
Consiglio Nazionale delle Ricerche
Istituto di Calcolo e Reti ad Alte Prestazioni

# Data Center Benchmarking: Stato dell'Arte e Best Practices

*Giovanni Massafra, Emanuele Damiano, Pier Giuseppe Meo, Mario Sicuranza*

***Rapporto Tecnico N:*** RT-ICAR-NA-2025-06 Data: Luglio 2025

Consiglio Nazionale delle Ricerche, Istituto di Calcolo e Reti ad Alte Prestazioni (ICAR)
Sede di Napoli, Via P. Castellino 111, I-80131 Napoli,
Tel: +39-0816139508, Fax: +39-0816139531,
e-mail: napoli@icar.cnr.it, URL: www.na.icar.cnr.it


Consiglio Nazionale delle Ricerche
Istituto di Calcolo e Reti ad Alte Prestazioni

# Data Center Benchmarking: Stato dell'Arte e Best Practices

*Giovanni Massafra, Emanuele Damiano, Pier Giuseppe Meo, Mario Sicuranza*

Istituto di Calcolo e Reti ad Alte Prestazioni del Consiglio Nazionale delle Ricerche
Via Pietro Castellino, 111 – 80131 Napoli, Italia

E-mail: {giovanni.massafra, emanuele.damiano, piergiuseppe.meo, mario.sicuranza}@icar.cnr.it

**Abstract**
*Il benchmarking rappresenta uno strumento essenziale per l'analisi delle prestazioni, dell'efficienza e della scalabilità nei moderni data center. Questo lavoro fornisce una panoramica dello stato dell'arte, analizzando in particolare l'evoluzione delle architetture di rete con modelli come Spine-Leaf All-Optical (AOSL) e l'integrazione con soluzioni innovative come Rockport Networking.*
*Viene inoltre esaminato il benchmarking applicativo per i sistemi Big Data, con particolare attenzione al paradigma DataFlow, e approfonditi aspetti trasversali come l'efficienza energetica e le metriche multi-criterio. Il rapporto propone infine una metodologia strutturata di benchmarking.*

**Keywords**: benchmark, data center, Spine-Leaf, AOSL, rockport, big data, efficienza energetica

# Sommario

# 1. Introduzione

Il benchmarking rappresenta uno strumento essenziale per l'analisi delle prestazioni, dell'efficienza e della scalabilità nei moderni data center. In un contesto tecnologico in continua evoluzione, caratterizzato dalla crescente complessità delle infrastrutture IT, dall'esplosione dei dati generati quotidianamente e dall'emergere di nuove esigenze computazionali legate a Big Data, Intelligenza Artificiale e calcolo distribuito, il benchmarking si conferma una metodologia indispensabile per supportare la progettazione, l'ottimizzazione e la valutazione critica delle architetture ICT.
Questo rapporto tecnico fornisce una panoramica aggiornata dello stato dell'arte del benchmarking, con particolare attenzione alle architetture emergenti come la rete Spine-Leaf All-Optical (AOSL) e all'integrazione di soluzioni innovative come Rockport Networking. L'obiettivo è offrire una visione integrata e multidisciplinare, che abbracci sia gli aspetti applicativi che quelli legati alla rete e ad altre componenti critiche dei data center, con un focus specifico su metriche multi-criterio, efficienza energetica e sostenibilità ambientale.
La prima parte del documento analizza lo stato dell'arte nel campo del benchmarking, partendo dall'esame approfondito delle sfide poste dai sistemi Big Data: vengono presentati i limiti degli approcci tradizionali basati su dataset sintetici e workload ridondanti, nonché le potenzialità offerte da nuovi paradigmi, come il passaggio dal modello ControlFlow al più flessibile DataFlow. Si discute inoltre il ruolo centrale dei benchmark standardizzati – tra cui TPC-DS, HiBench e TPCx-BigBench – nell'analisi comparativa tra diversi sistemi, come Apache Spark, Hive e Transwarp Inceptor. Particolare attenzione è dedicata ai tool innovativi per la generazione di dati realistici, come HPCAN, che permettono agli utenti di definire strutture e distribuzioni personalizzate, garantendo maggiore controllo e flessibilità nel testing.
Un ampio spazio è riservato al benchmarking di rete, un ambito cruciale per comprendere e migliorare le capacità comunicative dei data center: viene descritta l'evoluzione verso architetture avanzate come la rete Spine-Leaf All-Optical (AOSL), in grado di superare i limiti intrinseci delle reti tradizionali basate su switch elettrici. L'architettura AOSL presenta vantaggi significativi in termini di bassa latenza, alta larghezza di banda e gestione ottimizzata della congestione, grazie all'utilizzo di interconnessioni ottiche e di Wavelength Selective Switches (WSS). Vengono inoltre approfonditi anche gli scenari di riconfigurazione dinamica dei WSS, le problematiche relative alla congestione nei data center tradizionali e gli algoritmi proposti per l'ottimizzazione delle risorse di rete. In questo contesto viene presentata l'integrazione di tecnologie emergenti come Rockport Networking, che introduce un'architettura "switchless" distribuita direttamente negli endpoint, con benefici tangibili in termini di riduzione della latenza, controllo avanzato della congestione e semplificazione del cablaggio.
Non mancano approfondimenti su altri aspetti chiave del benchmarking: tra questi, il benchmarking applicato ai supercomputer, ai Warehouse-Scale Computers (WSC), ai data center geodistribuiti e alle architetture Software-Defined Networking (SDN). Per ciascuno di questi contesti vengono descritte le peculiarità, le metriche utilizzate e le principali soluzioni disponibili, come Fleetbench per i WSC o GDSim per la simulazione di ambienti geodistribuiti, mentre un paragrafo specifico è dedicato all'uso di algoritmi eseguiti su GPU per accelerare operazioni complesse come quelle legate all'allocazione di risorse e al calcolo parallelo.
Una sezione trasversale e altamente rilevante riguarda infine le metriche multi-criterio e l'efficienza energetica: il benchmarking non può più limitarsi a misurare solo le prestazioni in termini di throughput o latenza, ma deve tenere conto di parametri come consumo energetico, utilizzo dello spazio, impatto ambientale e costi operativi. Vengono introdotti indicatori avanzati come Performance per Watt, Total Cost of Ownership (TCO) e Power Usage Effectiveness (PUE), nonché benchmark specifici come SPECpower e Green500, finalizzati a misurare l'efficienza energetica dei sistemi. Si sottolinea inoltre l'importanza di una gestione intelligente dei carichi di lavoro, anche attraverso la virtualizzazione, per ridurre sprechi e migliorare la sostenibilità ambientale.
Nel terzo ed ultimo capitolo, infine, il rapporto propone una metodologia strutturata e modulare per il benchmarking, articolata in fasi ben definite: dalla definizione degli obiettivi e delle categorie di benchmarking, alla scelta delle metriche e degli strumenti, fino alla configurazione dell'ambiente di test e all'analisi dettagliata dei risultati. Ogni fase è corredata da esempi pratici e casi studio, che illustrano

l’applicazione di questa metodologia in contesti reali, come l’implementazione di cluster cloud, il testing di framework Big Data e l’ottimizzazione dell’efficienza energetica.
In sintesi, questo lavoro si rivolge a ricercatori, professionisti del settore e decisori tecnologici, fornendo un riferimento completo e critico sulle pratiche e le innovazioni nel campo del benchmarking dei data center. Grazie all’approccio interdisciplinare e all’attenzione ai dettagli tecnici, il rapporto si configura come un contributo significativo alla letteratura scientifica e una guida pratica per chiunque operi nella progettazione, gestione e valutazione di sistemi informatici ad alte prestazioni.

# 2. Stato dell'arte

Il presente capitolo pone le basi per la discussione sul benchmarking nei data center, sottolineando l'importanza di questo strumento per la misurazione e la valutazione delle prestazioni delle infrastrutture IT, permettendo per cui di individuare punti di forza e criticità.
Viene evidenziato come la crescita esponenziale delle applicazioni di Big Data e le esigenze di calcolo su larga scala abbiano trasformato i data center in sistemi estremamente complessi e diversificati: questa evoluzione richiede lo sviluppo di metodologie e strumenti di benchmarking sempre più avanzati, in grado di rappresentare accuratamente la varietà dei carichi di lavoro e di rispondere alle nuove necessità in termini di efficienza energetica e ottimizzazione delle risorse.
Viene infine definito il contesto critico in cui il benchmarking assume un ruolo centrale, delineando le sfide poste dalla complessità crescente delle architetture moderne e dai requisiti delle applicazioni contemporanee. L'obiettivo è offrire un'analisi approfondita degli aspetti fondamentali legati alla valutazione delle prestazioni nei data center.

## 2.2 Benchmarking applicativo

Negli ultimi anni, i big data hanno assunto un ruolo centrale nei dibattiti accademici, industriali e governativi, configurandosi come una forza propulsiva per l'innovazione nell'era dell'informazione. Tale rilevanza si fonda principalmente su due fattori: da un lato, la crescente velocità con cui i dati vengono generati; dall'altro, l'elevato valore informativo che i big data racchiudono, il quale ha catalizzato trasformazioni significative in settori strategici quali e-commerce, finanza, trasporti e sanità.

Le peculiarità intrinseche dei big data, comunemente sintetizzate nelle cosiddette "3V" (Volume, Varietà, Velocità), pongono sfide complesse in termini di acquisizione, gestione, elaborazione e analisi. Per far fronte a tali criticità, il mondo accademico e l'industria hanno sviluppato e reso disponibili numerosi sistemi di analisi dei big data, sia open source – tra cui Apache Hive e Apache Spark – sia commerciali, come Transwarp Inceptor, Cloudera Impala e IBM Big SQL. Tali sistemi sono oggi adottati in misura crescente da imprese e organizzazioni pubbliche, per supportare processi decisionali e realizzare applicazioni aziendali basate sui dati.

In questo contesto, il testing e la valutazione dei sistemi di analisi big data sono diventati ambiti di ricerca di fondamentale importanza. L'attività di testing si articola in tre obiettivi principali:

1. Verificare la correttezza funzionale e l'affidabilità del sistema prima della sua adozione operativa;
2. Consentire un confronto oggettivo delle prestazioni tra diversi sistemi;
3. Supportare l'ottimizzazione delle performance attraverso un'analisi sistematica.

Attualmente, il benchmarking rappresenta il metodo principale per la valutazione dei sistemi di analisi big data, consentendo l'analisi comparativa di funzionalità, performance, affidabilità e compatibilità.

### 2.1.1 Sfide del Testing dei Sistemi di Analisi Big Data

L'applicazione di tecniche di testing a sistemi di analisi big data comporta numerose criticità, derivanti tanto dalle caratteristiche dei dati quanto dalla complessità dei sistemi:

- **Complessità architetturale**: I sistemi si basano prevalentemente su architetture distribuite (master-slave o peer-to-peer), la cui performance è influenzata da molteplici fattori, tra cui configurazioni hardware, parametri di sistema (fino a centinaia in ambienti come Hadoop), rete e virtualizzazione.
- **Complessità dei dataset di test**: I dataset devono rispecchiare le "3V" e, al contempo, simulare scenari d'uso realistici e rappresentativi del dominio applicativo.

- **Limiti metodologici e strumentali**: Gli strumenti tradizionali di testing si rivelano spesso inadeguati. Mancano framework per il testing automatico e strumenti diagnostici flessibili. I vari moduli del sistema richiedono approcci di testing specifici: ad esempio, le prestazioni di Spark SQL si testano tramite query SQL, mentre throughput e latenza di Spark Streaming richiedono test basati su flussi di dati in tempo reale.
- **Competenze richieste**: I tester devono possedere una combinazione avanzata di competenze, che includa conoscenze in materia di testing, architetture big data e tecnologie di elaborazione distribuita (es. caricamento dati su HDFS e interrogazione tramite Hive).

### 2.1.2 Benchmark Testing dei Sistemi di Analisi Big Data

Il processo di benchmark testing si articola tipicamente in sei fasi:

1. **Analisi dei requisiti**: Definizione degli obiettivi, delle metriche, dell'ambiente, dei dataset, delle tecnologie e degli strumenti di testing. Le priorità includono prestazioni ed affidabilità.
2. **Preparazione dell'ambiente di test**: Configurazione di un cluster distribuito con adeguate risorse computazionali e spazio di archiviazione. È essenziale garantire un ambiente "pulito", libero da interferenze esterne.
3. **Preparazione dei dataset e dei workload**: I dati possono essere reali o generati con tool come TPC-DS o TPCx-BigBench. Il workload deve rappresentare fedelmente i casi d'uso analitici reali.
4. **Caricamento dei dati**: Validazione del corretto caricamento dei dati nei sistemi distribuiti o nei database.
5. **Esecuzione del testing**: Analisi delle funzionalità, performance, affidabilità e compatibilità del sistema.
6. **Analisi dei risultati**: Valutazione complessiva dei risultati e redazione del report di testing.

