VERTIV™

REPORT VERTIV

Archetipi Edge 2.0 Modelli di infrastruttura Edge pronti per l'implementazione

Progettare un'infrastruttura di edge computing ottimale per il tuo business

Sviluppato con STL Partners

STL PARTNERS

Sintesi

L'infrastruttura fisica è fondamentale in qualsiasi strategia di edge computing. Le apparecchiature di alimentazione, condizionamento e alloggiamento, così come il calcolo che esse supportano, forniscono le basi su cui le applicazioni possono essere eseguite e consentono innumerevoli casi d'uso Edge.

Scegliere la giusta infrastruttura fisica è ancora più importante quando si tratta di Edge, dato che molte implementazioni si trovano in luoghi in cui sono necessari un supporto e una protezione aggiuntivi. Anche la navigazione nell'infrastruttura Edge è resa più complicata dalle ampie e varie definizioni di Edge. Questi fattori creano difficoltà per il 49%[1] delle aziende che considerano l'implementazione dell'edge computing. Esse devono prendere decisioni su come utilizzare al meglio l'infrastruttura esistente e dove investire oggi per sostenere le esigenze di domani. Fortunatamente, esiste un ecosistema di fornitori, system integrator e altri partner di canale con esperienza e competenza nelle implementazioni di Edge pronti a fornire supporto.

Dopo il lavoro di Vertiv sugli archetipi Edge[2], che ha fornito una tassonomia per la categorizzazione dei casi d'uso Edge, questo report fa un ulteriore passo verso la definizione di quattro modelli di infrastruttura Edge distinti. Il framework è stato sviluppato sulla base di interviste con una serie di professionisti del settore, esperti di data center, provider di soluzioni e organismi del settore in tutte le applicazioni relative a Smart City, sanità, produzione e Retail. Con un'analisi approfondita delle esigenze di edge computing di diversi settori e casi d'uso, sono stati definiti i seguenti modelli di infrastruttura di edge computing:

[1] Indagine STL Partners con 699 professionisti a livello globale provenienti dai settori manifatturiero, retail, sanità e trasporti e logistica, maggio 2021
[2] Definizione dei quattro archetipi Edge e dei loro requisiti tecnologici

Device Edge	Micro Edge	Data center Edge distribuito	Data center Edge regionale
• Su dispositivo • Da collegare o integrato • All'esterno (ad es. lampioni) o all'interno (ad es. attrezzature di produzione)	• Numero ridotto di server o rack • 0-4 rack • Presso il sito aziendale (ad es. punto vendita, fabbrica, armadio IT)	• Piccolo data center • 5-20 rack • Sito aziendale (ad es. magazzino), sito di telecomunicazione, parcheggi	• Data center di medie dimensioni • Oltre 20 rack • Sede regionale

Risultati principali

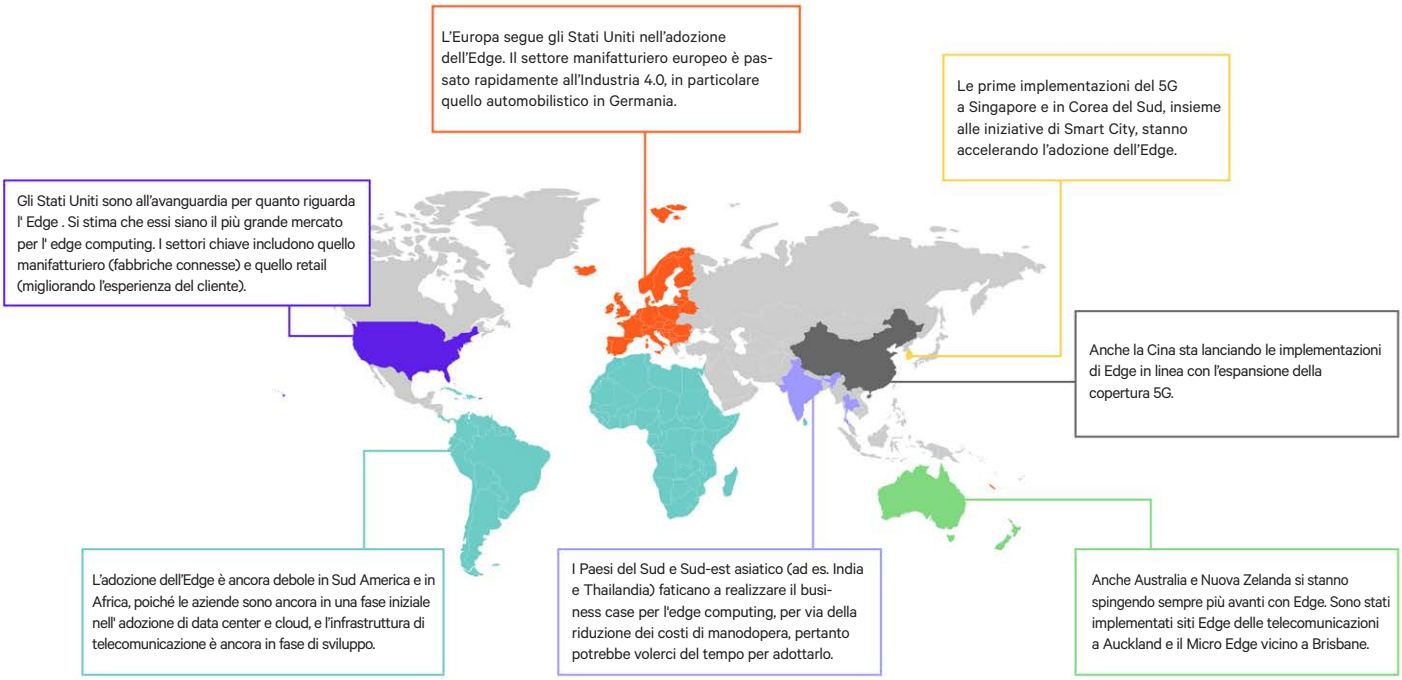
- L'infrastruttura di edge computing non fungerà da sostituto del cloud. Si stima che il numero totale di siti Edge crescerà del 226%[3] dal 2019 al 2025. Allo stesso modo, il cloud continuerà a crescere con un CAGR del 10%[4].
- Gli Stati Uniti sono all'avanguardia nelle iniziative Edge e si stima che rappresentino il mercato più grande di edge computing[5], guidato da settori chiave come quello manifatturiero.
- Le implementazioni di edge computing più sviluppate sono quelle allineate con l'archetipo Edge "Sensibile alla latenza umana" (ad es. giochi su cloud), seguito da "Data Intensive" (ad es. analisi video) e "Sensibili alla latenza da macchina a macchina" (ad es. il trading). I casi d'uso relativi all'archetipo "Life Critical" (ad es. auto a guida automatica) sono ancora in fase di esplorazione o di sperimentazione.
- La maggior parte dei casi d'uso dell'archetipo "Life Critical" utilizzerà il modello di infrastruttura "Device Edge" nel medio termine, mentre i casi d'uso degli archetipi "Data Intensive", "Sensibile alla latenza umana" e "Sensibile alla latenza da macchina a macchina" accelereranno la transizione dai modelli di infrastruttura di data center Edge regionale ai modelli di infrastruttura di data center Edge distribuito nel breve termine.
- Il coordinamento dei numerosi elementi di edge computing (software, hardware, infrastruttura, ecc.) è impegnativo e richiede che un ecosistema di partner supporti il 66% delle aziende che preferiscono avere un'intera soluzione Edge proveniente da un unico fornitore.

