

# Progetto di Architettura del Software e dei Dati

Andrea Metelli 745753  
Stefano Vittori 749976  
Simone Rossato 747072

Appello del 25/02/2015

# Testo del problema (1)

Si deve realizzare un sistema per l'osservazione della situazione idrogeologica del territorio e per la segnalazione di emergenze.

Il sistema deve supportare:

1. l'acquisizione in tempo reale di dati idrometrici DI (livello dei corsi d'acqua) attraverso opportuni sensori. La struttura della rete idrica, la localizzazione dei sensori e le serie storiche dei livelli osservati sono archiviati in una Base Dati della Rete Idrica (BRI), che fa parte del progetto;
2. l'acquisizione di segnalazioni di emergenze gravi SEG (esondazione in atto o a forte rischio) da parte di operatori a campo;
3. l'acquisizione di previsioni meteo sul medio termine relative a una Regione accedendo a una Base Dati Meteo (BDM) esterna preesistente. Si assuma che BDM fornisca previsioni per ogni ora delle prossime 36 ore, articolate per celle spaziali di dimensione 10X10 Km;
4. l'identificazione di situazioni di emergenza potenziali SEP a medio termine (alcune ore), attraverso l'incrocio delle informazioni BDM e DI. Le situazioni di emergenza potenziali devono essere rese visibili agli operatori di un Centro di Supervisione, ai Responsabili Territoriali della Protezione Civile e, in forma sintetica, a tutta la popolazione interessata;

# Testo del problema (2)

5. la pianificazione degli spostamenti delle Squadre di Emergenza in base alle informazioni SEP. La pianificazione degli spostamenti delle squadre deve essere notificata ai Responsabili Territoriali della Protezione Civile e alle Squadre di Emergenza coinvolte. La allocazione sul territorio delle squadre di emergenza deve essere memorizzata in una Base Dati Segnalazioni di Emergenza (BSE), che fa parte del progetto;
6. la notifica di emergenze gravi (SEG) alle Squadre di Emergenza più prossime.

# Testo del problema (3)

Si richiede di definire, utilizzando i formalismi opportuni:

1. l'architettura del problema in termini di informazioni e flussi informativi;
2. l'architettura logica in termini di componenti di elaborazione;
3. l'architettura concreta in termini di modalità di interazione fra componenti;
4. l'architettura di deployment;
5. le scelte tecnologiche (componenti hardware, reti di comunicazione, piattaforme software);
6. gli schemi logici relazionali delle basi di dati BRI e BDM;
7. gli schemi concettuali delle basi di dati BRI e BSE ottenuti con una attività di reverse engineering dagli schemi relazionali, avendo cura di prevedere in tali schemi concettuali almeno due eterogeneità e almeno una corrispondenza interschema;
8. le modalità e i problemi di integrazione concettuale fra BRI e BDM;
9. lo schema concettuale globale risultato della integrazione delle Basi Dati BRI e BDM;
10. assumendo di utilizzare una architettura di integrazione dati (virtual data integration), e assumendo di scegliere i mapping secondo la modalità Global as View, i mapping tra schema logico globale relazionale e schemi locali relazionali di BRI e BDM;
11. una interrogazione sullo schema globale che visiti ciascuno dei due schemi locali, con il suo unfolding sugli schemi locali;
12. Il sistema deve anche essere in grado di pubblicare parte dei dati contenuti nella architettura di integrazione (a scelta dello studente e tenuto conto di eventuali problemi di privacy) in formato aperto (indicando un insieme di metadati, tra cui il tipo di licenza con cui vengono rilasciati), indicando almeno una applicazione che potrebbe avvantaggiarsi dall'utilizzo di questi dati;

Le scelte architetturali dovranno essere discusse presentandone le motivazioni ed evidenziando, ove opportuno, possibili scelte alternative con i relativi vantaggi e svantaggi (ad esempio, per la architettura dati, una scelta di tipo base dati distribuita, con le relative problematiche di replicazione e distribuzione dei frammenti);

# Ambiguità nelle Specifiche

- ❑ Come vengono identificate le situazioni di emergenza:  
riguardano un tratto di fiume, tutto il fiume, una singola zona attraversata dal fiume, ...
- ❑ Come vengono gestite/pianificate le situazioni di emergenza:  
quali mezzi hanno a disposizione responsabili territoriali e le squadre di emergenza per risolvere il problema.
- ❑ Come avviene l'interazione con BDM esterna preesistente:  
è un web service, un collegamento SQL, un cavo fisico che viene collegato al server come periferica, ...