Le principali tipologie di test includono:

- **Test di funzionalità**: Verifica della correttezza dei processi di storage, I/O ed elaborazione dati.
- **Test di prestazioni**: Misurazione delle performance di interrogazione (es. SQL), elaborazione e analisi dati.
- **Test di affidabilità**: Valutazione dei meccanismi di fault tolerance, inclusa la migrazione automatica dei task.
- **Test di compatibilità**: Verifica dell'interoperabilità con diversi file system, formati di dati e sintassi SQL.

### 2.1.3 Categorie di Benchmark

I benchmark per sistemi di analisi big data si classificano in tre categorie principali:

1. **Micro benchmark**: Valutano componenti specifici (es. TeraSort per l'ordinamento, GridMax per MapReduce). Forniscono metriche granulari, ma non valutano il sistema nel suo complesso.
2. **Benchmark comprensivi**: Consentono test su più moduli del sistema. Esempio rappresentativo è HiBench, che include test per SQL, ricerca web, machine learning.
3. **Benchmark orientati all'applicazione**: Simulano scenari aziendali realistici. TPC-DS e TPCx-BigBench sono ampiamente adottati grazie alla loro standardizzazione e completezza funzionale.

### 2.1.4 Casi Studio: Benchmark Testing Applicato

**1. Transwarp Inceptor con TPC-DS**
Obiettivo: valutare funzionalità ETL, performance SQL, compatibilità con SQL standard. È stato utilizzato un cluster a quattro nodi. TPC-DS modella il supporto decisionale per un rivenditore multi-canale. Sono stati generati 500GB di dati e testate 99 query SQL complesse. I risultati hanno mostrato che 96 su 99 query sono state eseguite con successo, dimostrando una buona compatibilità con SQL 2003. Le performance sono state valutate per categorie di query (interattive, OLAP, analisi statistiche, data mining).

**2. Confronto tra Hive e Spark SQL con TPCx-BigBench**
Obiettivo: confronto delle performance e ottimizzazione dei parametri. Il test, condotto su un cluster Cloudera CDH con 4 nodi, ha coinvolto 30 query suddivise per tipo di dato (strutturato, semi-strutturato, non strutturato) e metodo analitico (HQL, MapReduce, Machine Learning, NLP). Spark SQL ha superato Hive nelle query SQL standard grazie al calcolo in memoria. Le performance erano simili per i workload di machine learning (entrambi usano Spark MLlib). Alcune query NLP hanno visto Hive in vantaggio, poiché le librerie Python non permettevano piena parallelizzazione. Ottimizzazioni specifiche (es. riduzione del parametro `spark.sql.shuffle.partition`) hanno ulteriormente migliorato le prestazioni di Spark SQL.

### 2.1.5 Uno Strumento per la Generazione di Dati di Benchmark: HPCAN

Le attuali tecnologie sono sempre più orientate a risolvere problemi che riguardano grandi quantità di dati. Sebbene numerosi sistemi di benchmark siano stati sviluppati per valutare la scalabilità, le prestazioni e l'efficienza di un sistema Big Data, la maggior parte di essi presenta una sfida di usabilità per gli utenti che non sono esperti in sistemi di high performance computing. Per confrontare un sistema esistente con dati del mondo reale, è fondamentale fornire all'utente un controllo granulare sulla struttura e sulla distribuzione dei dati di benchmark.

La ricerca descritta nei documenti si concentra su un **approccio incentrato sull'utente** per la creazione di tali strumenti e propone un framework **flessibile, estensibile e facile da usare** per supportare l'analisi delle prestazioni dei sistemi Big Data. I risultati di questa ricerca includono un framework per la generazione di dati di benchmark denominato **HPCAN** (High Performance Computing Analysis).

HPCAN è motivato dalla necessità di progettare uno strumento di benchmarking facile da usare. Può generare dati di benchmark in formato **RDF** e **grafo di attributi**. Il linguaggio di query di riferimento è **SPARQL** (Simple Protocol RDF Query Language).

Ciò che distingue HPCAN dai sistemi esistenti è la sua capacità di accettare una **specifica di struttura e distribuzione definita dall'utente** come input e produrre il set di dati RDF di benchmark. Adotta un approccio flessibile basato su schema per definire il modello dati e un modello gerarchico basato su XML per definire la struttura e la distribuzione dei dati. Sebbene offra maggiore flessibilità, HPCAN mira a mantenere l'usabilità, fornendo un'interfaccia utente semplice, leggera ed estensibile, oltre a utility da riga di comando per generare dati. L'utente può, ad esempio, governare la distribuzione dei dati.

Sono stati presentati casi d'uso per convalidare il framework HPCAN. Due domini specifici sono stati selezionati per questo scopo: **cyber security** e **social networks**.

Nel dominio della cyber security, HPCAN utilizza dati di **Netflow**, che consentono di raccogliere il traffico di rete IP. Un record Netflow include informazioni come IP e porta di origine e destinazione, protocollo, numero di pacchetti, tempo trascorso e ID delle macchine generatrici e di raccolta. Il framework HPCAN fornisce una distribuzione e una struttura predefinite dei dati, ma consente all'utente di modificarle specificando diversi tipi di protocolli, numeri di porta, indirizzi IP univoci, ecc. Il cambiamento nel numero di macchine generatrici influenza la topologia dei dati. HPCAN include anche scenari speciali come "port scan" e attacchi di denial-of-service (DOS) nella sua interfaccia utente e fornisce la distribuzione osservata per questi scenari, permettendo comunque all'utente di **affinarla**.

Nel dominio dei social network, HPCAN ha affrontato i dati di **comunicazione email**. La ricerca in questo ambito è motivata da sfide come l'elaborazione del linguaggio naturale, la privacy e la distribuzione temporale. HPCAN ha analizzato vari set di dati email anonimizzati e reali. Fornisce una distribuzione predefinita di mittenti e destinatari univoci, la frequenza di comunicazione oraria e la distribuzione della frequenza di comunicazione per mittente.

Il lavoro futuro per HPCAN prevede la creazione di set di query completi per sistemi specializzati, la pubblicazione di statistiche sulle prestazioni e il confronto dei tempi di esecuzione con altri sistemi. Si prevede anche l'implementazione di un componente di incertezza per aggiungere rumore ai dati di output con una soglia superiore. HPCAN includerà anche generatori di set di dati per applicazioni e paradigmi di computing specializzati come streaming, machine learning ed elaborazione di grafi eterogenei.

In sintesi, mentre esistono vari framework e strumenti per generare dati statici con struttura e distribuzione fisse, HPCAN si concentra sulla **flessibilità e l'usabilità** dei generatori di dati di benchmark, offrendo un nuovo approccio per costruire tali sistemi e fornendo un controllo più elevato agli utenti nella creazione dei set di dati. Sono stati presentati risultati iniziali da casi d'uso nella cyber security (usando dati Netflow) e nella comunicazione email.

### 2.1.6 Conclusioni

Il benchmark testing dei sistemi di analisi big data è un'attività sempre più strategica per supportare l'adozione di soluzioni affidabili e performanti. Tali test consentono non solo il confronto oggettivo tra diversi sistemi, ma anche l'ottimizzazione mirata dei parametri operativi. I casi analizzati dimostrano l'efficacia di benchmark come TPC-DS e TPCx-BigBench nella valutazione delle funzionalità chiave di soluzioni come Transwarp Inceptor, Hive e Spark SQL. Le prospettive future della ricerca includono lo sviluppo di benchmark orientati all'analisi in tempo reale (streaming) e all'elaborazione di grafi, al fine di affrontare le nuove sfide poste dall'evoluzione continua del panorama big data.

## 2.2. Benchmarking di rete

Il benchmarking di rete è un aspetto cruciale nella valutazione delle prestazioni dei data center, in quanto permette di analizzare l'efficacia e l'efficienza dell'infrastruttura di comunicazione. Con la crescita esponenziale del traffico dati, dovuta alle applicazioni di Big Data e all'elaborazione su larga scala, le reti dei data center si sono evolute verso architetture complesse.
In questo contesto, il benchmarking di rete diventa indispensabile per identificare potenziali colli di bottiglia, misurare la latenza, valutare la larghezza di banda e, in generale, comprendere le capacità della rete: tecnologie avanzate come le reti All-Optical con architettura Spine-Leaf e soluzioni innovative come Rockport Networking mirano a superare i limiti delle reti tradizionali basate su switch elettrici. In questo contesto il benchmarking di rete diventa uno strumento essenziale per validare i miglioramenti e ottimizzare le prestazioni complessive del data center: diversi strumenti come iperf, Netperf e Pktgen sono utilizzati per eseguire test specifici sulle prestazioni della rete.

### 2.2.1 Architettura Spine-Leaf All-Optical nei Data Center

Tra le architetture più utilizzate per le reti dei data center (DCN), l'architettura spine-leaf è divenuta uno standard per la sua capacità di costruire reti di commutazione non bloccanti su larga scala impiegando elementi di switching di dimensioni ridotte.

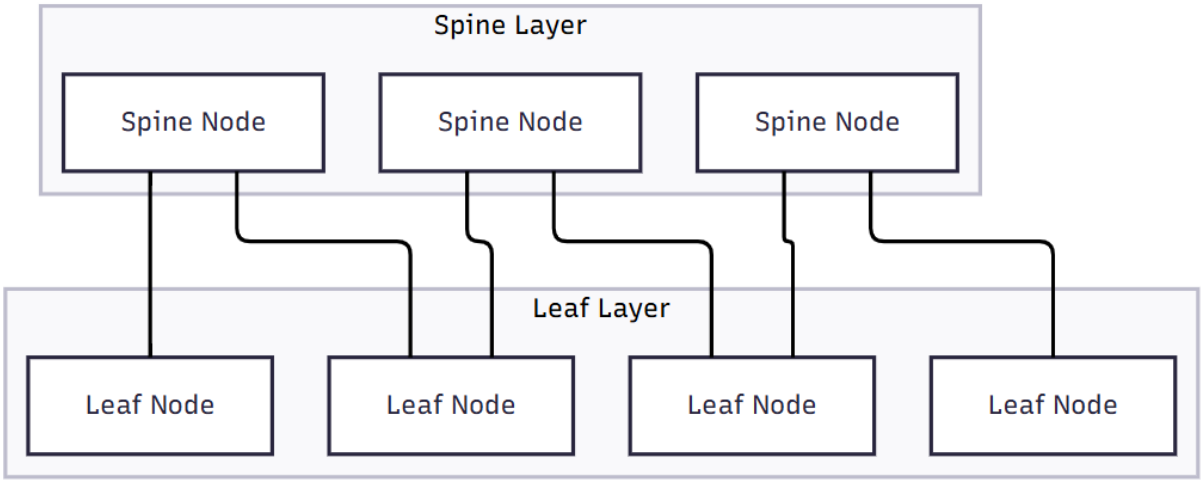


*Figura 1: Rappresentazione schematica dell'architettura gerarchica a due livelli di una rete Spine-Leaf All-Optical (AOSL).*

La Figura 1 illustra la struttura gerarchica a due livelli di una rete Spine-Leaf All-Optical (AOSL). Come evidenziato nel diagramma, il livello superiore è composto dai nodi Spine, mentre il livello inferiore è costituito dai nodi Leaf. I nodi Leaf si connettono direttamente a tutti i nodi Spine, garantendo percorsi multipli e ridondanza senza la necessità di stadi di aggregazione intermedi.
Le reti con architettura Spine-Leaf All-Optical rappresentano un'evoluzione avanzata delle infrastrutture di rete nei Data Center, offrendo prestazioni elevate, scalabilità intrinseca e bassa latenza. L'approccio All-Optical elimina le limitazioni intrinseche delle tradizionali reti basate su elettronica, utilizzando switch fotonici e interconnessioni in fibra ottica per ridurre significativamente la latenza e il consumo energetico. Questa soluzione è particolarmente adatta per ambienti ad alte prestazioni, quali il cloud computing, il machine learning distribuito e le applicazioni di big data, dove velocità, affidabilità e efficienza energetica della rete assumono un ruolo critico.

#### *Limiti delle Reti Basate su Switch Elettrici*

L'aumento esponenziale della domanda di traffico nei data center ha evidenziato i limiti delle reti basate su switch elettrici, tra cui bassa capacità di commutazione e consumo energetico elevato. Per superare queste difficoltà, le reti completamente ottiche (*all-optical DCNs*) sono considerate una soluzione promettente grazie alla loro elevata capacità e basso consumo energetico.

#### *Stato Attuale delle Reti All-Optical*

Le reti ottiche esistenti si concentrano principalmente su topologie piatte, come Torus e Dragonfly, che sono adatte per applicazioni specifiche, ad esempio l'apprendimento automatico distribuito. Tuttavia, queste topologie non sono sufficientemente flessibili per supportare una gamma più ampia di servizi generali nei data center: per affrontare questa problematica, si rende necessario progettare reti completamente ottiche con una topologia gerarchica, come l'architettura spine-leaf.