[3] Data Center 2025: Closer to the Edge
[4] Technavio, 2021
[5] Mercato edge computing: previsioni globali fino al 2025

Introduzione: lo stato attuale dell'infrastruttura Edge

Vent'anni fa, il pendolo del mercato dei data center era orientato verso il computing centralizzato per migliorare l'efficienza nell'elaborazione dei dati. Oggi il pendolo si sta muovendo dall'altra parte, verso l'edge computing. L'edge computing si riferisce all'elaborazione e all'archiviazione che si trovano tra data center centralizzati e utenti finali, dispositivi o fonti di dati. Da un lato, l'edge computing può essere considerato un'alternativa al cloud e ai data center centrali, quando tali opzioni non sono in grado di soddisfare i requisiti di latenza o quando è troppo costoso trasferire volumi elevati di dati su lunghe distanze. D'altro canto, l'edge computing è anche un fattore trainante per l'adozione del cloud. Un sito Edge può fungere da area di raggruppamento per i dati che alla fine vengono inviati al cloud per l'elaborazione, l'archiviazione o l'analisi a lungo termine.

Negli ultimi due anni, l'adozione dell'edge computing è aumentata significativamente, di pari passo con la continua crescita del cloud. Secondo una recente indagine condotta da STL Partners, il 49% delle aziende di settori specifici sta considerando attivamente l'edge computing[6] e si stima che il numero totale di siti Edge crescerà del 226% dal 2019 al 2025[7]. Tuttavia, l'adozione varia in base alle aree geografiche. Ciò è dovuto al livello di maturità delle tecnologie adiacenti (ad es. intelligenza artificiale), alle infrastrutture di telecomunicazione esistenti, alle politiche governative e alle dimensioni di alcuni settori del Paese. Per esempio, il settore manifatturiero sta guidando l'adozione dell'edge computing negli Stati Uniti e in Germania e si prevede che rappresenti la più grande quota di spesa europea per l'Edge nelle imprese nel 2021[8].



Le imprese vedono l'edge computing come un fattore chiave per superare le sfide legate alla sicurezza e all'affidabilità dei dati, oltre a migliorare le prestazioni delle applicazioni. Ad esempio, grandi cluster di data center potrebbero diventare obiettivi primari per gli attacchi. La suddivisione del core in più siti Edge può costare di più per kW, ma elimina la minaccia di negazione simultanea del servizio. L'edge computing promette inoltre di avvantaggiare un'ampia gamma di settori in una serie diversificata di casi d'uso. Dal cloud gaming alle smart grid per reti di distribuzione elettrica ai robot autonomi in ambito industriale, tutti questi casi d'uso hanno qualcosa da guadagnare dall'elaborazione dei dati più vicino al dispositivo finale.

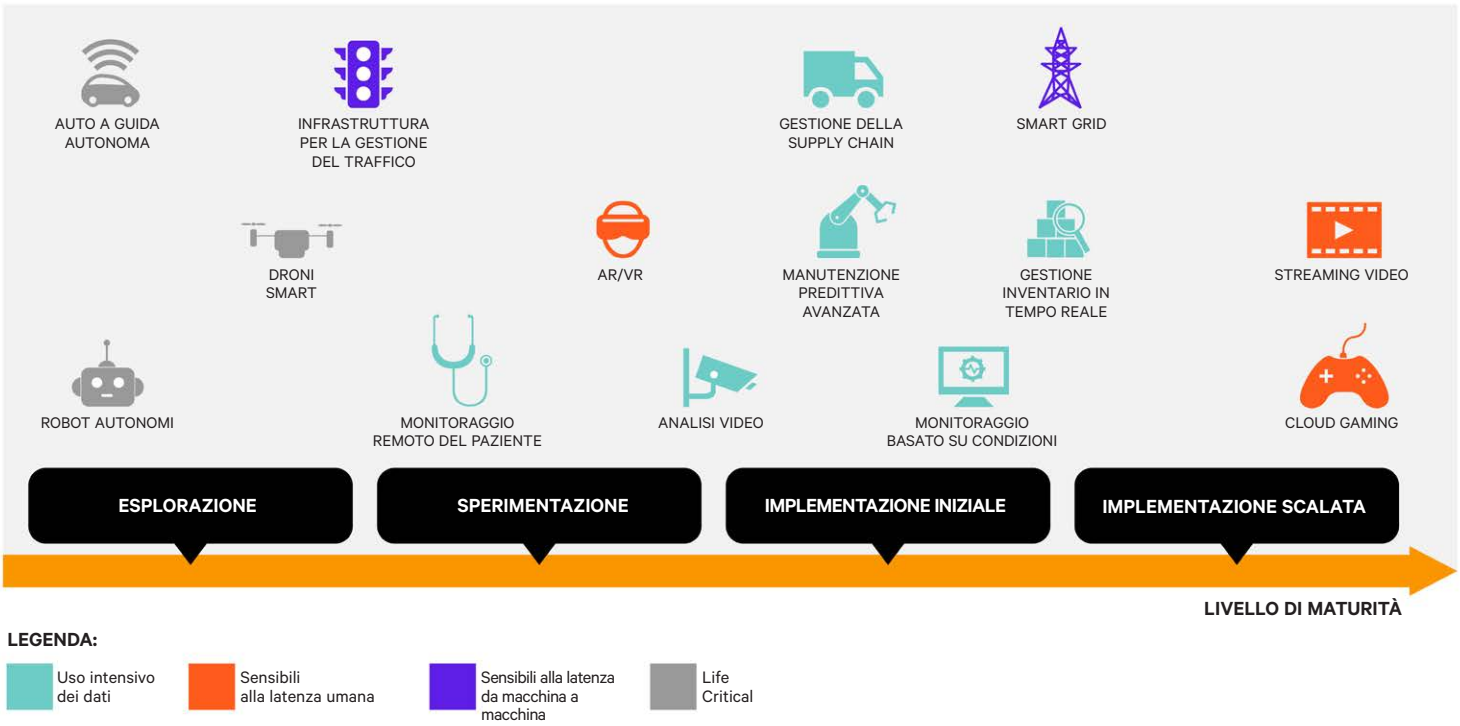
I primi utilizzatori stanno già implementando soluzioni innovative, superando la fase di sperimentazione e portando i progetti pilota iniziali verso implementazioni multi-sito su larga scala. Un esempio è Lloyds Register, una società di servizi marittimi che ha implementato L'edge computing su flotte di navi[9] per ottimizzare il consumo di carburante attraverso l'analisi dei dati. L'adozione dell'edge computing sarà supportata anche da un crescente ecosistema di fornitori, system integrator e altri operatori di canale. La natura distribuita dell'edge computing richiede una rete di operatori Edge con la portata e la funzionalità di implementare, fornire assistenza e supportare l'infrastruttura Edge.

[6] Indagine STL Partners con 699 professionisti a livello globale provenienti dai settori manifatturiero, retail, sanità e trasporti e logistica, maggio 2021
[7] Data Center 2025: Closer to the Edge
[8] Guida alla spesa Edge globale di International Data Corporation (IDC)

Nel 2018, Vertiv ha pubblicato il report, "Defining Four Edge Archetypes and Their Technology Requirements" (Definizione di quattro archetipi Edge e dei loro requisiti tecnologici), che ha fornito un quadro per la categorizzazione dei casi d'uso, il primo del settore. Questi quattro archetipi hanno aiutato le aziende e gli operatori di data center Edge a comprendere meglio i requisiti comuni sottostanti in casi d'uso simili. I quattro archetipi sono:

- **Uso intensivo dei dati:** I casi d'uso in cui la quantità di dati rende non pratico il trasferimento dalla rete direttamente al cloud, oppure dal cloud al punto di utilizzo per via del volume dei dati, del costo o di problemi di larghezza di banda.
- **Sensibili alla latenza umana:** I casi d'uso in cui i servizi sono ottimizzati per il consumo umano o per migliorare l'esperienza umana con i servizi basati sulla tecnologia. La velocità è la caratteristica distintiva di questo caso d'uso, visto che i ritardi nella fornitura dei dati influiscono direttamente sull'esperienza dell'utente.
- **Sensibili alla latenza da macchina a macchina (M2M):** I casi d'uso in cui i servizi sono ottimizzati per il consumo da macchina a macchina. Poiché le macchine sono in grado di elaborare i dati molto più velocemente rispetto agli esseri umani, la velocità è la caratteristica che definisce questo aspetto (e le conseguenze del mancato invio dei dati entro il tempo richiesto, il "budget" può essere superiore rispetto ai casi d'uso sensibili alla latenza umana).
- **Life Critical:** I casi d'uso che influiscono direttamente sulla salute e sulla sicurezza umana. La velocità e l'affidabilità sono quindi di primaria importanza.

