

S3 federato per la ricerca: design e confronto con storage geo-distribuito

Stefano Diciotti

Alma Mater Studiorum – Università di Bologna

I Big Data della ricerca

Dati scientifici

- Acquisizione strumentale
- Simulazioni HPC
- AI e analisi dati
- Imaging medico

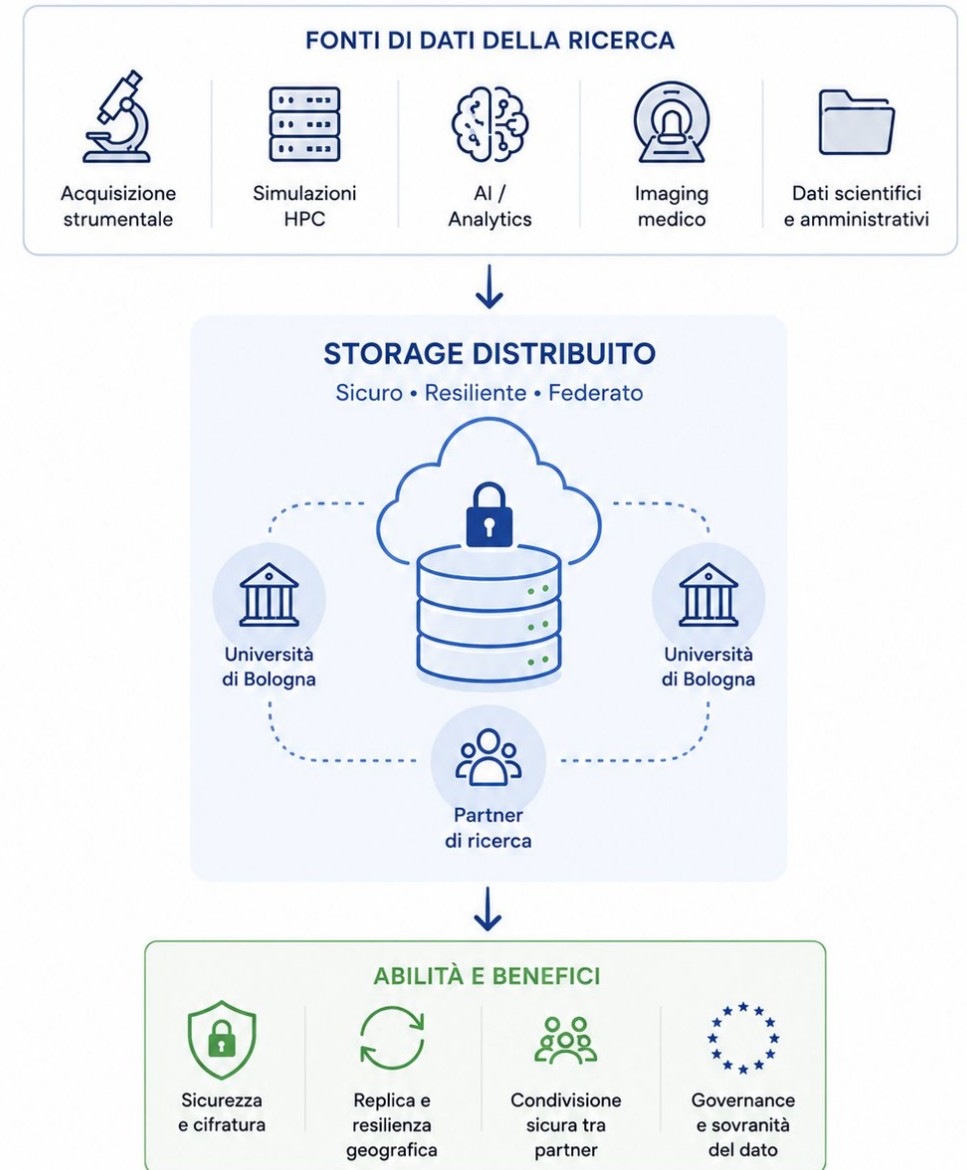
Esigenze

- Archiviazione sicura e cifrata
- Replica e resilienza geografica
- Condivisione tra partner di progetto
- Governance e sovranità del dato