### 2.2.2 Architettura Spine-Leaf All-Optical (AOSL)

L'architettura **AOSL** (All-Optical Spine-Leaf) è una rete ottica avanzata progettata per i data center (DC). È caratterizzata da quattro parametri principali:

- **l**: numero di nodi *leaf* (Wavelength Selective Switch, WSS), ossia i punti di accesso a cui i server sono collegati.
- **s**: numero di nodi *spine* (WSS), che gestiscono la commutazione tra i leaf WSS.
- **p**: numero di fibre parallele tra un nodo leaf e un nodo spine, utilizzate per ridurre la congestione.
- **f**: numero di server collegati a ciascun nodo leaf.

Alcuni studi presenti in letteratura propongono un'architettura spine-leaf completamente ottica (*AOSL*), in cui gli switch elettrici tradizionali sono sostituiti da *wavelength selective switches* (WSS). L'obiettivo principale

è ottimizzare la fornitura di servizi nei data center, assegnando in modo efficiente percorsi, lunghezze d'onda e slot temporali.

### *Problema della congestione nei data center tradizionali*

Nei data center convenzionali, ogni collegamento tra un nodo leaf e uno spine è costituito da una singola fibra. Questo approccio può causare **limitazioni di capacità** a causa di un elevato rapporto di oversubscription (troppi server condividono poche risorse di rete). Per affrontare il problema, l'architettura AOSL implementa **fibre parallele** tra i nodi leaf e spine, riducendo i conflitti di lunghezza d'onda (wavelength conflict) e migliorando le prestazioni complessive.

### *Portate e capacità della rete*

I parametri l,s,p,f consentono di calcolare:

- Il numero totale di porte dei **leaf WSS**: $p \times s + f$.
- Il numero totale di porte dei **spine WSS**: $p \times l$.
- Il numero massimo di **server** che la rete può supportare: $l \times f$.

### *Architettura interna dei server*

Ogni nodo server comprende:

1. **Accoppiatore**: aggrega i segnali.
2. **Arrayed Waveguide Grating (AWG)**: demultiplexa i segnali ottici.
3. **Trasmettitori/ricevitori (W)**: per la trasmissione e ricezione dei dati.
4. **Risorse di calcolo**: per gestire il carico di lavoro dei server.

### *Supporto ai servizi unicast*

Per il supporto ai servizi unicast, tipici degli ambienti data center, si rende necessaria l'instaurazione di lightpath, ovvero percorsi ottici dedicati, tra specifiche coppie di nodi sorgente e destinazione per la durata del loro tempo di utilizzo (holding time), misurato in slot temporali. La corretta ed efficiente configurazione di tali lightpath è gestita attraverso processi di ottimizzazione che riguardano l'instradamento, l'assegnazione delle lunghezze d'onda e l'assegnazione dei time slot (Routing, Wavelength, Time Slot Assignment - RWTA).


*Figura 2: Diagramma di flusso che illustra le fasi del processo RWTA (Routing, Wavelength, Time Slot Assignment) per l'instaurazione di servizi unicast in una rete Spine-Leaf All-Optical (AOSL).*

La Figura 2 illustra in dettaglio le fasi principali del processo RWTA per l'attivazione di un servizio unicast in una rete AOSL. Come descritto nel diagramma di flusso, a partire dalla richiesta iniziale di servizio, il processo

procede attraverso l'identificazione dei nodi sorgente e destinazione, la determinazione del percorso ottico ottimale (Routing), seguita dall'assegnazione di una lunghezza d'onda disponibile (Wavelength Assignment) e di uno slot temporale disponibile (Time Slot Assignment) lungo il percorso individuato. La fase conclusiva consiste nella configurazione dinamica dei componenti ottici, in particolare dei Wavelength Selective Switches (WSS), per stabilire fisicamente il lightpath e rendere il servizio unicast attivo.

L'efficienza del processo RWTA, come schematizzato, dipende crucialmente dalla capacità di riconfigurazione dinamica dei WSS, che consente un'allocazione flessibile delle risorse di rete in base alle richieste. Tuttavia, è fondamentale considerare che la riconfigurazione dei WSS introduce un ritardo intrinseco, tipicamente dell'ordine dei 10 ms. Durante questo intervallo, il dispositivo non è in grado di elaborare nuove richieste di servizio, il che impone una limitazione significativa sulla velocità di setup di nuovi lightpath e, di conseguenza, sulla reattività complessiva della rete.

### *Scenari di riconfigurazione WSS*

La gestione efficiente delle risorse ottiche in una rete Spine-Leaf All-Optical (AOSL) si basa sulla riconfigurazione dinamica dei Wavelength Selective Switches (WSS). Esistono diversi scenari per l'implementazione di tale riconfigurazione, ciascuno con implicazioni significative sulla disponibilità del servizio e sulla complessità gestionale della rete.

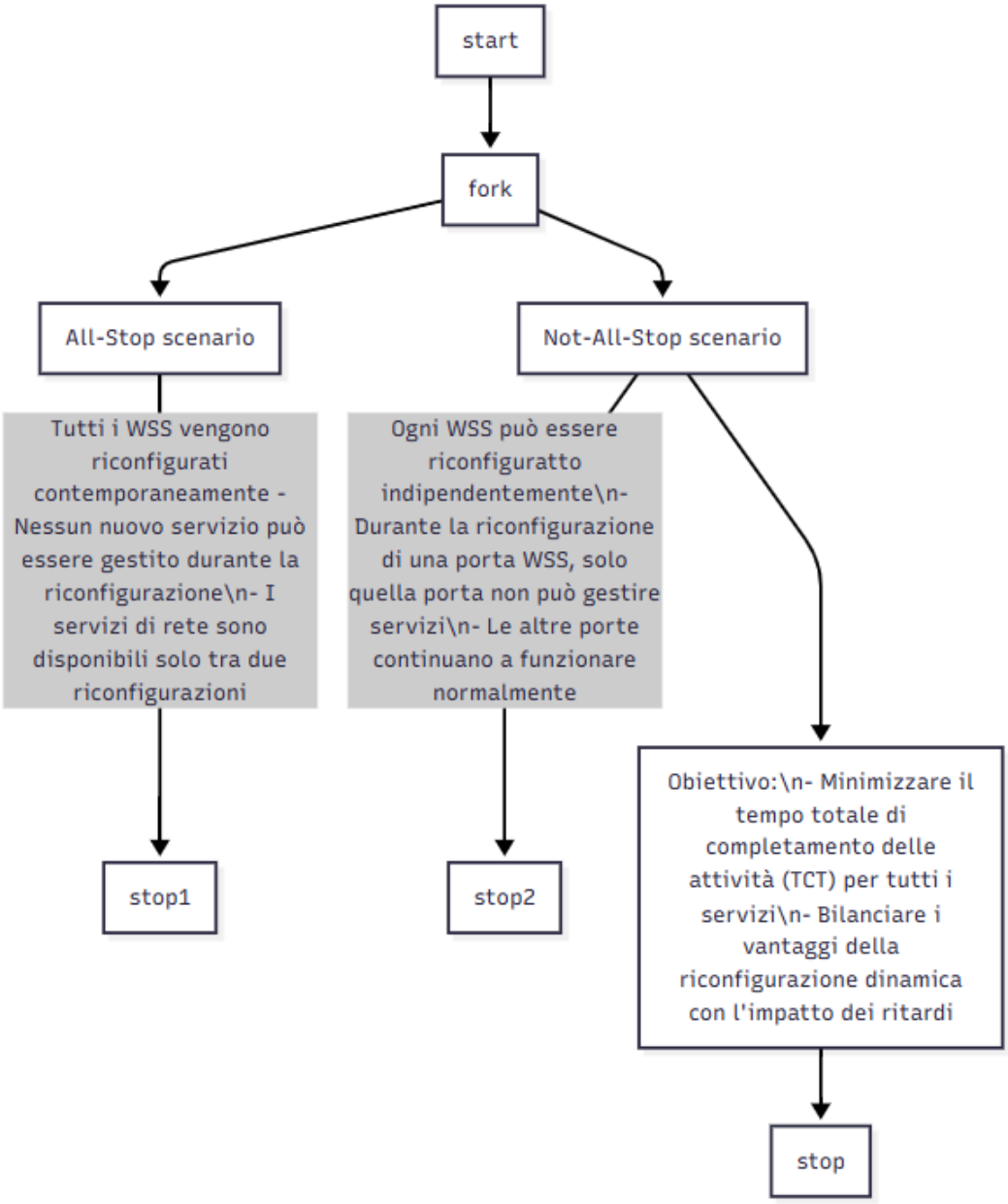


*Figura 3: Confronto schematico degli scenari di riconfigurazione dei WSS: All-Stop (AS) e Not-All-Stop (NAS).*

La Figura 3 illustra schematicamente due approcci principali alla riconfigurazione dei WSS: lo scenario All-Stop (AS) e lo scenario Not-All-Stop (NAS). Come evidenziato nel diagramma, nello scenario All-Stop, tutti i WSS della rete vengono riconfigurati simultaneamente in istanti temporali predefiniti. Durante il periodo di riconfigurazione collettiva, l'intera rete di commutazione ottica non è disponibile per l'instaurazione di nuovi lightpath o la gestione di servizi, rendendo i servizi di rete utilizzabili unicamente negli intervalli tra due riconfigurazioni successive. Questo approccio presenta una minore complessità implementativa a livello di controllo distribuito, ma richiede un'accurata sincronizzazione a livello dell'intera infrastruttura di rete.

Al contrario, nello scenario Not-All-Stop, la riconfigurazione può avvenire in modo indipendente per ciascun WSS o anche per singole porte di un WSS, tipicamente allocando uno specifico time slot per tale operazione. Come mostrato nel diagramma, durante la riconfigurazione di una singola porta WSS, quella specifica porta non è in grado di gestire servizi, ma le altre porte dello stesso WSS e gli altri WSS non coinvolti continuano a operare normalmente. Questo garantisce una maggiore flessibilità operativa e una disponibilità di rete potenzialmente continua, sebbene la sua gestione e pianificazione risultino intrinsecamente più complesse. Il

diagramma evidenzia inoltre che, in particolare nello scenario NAS, l'obiettivo primario è la minimizzazione del tempo totale di completamento delle attività (Total Completion Time - TCT) per tutti i servizi, ricercando un bilanciamento ottimale tra i vantaggi derivanti dalla riconfigurazione dinamica e l'impatto dei ritardi introdotti dal processo stesso.

#### *Obiettivo della rete AOSL*

Per entrambi gli scenari (AS e NAS), l'obiettivo è:

- **Minimizzare il tempo totale di completamento delle attività (TCT)** per tutti i servizi.
- Equilibrare i vantaggi della riconfigurazione dinamica dei WSS (migliore utilizzo della rete) con l'impatto dei ritardi introdotti dal processo di riconfigurazione.

In sintesi, l'architettura AOSL offre un approccio scalabile ed efficiente per superare i limiti di capacità dei data center tradizionali, migliorando l'utilizzo delle risorse ottiche e riducendo la congestione tramite un'ottimale gestione dei lightpath e delle risorse di rete.

### 2.2.3 Approcci di Ottimizzazione e Algoritmi

Per lo scenario AS, lo studio sviluppa un modello di ottimizzazione basato sulla programmazione lineare intera (ILP) per minimizzare il tempo di completamento dei task (*Task Completion Time*, TCT). Inoltre, per entrambi gli scenari, vengono proposti algoritmi euristici per la gestione dei servizi nella rete AOSL. I risultati delle simulazioni evidenziano che:

- Gli algoritmi euristici proposti si avvicinano alle prestazioni del modello ILP, pur essendo più efficienti in termini computazionali.
- Lo scenario NAS riduce significativamente il TCT rispetto allo scenario AS, dimostrando una maggiore flessibilità ed efficienza nella riconfigurazione dei WSS.

### 2.2.4 Integrazione di Rockport Networking in un'Architettura All Optical Spine-Leaf

Rockport Networks ha sviluppato una tecnologia di rete di nuova generazione incentrata sulle prestazioni su larga scala. La sua architettura distintiva, definita "switchless", distribuisce le funzioni di switching direttamente agli endpoint di rete, come i server e i sistemi di storage. Questa tecnologia si basa su interconnessioni ottiche, controllo avanzato della congestione e routing adattivo multi-path per migliorare le prestazioni delle reti di grandi data center e aziendali, con un focus particolare sulle applicazioni AI e HPC. L'esplorazione della potenziale integrazione del networking Rockport in un'architettura All Optical Spine-Leaf potrebbe portare a vantaggi significativi in termini di prestazioni, efficienza e scalabilità per carichi di lavoro esigenti.