Le interviste con esperti del settore dei data center hanno rivelato che gli archetipi variano in base al livello di maturità. I casi d'uso Edge **sensibili alla latenza umana** (ad es. giochi cloud) sono i più maturi e già in grado di raggiungere implementazioni scalate. La crescita del 5G e l'aumento delle implementazioni della fibra accelereranno ulteriormente questa maturità. Al contrario, i casi d'uso **Life Critical** richiederanno molto più tempo per adottare l'edge computing. Questo perché hanno requisiti rigorosi di latenza e affidabilità e spesso richiedono l'implementazione di modifiche normative su larga scala. I droni intelligenti ne sono un esempio. I governi devono essere certi che i droni autonomi non imporranno alcuna minaccia alla vita umana prima di ridurre le restrizioni dello spazio aereo. Allo stesso modo, l'infrastruttura del traffico connessa è ancora in fase iniziale. Soltanto negli Stati Uniti, solo il 7%[10] dei semafori è intelligente.



[9] WWT, 2020: Tre case study reali su come i produttori possono massimizzare l'edge computing
[10] Vertiv Interview Program: citazione dell'intervistato (Director – Esperienza AI, produttore automobilistico).

Passaggio dai casi d'uso all'infrastruttura

Tre anni dopo il lancio del report originale sugli archetipi, il mercato dell'edge computing è ancora in evoluzione e le aziende continuano a sviluppare le proprie soluzioni. I casi d'uso sono passati da applicazioni concettuali ad applicazioni reali implementate sul campo. Tali applicazioni software necessitano di un'infrastruttura adeguata in grado di supportare l'elaborazione dei dati a bassa latenza e larghezza di banda elevata a livello di Edge.

Il termine "infrastruttura Edge" si riferisce all'infrastruttura fisica di elaborazione (server, alimentazione, condizionamento, enclosure) che è deliberatamente posizionata in qualsiasi punto tra il dispositivo finale e i data center centrali. Ciò include anche funzionalità di hosting on-premise, un aspetto ovviamente non nuovo per molte aziende. In realtà, alcuni stanno reinvestendo nell'infrastruttura esistente in loco (ad es. server, armadi di rete o data center) per ottimizzare le applicazioni e implementare nuovi casi d'uso. Ad esempio, un produttore multinazionale di cellulosa e carta[11] sta abilitando applicazioni che richiedono un uso intensivo dei dati, come la manutenzione predittiva avanzata, sfruttando i data center delle sue cartiere più grandi.

Lavorando in base a una definizione rigorosa, l'infrastruttura Edge vera e propria deve utilizzare un'infrastruttura IT standard pronta all'uso ed essere impostata in base ai principi del cloud per ospitare applicazioni e carichi di lavoro nativi del cloud. In questa definizione, l'infrastruttura legacy on-premise monolitica o basata su hardware proprietario non è considerata "edge computing".

Ad oggi, il mercato non ha chiarito ciò che costituisce l'infrastruttura Edge.[12] I clienti aziendali desiderano adottare soluzioni Edge con la certezza che queste soluzioni soddisferanno le esigenze future.
Allo stesso modo, gli operatori di data center Edge oggi devono investire in infrastrutture che supporteranno le applicazioni di domani. Entrambe le parti hanno bisogno di risposte alle domande chiave sull'infrastruttura di edge computing:

- Come si presenta l'Edge in termini di infrastruttura fisica?
- Quali saranno i vantaggi misurabili dell'implementazione dell'IT più vicino alle applicazioni?
- Chi possiederà e gestirà l'infrastruttura di edge computing?
- Come possiamo implementarlo in modo efficace e su larga scala?

In questo documento, esploreremo i fattori chiave che influenzano l'infrastruttura Edge, tra cui il caso d'uso, il settore e l'ambiente esterno. Nell'ambito di questa ricerca, abbiamo condotto 22 interviste con una serie di professionisti del settore, tra cui imprese, esperti di data center, provider di soluzioni e organismi industriali.

Sviluppare il proprio Edge: i quattro modelli di infrastruttura Edge forniscono le fondamenta

Vertiv ha sviluppato un framework innovativo per categorizzare l'infrastruttura Edge in modelli specifici per aiutare le organizzazioni a prendere decisioni pratiche sull'implementazione dell'infrastruttura fisica e dell'elaborazione a livello di Edge. Il termine "infrastruttura" viene utilizzato al posto di data center, in quanto non tutte le implementazioni di Edge possono essere descritte come fattore di forma di un data center di per sé.[13] I modelli aiutano ad allineare la terminologia utilizzata quando si parla di edge computing. Essi comprendono la varietà di implementazioni di Edge viste oggi, nonché l'evoluzione delle stesse prevista nei prossimi anni.

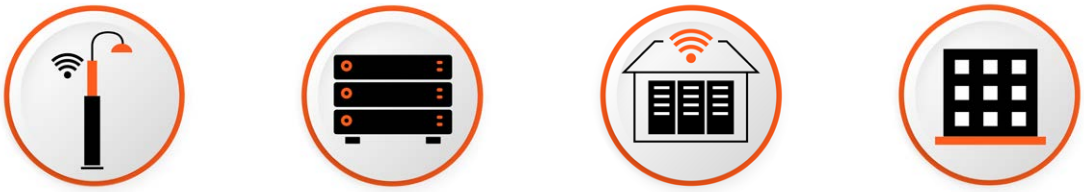
I quattro modelli di infrastruttura Edge sono i seguenti:

- **Device Edge:** L'elaborazione avviene sul dispositivo finale. È integrato nel dispositivo (ad es. una videocamera smart con funzionalità di intelligenza artificiale) o è un fattore di forma indipendente "add-on Edge" che si collega direttamente al dispositivo (ad es. un computer Raspberry Pi collegato a un veicolo a guida automatica). Quando il calcolo è integrato, l'hardware IT è completamente racchiuso all'interno del dispositivo, quindi non deve essere progettato per resistere ad ambienti difficili. Ad esempio, quando il calcolo è collegato all'esterno di una telecamera, deve essere irrobustito, ma se è integrato nella telecamera, si trova all'interno di un ambiente controllato, quindi l'irrobustimento non è necessario.

- **Micro Edge:** Una soluzione piccola e autonoma che può essere di dimensioni variabili da uno o due server fino a quattro rack. Viene spesso impiegato presso la sede di un'azienda (ad es. per un produttore, potrebbe trovarsi nel reparto di produzione della fabbrica o nel back office). Può anche trovarsi presso un sito di telecomunicazioni (ad es. un rack di server situato presso una stazione base di telecomunicazioni). Il Micro Edge può essere implementato sia in ambienti condizionati che non condizionati. Negli ambienti condizionati (ad es. armadio IT), il Micro Edge non richiede il condizionamento e la filtrazione avanzati, poiché fattori esterni come temperatura e qualità dell'aria sono stabili. Negli ambienti non condizionati (ad es. il reparto di produzione di una fabbrica), il calcolo è irrobustito e il Micro Edge richiede un condizionamento e una filtrazione specializzati per tenere conto dei fattori esterni più aggressivi (ad es. temperature elevate e polvere).

[11] Intervistato dal programma di ricerca Vertiv, 2021
[12] L'infrastruttura di edge computing si riferisce allo stack Edge informatico e alle strutture fisiche che lo supportano (ad es. alimentazione, condizionamento, sicurezza, alloggiamenti).
[13] Un tipico ambiente di data center include normalmente: connessione in fibra, UPS, raffreddamento, sicurezza, cablaggio, pavimento rialzato.