# Assunzioni - Struttura della rete idrica

- ❑ Un fiume è un insieme di tratti di fiume caratterizzati da due nodi idrici, uno di inizio e uno di fine; l'insieme dei tratti di fiume costituisce un fiume e l'insieme dei fiumi genera una rete idrica;  
Un fiume è inoltre diviso in tratti fiume in corrispondenza di curve, affluenti, emissari, variazioni di pendenza o passaggio in prossimità di centri abitati.
- ❑ In corrispondenza di affluenti/emissari è presente un nodo idrico;
- ❑ I Sensori sono posizionati nei tratti di fiume in prossimità dei centri urbani o/e nei nodi idrici di interesse; ogni sensore è un idrometro;
- ❑ Nodi Idrici e Sensori sono stati Geo-localizzati;

# Assunzioni - Situazioni di Emergenza

- ❑ Una **Segnalazione di Emergenza Potenziale (SEP)** è una situazione di pericolo potenziale che riguarda un singolo tratto di fiume (e.g. l'intero fiume Adda in una situazione di pericolo potenziale produrrà circa 30 SEP in 30 tratti di fiume).

Ad ogni SEP è associato un livello di pericolo potenziale (Basso, Medio-Basso, Medio, Medio-Alto, Alto) che indica la situazione relativa al tratto di fiume.

Il livello di pericolo potenziale è calcolato in base alla crescita del livello d'acqua del fiume (0~10 cm, 10~50 cm, > 50 cm), alle strutture di contenimento preesistenti (argine 0~1 m, 1~5 m, > 5 m) e alla presenza di centri abitati nel territorio attraversato.

Le SEP vengono identificate su scala Nazionale.

- ❑ Una **Segnalazione di Emergenza Grave (SEG)** è data da una notifica da parte di un operatore di campo che segnala un'emergenza imminente/in corso relativa ad un tratto di fiume.

# Assunzioni - Squadre di Emergenza

- ❑ Gli spostamenti delle Squadre di Emergenza vengono effettuati prevalentemente via terra.
- ❑ Ogni squadra di emergenza è localizzata tramite coordinate GPS.
- ❑ Si assume di avere accesso ad un servizio esterno che calcola la distanza stradale tra due punti geo-localizzati (e.g. Google Maps).
- ❑ Si assume che le Squadre di Emergenza siano egualmente equipaggiate; le squadre speciali se esistono, sono in un numero irrilevante e vengono impiegate esclusivamente nelle situazioni di emergenza grave.
- ❑ Si assume che per ogni Squadra sia presente un Responsabile.
- ❑ A seguito di una pianificazione, una Squadra di Emergenza è associata ad un tratto di fiume.
- ❑ Tutte le notifiche da parte del sistema vengono effettuate mediante l'invio di SMS al Responsabile di Squadra.



# Assunzioni - Servizio Previsioni Meteo

- ❑ Assumiamo che il database meteo BDM sia direttamente accessibile mediante la rete internet e quindi sia possibile interrogare il database.
- ❑ Assumiamo che BDM mantenga lo storico delle previsioni meteo.

# Requisiti Funzionali (1)

- ❑ **Rilevare i dati idrometrici (DI):** consiste nell'acquisizione dei dati riguardanti il livello dei corsi d'acqua dagli opportuni sensori posti sui tratti di fiume ed aggiornare il database BRI (Base Dati Rete Idrica).
- ❑ **Consultare le Previsioni Meteo:** reperire le informazioni meteo contattando il servizio esterno (BDM).
- ❑ **Identificare SEP:** incrociando le informazioni contenute in BRI e quelle ricevute da BDM, si identificano e si memorizzano le potenziali situazioni di emergenza relative ai tratti di fiume.
- ❑ **Pianificare Squadre di Emergenza:** gli operatori del Centro di Supervisione avviano l'algoritmo per la pianificazione delle Squadre di Emergenza. Questo algoritmo specifico prende in input le SEP generate dal sistema e pianifica le Squadre a seconda dei livelli di pericolo e della vicinanza di ciascuna Squadra alle zone identificate dalle SEP (il calcolo del percorso viene effettuato mediante un servizio esterno).