#### *Fondamentali del Networking Rockport*

L'architettura "switchless" sviluppata da Rockport Networks si discosta radicalmente dalle tradizionali topologie di rete per data center basate su switch centralizzati (come ad esempio le architetture a tre livelli o Spine-Leaf tradizionali). Invece di affidarsi a switch spine e leaf dedicati come punti di aggregazione e smistamento del traffico, Rockport distribuisce la funzionalità di switching direttamente agli endpoint di rete, integrando tale capacità nelle schede di rete (Network Interface Cards - NIC), identificate come NC1225, installate nei server e nei sistemi di storage. In questa configurazione, ogni nodo della rete Rockport è direttamente interconnesso con altri nodi attraverso collegamenti punto-punto, formando logicamente una topologia a mesh completa o parziale. Questo approccio elimina la necessità di switch di rete esterni all'interno del cluster di calcolo, spostando l'intelligenza della rete più vicino al carico di lavoro.
La Figura 4 mostra come l'approccio di Rockport si discosti dalle tradizionali topologie di rete basate su switch centralizzati. Evidenzia i componenti chiave come l'NIC NC1225, il FPGA aggiornabile, il sistema operativo rNOS e l'Autonomous Network Manager (ANM), che insieme abilitano le funzionalità avanzate di questa architettura "switchless".

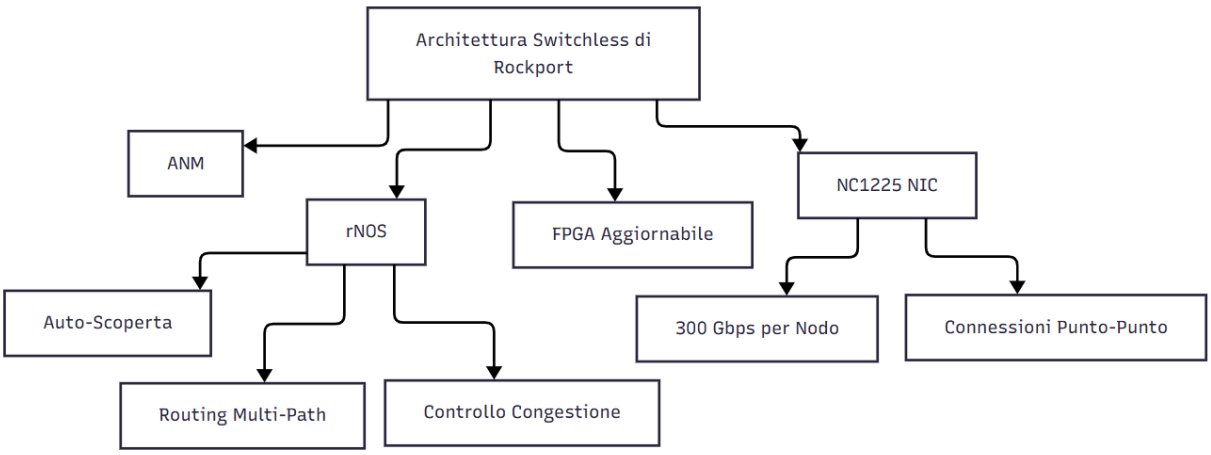


*Figura 4: Diagramma dell'architettura 'switchless' di Rockport Networks, che mostra i principali componenti e le loro interconnessioni. Questa architettura distributa il processo di switching direttamente agli endpoint della rete, eliminando la necessità di switch centralizzati..*

Il cuore della soluzione Rockport è rappresentato dalla scheda di rete NC1225. Questa scheda costituisce il nodo fondamentale del fabric di rete, fornendo una capacità di interconnessione pari a 300 Gbps per nodo. La connettività ottica è garantita da 12 porte in fibra ottica, ciascuna operante a una velocità di 25 Gbps. Nonostante la sua architettura interna ottica avanzata, la NC1225 presenta un'interfaccia host standard (tipicamente Ethernet), il che garantisce una compatibilità elevata con le applicazioni e gli stack software di rete esistenti, facilitando l'integrazione in ambienti IT preesistenti.
Un elemento cruciale che abilita le funzionalità avanzate della rete Rockport è la presenza di un FPGA (Field-Programmable Gate Array) aggiornabile sul campo. Questo componente hardware configurabile esegue il Rockport Network Operating System (rNOS), il sistema operativo proprietario della rete.
rNOS è responsabile dell'orchestrazione e del controllo dell'intera fabric di rete Rockport, definendone le capacità e le caratteristiche operative. Il sistema operativo di rete è progettato per consentire alla fabric di auto-scoprirsi (identificare automaticamente i nodi e i collegamenti presenti), auto-configurarsi (ottimizzare le impostazioni di rete in base alla topologia rilevata) e auto-ripararsi (gestire autonomamente guasti e malfunzionamenti per mantenere la connettività), semplificando notevolmente le operazioni di gestione e manutenzione della rete su larga scala.
Un'altra funzionalità chiave implementata da rNOS è il routing adattivo multi-path. Questo meccanismo consente al traffico di essere distribuito dinamicamente su fino a 8 percorsi ottimali e fisicamente indipendenti disponibili tra sorgente e destinazione. Tale capacità assicura un utilizzo altamente efficiente della larghezza di banda totale del fabric e migliora significativamente la tolleranza ai guasti, poiché il traffico può essere reindirizzato rapidamente in caso di interruzioni su un percorso. Inoltre, Rockport impiega una tecnologia brevettata di switching distribuito basato su unità di controllo di flusso (FLIT - Flow Control Unit), che scompone i pacchetti di dati in unità più piccole per una trasmissione e una commutazione più rapide e a bassissima latenza all'interno della rete.
Il sistema Rockport include anche controlli avanzati della congestione e meccanismi di monitoraggio del percorso end-to-end. Questi strumenti sono fondamentali per ottimizzare continuamente le prestazioni della rete sotto carico, prevenendo colli di bottiglia e garantendo un flusso di dati efficiente. Per i messaggi più critici e sensibili alla latenza, Rockport offre una priorità ultra elevata. Questa funzionalità assicura che tali messaggi siano trattati con la massima urgenza, rendendoli virtualmente immuni agli effetti della congestione di rete. A garanzia di un funzionamento affidabile e privo di stalli, l'algoritmo di routing Deadlock-Free (DFR) implementato da rNOS assicura la corretta gestione del traffico su varie topologie di rete.
L'interconnessione ottica Rockport SHFL (Shuffle) gioca un ruolo fondamentale nella semplificazione della distribuzione fisica delle reti Rockport, in particolare per topologie complesse come quelle richieste dagli ambienti di supercalcolo (HPC). SHFL è una soluzione di cablaggio ottico passiva e pre-cablata che riduce drasticamente la complessità e il tempo necessari per l'installazione rispetto ai cablaggi tradizionali su larga scala, offrendo un'esperienza quasi plug-and-play. Essendo un componente passivo, SHFL non richiede

alimentazione elettrica o raffreddamento dedicato, contribuendo all'efficienza energetica complessiva del data center.
Per la gestione e il monitoraggio dell'infrastruttura di rete, Rockport fornisce l'Autonomous Network Manager (ANM). Questa piattaforma software autonoma offre agli amministratori una visibilità completa e centralizzata sull'intera rete Rockport. ANM consente di monitorare in tempo reale le prestazioni dei singoli nodi, tracciare i flussi di traffico end-to-end, gestire le impostazioni a livello di sistema e ottimizzare le funzionalità avanzate, semplificando notevolmente la gestione operativa di reti distribuite complesse.

### *Integrazione e Potenziali Sinergie: Rockport in un Contesto All Optical Spine-Leaf*

L'integrazione di Rockport in un'architettura All Optical Spine-Leaf implica la sostituzione dei tradizionali switch leaf e spine con la fabric distribuita e switchless di Rockport. In questo scenario, ogni server (o gruppo di server) potrebbe fungere da nodo "leaf" in una topologia Spine-Leaf logica, interconnessa utilizzando la tecnologia Rockport. La funzione di "spine" verrebbe distribuita tra i nodi Rockport interconnessi.
Una potenziale sinergia significativa risiede nella possibilità di ottenere una latenza ancora più bassa. Lo switching FLIT e il routing adattivo di Rockport sono progettati per una latenza estremamente bassa. L'eliminazione degli hop degli switch tradizionali potrebbe ridurre ulteriormente la latenza rispetto a un'architettura All Optical Spine-Leaf con switch convenzionali. Rockport afferma di offrire una latenza significativamente inferiore rispetto alle reti tradizionali e persino a InfiniBand.
Un altro vantaggio potenziale è il controllo della congestione e l'utilizzo delle risorse migliorati: meccanismi avanzati di controllo della congestione di Rockport potrebbero prevenire i colli di bottiglia all'interno della fabric Spine-Leaf formata logicamente.
Il routing adattivo per pacchetto garantisce un utilizzo efficiente di tutta la larghezza di banda disponibile.[9] Ciò si traduce in una maggiore sfruttamento delle risorse di calcolo e storage, poiché non rimangono inattive in attesa dei dati.
La complessità del cablaggio, una delle sfide delle architetture All Optical Spine-Leaf, potrebbe essere semplificata con l'interconnessione ottica Rockport SHFL, la quale offre una soluzione di interconnessione ottica plug-and-play e pre-cablata, semplificando potenzialmente il complesso cablaggio di una Spine-Leaf All Optical su larga scala: ciò potrebbe anche portare a tempi di implementazione più rapidi rispetto alle reti tradizionali basate su switch.
L'integrazione di Rockport in un contesto All Optical Spine-Leaf potrebbe anche migliorare significativamente le prestazioni per carichi di lavoro HPC e AI/ML: Rockport è specificamente progettato per questi carichi di lavoro esigenti, offrendo funzionalità come la priorità ultra-elevata per i messaggi critici.[9] Questa combinazione potrebbe superare i limiti di prestazioni delle tradizionali architetture Spine-Leaf in ambienti di scala estrema.
Infine, l'Autonomous Network Manager (ANM) di Rockport potrebbe semplificare la gestione della fabric Rockport sottostante che forma la Spine-Leaf All Optical, offrendo funzionalità di monitoraggio, ottimizzazione e auto-riparazione.

La tabella seguente riassume un confronto tra un'architettura All Optical Spine-Leaf tradizionale e una che integra la tecnologia Rockport:

| **Caratteristica** | **Spine-Leaf All Optical Tradizionale** | **Spine-Leaf All Optical Integrata con Rockport** |
|---|---|---|
| **Switching** | Centralizzato (Switch Spine e Leaf) | Distribuito (all'interno delle schede di rete dei server) |
| **Latenza** | Bassa (rispetto a 3-tier), prevedibile | Ultra-bassa, potenzialmente inferiore per via di un minor numero di hop |

| | | |
|---|---|---|
| **Controllo Congestione** | Basato su protocolli come ECMP, TRILL, SPB | Avanzato, routing adattivo per pacchetto, switching FLIT |
| **Complessità Cablaggio** | Può essere complessa, alta densità di cavi ottici | Potenzialmente semplificata con Rockport SHFL |
| **Scalabilità** | Buona, tramite aggiunta di switch spine e leaf | Eccellente, scalabilità lineare con l'aggiunta di nodi di calcolo |
| **Gestione** | Strumenti di gestione di rete standard | Rockport Autonomous Network Manager (ANM) |
| **Focus Primario** | Alta larghezza di banda, bassa latenza per esigenze generali del data center | Prestazioni estreme per HPC, AI/ML e applicazioni sensibili alla latenza |

*Tabella 1: Confronto tra Spine-Leaf All Optical Tradizionale e Spine-Leaf All Optical Integrata con Rockport*

La tabella seguente evidenzia le caratteristiche e i vantaggi chiave del networking Rockport:

| Caratteristica | Descrizione | Vantaggio |
|---|---|---|
| **Architettura Switchless** | Funzione di switching distribuita alle schede di rete nei server | Elimina i colli di bottiglia degli switch tradizionali, riduce la latenza |
| **Scheda di Rete NC1225** | Fabric da 300 Gbps per nodo, 12 connessioni ottiche | Elevata larghezza di banda per nodo per applicazioni esigenti |
| **rNOS** | Sistema operativo auto-scoprente, auto-configurante, auto-riparante | Gestione semplificata, resilienza |
| **Routing Adattivo Multi-Path** | Traffico distribuito su 8 percorsi ottimali | Utilizzo efficiente della larghezza di banda, tolleranza ai guasti |
| **Switching FLIT** | Pacchetti suddivisi in piccole unità per l'inoltro rapido | Latenza ultra-bassa, anche sotto carico |
| **Priorità Ultra-Elevata** | Messaggi critici prioritizzati | Garantisce bassa latenza per il traffico sensibile alla latenza |
| **Interconnessione Ottica SHFL** | Cablaggio ottico passivo e pre-cablato | Distribuzione semplificata, complessità di cablaggio ridotta |
| **Autonomous Network Manager (ANM)** | Strumento di gestione di rete autonomo | Visione olistica, monitoraggio e ottimizzazione |

*Tabella 2: Caratteristiche e Vantaggi Chiave del Networking Rockport*

### 2.1.5 Le reti RDMA

La tecnologia Reliable Data Memory Access (RDMA) si è affermata come un elemento cruciale negli ambienti di high-performance computing (HPC) e nei moderni servizi di cloud computing. Questa popolarità deriva dalla sua capacità fondamentale di supportare trasferimenti di dati a bassissima latenza e ad alta larghezza di banda direttamente tra la memoria dei dispositivi connessi in rete, bypassando significativamente lo stack di rete tradizionale (ad esempio, TCP/IP) e riducendo il carico computazionale sui processori dei server. Tale efficienza rende RDMA particolarmente adatto e vantaggioso per un'ampia gamma di applicazioni esigenti e sensibili alla latenza, tra cui database distribuiti, sistemi di machine learning, analisi di big data e simulazioni

scientifiche complesse. Nonostante i notevoli benefici in termini prestazionali, garantire l'affidabilità operativa delle reti RDMA rappresenta una sfida ingegneristica e di progettazione cruciale, che richiede un'attenta valutazione prima di poter implementare tali tecnologie su larga scala negli ambienti di data center di produzione.