- **Data center Edge distribuito:** Un piccolo data center con meno di 20 rack che si trova presso il sito aziendale, le strutture di rete delle telecomunicazioni o in un sito regionale (ad es. in fabbriche moderne o grandi proprietà commerciali).
- **Data center Edge regionale:** Una struttura di data center situata al di fuori degli hub di data center principali. Poiché si tratta in genere di una struttura costruita appositamente per ospitare l'infrastruttura di elaborazione, condivide molte caratteristiche dei data center hyperscale (ad es. è condizionata e controllata, ha un'elevata sicurezza e un'elevata affidabilità).

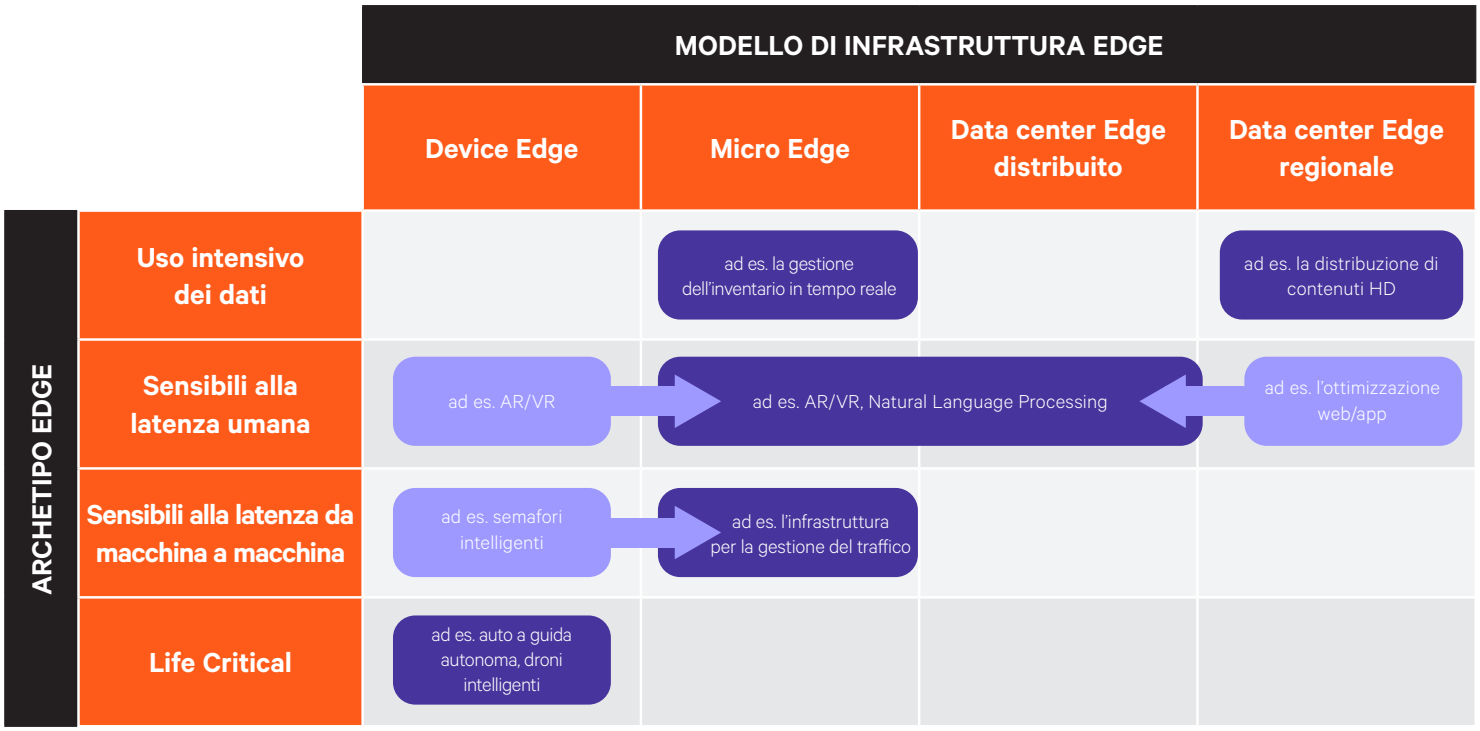


		MODELLO DI INFRASTRUTTURA EDGE			
		Device Edge	**Micro Edge**	**Data center Edge distribuito**	**Data center Edge regionale**
CARATTERISTICHE	**Posizionamento**	Dispositivi Smart (ad es. veicolo, lampione stradale, IoT)	Sito aziendale (ad es. retail, reparto produzione, armadio IT, Comuni)	Sito aziendale (ad es. magazzino, ufficio), sito di telecomunicazioni, parcheggio, livello 2/3 della città	Livello 2/3 della città[14]
	Numero di rack	0	0-4 rack	5-20 rack	Oltre 20 rack
	Potenza	Fino a 1 kW	Fino a 20 kW	Fino a 200 kW	Fino a 4000 kW
	Tenancy	Tenant singolo	Tenant singolo	Tenant singolo/Multi-tenant	Multi-tenant
	Ambiente esterno	Controllato (all'interno del dispositivo), irrobustito e resistente	Armadio IT, commerciale e ufficio, resistente e irrobustito	Resistente e irrobustito, commerciale e ufficio, condizionato e controllato	Condizionato e controllato
	Infrastruttura passiva	Può avere o meno potenza e filtrazione, assenza di condizionamento, ecc.	Ha potenza con condizionamento e filtrazione limitati, ecc.	Livello 1 e oltre	Livello 3 e oltre
	Provider di infrastruttura Edge	Produttore di dispositivi o soluzioni all'interno di aziende/governi	OEM di hardware, provider di data center, operatore di telecomunicazioni o soluzione all'interno di aziende/governi	Provider di Colocation, provider di cloud hyperscale (cloud pubblico), operatore di telecomunicazioni	Provider di Colocation, provider di cloud hyperscale (cloud pubblico)
	Implementazioni previste	Milioni	Centinaia di migliaia	Migliaia	Centinaia

** entro il 2030 per regione principale*

L'identificazione del modello di infrastruttura Edge appropriato dipende dal caso d'uso che si sta implementando. Poiché casi d'uso simili spesso hanno requisiti simili, può essere utile iniziare identificando l'archetipo Edge.

In genere, minore è la latenza richiesta, più vicina deve essere l'infrastruttura Edge al dispositivo finale. Per questo motivo, i casi d'uso **Life Critical** spesso devono essere ospitati nel **Device Edge**, mentre i casi d'uso **Data Intensive** sono spesso ospitati on-premise nel **Micro Edge**.



LEGENDA:

Modello di infrastruttura che viene generalmente implementato oggi

Modello di infrastruttura che secondo noi sarà più implementato in futuro

- **Uso intensivo dei dati:** Poiché i casi d'uso Data Intensive richiedono che l'Edge sia vicino all'origine dei dati per prevenire costi elevati della larghezza di banda, sono auspicabili implementazioni on-premise. Un Micro Edge fornisce un buon equilibrio tra la distanza di trasmissione dei dati breve (limitando quindi i costi di banda) e le maggiori funzionalità di elaborazione rispetto a un Device Edge.
- **Sensibili alla latenza umana:** L'archetipo sensibile alla latenza umana è dominato dalle applicazioni consumer (ad es. l'ottimizzazione della velocità del sito Web[15]) per le quali una soluzione Edge on-premise non è un'opzione. Per questo motivo, la maggior parte dei casi d'uso sensibili alla latenza umana oggi è ospitata presso i data center Edge regionali. Tuttavia, man mano che le esigenze di latenza si spostano nel range inferiore a 10 millisecondi e i data center Edge diventano più disponibili nelle posizioni di accesso[16], i data center Edge distribuiti saranno un'opzione favorevole. Le applicazioni sensibili alla latenza umana per le aziende (ad es. AR/VR) sono generalmente ospitate oggi sul Device Edge per soddisfare i requisiti di latenza, ma passeranno a Micro Edge on-premise poiché sono sempre più implementate dalle aziende.
- **Sensibili alla latenza da macchina a macchina:** Le macchine sono in grado di elaborare i dati molto più velocemente rispetto agli esseri umani, pertanto la velocità è il requisito fondamentale delle applicazioni sensibili alla latenza da macchina a macchina. Il Device Edge soddisfa questi requisiti di latenza, ma ci sarà un passaggio al Micro Edge man mano che l'adozione dell'Edge aziendale diventa più diffusa, in particolare per i dispositivi da macchina a macchina che sono troppo piccoli o a basso costo per giustificare un Device Edge. Ad esempio, nel settore manifatturiero, i fornitori stanno mettendo il computing nel reparto produzione. Un piccolo dispositivo Edge in un alloggiamento self-contained con alimentazione e condizionamento integrati.
- **Life Critical:** La bassa latenza è fondamentale per questi casi d'uso, in quanto influiscono direttamente sulla salute e sulla sicurezza degli esseri umani. Il Device Edge fornisce la latenza più bassa; pertanto, molti casi d'uso Life Critical si basano su questo modello.

[14] Le città di livello 2 e 3 hanno spesso una popolazione inferiore a 1 milione e raramente avranno ancora un punto di interscambio/peering del traffico Internet al loro interno né un data center hyperscale. Tra gli esempi vi sono Austin negli Stati Uniti o Berlino e Milano in Europa.
[15] L'ottimizzazione della velocità dei siti Web utilizza edge computing per ridurre il tempo di caricamento di una pagina. Molti provider di e-commerce hanno avuto un impatto negativo sulle entrate quando i siti sono più lenti, e Google ha osservato che un ritardo di 500 millisecondi nella risposta alla paginazione ha comportato una diminuzione del 20% del traffico.
[16] Quando l'Edge si trova nelle posizioni di accesso, si trova in siti o punti di presenza di proprietà di un operatore di telecomunicazioni (ad es. celle radio, uffici centrali o nodo di un ISP). LF Edge ha stabilito l'Edge di accesso nel suo Edge Continuum.

In pratica, le aziende prendono in considerazione altri fattori insieme ai loro requisiti di casi d'uso quando prendono decisioni sull'infrastruttura. Queste importanti considerazioni includono:

- **Ambiente:** Temperatura, inquinamento e presenza di particolati hanno tutti un impatto sull'infrastruttura richiesta (ad es. il grado di condizionamento e filtrazione). Il rumore prodotto, compreso quello elettrico, deve essere preso in considerazione soprattutto se lo spazio raddoppia come ufficio. Ad esempio, i cavi di comunicazione non possono essere fatti passare vicino alla tromba dell'ascensore.
- **Caso d'uso:** La quantità e la velocità con cui i dati devono essere elaborati influenzano la vicinanza al dispositivo finale in cui deve avvenire il calcolo. Anche il tipo di carico di lavoro (ossia, intensivo in termini di elaborazione rispetto a intensivo in termini di archiviazione) influisce sull'infrastruttura Edge, in quanto i carichi di lavoro più intensivi in termini di calcolo (ad es. video ad alta definizione) richiedono più potenza e quindi un maggiore condizionamento.

Il 5G accelererà l'adozione dell' Edge

Il 5G sarà un fattore significativo nella determinazione dell'adozione dell'Edge, poiché l'implementazione del 5G funge da catalizzatore per il passaggio all'Edge. Le aree geografiche che sono più avanti con il lancio del 5G (America del Nord, Europa e Asia Orientale) saranno quindi in prima linea nell'adozione dell'Edge. Per saperne di più su come i casi d'uso Edge trarranno vantaggio dal 5G, consulta la precedente ricerca di Vertiv.

> Il 5G sta iniziando ora e il suo consolidamento richiederà 3-5 anni sui grandi mercati sviluppati. Riteniamo che ciò accelererà il percorso verso l'Edge.
>
> **VP Innovation,**
> **Tower company leader del settore**

> È una sfida perché questi uffici non sono mai stati pensati per ospitare attrezzature informatiche, quindi dobbiamo andare lì e aggiornare l'impianto elettrico. Ora produciamo calore nello spazio, quindi dobbiamo parlare di raffreddamento. Soprattutto se si tratta di uno spazio in cui le persone stanno lavorando, non vogliamo rendere l'ambiente troppo rumoroso o surriscaldarlo.
>
> **Technical Solutions Architect,**
> **World Wide Technology**

- **Attrezzature/infrastrutture legacy:** La decisione di implementare un'infrastruttura Edge in un data center esistente, rispetto alla creazione di una nuova implementazione autonoma, dipende in ultima analisi dalla presenza di un data center legacy esistente. Per un Micro Edge, la forma specifica dell'infrastruttura è guidata dallo spazio in cui deve essere inserita (ad es. se lo spazio è insufficiente, l'infrastruttura deve essere montata a parete).
- **Operazioni aziendali:** La scelta tra l'aggiornamento di un data center on-premise esistente e l'introduzione di una nuova implementazione autonoma dipende anche dal fatto che l'azienda possa permettersi i tempi di inattività necessari per aggiornare l'infrastruttura esistente. Le aziende per le quali i tempi di inattività sono costosi possono trarre vantaggio dal pagamento di un premio per un data center prefabbricato, che può essere costruito fuori sede e quindi implementato rapidamente.
- **Sicurezza e manutenzione:** Se l'infrastruttura Edge si trova in un luogo esposto dove le persone potrebbero danneggiarla, l'alloggiamento deve essere progettato con ulteriore sicurezza. Se i dipendenti dovranno mantenere o interagire regolarmente con l'attrezzatura informatica, questa deve essere facilmente accessibile (ad es. non fuori portata sul soffitto).
- **Infrastruttura delle comunicazioni:** Se l'Edge si trova in una posizione remota e l'infrastruttura non è presente per trasportare i dati sulla rete (ad es. nei settori minerario e agricolo), è necessaria una soluzione on-premise più robusta.

Navigazione nei modelli di infrastruttura Edge: Raccomandazioni principali

Device Edge

I casi d'uso che sfruttano un Device Edge includono quelli che rientrano nell'archetipo Life Critical, come droni, veicoli autonomi, chirurgia robotica e monitoraggio dei pazienti in ospedale. Il Device Edge è adatto perché può soddisfare i requisiti di mobilità di un dispositivo, come un drone, per navigare autonomamente nel contesto dell'ambiente in cui si trova. Esso fornisce inoltre una latenza estremamente bassa, necessaria per i casi d'uso Life Critical. Infine, consente ad alcuni aspetti del caso d'uso di funzionare (ad es. navigazione, allarmi locali) anche quando la connettività non è disponibile a causa di una copertura limitata o di un guasto della rete.

Di conseguenza, l'assistenza sanitaria è uno dei settori con un elevato tasso di adozione di Device Edge, in quanto molti casi d'uso dovranno rilevare situazioni potenzialmente letali in modo rapido e affidabile, sia che il paziente si trovi in ospedale sia che venga assistito da remoto. Il settore manifatturiero dispone anche di casi d'uso Life Critical, ed è per questo che i sistemi di controllo delle macchine funzionano principalmente sulle apparecchiature stesse (una forma di Device Edge).

Raccomandazioni principali per la distribuzione di un Device Edge:

- Il componente aggiuntivo Device Edge è più appropriato per il retrofitting di apparecchiature legacy, ma le implementazioni in un ambiente totalmente nuovo potrebbero considerare l'inclusione del calcolo nel dispositivo. Tuttavia, questi sono spesso dispositivi proprietari che non si prestano all'integrazione con funzionalità di edge computing generiche.
- La potenza di elaborazione di Device Edge è limitata. L'aggiunta di più calcoli renderà i dispositivi finali molto più pesanti, quindi considera sempre il compromesso potenza/peso[17] (che è fonte di maggiore preoccupazione nei casi in cui il dispositivo è alimentato a batteria o non ha accesso a un'alimentazione elettrica).
- Fai attenzione ai dati raccolti dal dispositivo finale. I casi d'uso, come le telecamere di sicurezza Smart, le infrastrutture di traffico connesse e i droni, raccolgono dati visivi o di localizzazione relativi alle persone. È quindi importante essere consapevoli delle potenziali sfide relative alla privacy e alla condivisione dei dati, poiché questo potrebbe essere un problema controverso.