# Requisiti Funzionali (2)

- ❑ **Acquisizione SEG:** l'acquisizione delle Situazioni di Emergenza Gravi viene fatta da un operatore di campo che, in quel momento, comunica che la situazione in un dato tratto di fiume sta degenerando.
- ❑ **Notificare le SEG alle Squadre:** consiste nella notifica alle Squadre più prossime di una SEG in corso.
- ❑ **Pubblicare SEP:** è l'azione di rendere visibili agli operatori del Centro di Supervisione e ai Responsabili Territoriali della Protezione Civile i dati dettagliati relativi alle SEP identificate dal sistema.
- ❑ **Pubblicare SEP sintetiche:** consiste nel pubblicare le informazioni sintetiche delle SEP relative ai tratti di fiume sul portale web accessibile alla popolazione.

# Stime

- ❑ Circa 100 sensori per Regione\*
- ❑ Circa 50 Squadre per Regione
- ❑ Circa 20 Operatori di Campo per Regione
- ❑ Un Responsabile Territoriale per ogni Regione
- ❑ I dati idrometrici (DI) vengono acquisiti ogni ora
- ❑ L'identificazione delle SEP viene svolta ogni ora
- ❑ Ogni Squadra impiega dai 45~60 minuti per raggiungere la zona di intervento, che è stata calcolata in base alle SEP identificate dal sistema
- ❑ Ogni Squadra impiega 1~4 ore per mettere in sicurezza un argine

\*La stima dei Sensori (idrici) presenti in ciascuna regione è basata sul numero di sensori presenti in Lombardia.