### *Approcci Esistenti per la Valutazione dell'Affidabilità RDMA*

La valutazione sistematica dell'affidabilità delle infrastrutture di rete che utilizzano la tecnologia RDMA è un prerequisito essenziale per la loro adozione in contesti critici. Attualmente, le metodologie impiegate per analizzare e quantificare l'affidabilità dei sistemi RDMA possono essere ricondotte principalmente a due categorie distinte: i metodi sperimentali e i metodi basati sulla simulazione.

La Figura 5 fornisce una rappresentazione strutturale di questi due approcci principali per la valutazione dell'affidabilità RDMA. Come illustrato nel diagramma, le metodologie si ramificano nelle categorie degli approcci Sperimentali e delle metodologie di Simulazione, ciascuna caratterizzata da specifici processi, strumenti e intrinseche limitazioni.

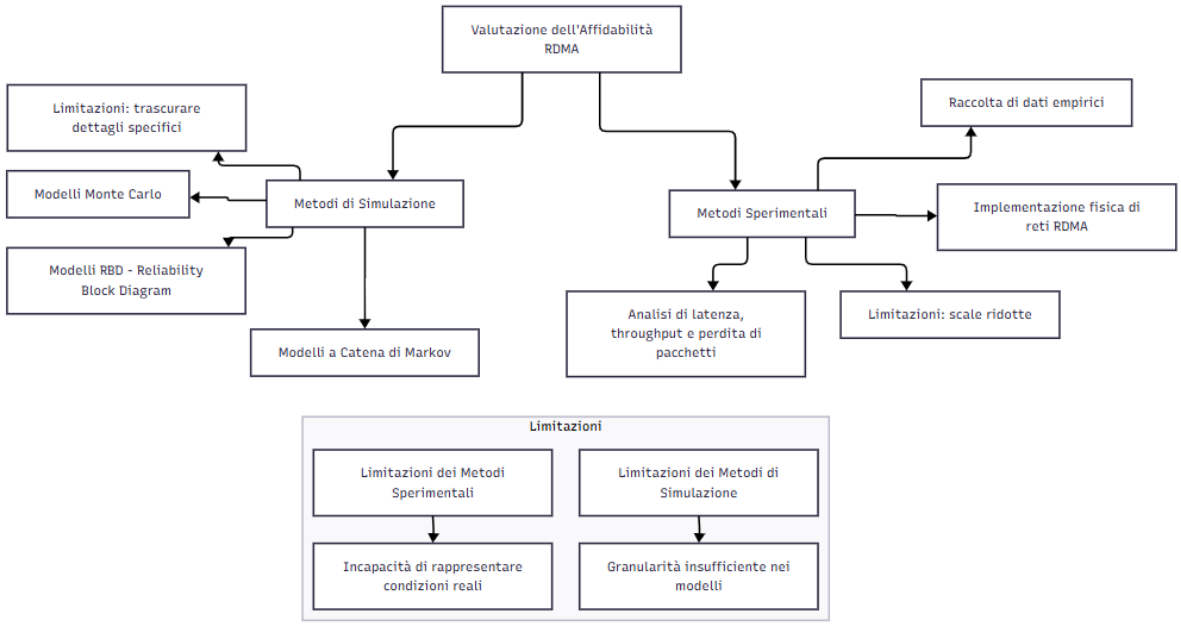


*Figura 5: Panoramica delle metodologie esistenti per la valutazione dell'affidabilità delle reti RDMA, distinguendo tra approcci Sperimentali e di Simulazione e indicandone caratteristiche e limitazioni principali.*

Nel dettaglio, i Metodi Sperimentali, come schematizzato nel ramo corrispondente della Figura 5, si basano sull'implementazione fisica e sull'osservazione diretta di reti RDMA reali, tipicamente configurate come piccoli cluster di dimensioni limitate (generalmente comprendenti tra 6 e 20 server). L'analisi sperimentale si concentra sulla raccolta e misurazione empirica di parametri di performance chiave in condizioni operative controllate, quali la latenza di trasmissione end-to-end, la larghezza di banda effettivamente raggiungibile (throughput) e i tassi di perdita di pacchetti in diverse condizioni di carico. Diversi studi preesistenti rientrano in questa metodologia; ad esempio, ricerche hanno impiegato cluster di 16 server per investigare le dinamiche della latenza nei flussi di trasmissione RDMA al fine di inferire conclusioni sull'affidabilità complessiva del sistema, mentre altri lavori condotti su cluster di 10 server hanno quantificato la larghezza di banda massima e media trasmissibile, nonché i tassi di perdita di pacchetti. Sebbene esperimenti prolungati su scale ridotte possano fornire dati empirici preziosi su metriche come latenza e perdita, è fondamentale riconoscere che tali studi spesso non riescono a catturare o tenere adeguatamente conto di fenomeni complessi e critici che si manifestano nelle reti su larga scala e negli ambienti di produzione, quali i ritardi a lunga coda (long-tail delays) indotti da micro-burst o la gestione distribuita dei guasti hardware. Come indicato nel ramo

'Limitazioni' degli approcci Sperimentali nella Figura 5, il limite intrinseco e principale di queste metodologie risiede nell'incapacità di rappresentare in modo pienamente accurato e predittivo le condizioni operative reali e le sfide di affidabilità che caratterizzano le reti RDMA implementate in grandi data center, a causa delle inevitabili limitazioni di scala degli ambienti di test sperimentali.

Per ovviare alle limitazioni inerenti alle valutazioni sperimentali confinate a scale ridotte, i ricercatori hanno ampiamente adottato i Metodi di Simulazione. Questi approcci permettono di modellare digitalmente e analizzare l'affidabilità di reti RDMA su scale molto più ampie, difficilmente replicabili in laboratorio, e in condizioni operative variabili. Come descritto nel ramo 'Simulazione' della Figura 5, all'interno di questa categoria rientrano diverse tecniche modellistiche: i Modelli RBD (Reliability Block Diagram), che consentono di simulare i processi di trasmissione su percorsi di rete RDMA predefiniti, sebbene possano presentare condizioni applicative rigide e una complessità elevata nella loro modellazione strutturale; i Modelli a Catena di Markov o approcci basati su processi stocastici simili, utilizzati per modellare l'evoluzione degli stati della rete RDMA in diversi istanti temporali e valutarne l'affidabilità attraverso l'analisi delle probabilità di transizione tra gli stati; e infine, i Modelli Monte Carlo, un approccio stocastico che consente di effettuare un test della trasmissione RDMA simulando i moduli funzionali del protocollo e dell'hardware di rete come unità atomiche all'interno di un modello computazionale. Tuttavia, anche i metodi di simulazione presentano delle sfide e delle limitazioni intrinseche. Come riportato nelle 'Limitazioni' del ramo Simulazione nella Figura 5, i modelli, in particolare quelli Monte Carlo su larga scala, tendono spesso a trascurare i dettagli specifici del protocollo RDMA e le peculiarità implementative dell'hardware del server al fine di mantenere la complessità computazionale gestibile. Ciò può comportare una granularità insufficiente del modello e una rappresentazione non pienamente fedele del comportamento dinamico della rete, specialmente in contesti di reti su larga scala caratterizzate da interazioni complesse e fenomeni non lineari.

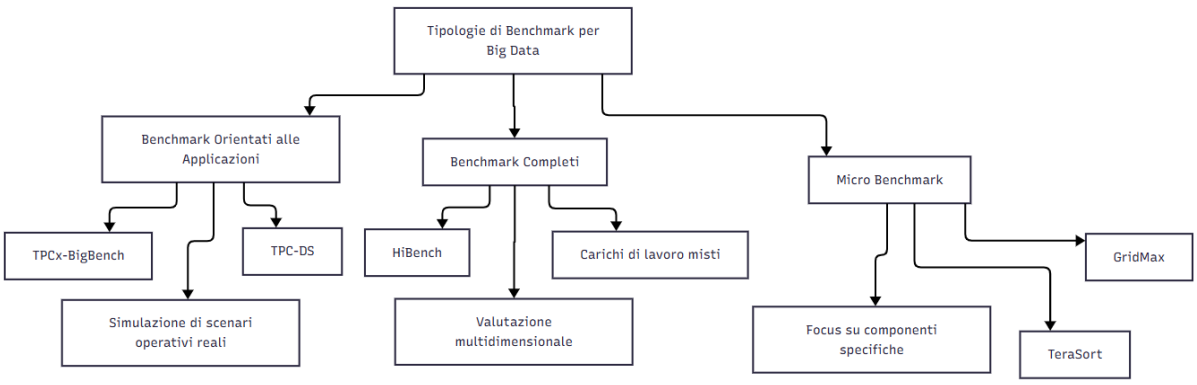


*Figura 6: Classificazione delle principali tipologie di benchmark impiegate per la valutazione delle prestazioni dei sistemi di analisi dei Big Data, con esempi rappresentativi per ciascuna categoria.*

## 2.3 Altri aspetti del benchmarking

Questa sezione amplia la prospettiva introdotta finora, esplorando come le metodologie di benchmarking si applichino a contesti specifici e considerazioni emergenti nel mondo dei data center. Viene approfondito il benchmarking per i centri di supercalcolo, evidenziandone le esigenze uniche in termini di capacità computazionale ed efficienza energetica.
Segue l'analisi del benchmarking dei *Warehouse-Scale Computers (WSC)*, sistemi caratterizzati dalla diversità delle applicazioni e dalla dipendenza dai dati di input, per i quali sono state sviluppate suite di benchmark dedicate, come *Fleetbench*.
Viene inoltre considerata l'importanza della simulazione nei data center geodistribuiti per testare e valutare gli schedulatori di lavoro in ambienti complessi.

Un ulteriore ambito trattato è l'utilizzo del *Software-Defined Networking (SDN)* nei data center, impiegato per ottimizzare il traffico generato da applicazioni distribuite.
La sezione esamina anche l'impiego di algoritmi eseguiti su GPU per accelerare compiti computazionalmente intensivi legati all'allocazione delle risorse.
Infine, viene introdotto il concetto di *metriche multi-criterio* per il benchmarking, che superano la sola velocità di esecuzione per includere parametri come l'efficienza energetica, l'utilizzo dello spazio e i costi operativi, offrendo una visione più olistica e sostenibile della valutazione delle prestazioni dei data center.