Criticità

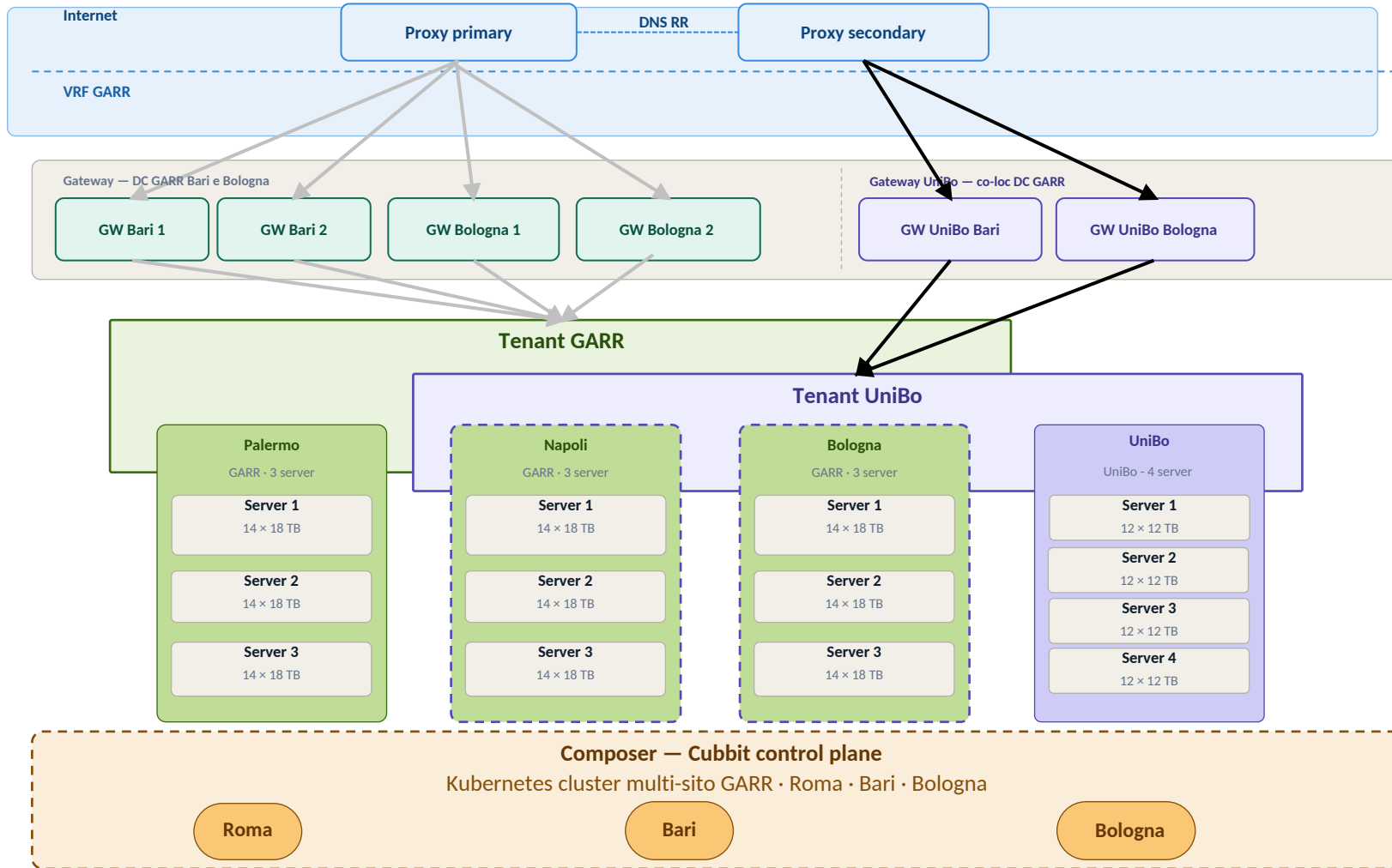
- Dati sensibili
- Collaborazione multi-ente
- Compliance GDPR
- Complessità strutturale

Le infrastrutture di storage distribuite richiedono competenze specialistiche



GARR EduStorage

Infrastruttura federata S3 multi-sito per la ricerca



GET Throughput / velocità lettura

119.8 MiB/s

PUT Throughput / velocità scrittura

40.9 MiB/s

Time To First Byte / latenza iniziale

215 ms

Storage distribuito multi-sito

Infrastruttura presso Università di Bologna

Sicurezza e governance

- Conformità GDPR
- Certificazione ISO/IEC 27001 in corso
- Auditing e controllo accessi

Architettura

- Compatibilità S3
- Replica geografica resiliente
- 3 nodi distribuiti geograficamente (2× Bologna + Cesena)



GET Throughput / velocità lettura

81.6 MiB/s

PUT Throughput / velocità scrittura

26.7 MiB/s

Time To First Byte / latenza iniziale

122 ms

Risultati e conclusioni

Pilot validato con successo

- Accesso S3 distribuito multi-utente
- Resilienza dei dati
- Collaborazione multi-ente
- Compatibilità con workflow esistenti

Benefici osservati

- Interoperabilità S3
- Sovranità digitale
- Scalabilità multi-sito
- Gestione distribuita

Visione

Un'infrastruttura nazionale condivisa
per i dati della ricerca