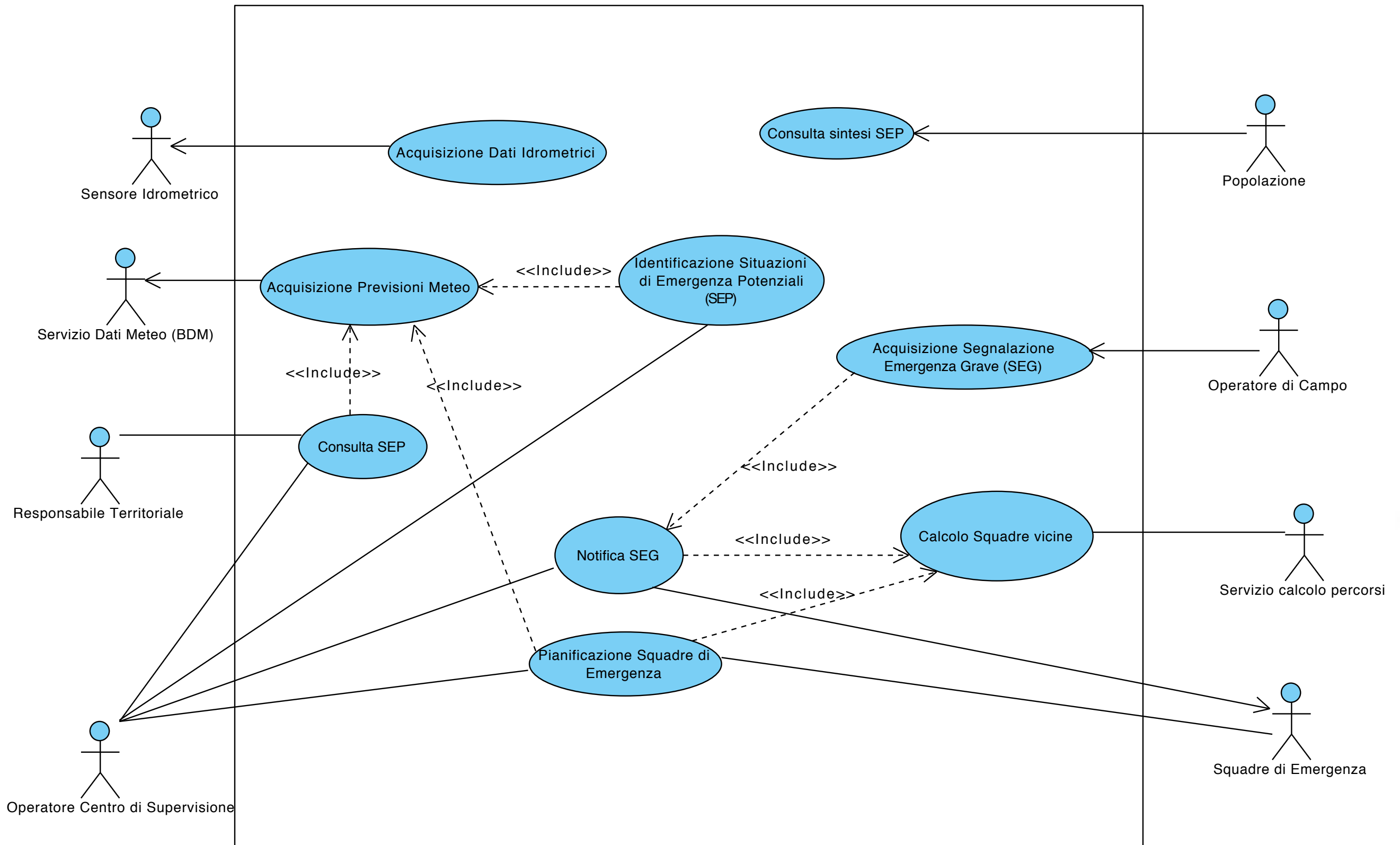
# Requisiti non Funzionali

- ❑ Il sistema deve essere in grado di acquisire ogni ora tutte le rilevazioni dei sensori.
- ❑ Il sistema deve essere in grado di processare la pianificazione delle squadre in un tempo limite di 30 minuti.
- ❑ Il sistema deve produrre delle pianificazioni affidabili nel 75% dei casi, l'affidabilità di una pianificazione è il numero di assegnamenti corretti diviso il numero di assegnamenti superflui.
- ❑ In caso di malfunzionamento il sistema deve essere riparabile in meno di 3 giorni.

