dispositivi. Mirai ha raddoppiato le sue dimensioni in 75 minuti. La maggior parte dei dispositivi infetti, che si sono trasformati in botnet, si trovavano in Brasile (15,0%), Colombia (14,0%) e Vietnam (12,5%), sebbene gli obiettivi degli attacchi DDoS fossero anche in altre regioni del pianeta.

Il danno si è limitato ad interruzioni di servizio (attacco DDoS). Gli attacchi si sono presentati sotto forma di SYN flood, GRE IP network flood, STOMP flood e DNS flood.

Nel corso di cinque mesi, 15.194 singoli comandi di attacco sono stati emessi dai server C2 e hanno colpito 5.042 siti Internet.

Il 21 settembre 2016, la botnet Mirai ha scatenato un massiccio attacco DDoS al sito del blog Krebs on Security e ha generato 623 Gbps di traffico.

Ha rappresentato il peggior attacco DDoS di tutti i tempi. Quello che segue è uno screenshot in tempo reale catturato durante l'attacco Mirai utilizzando www.digitalattackmap.com.

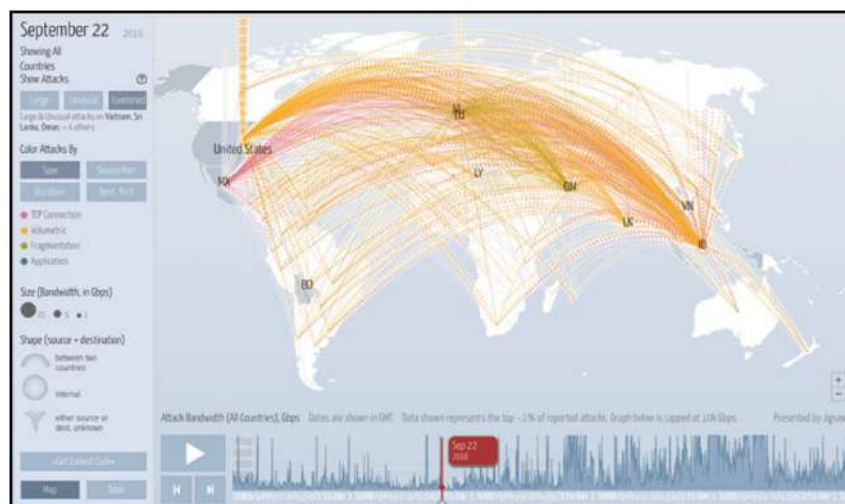


Figura 9.1: Una vista dell'attacco DDoS Mirai al sito Web di Krebs on Security (fonte DigitalAttackMap.com)

9.3.2 STUXNET

Stuxnet è stata la prima arma informatica documentata nota rilasciata per danneggiare permanentemente i beni dell'Iran.

Si trattava di un worm che veniva rilasciato per danneggiare i controllori logici programmabili (PLC) Siemens basati su SCADA e utilizzava un rootkit per modificare la velocità di rotazione dei motori sotto il controllo diretto del PLC.

I progettisti hanno fatto di tutto per garantire che il virus prendesse di mira solo i dispositivi collegati a PLC Siemens S7-300 con velocità di rotazione a frequenza variabile tra 807 Hz e 1210 Hz, poiché sono generalmente utilizzati per pompe e centrifughe a gas per l'arricchimento di uranio.

L'attacco presumibilmente è iniziato ad aprile o marzo 2010. Il processo di infezione ha seguito questi passaggi:

- **Infezione iniziale:** il worm è iniziato infettando un computer Windows host sfruttando le vulnerabilità rilevate in precedenti attacchi di virus.
Si pensa che si sia diffuso inserendo una chiavetta USB nella macchina iniziale.
Ha utilizzato quattro exploit zero-day contemporaneamente (un livello di sofisticazione senza precedenti).
Gli exploit hanno utilizzato un attacco rootkit utilizzando codice in modalità utente e modalità kernel e hanno installato un driver di dispositivo rubato ma correttamente firmato e certificato da Realtek.
Questo driver firmato in modalità kernel era necessario per nascondere Stuxnet ai pacchetti anti-virus diffidenti.
- **Attacco e diffusione di Windows:** una volta installato tramite il rootkit, il worm ha iniziato a cercare nel sistema Windows i file tipici di un controller SCADA Siemens, WinCC/PCS 7 SCADA, noto anche come Step-7.
Se il worm trovava il software di controllo SCADA Siemens, tentava di accedere a Internet tramite un C2 utilizzando URL non corretti (www.mypremierfutbol.com e www.todaysfutbol.com) per scaricare sue versioni più recenti.
Cercava poi ulteriormente nel filesystem per individuare un file chiamato `s7otbdc.dll`, che fungeva da libreria di comunicazione critica tra la macchina Windows e il PLC.
Stuxnet si è inserito tra il sistema WinCC e `s7otbdc.dll` per agire come un attaccante man-in-the-middle.
Il virus ha iniziato la sua attività registrando il normale funzionamento delle centrifughe.
- **Distruzione:** quando ha deciso di coordinare l'attacco, ha riprodotto i dati preregistrati ai sistemi SCADA, che non avevano motivo di credere che qualcosa fosse compromesso o che si comportasse in modo irregolare.
Stuxnet ha inflitto il suo danno manipolando i PLC con due diversi attacchi coordinati per danneggiare l'intera schiera della struttura iraniana.
Il danno ai rotori della centrifuga si è verificato lentamente nel tempo, con incrementi di 15 o 50 minuti separati da 27 giorni di normale funzionamento.
Ciò ha prodotto uranio arricchito in modo improprio, perni dei rotori delle centrifughe incrinati o addirittura distrutti.
Si ritiene che oltre 1.000 centrifughe di arricchimento dell'uranio siano state paralizzate e danneggiate da questo attacco al principale impianto di arricchimento iraniano a Natanz, in Iran.

