

Progetto GAINS

(Generative AI for Network Sustainability)

Deliverable D2.1

Rapporto su obiettivi, risultati, e lessons learned relative al Task T2.1
Progettazione del Sistema

Giordano, A. - Basile, L. - De Giacomo, A. - Haider, B. - Fruncillo, D. - McGann, O.

2026-01

Work Package: WP2 Progettazione del Sistema, Integrazione degli Agenti e Validazione del Proof-of-Concept

Stato: Definitivo | Versione: 1.0

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

Abstract

In questo report si descrive in dettaglio la progettazione del sistema di automazione e supporto alle decisioni del progetto GAINS. Vengono presentati obiettivi, risultati di design e lessons learned relativi all'architettura ad agenti e ai principali artefatti progettuali (requisiti, modello dati, contratti di interazione e flussi operativi) che abilitano l'integrazione dei componenti e la successiva validazione del proof-of-concept. Il sistema è concepito per fornire supporto decisionale automatizzato lungo un flusso di lavoro end-to-end che include ingestione e normalizzazione di dati eterogenei, analisi di stato e trend, generazione di modelli di deployment standardizzati e ottimizzati per la sostenibilità ambientale, e produzione di documentazione tecnica e procedurale a corredo. L'architettura adotta un design modulare che separa le funzioni di osservabilità, inferenza e controllo e le organizza in tre agenti principali, interoperanti: (i) Network Monitor, per raccolta e analisi in sola lettura della telemetria, costruzione di baseline storiche, rilevazione di anomalie e stime di impatto; (ii) Knowledge Manager, per ingest, indicizzazione e sintesi di conoscenza tecnico-operativa e vincoli (vendor/ruolo/versione), con produzione di blueprint riusabili; (iii) Network Orchestrator, per traduzione dell'intent in candidate configuration governate, corredate da change plan, rollback e valutazione dell'impatto atteso. La componente Climate è integrata come oggetto canonico del workflow e abilita una valutazione attribuibile per device/servizio/configurazione, distinguendo esplicitamente tra misure dirette e stime model-based/hybrid, con assunzioni e livelli di confidenza dichiarati. La progettazione include inoltre presidi di governance e sicurezza (audit/traceability, human-in-the-loop, controlli di consistenza e politiche di accesso) necessari a garantire coerenza operativa e riduzione del rischio, mantenendo la modularità del sistema e la possibilità di evoluzione indipendente dei singoli agenti nelle fasi successive della progettazione.

Indice

1. Introduzione

1.1 Scopo e obiettivi del sistema

1.2 Perimetro e contesto di riferimento

1.3 Architettura ad agenti e paradigma divide et impera

1.4 Replicabilità industriale e vincoli progettuali

2. Requisiti del sistema

2.1 Requisiti funzionali

2.1.1 Network Monitor

2.1.2 Knowledge Manager

2.1.3 Network Orchestrator

2.1.4 Interazione utente e integrazione

2.2 Requisiti non funzionali

2.2.1 Sicurezza e segregazione

2.2.2 Auditabilità, osservabilità e riproducibilità

2.2.3 Prestazioni e affidabilità

2.3 Requisiti footprint/climate

2.4 Tabella requisiti e criteri di verifica

2.5 Matrice di tracciabilità Requisiti/Interfacce/Artefatti e Test case

3. Architettura del sistema

3.1 Principi architetturali

3.2 Vista logica: componenti e flussi

3.3 Component model: componenti runtime

3.3.1 GAINS Control Plane

3.3.2 Network Monitor

3.3.3 Knowledge Manager

3.3.4 Network Orchestrator

3.3.5 Datastore e artifact management

3.4 Interfacce e confini di responsabilità

3.5 Workflow architetturali principali

3.6 Considerazioni di deployment e profilo prototipale

4. Componenti e interfacce

4.1 Convenzioni di contratto versioning, refs, error model, RBAC

4.2 Interfacce northbound del Control Plane

4.2.1 Monitor API

4.2.2 Knowledge API

4.2.3 Orchestration API

4.3 Interfacce interne Control Plane - agenti

4.4 Confini di responsabilità contrattualizzati

4.5 Nota di integrazione UI user-friendly

5. Data Model di dettaglio

5.1 Principi del modello dati

5.2 Entità canoniche e campi minimi

5.3 Knowledge store: classi/oggetti proposti

5.3.1 DocSource

5.3.2 DocChunk

5.3.3 Policy

5.3.4 ServiceDefinition

5.3.5 ConfigSnippet

5.3.6 Blueprint

5.3.7 ConfigSnapshotSanitized

5.4 Artifact store: oggetti proposti e campi minimi

5.4.1 OrchestrationJob

5.4.2 CandidateConfigurationArtifact

5.4.3 ChangePlanArtifact

5.4.4 FootprintEstimateArtifact

5.4.5 ApprovalDecisionArtifact

5.5 Regole di indicizzazione knowledge store

5.6 Regole di indicizzazione artifact store

6. Sequence diagram dei flussi principali

6.1 Sequence overview (flussi A-E)

6.2 Sequence A Monitor query

6.3 Sequence B Knowledge query + blueprint

6.4 Sequence C Intent-to-config job

6.5 Sequence D Approval

6.6 Sequence E Ingest

7. Razionale delle scelte tecnologiche e dei tool

7.1 Criteri di selezione

7.2 Stack selezionato e ruolo nel sistema

7.3 Coerenza delle scelte per agente

7.4 Alternative considerate e motivazioni di scarto

7.5 Conclusioni

Bibliografia