

Progetto GAINS

(Generative AI for Network Sustainability)

Deliverable D1.2

Report sullo Stato dell'Arte su Impatto Energetico della Infrastruttura di Rete e Cloud

McGann, O. - Giordano A. - Haider, B. - Paesani, M. | 2025-11

Work Package: WP1 Problem Framing ed Analisi dello Stato dell'Arte

Stato: Definitivo | Versione: 1.0

Classificazione: uso interno di progetto; condivisione controllata con partner su necessità

Abstract

Questo report presenta lo stato dell'arte sull'impatto energetico delle infrastrutture di rete e cloud, con un focus integrato su consumi, emissioni e metodi di misura, modellazione e predizione applicabili lungo l'intera catena rete-data center-piattaforma. In un contesto in cui la domanda di servizi digitali cresce rapidamente e l'energia diventa un vincolo di progetto, la letteratura più recente mette in risalto due evidenze: (i) l'aumento di traffico e capacità non si traduce automaticamente in crescita proporzionale dei consumi grazie ai progressi di efficienza, ma (ii) l'adozione di workload AI e l'incremento della densità di calcolo stanno riaprendo un fronte critico su potenza disponibile, raffreddamento e carbon footprint, richiedendo misure operative più solide e attribuibili per servizio. La revisione integra contributi scientifici, standard tecnici e indicazioni industriali per costruire un quadro coerente di metriche e metodi: dai KPI di sito per data center, alle metriche per apparati e reti. L'evidenza che emerge è che l'efficienza non può essere valutata solo a livello di struttura o datasheet: la non energy-proportionality, l'elevato consumo in idle e l'eterogeneità multi-vendor impongono una catena di telemetria che unisca misure fisiche e logiche, abiliti benchmarking comparabile e supporti l'attribuzione per workload/slice/servizio. Sul piano metodologico, il report sintetizza survey e tassonomie recenti sul monitoraggio e la predizione dell'impatto energetico nelle reti e nei data center, evidenziando l'evoluzione da approcci prevalentemente qualitativi

e scenario-based verso modelli data-driven e capacità predittive integrate in architetture softwarizzate (SDN/NFV/Network Slicing). Vengono inoltre richiamate evidenze sperimentali che dimostrano risparmi energetici significativi tramite gestione dinamica di lightpath in reti ottiche elastiche e strategie SDN basate su consolidamento dei flussi e minimizzazione delle transizioni di stato, con parametri di consumo e assunzioni operative. Infine, lo stato dell'arte indica che l'automazione energy-aware e carbon-aware, supportata da telemetria fine-grained e da strumenti interoperabili, è la leva più promettente per ridurre over-provisioning e emissioni mantenendo SLA, auditabilità e governance. Il report traduce tali evidenze in criteri replicabili per il progetto, proponendo un impianto "measurement-first" e un set di KPI e guardrail operativi per valutare soluzioni di rete e cloud orientate alla sostenibilità.

Indice

1. Introduzione

- 1.1 Contesto industriale e motivazione
- 1.2 Obiettivo report e perimetro del deliverable
- 1.3 Domande di ricerca operative

2. Metodi

- 2.1 Strategia di ricerca e fonti
- 2.2 Criteri di selezione dello stato dell'arte
- 2.3 Disegno sperimentale replicabile

3. Risultati

- 3.1 Evidenze macro: crescita della domanda e vincoli di potenza
- 3.2 Sfide e obiettivi per reti energy-aware
- 3.3 Metriche e standard per data center e cloud
- 3.4 Metriche e standard per apparati e reti
- 3.5 Driver tecnici del consumo: perché l'over-provisioning pesa più del previsto
- 3.6 Virtualizzazione, Network Function Virtualisation e Network Slicing

3.7 Telemetria e accounting energetico: da misura di sito a misura per servizio

3.8 AI per allocazione dinamica e riduzione delle emissioni: energy-aware e carbon-aware

3.9 Trend industriali: AI factories, sovranità del dato e sostenibilità come vincolo

3.10 Survey e rassegne correlate su monitoraggio e predizione dell'impatto energetico dei componenti di rete

3.11 Approfondimenti: monitoraggio, ottimizzazione, modelli di potenza e gestione energy-aware

3.11.1 Monitoraggio della rete

3.11.2 Ottimizzazione dei consumi energetici

3.11.3 Modello di potenza

3.11.4 Gestione energy-aware

4. Discussione

4.1 Sintesi critica: misurare, attribuire, controllare

4.2 Leve ad alto impatto e basso rischio strategico

4.3 Implicazioni per l'intero progetto: integrazione con automazione agentica

4.4 Replicabilità industriale e licenze: evitare lock-in e massimizzare trasferibilità

4.5 Lezioni apprese e raccomandazioni operative

4.6 Sintesi integrata delle evidenze e implicazioni operative

4.7 Gap di ricerca e opportunità per GAINS

4.8 Scelte di progetto

Bibliografia