### 2.3.1 Benchmarking per Centri di Supercalcolo

Il *National Center for Atmospheric Research* (NCAR) ha tradizionalmente utilizzato il Mesa Lab Computing Facility (MLCF) a Boulder, Colorado, come principale infrastruttura di supercalcolo. Nonostante gli aggiornamenti significativi effettuati nel 2005, che includevano miglioramenti ai sistemi di alimentazione elettrica, l'installazione di un generatore di emergenza e l'ottimizzazione del sistema di raffreddamento per supportare fino a 1,2 MW di apparecchiature informatiche, l'MLCF non è più in grado di soddisfare le esigenze delle prossime generazioni di apparecchiature di supercalcolo.
Per rispondere a queste sfide, NCAR sta costruendo il *NCAR-Wyoming Supercomputing Center* (NWSC) a Cheyenne, Wyoming, in collaborazione con diverse istituzioni pubbliche e private, tra cui la National Science Foundation e l'Università del Wyoming. La costruzione è iniziata nel 2010 e il centro, progettato per essere operativo entro giugno 2012, si distingue per il suo focus su efficienza energetica e riduzione dell'impronta di carbonio. Il NWSC utilizzerà risorse rinnovabili e sistemi di raffreddamento naturali, con una previsione di efficienza energetica quattro volte superiore rispetto all'MLCF e superiore del 10% rispetto ai data center all'avanguardia esistenti.
Quando diventerà operativo, il NWSC ospiterà il primo sistema di supercalcolo petascale di NCAR, offrendo una capacità computazionale 15-20 volte superiore rispetto alle risorse attuali. Tuttavia, la complessità della sua implementazione va oltre la semplice installazione di apparecchiature. NCAR sta sviluppando un approccio integrato per procurare simultaneamente risorse di calcolo, archiviazione, analisi dei dati e visualizzazione. Questa metodologia è supportata da una suite di benchmark progettata per valutare le capacità dei sistemi, con particolare attenzione all'efficienza energetica e alla capacità di gestire applicazioni scientifiche complesse, in particolare nella ricerca sulle scienze della Terra.
Lo studio descrive inoltre come il processo di benchmarking del NWSC possa essere applicato ad altri centri di calcolo ad alte prestazioni (HPC), promuovendo standard per l'analisi e il confronto delle prestazioni nei centri che supportano la ricerca scientifica avanzata. Questo approccio contribuisce a migliorare la progettazione e la gestione delle infrastrutture di supercalcolo, garantendo che siano in grado di soddisfare le esigenze crescenti della comunità scientifica globale.

### 2.3.2 Benchmarking dei Warehouse-Scale Computers (WSC)

I benchmark rappresentano strumenti fondamentali per valutare le prestazioni dei sistemi, permettendo confronti standardizzati tra configurazioni hardware e software, identificando colli di bottiglia e guidando gli sforzi di ottimizzazione. Tuttavia, i benchmark esistenti, come SPEC CPU, sono progettati principalmente per carichi di lavoro tipici di desktop o server tradizionali e non tengono conto delle caratteristiche uniche dei Warehouse-Scale Computers (WSC).
I WSC, sempre più utilizzati dai fornitori di servizi cloud e dalle grandi aziende internet, mostrano caratteristiche prestazionali significativamente diverse rispetto ai sistemi tradizionali. Due sfide principali ostacolano la creazione di benchmark realistici per questi sistemi:

1. **Diversità delle applicazioni**: I WSC eseguono un'ampia varietà di applicazioni, spesso proprietarie, il cui codice sorgente non viene condiviso pubblicamente per motivi di riservatezza.
2. **Dipendenza dai dati di input**: Le prestazioni delle applicazioni nei WSC dipendono fortemente dai dati utilizzati, che spesso contengono informazioni sensibili di aziende o clienti e non possono essere divulgate.

Per affrontare queste limitazioni, lo studio introduce *Fleetbench*, una nuova suite open-source di benchmark progettata per WSC. Questa soluzione si basa su due principi chiave:

- **Codice sorgente rappresentativo**: Fleetbench si concentra su componenti software comuni ai WSC, definiti come il "datacenter tax" (DC tax), che includono attività come allocazione di memoria, trasferimento di dati, compressione e hashing. Molte di queste componenti sono open-source, rendendo Fleetbench rappresentativo senza richiedere codice proprietario.
- **Dati di input realistici**: Per generare input rappresentativi, Fleetbench utilizza un profiler che raccoglie in modo continuo e anonimo i parametri delle chiamate di funzione dai WSC. I dati vengono aggregati in istogrammi per garantire la riservatezza, minimizzando il rischio di esposizione di informazioni sensibili.

Fleetbench rappresenta un passo significativo verso la creazione di benchmark accessibili e realistici per i WSC, superando le limitazioni dei framework tradizionali.

### 2.3.3 Simulazione nei Data Center Geodistribuiti

Con la diffusione globale dei data center, è diventato cruciale sviluppare strumenti in grado di simulare scenari geodistribuiti. GDSim è un esempio rilevante di piattaforma open-source progettata per testare e valutare schedulatori di lavoro in ambienti geodistribuiti. Questa piattaforma permette di simulare condizioni operative realistiche, riproducendo esperimenti con diversi carichi di lavoro e configurazioni topologiche.
GDSim include un generatore di carichi di lavoro che combina dati reali e sintetici per creare scenari complessi. Ad esempio, può generare tracce sintetiche che riproducono caratteristiche osservate in dataset di Facebook, Google e Alibaba, consentendo una valutazione approfondita degli schedulatori in condizioni che altrimenti sarebbero difficili da replicare. Questo approccio è particolarmente utile per esaminare come gli schedulatori gestiscano la latenza di rete e il trasferimento di dati tra data center geograficamente distanti.

### 2.3.4 Utilizzo di SDN nei Data Center per Applicazioni Distribuite

I significativi progressi nelle tecnologie di rete hanno consolidato il ruolo cruciale delle infrastrutture di comunicazione ad alte prestazioni per supportare le operazioni aziendali e abilitare l'erogazione efficiente di servizi digitali, in particolare nell'ambito dei data center. Nei data center moderni, le attività di storage, elaborazione e trasferimento dei dati costituiscono le funzioni fondamentali. La rilevanza del traffico di rete è tale che in ambienti su larga scala, come ad esempio i data center di Facebook, una frazione considerevole del tempo totale di esecuzione di un'applicazione (riportata essere fino al 33%) è consumata specificamente dalle operazioni di trasferimento dati tra i nodi computazionali.

L'esigenza di una gestione più flessibile e programmabile delle risorse di rete ha portato all'emergere del paradigma Software-Defined Networking (SDN). L'SDN ha rivoluzionato le architetture di rete tradizionali introducendo una netta separazione tra il piano di controllo (control plane) e il piano dati (data plane). Questo disaccoppiamento architetturale abilita un controllo logicamente centralizzato della rete, tipicamente gestito da un controller SDN. Il controller possiede una visione globale e in tempo reale dello stato della rete, incluse le informazioni sulle risorse disponibili (come switch e router) e sul carico di traffico istantaneo. Tale visibilità completa consente l'applicazione di algoritmi di routing avanzati e altamente specializzati, capaci di identificare dinamicamente i percorsi ottimali per i diversi flussi di traffico basandosi su politiche predefinite, requisiti delle applicazioni o condizioni di rete attuali. Le decisioni di instradamento e le politiche definite dal controller vengono quindi propagate ai dispositivi di rete del piano dati (switch, router) per aggiornare le loro tabelle di inoltro (flow tables), influenzando così il modo in cui il traffico viene effettivamente inoltrato. L'adozione di protocolli standardizzati per la comunicazione tra il controller e i dispositivi del piano dati, come OpenFlow, ha giocato un ruolo chiave nello standardizzare e facilitare l'implementazione dell'SDN, contribuendo alla sua ampia diffusione, in particolare nella comunità accademica e nelle implementazioni di data center.

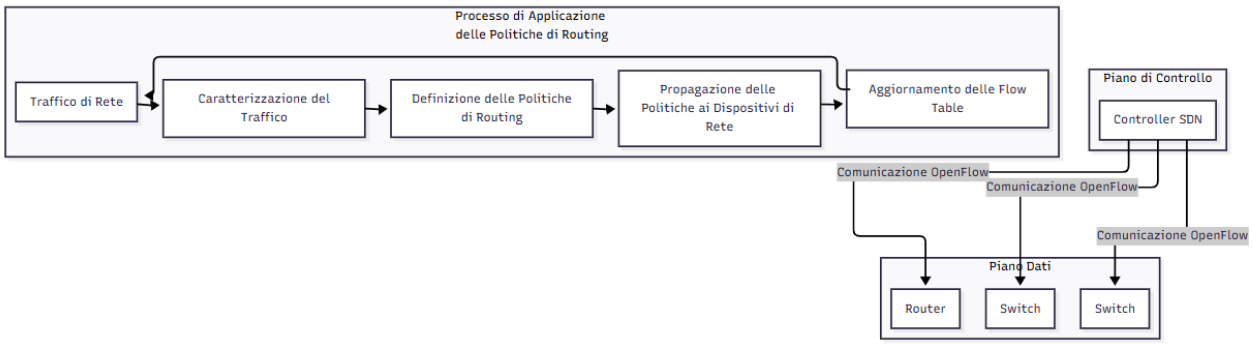


*Figura 7: Processo di applicazione delle politiche di routing in un ambiente Software-Defined Networking (SDN), con evidenza dell'interazione tra piano di controllo e piano dati tramite protocolli standardizzati come OpenFlow.*

La Figura 7 illustra le fasi principali del processo di applicazione delle politiche di routing in un ambiente Software-Defined Networking, evidenziando l'interazione tra il piano di controllo e il piano dati. Come mostrato nel diagramma di flusso, il processo tipicamente parte dalla raccolta e caratterizzazione del traffico di rete per informare la definizione delle politiche di routing a livello del piano di controllo. Queste politiche vengono quindi propagate ai dispositivi di rete nel piano dati attraverso un canale di comunicazione (come OpenFlow), determinando l'aggiornamento delle tabelle di flusso e influenzando l'inoltro del traffico effettivo.

L'applicazione dell'SDN nei contesti dei data center ha stimolato in modo significativo lo sviluppo di tecniche avanzate di traffic engineering e ha promosso una revisione profonda delle politiche di instradamento tradizionali, orientandole maggiormente alle esigenze specifiche delle applicazioni distribuite ospitate. Attraverso la caratterizzazione dettagliata del traffico generato dalle diverse applicazioni (analizzandone pattern di comunicazione, volumi, sensibilità a latenza e perdita pacchetti, ecc.), gli amministratori di rete, supportati dalle capacità di elaborazione e gestione del controller SDN, possono sviluppare e implementare politiche di routing altamente specifiche e granulari. Queste politiche possono includere l'applicazione dinamica di meccanismi avanzati come il bilanciamento del carico (load balancing) basato sul contenuto o sul contesto dell'applicazione, l'implementazione di regole di firewall basate sull'identità dell'applicazione o sul tipo di traffico, e l'adozione di strategie di multi-path routing adattivo per sfruttare al meglio le risorse di banda disponibili e migliorare la resilienza. L'obiettivo primario di tale ottimizzazione del profilo di comunicazione è garantire il soddisfacimento rigoroso dei requisiti di Quality of Service (QoS) specifici per ogni applicazione in termini di banda, latenza, jitter e perdita pacchetti, con la conseguente finalità di migliorare la Quality of Experience (QoE) percepita dagli utenti finali.

### *Applicazione dell'SDN nei Data Center*

Nel contesto di questa analisi, lo studio si è focalizzato sull'evidenziare in modo tangibile i vantaggi derivanti dall'utilizzo del paradigma SDN per l'ottimizzazione del traffico generato da applicazioni distribuite, particolarmente rilevanti in ambienti ad alte prestazioni. In particolare, è stato condotto un'analisi approfondita sul traffico prodotto dalle applicazioni del benchmark NAS (Numerical Aerodynamic Simulation), un set standardizzato di workload computazionali ampiamente impiegato per valutare le performance in ambienti di high-performance computing (HPC). È noto che ciascuna applicazione NAS presenta profili di rete e di elaborazione intrinsecamente differenti, rendendole un caso di studio rappresentativo e stimolante per l'applicazione di tecniche di ottimizzazione del traffico orientate all'applicazione tramite SDN. Lo studio ha investigato due aspetti principali correlati all'utilizzo di SDN:

- **Caratterizzazione del Traffico:** Utilizzando switch SDN capaci di fornire dati telemetrici dettagliati sui flussi, sono stati raccolti e analizzati i dati relativi al traffico di rete generato dalle applicazioni NAS. L'obiettivo era ottenere una comprensione dettagliata e quantitativa dei loro profili di comunicazione sottostanti, essenziale per informare le decisioni di policy.

- **Confronto delle Politiche di Instradamento:** Sono state implementate e valutate le prestazioni di tre diverse politiche di routing all'interno di un ambiente SDN. Il confronto è stato condotto in termini di impatto sul tempo totale di esecuzione (completion time) delle applicazioni NAS, al fine di identificare l'approccio di routing più efficiente nel contesto dell'architettura SDN.

I risultati ottenuti da questa investigazione confermano che l'SDN si rivela una tecnologia estremamente promettente e performante per migliorare l'efficienza, la flessibilità e l'adattabilità della gestione del traffico all'interno dei data center moderni. La sua capacità di centralizzare il controllo logico e di adattarsi dinamicamente alle esigenze specifiche delle applicazioni distribuite, resa possibile dalla caratterizzazione granulare del traffico e dall'ottimizzazione mirata delle politiche di instradamento, consente di sfruttare appieno il suo potenziale, con ricadute positive significative sia sulla Quality of Service che sulla Quality of Experience. Questo studio fornisce un quadro analitico e pratico delle prestazioni ottenibili con l'impiego di SDN, attraverso un'analisi dettagliata basata su applicazioni HPC reali e un confronto empirico di diverse politiche di routing, contribuendo così allo sviluppo di soluzioni avanzate per la gestione intelligente, programmabile e ottimizzata del traffico nei data center di nuova generazione.