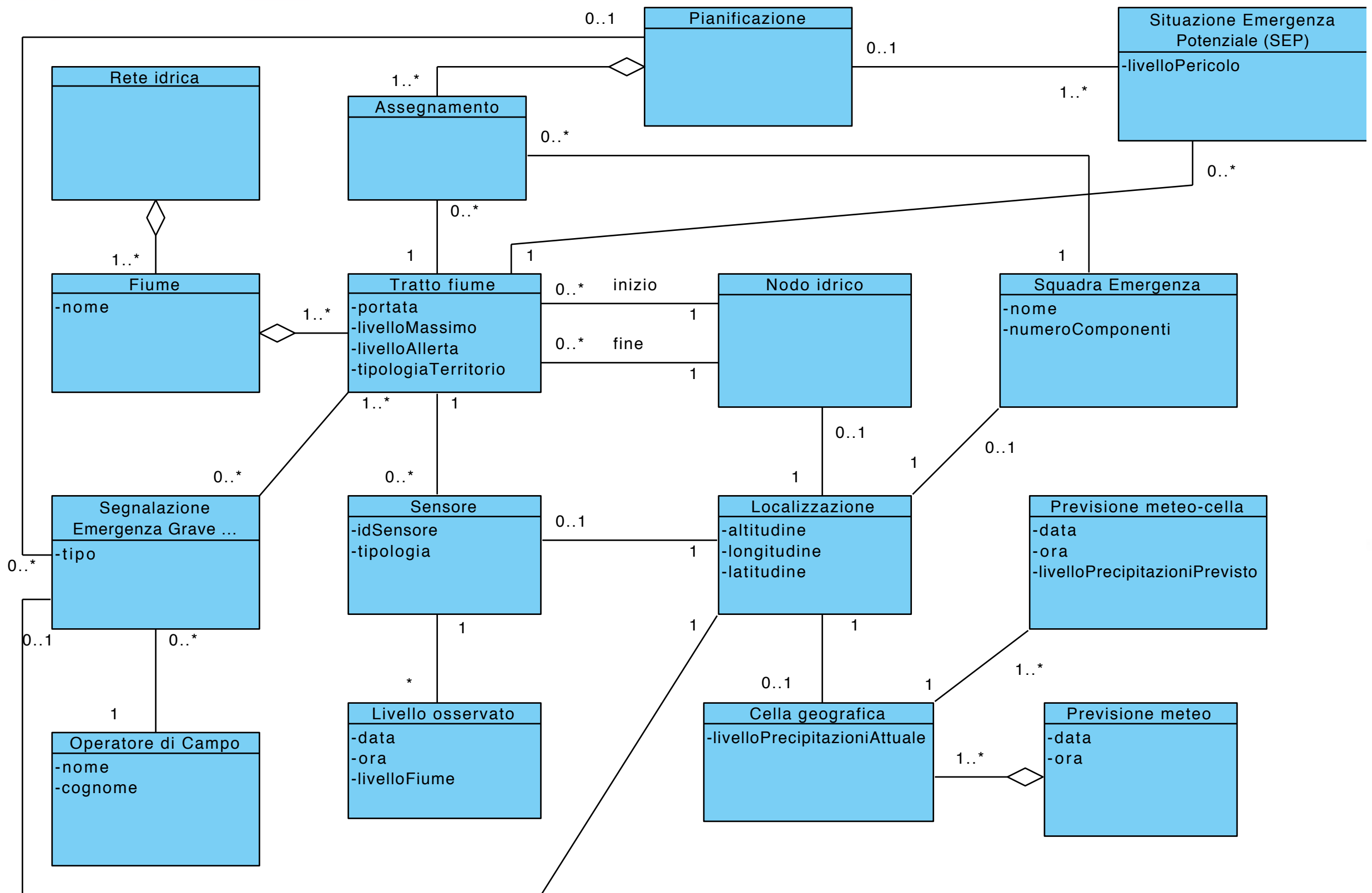


# **Architettura del problema**

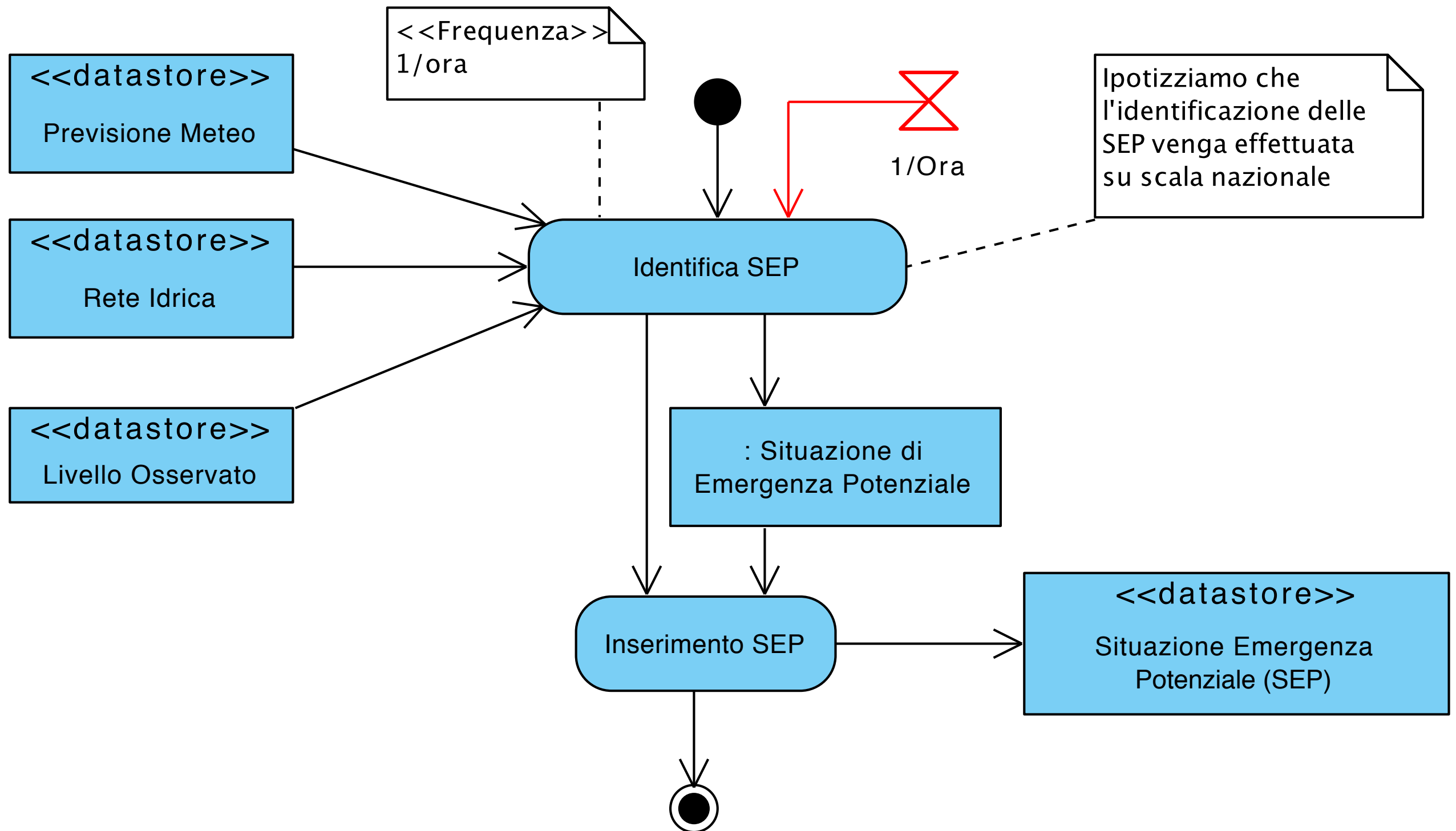