[17] Il compromesso tra potenza, peso e costo per le cuffie AR/VR viene esplorato in Apple Glass: Un momento iPhone per il 5G?

Micro Edge

- Se le decisioni relative a software, hardware e infrastruttura vengono prese da diversi soggetti interessati, mantenere l'allineamento tra questi soggetti in modo che le decisioni vengano prese in parallelo, non in sequenza, in quanto ciò si traduce in una soluzione più efficace.
- Seleziona il tipo di apparecchiatura. Le apparecchiature temprate sono realizzate per ambienti meno controllati, quindi possono resistere a 50 gradi Celsius. Le aziende possono invece utilizzare server generici e commerciali pronti all'uso (COTS[18]), che sono più economici, ma la durata di conservazione di questi server è notevolmente ridotta quando operano a temperature superiori a 30 °C. Sebbene entrambi i tipi di hardware richiedano un alloggiamento, l'infrastruttura di supporto per i server COTS deve offrire un maggiore controllo di temperatura, umidità e potenza. È necessario un equilibrio economico tra standardizzazione e personalizzazione specifica per sede.

[18] COTS: prodotti commerciali pronti all'uso che sono facilmente disponibili per la vendita e progettati per integrarsi facilmente con i sistemi esistenti (piuttosto che essere personalizzati o su misura).

Un Micro Edge può essere posizionato vicino alla fonte dei dati grazie alle sue dimensioni ridotte e alla relativa facilità di implementazione (rispetto a un data center più grande). Per questo offre una bassa latenza e riduce il costo della trasmissione dei dati, diventando un modello di infrastruttura adatto per i casi d'uso nei seguenti tre archetipi: uso intensivo di dati, sensibile alla latenza umana e sensibile alla latenza da macchina a macchina. In settori con vincoli di spazio, come il retail o l'istruzione, Micro Edge è una soluzione interessante, in quanto limita l'immobiliare richiesto, consentendo l'implementazione del calcolo in un ingombro a pavimento ridotto. Ad esempio, una grande catena di supermercati con 16.000 sedi in Europa sta implementando un Micro Edge nei negozi per la raccolta e l'elaborazione di dati locali e sta anche aggiungendo data center centrali per l'aggregazione e la gestione IT generale.

Raccomandazioni principali per l'implementazione di Micro Edge:

- Considerare lo spazio disponibile (potrebbe essere necessario fissare alle pareti o al soffitto), la funzione dello spazio (se saranno presenti clienti o dipendenti) e i requisiti di sicurezza (quando l'infrastruttura è facilmente accessibile, è necessario un livello fisico di sicurezza). Le implementazioni di Micro Edge spesso coprono aree con diversità di alimentazioni elettriche, normative, accesso al sito (ad es. altezza dell'ascensore), controllo del sito (Store Manager, Plant Manager) e competenze tecniche.

> L'infrastruttura fisica e virtuale deve essere coordinata insieme, altrimenti semplicemente non funzionerà.
>
> **Jon Abbott, Technologies Director, Vertiv**

Data center Edge distribuito

Come un Micro Edge, i data center Edge distribuiti si trovano presso la sede aziendale e sono adatti a molti casi d'uso del settore, in quanto offrono bassa latenza e costi di banda ridotti. La ricerca ha scoperto che le società di telecomunicazioni utilizzano data center Edge distribuiti per ospitare sia le applicazioni consumer che le funzioni di rete interne, che sono sensibili alla latenza da macchina a macchina. Allo stesso modo, i produttori di medie e grandi dimensioni utilizzeranno questi data center più piccoli per i loro casi d'uso di Internet of Things (IoT). Per gli impianti di produzione di medie dimensioni, la maggior parte dell'infrastruttura Edge risiederà all'interno di un data center a otto rack.

Raccomandazioni fondamentali per l'implementazione di un data center Edge distribuito:

- Potrebbe essere necessario un investimento per aggiornare un data center o una sala di rete esistente e il tempo necessario per implementare le modifiche potrebbe avere un costoso impatto sulle operazioni. Questo impatto sui tempi di inattività deve essere ponderato rispetto al costo di acquisto di un nuovo data center prefabbricato, che può essere implementato rapidamente in sede.
- Si consiglia di creare un backup nel data center per mantenere la flessibilità in futuro, ma occorre notare che il sovradimensionamento per prepararsi a tutti i risultati è costoso e potrebbe non essere necessario. Per trovare l'equilibrio tra ciò che è necessario oggi e ciò che sarà richiesto domani, gli utenti devono considerare l'evoluzione del loro caso d'uso Edge nel loro settore.
- Quando si crea ridondanza nel data center, considerare sia il valore delle applicazioni in esecuzione sia la stabilità dell'ambiente (ad esempio, in alcuni Paesi la rete non è affidabile, quindi il rischio di perdita di potenza è significativo).
- A volte non è necessario implementare un data center Edge distribuito presso la sede aziendale, poiché un'implementazione "near-premise" soddisfa requisiti come quelli relativi alla latenza o alla sicurezza. Questo potrebbe essere ancora di proprietà dell'azienda o potrebbe trattarsi di una struttura multi-tenant che serve più aziende.
- Se un data center Edge distribuito viene utilizzato come struttura di Colocation, deve disporre di livelli di sicurezza e isolamento per fornire questa esperienza di edge computing multi-tenant. Ciò può includere cancelli, serrature e telecamere.

Data center Edge regionale

Un data center Edge regionale agisce come sito di edge computing o come sito intermediario in cui i dati Edge vengono inviati per la pre-elaborazione prima di essere inviati al cloud. Esso soddisfa sia i casi d'uso a bassa latenza che quelli che richiedono un uso intensivo dei dati, pertanto tutti gli archetipi Edge sfruttano i data center Edge regionali. I casi d'uso sensibili alla latenza umana per i consumatori, in particolare, si basano su data center Edge regionali, poiché le implementazioni di Edge on-premise (ad es. streaming multimediale a bassa latenza o giochi immersivi) non sono un'opzione.

I data center Edge regionali sono spesso adottati nel settore retail, in quanto possono ridurre la necessità di implementare l'infrastruttura di elaborazione tra i singoli punti vendita al dettaglio.[19] Nei casi in cui il retailer abbia investito in implementazioni on-premise individuali, il data center può fungere da sito di elaborazione dati intermediario.