### 2.3.5 Algoritmi eseguiti su GPU

L'allocazione delle risorse nei data center è un problema complesso e computazionalmente intensivo, soprattutto nei contesti di infrastrutture virtuali. Gli algoritmi accelerati da GPU offrono una soluzione innovativa, consentendo di sfruttare il parallelismo massivo delle GPU per eseguire calcoli complessi in modo più rapido rispetto ai tradizionali approcci basati su CPU.
Ad esempio, algoritmi di grafi come Dijkstra, K-Means e PageRank sono stati adattati per funzionare su GPU utilizzando tecniche di rappresentazione sparse, come la Compressed Sparse Row. I risultati dimostrano che questi algoritmi possono ridurre significativamente i tempi di esecuzione, migliorando la scalabilità nei data center di grandi dimensioni.

### 2.3.6 Metriche Multi-criterio per Benchmarking

Tradizionalmente, il benchmarking si è concentrato sulla velocità di esecuzione come metrica primaria. Tuttavia, con l'aumento delle preoccupazioni per l'efficienza energetica e la sostenibilità, sono emerse metriche olistiche che tengono conto di tempo, consumo energetico e utilizzo dello spazio.
Una metodologia proposta integra questi fattori in un unico indicatore di performance, fornendo una valutazione più completa delle piattaforme. Ad esempio, il parametro H combina il tempo di esecuzione con il numero di unità fisiche necessarie, rappresentando un approccio più realistico per confrontare diverse architetture hardware.
Metriche come il rapporto tra prestazioni e consumo energetico (*Performance per Watt*) e il *Total Cost of Ownership* (TCO) sono sempre più adottate per fornire un'analisi più dettagliata dell'impatto operativo delle infrastrutture IT.
Inoltre, la crescente attenzione alla sostenibilità ha portato allo sviluppo di benchmark specifici per misurare l'efficienza energetica dei data center, come SPECpower e Green500, che valutano il consumo energetico rispetto alla potenza computazionale erogata. Questi strumenti permettono di individuare le migliori strategie di ottimizzazione per ridurre il consumo energetico senza compromettere le prestazioni complessive del sistema.

#### *Efficienza Energetica nei Data Center e nella Gestione dei Big Data*

L'efficienza energetica nei data center è stata al centro della ricerca per diversi anni, con particolare attenzione alla gestione energetica delle tecnologie e delle infrastrutture IT, inclusi server, sistemi di storage e reti. I sistemi di riscaldamento, ventilazione e condizionamento dell'aria (HVAC) costituiscono una componente significativa nei data center moderni, influenzando profondamente il design di nuove tecnologie e topologie energeticamente consapevoli.

I data center devono evolvere per operare con maggiore efficienza energetica, adattando dinamicamente la domanda energetica in base alla disponibilità di fonti di energia rinnovabile, come il sole e il vento. Questo approccio, che definisce i data center eco-aware, consente di:

- Integrare flessibilmente nuove risorse nell'infrastruttura esistente.
- Ridurre i picchi di domanda intensivi di CO2.
- Sostenere un utilizzo adattivo dell'energia, un elemento chiave per integrare fonti rinnovabili nella produzione energetica.

L'efficienza energetica non è un obiettivo finale ma una tappa evolutiva per i data center. Le infrastrutture devono supportare una gestione energetica dinamica e adattabile, che introduce sfide significative nella progettazione dell'architettura, della comunicazione e dell'infrastruttura dei data center, inclusi quelli tradizionali, basati su cloud e destinati al calcolo ad alte prestazioni (HPC).

### *Gestione dei Carichi di Lavoro e Virtualizzazione*

Un'area di ricerca fondamentale rimane la gestione energetica dei carichi di lavoro. La virtualizzazione svolge un ruolo cruciale facilitando l'integrazione di diverse tecnologie per la fornitura di risorse e servizi energeticamente efficienti. Tuttavia, l'elaborazione dei big data presenta sfide uniche, poiché i set di dati di grandi dimensioni e complessi non possono essere gestiti efficacemente con strumenti tradizionali di gestione dei database. Tra le aree critiche del calcolo dei big data si annoverano la cattura, la conservazione, lo storage, le operazioni e l'analisi dei dati.

### *Consumo Energetico nei Data Center*

Uno studio sul consumo energetico nei data center rivela che il carico computazionale rappresenta il 63% del consumo totale, mentre i sistemi HVAC utilizzano il 21%, e il restante 14% è destinato a illuminazione e perdite dei sistemi UPS. Questo sottolinea la necessità di definire metriche coerenti per la gestione energetica dei data center e di migliorare la comprensione del consumo energetico dei big data.

È bene quindi prendere in considerazione i vari aspetti sopra citati, sottolineando:

- L'importanza dell'efficienza energetica e della gestione adattiva della domanda.
- Le opportunità offerte dall'integrazione delle fonti di energia rinnovabile.
- Le sfide legate alla gestione dei big data in termini di consumo energetico e progettazione di infrastrutture sostenibili.

Lo sviluppo di metriche per valutare e gestire l'efficienza energetica nei data center rimane una priorità, offrendo un quadro per migliorare la sostenibilità e ridurre l'impatto ambientale.

# 3. Obiettivi e metodi di benchmarking

## 3.1 Definizione degli obiettivi

Definire chiaramente gli obiettivi del benchmarking è essenziale per garantire che le metriche e i risultati ottenuti siano rilevanti e applicabili al contesto specifico. Questo capitolo esplora gli obiettivi principali del benchmarking nei data center, le categorie di analisi e le implicazioni per il miglioramento delle infrastrutture IT, tenendo conto tuttavia che spesso le valutazioni sperimentali non catturano adeguatamente i fenomeni che emergono nei grandi data center, come guasti hardware o ritardi su scala elevata. Anche i modelli di simulazione attuali presentano limitazioni legate alla rigidità dei percorsi di trasmissione e alla modellazione superficiale dei sistemi coinvolti

### 3.1.1 Obiettivi Generali

Gli obiettivi del benchmarking possono essere suddivisi in diverse categorie, a seconda del contesto applicativo. Le finalità principali includono:

- **Valutazione delle Prestazioni**: Misurare le capacità computazionali, di storage e di rete di un data center rispetto a standard di riferimento.
- **Ottimizzazione dell'Efficienza Energetica**: Ridurre il consumo energetico e migliorare la sostenibilità ambientale.
- **Confronto tra Tecnologie**: Analizzare diverse configurazioni hardware e software per determinare le soluzioni più performanti.
- **Scalabilità e Affidabilità**: Valutare la capacità di un sistema di adattarsi a carichi di lavoro variabili mantenendo elevati livelli di affidabilità.
- **Costi e Rendimento**: Ottimizzare il rapporto tra costi operativi e prestazioni per garantire un uso efficiente delle risorse.

### 3.1.2 Categorie di Benchmarking

A seconda degli obiettivi, il benchmarking può essere suddiviso nelle seguenti categorie:

- **Benchmarking Funzionale**: Misura le prestazioni di specifiche funzionalità di un sistema, come il throughput o la latenza.
- **Benchmarking Competitivo**: Confronta le prestazioni di un data center con quelle di concorrenti o di best practice di settore.
- **Benchmarking Interno**: Analizza e confronta le prestazioni di diversi sistemi all'interno della stessa organizzazione per identificare aree di miglioramento.
- **Benchmarking Collaborativo**: Consente di condividere dati prestazionali tra organizzazioni per migliorare collettivamente l'efficienza dei data center.

## 3.2 Definizione dei test e tool di benchmark

Il presente paragrafo definisce un piano di test e benchmark per la valutazione delle prestazioni dei data center. L'obiettivo è misurare e confrontare diverse configurazioni hardware e software al fine di ottimizzare l'efficienza operativa, il consumo energetico e le prestazioni generali dei sistemi di elaborazione dati.

### 3.3.1 Metriche di Valutazione

- **Prestazioni computazionali:** tempo di esecuzione (ms), throughput (TPS - Transactions Per Second), latenza media (ms).
- **Efficienza energetica:** consumo medio di energia (W), PUE (Power Usage Effectiveness), efficienza CPU/GPU (% utilizzo).

- **Scalabilità:** tempo di provisioning di nuove istanze, variazione del throughput con l'aumento del carico.
- **Affidabilità:** tempo medio tra guasti (MTBF), tempo medio di ripristino (MTTR).
- **Utilizzo delle risorse:** percentuale di utilizzo di CPU, RAM e storage.

### 3.3.2 Tool di benchmark

L'utilizzo di benchmark consolidati permette di ottenere dati comparabili nel tempo e tra diverse configurazioni hardware o software, fornendo una base oggettiva su cui effettuare scelte progettuali informate. Di seguito vengono descritti i principali strumenti utilizzati per ciascuna categoria critica all'interno dell'ecosistema dei data center.

#### *Prestazioni Computazionali*

La potenza elaborativa dei server rappresenta uno degli aspetti più importanti nella definizione delle capacità di un data center. I tool di benchmarking dedicati alla CPU consentono di valutare il comportamento del processore sotto carichi tipici e di confrontare soluzioni diverse sulla base di metriche comuni.

- **SPEC CPU**
  Questo benchmark, sviluppato dal consorzio Standard Performance Evaluation Corporation (SPEC), si concentra sulle prestazioni della CPU in scenari realistici. Include sia workload orientati al calcolo scientifico che applicazioni commerciali, ed è suddiviso in due macrocategorie: SPECint (per operazioni intere) e SPECfp (per operazioni in virgola mobile). È comunemente utilizzato per testare l'efficienza dei server in ambienti data-intensive.
- **TPC (Transaction Processing Performance Council)**
  Il TPC propone una serie di benchmark specifici per il mondo delle transazioni e dell'analisi dei dati. Tra i più noti:
  - TPC-C , incentrato sui sistemi OLTP (Online Transaction Processing), simula un ambiente aziendale con molteplici operazioni concorrenti, come ordini, pagamenti e consultazioni.
  - TPC-H , invece, si rivolge ai sistemi OLAP (Online Analytical Processing), valutando la capacità di eseguire query complesse su grandi dataset.
- **Linpack**
  Originariamente sviluppato per il calcolo numerico, Linpack è ampiamente utilizzato per misurare le performance dei supercomputer e dei cluster HPC (High Performance Computing). La sua versione parallela, MPI Linpack, è impiegata per classificare i sistemi presenti nella lista TOP500.
- **NAS Parallel Benchmarks (NPB)**
  Sviluppati dalla NASA, questi benchmark sono specificatamente pensati per valutare le performance di sistemi paralleli nell'ambito del calcolo scientifico. Simulano applicazioni HPC tipiche, come fluidodinamica computazionale e simulazioni climatiche.

#### *Big Data e Intelligenza Artificiale*

Con l'aumento esponenziale di dati generati ogni giorno, i data center devono affrontare carichi di lavoro sempre più complessi, spesso basati su framework distribuiti e tecnologie di machine learning. Per misurare le capacità di tali architetture, sono stati sviluppati benchmark ad hoc.

- **HiBench**
  HiBench è uno strumento open-source che permette di valutare le performance di framework Big Data come Apache Hadoop, Spark e Flink. Include workload di sorting, analisi di testo e operazioni di aggregazione, offrendo metriche come throughput e latenza media.
- **TPCx-BB (Benchmark for Big Data)**
  Sviluppato da TPC, questo benchmark simula un ambiente reale di analisi dati basato su piattaforme

come Apache Spark e Hadoop. Misura la capacità del sistema di gestire pipeline di dati complesse e di rispondere a interrogazioni sofisticate su dataset molto grandi.

- **DeepBench**
  Creato da NVIDIA, DeepBench è focalizzato sull'efficienza delle operazioni fondamentali utilizzate nelle reti neurali. Misura il tempo necessario per eseguire operazioni come convoluzioni e prodotti matriciali su hardware accelerato (GPU, TPU), permettendo di confrontare differenti piattaforme per l'AI.

### *Storage*

Le prestazioni dello storage giocano un ruolo chiave nel determinare l'efficienza complessiva di un sistema, soprattutto in contesti in cui avvengono accessi frequenti a grandi quantità di dati. Esistono diversi strumenti per valutare la velocità, la capacità e l'affidabilità dei dispositivi di memorizzazione.

- **fio (Flexible I/O Tester)**
  fio è uno strumento estremamente versatile che permette di simulare quasi ogni tipo di carico di lavoro di lettura/scrittura su storage locali o remoti. Può essere utilizzato per testare SSD, HDD, NAS e array SAN, fornendo dettagli precisi su IOPS, velocità di trasferimento e latenze.
- **IOMeter**
  Questo benchmark è stato storicamente utilizzato per misurare le prestazioni di sistemi di storage in ambienti enterprise. Fornisce una visione completa sul throughput di I/O e sull'adattamento del sistema a carichi multi-threaded.
- **SPECstorage**
  Parte del portfolio SPEC, questo benchmark valuta le prestazioni complessive di sistemi di storage distribuiti. Misura il numero di operazioni al secondo (IOPS) e la capacità di scalare in presenza di carichi crescenti, risultando particolarmente utile per il testing di soluzioni cloud-native.