Oggi il codice Stuxnet è disponibile online ed è essenzialmente un campo da gioco open source per creare exploit derivati (<https://github.com/micriccator/stuxnet>).

9.3.3 REAZIONE A CATENA

Chain Reaction è uno studio accademico che mostra una nuova generazione di attacchi informatici incentrati sulle reti mesh PAN che possono essere eseguiti senza alcun collegamento a Internet.

Mostra inoltre quanto possono essere vulnerabili i sensori e i sistemi di controllo IoT remoti.

Il vettore di attacco è costituito dalle lampadine Philips Hue che si trovano generalmente nelle case dei consumatori e che possono essere controllate da Internet e dalle app per smartphone.

L'exploit può essere esteso agli attacchi delle smart city e avviato semplicemente inserendo una singola

lampada intelligente infetta.

Le luci Philips Hue utilizzano il protocollo Zigbee per stabilire una mesh.

I sistemi di illuminazione Zigbee rientrano in un programma chiamato Zigbee Light Link (ZLL) per imporre un metodo standard per l'interoperabilità dell'illuminazione.



Figura 9.2: Luci Philips Hue

I messaggi ZLL non sono crittografati o firmati ma la crittografia viene utilizzata per proteggere le chiavi scambiate se viene aggiunta una lampada alla rete mesh.

Questa chiave principale è nota a tutti nell'alleanza ZLL e purtroppo è trapelata all'esterno dell'alleanza. ZLL impone anche che le lampade che si uniscono alla rete debbano essere molto vicine all'iniziatore per impedire che ci si possa impossessare delle luci del vicino.

Zigbee offre anche un metodo di riprogrammazione over-the-air (OTA) e i pacchetti del firmware sono crittografati e firmati.

I ricercatori hanno utilizzato un piano di attacco in quattro fasi:

1. "Rompe" la crittografia e la firma del pacchetto firmware OTA.
2. Distribuisce un aggiornamento del firmware malevolo su una singola lampadina usando la crittografia rotta e le chiavi di firma
3. La lampadina compromessa si unisce alla rete in base alla chiave principale rubata e sfrutta la sicurezza di prossimità attraverso un difetto zero-day trovato nella parte Atmel AtMega comunemente utilizzata
4. Dopo essersi unita con successo a una rete Zigbee, invia il suo carico utile alle luci vicine, infettandole rapidamente. Il processo si espande in base alla teoria della percolazione e infetta interi sistemi di illuminazione cittadini.

Zigbee utilizza la crittografia AES-CCM (parte dello standard IEEE 802.15.4) per crittografare gli aggiornamenti del firmware OTA.

Per violare la crittografia del firmware, gli aggressori hanno utilizzato l'analisi della potenza di correlazione (CPA) e l'analisi della potenza differenziale (DPA).

L'analisi della potenza di correlazione (CPA) e l'analisi della potenza differenziale (DPA) sono forme di attacco sofisticate, in cui un dispositivo come l'hardware del controller della lampadina viene posizionato su un banco e viene misurata la potenza che consuma.

Per mezzo di strumenti sofisticati, è possibile misurare la potenza dinamica utilizzata da una CPU che esegue un'istruzione o sposta dati (ad esempio, quando viene eseguito un algoritmo di crittografia).

Questo metodo è chiamato analisi della potenza semplice (SPA), con il quale però risulta ancora molto difficile decifrare la chiave. CPA e DPA estendono le capacità oltre SPA utilizzando una correlazione statistica.

Utilizzando sia DPA che CPA, i ricercatori hanno corrotto il sistema di illuminazione Philips Hue come segue:

- Hanno utilizzato il CPA per decifrare l'AES-CBC. Gli aggressori non avevano alcuna chiave, nessun nonce, nessun vettore di inizializzazione
- Hanno usato DPA per decifrare la modalità contatore AES-CTR per rompere la crittografia del bundle del firmware ed hanno trovato 10 posizioni che l'AES-CTR sembrava eseguire, il che ha esposto 10 possibilità

- Si sono quindi concentrati sulla violazione della protezione di prossimità Zigbee per l'adesione a una rete

L'exploit zero-day è stato il risultato trovato ispezionando il codice sorgente di Atmel per il bootloader sul SOC. Dopo aver esaminato il codice, hanno scoperto che il controllo di prossimità era valido quando si avviava una richiesta di scansione in Zigbee. Se iniziavano con un altro messaggio, il controllo di prossimità veniva ignorato. Ciò ha permesso loro di unirsi a qualsiasi rete.

9.3.3.1 Conclusioni

Un vero attacco potrebbe costringere una lampadina infetta a infettarne altre entro poche centinaia di metri con un carico utile per rimuovere la capacità di aggiornamento del firmware di ciascuna lampadina in modo che non possano mai più essere recuperate.

Le lampadine sarebbero effettivamente sotto un controllo malevolo e dovrebbero essere distrutte.

I ricercatori sono stati in grado di costruire un sistema di attacco completamente automatizzato e di collegarlo a un drone che volava sistematicamente entro la portata delle luci Philips Hue in un ambiente universitario.