Raccomandazioni fondamentali per l'implementazione di un data center Edge regionale:

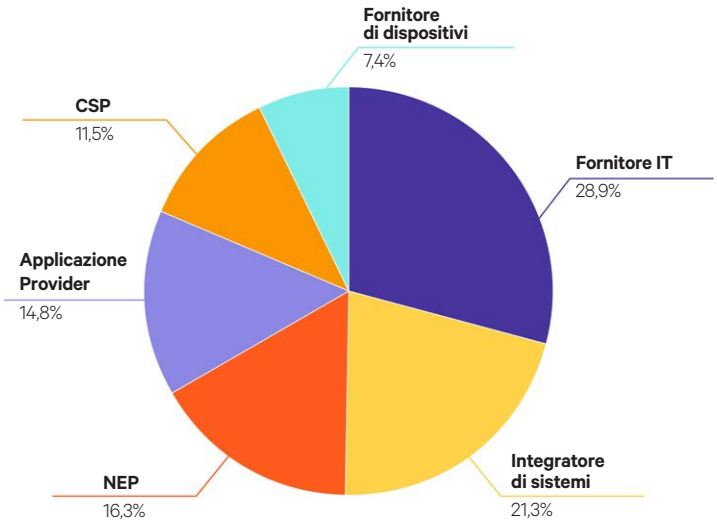
- La sicurezza e l'isolamento sono elementi essenziali (poiché molti data center regionali sono strutture multi-tenant). I clienti devono garantire che il data center disponga di meccanismi adeguati per proteggere l'infrastruttura e i dati del tenant.
- Considera il caso d'uso specifico quando progetti l'infrastruttura di edge computing (ad es. i carichi di lavoro con calcolo più intensivo richiederanno probabilmente più alimentazione e quindi più raffreddamento).
- La posizione è una considerazione chiave. Se la sovranità dei dati è un fattore, potrebbe essere necessario archiviare i dati all'interno della giurisdizione dei clienti finali. Tuttavia, se il fattore chiave è la latenza (< 50 millisecondi), individua una posizione strategicamente importante che riduca la latenza su più siti finali possibile. Questo sarà spesso un data center che si trova presso o in prossimità di un importante interscambio Internet.
- I principali provider di cloud pubblici stanno estendendo il loro servizio ai data center locali (ad es. AWS Local Zones), che consentiranno alle aziende di distribuire le loro applicazioni cloud più facilmente. Tuttavia, ci sono due considerazioni importanti: i provider di cloud pubblici sono in una fase iniziale di distribuzione di questi cloud locali e alcune applicazioni (e i dati) non saranno adatte per l'archiviazione e l'elaborazione su un cloud pubblico (in parte, a causa delle normative governative).

[19] Secondo la Guida alla spesa Edge internazionale di International Data Corporation (IDC), il retail è il secondo settore più grande e a crescita più rapida nel mercato Edge delle imprese europee.

È necessario un ecosistema per sviluppare l'Edge

L'infrastruttura è solo una parte del puzzle per qualsiasi organizzazione che desideri implementare soluzioni abilitate all'Edge. Ci sono molti elementi che influenzano la costruzione dell'Edge (software, hardware, infrastruttura, orchestrazione, gestione, ecc.) e le aziende avranno difficoltà a coordinare questi elementi da sole.

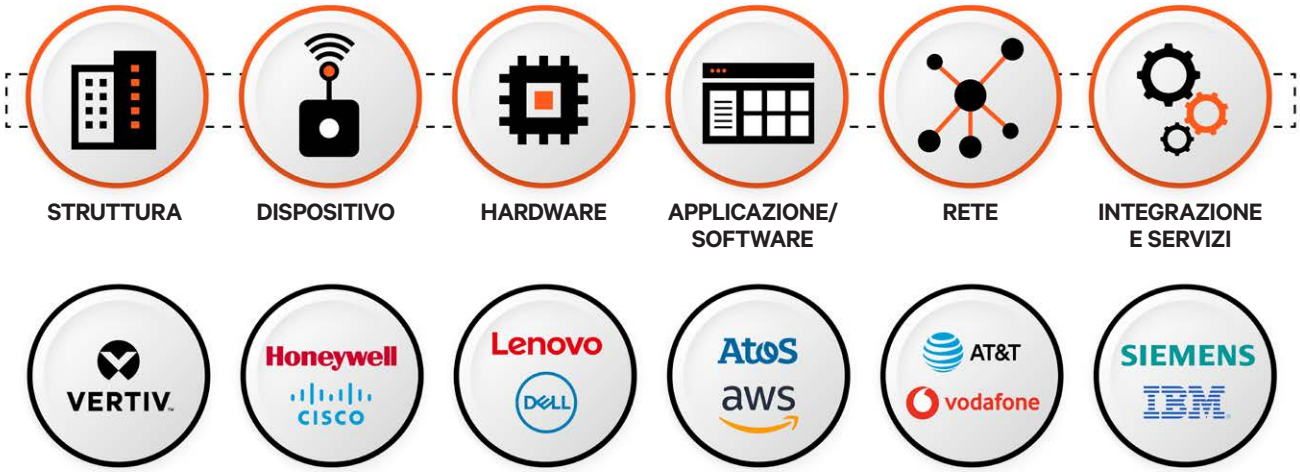
Non esiste una soluzione unica per tutte le aziende per quanto riguarda il modo in cui desiderano acquistare queste soluzioni di prossima generazione basate su tecnologia informatica e comunicazione (ICT) ed Edge. Le ricerche dimostrano che il 34% delle aziende preferisce un approccio fai da te in cui selezionano diversi componenti da ciascun fornitore. Il 66% che preferisce che intere soluzioni provengano da un fornitore principale varia in base a chi è il fornitore principale: fornitore IT rispetto a system integrator rispetto a provider di apparecchiature di rete, ecc.



Fonte: Indagine di STL Partners con 699 professionisti del settore a livello globale, maggio 2021

Sia che si tratti di un'azienda che costruisce il proprio Edge o di un service provider che implementa un'infrastruttura Edge per eseguire applicazioni o che consente ad altri di eseguire carichi di lavoro, la collaborazione con altri attori nell'ecosistema di edge computing è essenziale per il successo. La costruzione di solide relazioni con specialisti del settore (ad es. Siemens o Honeywell nella produzione) garantisce che le soluzioni soddisfino le esigenze specifiche verticali e possano integrarsi con successo con i sistemi e le infrastrutture esistenti.

La catena di valore dell'edge computing:



Conclusioni e raccomandazioni

Le infrastrutture Edge rimangono un argomento complicato (come dimostrano le interviste con i professionisti del settore). Ciononostante, il modello di infrastruttura Edge definito in questo report può aiutare le aziende a navigare nella gamma di soluzioni Edge disponibili e fornire indicazioni sulle scelte di infrastruttura appropriate.

Guardando oltre i modelli di infrastruttura Edge, Vertiv riconosce che ci sono complessità associate al compito pratico di costruire un'infrastruttura Edge che sia unica per ogni azienda. È stato sviluppato uno strumento Web interattivo per consentire alle aziende e ad altri operatori di data center di esplorare in modo approfondito i principali casi d'uso. Le organizzazioni saranno in grado di comprendere meglio il carico di lavoro associato e le caratteristiche dell'infrastruttura e di prendere decisioni informate sulla progettazione, sulla realizzazione e sull'implementazione dell'infrastruttura.

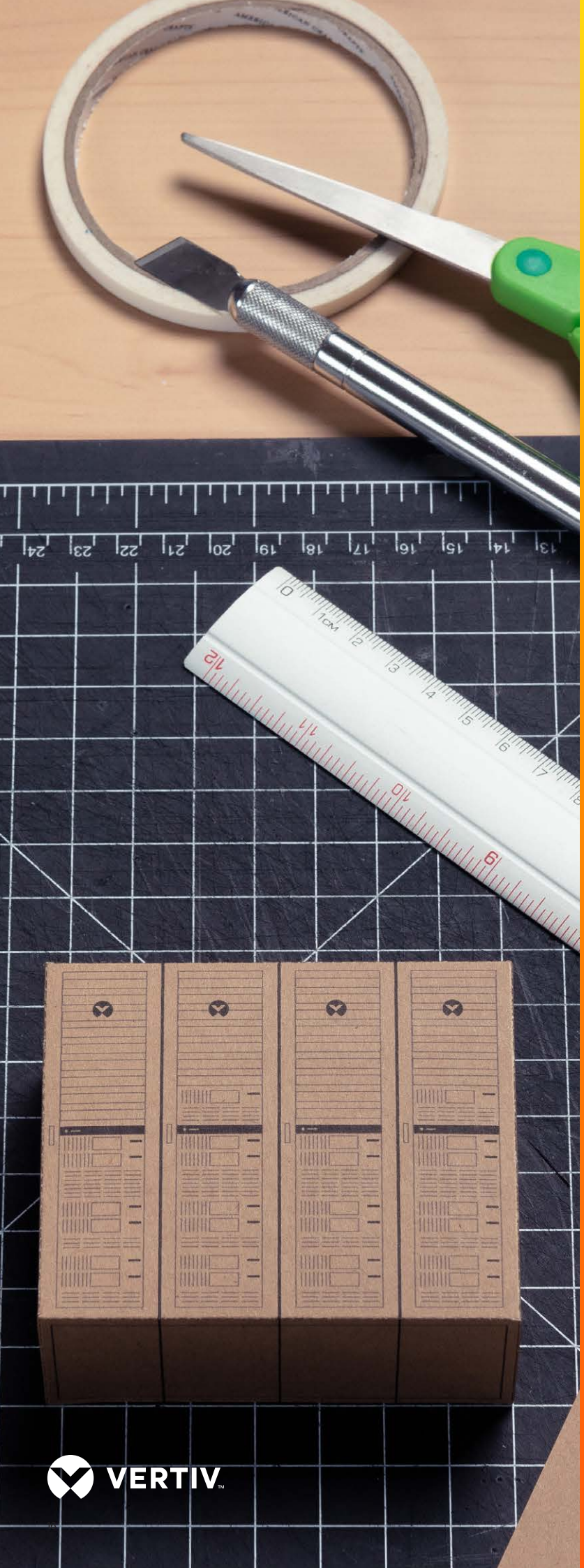
Altre raccomandazioni fondamentali includono:

Aziende

- **Identificazione di un caso d'uso di ancoraggio.** C'è ancora incertezza sulla natura dei casi d'uso futuri e sulle loro esigenze di edge computing. Il primo caso d'uso deve giustificare il business case per uno sviluppo iniziale, quindi assicurati di avere una buona comprensione del motivo per cui il tuo caso d'uso necessita di edge computing. Anche comprendere quali caratteristiche del carico di lavoro sono il fattore chiave per l'implementazione di Edge (ad es. latenza, larghezza di banda, sicurezza) aiuterà a prendere decisioni informate sull'infrastruttura.
- **Preparati a gestire una varietà di modelli di infrastruttura Edge**. Ad esempio, molti retailer optano per un Micro Edge nei loro negozi, quindi integrano con un data center Edge distribuito, vicino ai punti vendita, che può filtrare e aggregare i dati da più sedi, inviando solo le informazioni necessarie al cloud.
- **Cerca di non definire un singolo piano per tutti gli scenari.** Anche all'interno dei tipi di modello vi saranno delle variazioni, dato che sedi diverse hanno ambienti legacy diversi. Le aziende con copertura multinazionale affronteranno le differenze geografiche in termini di clima, inquinamento, alimentazione elettrica, normative, ecc. (ad es. l'UE regola il numero di decibel consentiti che potrebbero limitare l'infrastruttura con ventilatori o switch).

Solution Providers

- **Infrastruttura Edge a prova di obsolescenza.** Comprendi i casi d'uso che i clienti adottano ora e pianifica di adottare in futuro e integrare la potenza di riserva (archiviazione, computing, ecc.) in modo appropriato. L'adozione di modelli di implementazione più flessibili ridurrà i rischi.
- **Lavora con l'ecosistema.** Edge non è un singolo prodotto che deve essere venduto da un unico fornitore, ma una soluzione che più operatori dell'ecosistema creano insieme. Le soluzioni devono essere standardizzate, quindi è facile per i clienti utilizzare una soluzione come componente. Anche la collaborazione è importante, in particolare quando si cerca di soddisfare esigenze altamente specifiche del settore.
- **Considera nuovi modelli economici.** Non è possibile ripetere quanto è stato fatto con il cloud. L'infrastruttura Edge ha esigenze specifiche, quindi è importante considerare i modelli economici che garantiscono che alimentazione, condizionamento, sicurezza e spazio siano ottimizzati con nuovi modi di garantire economie di scala.



Appendice: Glossario

ACCESS EDGE	Una location di Edge all'interno della rete delle telecomunicazioni, che connette gli abbonati alla dorsale dell'operatore principale e poi ad altre reti, a Internet e agli hyperscaler cloud.
STRUTTURA O SERVIZIO DI COLOCATION	Una struttura di Colocation, o "Colo", è una struttura di data center in cui un'azienda può affittare uno spazio per server e altri hardware informatici. In genere, un Colo fornisce l'edificio, il condizionamento, l'alimentazione, la connettività ad altri utenti o a Internet e la sicurezza fisica, mentre il cliente fornisce server e archiviazione.
AMBIENTE CONDIZIONATO E CONTROLLATO	Ambienti con sistemi dedicati per il controllo di vari fattori, tra cui temperatura e umidità, particelle di polvere, inquinamento, ecc.
DATA CENTER	Una struttura fisica che le organizzazioni utilizzano per ospitare le applicazioni e i dati critici. Il design di un data center si basa su una rete di risorse di computing e archiviazione, che consentono la distribuzione di applicazioni e dati condivisi. I componenti chiave di un design di data center includono router, switch, firewall, sistemi di archiviazione, server e controller di distribuzione delle applicazioni.
EDGE COMPUTING	Questa infrastruttura di calcolo fisico è posizionata tra il dispositivo e il cloud hyperscale, supportando vari carichi di lavoro. L'edge computing avvicina le funzionalità di elaborazione all'utente finale/dispositivo/origine dei dati, eliminando il passaggio ai data center dei provider cloud e riducendo la latenza.
FATTORE DI FORMA	Design e funzionalità generali dei sistemi hardware.
HYPERSCALE	Nel computing, l'hyperscale è la capacità di raggiungere una scalabilità massiccia, specialmente per i big data e il Cloud computing. Oggi, AWS, Azure e Google Cloud sono considerati "hyperscaler".
ARMADIO IT/DI RETE	Un armadio o una piccola sala in cui sono installati i cavi elettrici e l'hardware di rete del computer.
EDGE COMPUTING MULTI-ACCESSO (MEC)	Tipo di architettura di rete che fornisce funzionalità di Cloud computing e un ambiente di servizi IT nell'Edge della rete.
ON-PREMISE	Conosciuto anche con i termini "on-premises" o "on-prem", si tratta della tecnologia ospitata all'interno dei confini fisici del sito di un'azienda.
MANUTENZIONE PREDITTIVA	Processo di monitoraggio dei dati dai sensori delle apparecchiature per garantire che siano in buone condizioni e per segnalare preventivamente se è necessario ripararli, eliminando eventualmente la necessità di manutenzione programmata.
HARDWARE IRROBUSTITO	Hardware progettato specificamente per resistere ad ambienti difficili, come inquinamento esterno, temperature alte o basse, umidità, ecc.
STAND ALONE	In grado di funzionare indipendentemente da altro hardware o software.
STAZIONE BASE TELCO	Stazione di trasmissione e ricezione in una posizione fissa, costituita da una o più antenne di ricezione/trasmissione, antenna a microonde e circuiteria elettronica, utilizzata per gestire il traffico dei cellulari.

PARTNERS Questo report di ricerca è stato sviluppato con il supporto di STL Partners

Vertiv.com | Vertiv S.r.l., Via Leonardo da Vinci 16-18, 35028 Piove di Sacco (PD), Italia

© 2021 Vertiv Group Corp. Tutti i diritti riservati. Vertiv™ e il relativo logo sono marchi o marchi registrati di Vertiv Group Corp. Tutti gli altri nomi e loghi menzionati sono nomi commerciali, marchi o marchi registrati dei rispettivi titolari. Anche se sono state adottate tutte le precauzioni per garantire la precisione e la completezza di questa documentazione, Vertiv Group Corp. declina qualsiasi responsabilità in generale e in relazione a eventuali danni risultanti dall'uso di queste informazioni, oltre che in relazione a eventuali errori e omissioni. Le specifiche, i rebate e le altre offerte promozionali sono soggetti a modifiche a sola discrezione di Vertiv, previa notifica.

SL-70825 (R10/21)