### *Rete*

La rete rappresenta il collegamento vitale tra tutti i componenti di un data center. Una corretta valutazione delle sue capacità consente di evitare colli di bottiglia e garantire comunicazioni efficienti tra nodi distribuiti.

- **iperf**
  iperf è uno strumento leggero ma efficace per testare la larghezza di banda tra due host. Consente di valutare la capacità massima di trasferimento dati su reti cablate e wireless, nonché di identificare eventuali problemi di congestione.
- **Netperf**
  Netperf offre una visione più approfondita delle performance di rete, includendo la misurazione della latenza, del throughput e del comportamento under load. È comunemente utilizzato per analizzare le prestazioni di stack TCP/IP e protocolli alternativi.
- **Pktgen**
  Pensato principalmente per ambienti SDN (Software Defined Networking), Pktgen genera traffico di rete sintetico ad alte prestazioni. È ideale per testare la capacità di switching e routing in contesti virtualizzati o containerizzati.

### *Efficienza Energetica*

Negli ultimi anni, l'attenzione verso l'efficienza energetica nei data center è cresciuta considerevolmente, anche in relazione agli obiettivi di sostenibilità ambientale e alla riduzione dei costi operativi. Gli strumenti di monitoraggio energetico permettono di valutare il consumo di energia in relazione alle prestazioni erogate.

- **SPECpower_ssj2008**
  Questo benchmark misura la relazione tra il consumo energetico e le prestazioni di un server. Durante

l'esecuzione di un carico di lavoro scalabile, registra il consumo in Watt e lo rapporta al livello di attività del sistema, fornendo un'indicazione precisa del rendimento energetico.

- **Green500**
  Sebbene non sia uno strumento di benchmarking autonomo, Green500 è una classifica annuale dei supercomputer più efficienti energeticamente. Utilizza i dati del benchmark Linpack e li integra con i consumi registrati, permettendo un confronto diretto tra efficienza e potenza.
- **PowerAPI**
  PowerAPI è un framework open-source che permette di monitorare il consumo energetico dei componenti del sistema (CPU, GPU, ecc.) direttamente da software. Si integra facilmente in ambienti di testing automatizzati e può essere utilizzato per tracciare il consumo durante l'esecuzione di qualsiasi workload.

### 3.3.3 Analisi delle prestazioni e dell'efficienza

Nel contesto della valutazione complessiva delle prestazioni di un data center destinato al trattamento di carichi di lavoro avanzati, come quelli tipici degli ambienti Big Data e cloud computing, si rende necessaria una serie di test strutturati. Questi test permettono di misurare le capacità del sistema in termini di elaborazione, efficienza energetica, scalabilità, affidabilità e utilizzo ottimale delle risorse hardware. Di seguito vengono descritti i principali test effettuati, con particolare attenzione alle metodologie adottate, alle metriche utilizzate e a esempi rappresentativi per ciascun caso.

#### *Test di Performance Computazionale*

- **Obiettivo:** misurare il tempo di elaborazione richiesto per completare workload tipici di tipo Big Data, al fine di valutare la capacità computazionale del sistema sottoposto a diversi scenari operativi.
- **Metodologia:** i test vengono condotti utilizzando HiBench, uno strumento di benchmarking comunemente impiegato per valutare le performance di framework distribuiti come Hadoop e Spark. Vengono configurate diverse macchine virtuali con differenti specifiche hardware (ad esempio, numero di core CPU, quantità di RAM e velocità di rete). Su ciascuna configurazione si esegue lo stesso set di workload, comprendente operazioni di sorting, wordcount e TeraSort.
- **Metriche misurate:**
  - Throughput (numero di record processati al secondo)
  - Latenza media per task
  - Tempo totale di esecuzione
- **Esempio:** in un ambiente dotato di 8 core CPU e 32 GB di RAM, l'esecuzione del workload "WordCount" su un dataset da 10 GB ha richiesto un tempo medio di 78 secondi, con un throughput di circa 128 MB/s. In un sistema meno performante (4 core e 16 GB di RAM), lo stesso workload ha richiesto 125 secondi, evidenziando una riduzione significativa nell'efficienza computazionale.

#### *Test di Efficienza Energetica*

- **Obiettivo:** analizzare il consumo energetico del sistema in relazione al carico di lavoro elaborativo, al fine di valutare l'efficienza dal punto di vista dei costi operativi e della sostenibilità ambientale.
- **Metodologia:** per questi test occorre utilizzare uno strumento come SPECpower_ssj2008, che permette di monitorare il consumo energetico in relazione alla potenza elaborativa erogata. I server vengono sottoposti a livelli crescenti di carico (dal 10% al 100%) e si registrano i consumi elettrici associati. L'analisi va effettuata su sistemi con diverse architetture hardware (ad esempio, processori Intel Xeon vs AMD EPYC).
- **Metriche misurate:**
  - PUE (Power Usage Effectiveness): rapporto tra l'energia totale consumata dal data center e quella effettivamente utilizzata dai server.

    - Consumo energetico per unità di elaborazione (Watt per operazione)
- **Esempio:** un server equipaggiato con un processore Intel Xeon Silver ha mostrato un consumo medio di 220 W durante l'esecuzione di un workload medio-intensivo, con un PUE di 1.45. Un server equivalente basato su tecnologia ARM ha invece mostrato un consumo di 180 W nello stesso scenario, con un miglioramento del 18% in termini di efficienza energetica.

### *Test di Scalabilità*

- **Obiettivo:** verificare la capacità del sistema di adattarsi all'aumentare del carico operativo, valutando come varia il throughput in seguito all'aggiunta di nuove istanze all'interno di un ambiente cloud.
- **Metodologia:** occorre creare cluster sulla piattaforma cloud configurata con immagini di macchine virtuali standard. Si parte da un minimo di 2 nodi fino ad arrivare a un massimo di 10 nodi. Per ogni livello di scalabilità, viene eseguito lo stesso workload rappresentativo, quale un job MapReduce su dati sintetici di grandi dimensioni. Dopo ogni aggiunta di nodi, vanno misurate le variazioni nel throughput e il tempo necessario per il provisioning delle nuove risorse.
- **Metriche misurate:**
    - Tempo di provisioning (tempo richiesto per avviare e integrare nuovi nodi al cluster)
    - Aumento percentuale del throughput in relazione al numero di nodi aggiunti
- **Esempio:** con una configurazione iniziale composta da 2 nodi, il sistema registrava un throughput medio di 300 MB/s. Incrementando progressivamente il numero di nodi fino a 6, si osserva un aumento lineare del throughput fino a raggiungere i 900 MB/s. Tuttavia, aggiungendo ulteriori nodi oltre questa soglia, l'aumento del throughput rallenta, indicando un incremento dell'overhead legato alla comunicazione inter-nodo e alla gestione distribuita del carico.

### *Test di Affidabilità*

- **Obiettivo:** valutare la resilienza del sistema in presenza di guasti hardware o software, e verificare la sua capacità di ripristinare automaticamente i servizi senza interruzioni critiche.
- **Metodologia:** occorre simulare guasti volontari, come il crash improvviso di un nodo master o lo spegnimento forzato di un disco rigido virtuale, all'interno di un ambiente cluster gestito da Kubernetes o Hadoop. Occorre misurare il tempo trascorso tra il rilevamento del guasto e il ripristino completo del servizio.
- **Metriche misurate:**
    - MTBF (Mean Time Between Failures): tempo medio tra due guasti consecutivi
    - MTTR (Mean Time To Repair): tempo medio necessario per riparare il guasto e riprendere l'operatività
- **Esempio:** durante una simulazione di crash del nodo master in un cluster Hadoop, il sistema ha impiegato circa 25 secondi per rilevare il guasto e promuovere un nuovo nodo standby. Il servizio è tornato pienamente operativo dopo ulteriori 15 secondi, risultando in un MTTR complessivo di 40 secondi.

### *Test di Utilizzo delle Risorse*

- **Obiettivo:** misurare l'utilizzo effettivo delle risorse hardware fondamentali (CPU, RAM e storage) durante l'esecuzione di workload tipici, al fine di identificare eventuali colli di bottiglia o sprechi di risorse.
- **Metodologia:** durante l'esecuzione di diversi workload (sorting, analisi dati, machine learning), vengono raccolti dati continui sull'utilizzo della CPU, della memoria RAM e dello storage tramite strumenti come htop , iostat e vmstat . Le informazioni vanno registrate sia in tempo reale che aggregate per fornire un'immagine completa del comportamento del sistema.
- **Metriche misurate:**

    - Percentuale di utilizzo della CPU
    - Occupazione della memoria RAM (MB/GB utilizzati)
    - IOPS (Input/Output Operations Per Second) e capacità residua dello storage
- **Esempio:** durante l'esecuzione di un job di clustering K-Means su un dataset da 50 GB, si è osservato un picco di utilizzo della CPU intorno al 92%, mentre la RAM ha raggiunto un uso massimo del 78% (di un totale di 64 GB disponibili). Lo storage, invece, ha mostrato un'intensa attività di lettura/scrittura con picchi di 1200 IOPS, causando occasionali rallentamenti nella fase di persistenza dei risultati intermedi.

# 4. Conclusioni

Il presente rapporto ha messo in evidenza come il benchmarking rappresenti un pilastro essenziale per l'analisi, la valutazione e il miglioramento delle prestazioni nei data center di nuova generazione. In un contesto in cui le infrastrutture digitali si evolvono rapidamente per rispondere alle crescenti richieste derivanti dai Big Data, dall'Intelligenza Artificiale e dai carichi computazionali su larga scala, emerge con forza la necessità di disporre di strumenti e metodologie di benchmarking sempre più sofisticati, adattabili e multidimensionali.

L'analisi ha evidenziato l'importanza delle innovazioni architetturali in ambito networking, in particolare l'adozione del modello *Spine-Leaf All-Optical (AOSL)*, che offre vantaggi significativi in termini di larghezza di banda, latenza e gestione della congestione. L'integrazione con soluzioni emergenti come *Rockport Networking* apre ulteriori prospettive per l'ottimizzazione delle prestazioni, grazie a un'architettura switchless, all'utilizzo di interconnessioni ottiche avanzate e a un miglior controllo distribuito del traffico.

Sul versante applicativo, il documento ha affrontato in maniera critica il ruolo dei benchmark standardizzati per l'analisi dei Big Data, sottolineando le attuali limitazioni legate all'uso di dataset sintetici e all'inefficienza dei carichi di lavoro ridondanti. In tale contesto, l'emergente paradigma *DataFlow* si configura come un approccio promettente per il superamento delle rigidità tradizionali, aprendo nuove possibilità per la valutazione realistica e mirata delle applicazioni.

Altri ambiti di rilievo approfonditi nel rapporto includono il benchmarking per sistemi di supercalcolo, *Warehouse-Scale Computers* (WSC), ambienti geodistribuiti, infrastrutture SDN e algoritmi di allocazione su GPU, con particolare attenzione alle metriche multi-criterio e all'efficienza energetica. Quest'ultima si conferma come parametro prioritario per i data center orientati alla sostenibilità (*eco-aware*), i quali richiedono soluzioni in grado di bilanciare carichi di lavoro, ottimizzare il consumo energetico e integrare fonti rinnovabili.

Il documento ha infine delineato una metodologia operativa per l'implementazione del benchmarking, articolata in fasi ben definite che comprendono l'identificazione degli obiettivi, la selezione delle metriche, la configurazione dell'ambiente di test e l'impiego di strumenti dedicati per la valutazione delle principali dimensioni prestazionali.

In conclusione, questo rapporto ha fornito un quadro esaustivo e aggiornato del benchmarking nei data center, mettendo in luce le principali sfide tecniche, le opportunità emergenti e le direzioni evolutive per il futuro. La continua innovazione nelle architetture, nelle metriche e negli strumenti di valutazione costituirà un fattore abilitante per data center sempre più performanti, scalabili, sostenibili e in grado di rispondere alle esigenze complesse e dinamiche dell'ecosistema digitale contemporaneo.

# *Ringraziamenti*

*Desideriamo esprimere la nostra sincera gratitudine a Simona Sada per il prezioso supporto tecnico e amministrativo fornito durante lo svolgimento delle attività descritte nel presente rapporto tecnico.*

*Un sentito ringraziamento va inoltre al progetto FOSSR (https://www.fossr.eu/) per averci offerto l'opportunità di approfondire le nostre competenze nell'ambito del benchmarking dei data center e di contribuire allo sviluppo di buone pratiche nel settore.*