# Casi d'Uso



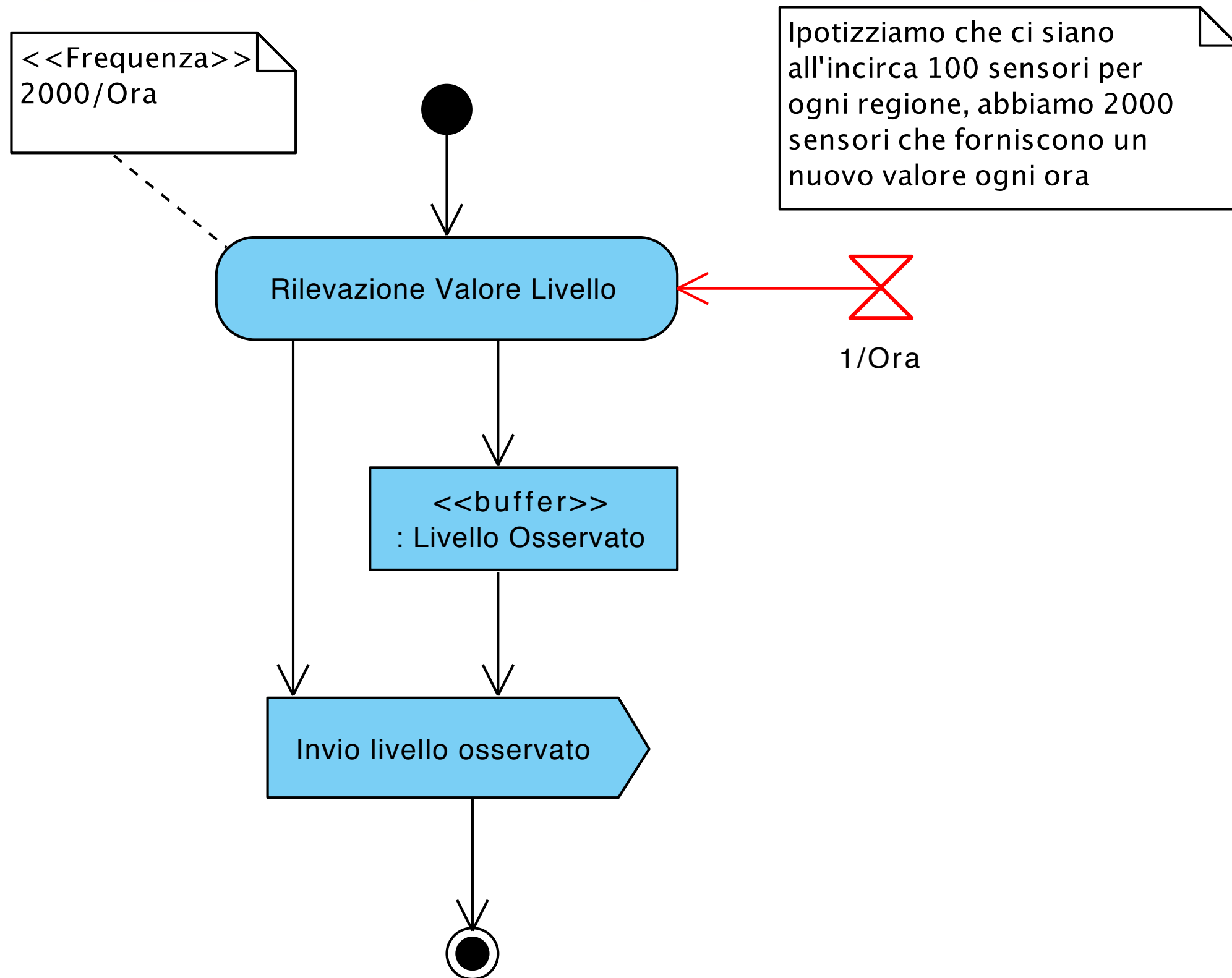
# Modello dei Dati



# Identificazione SEP

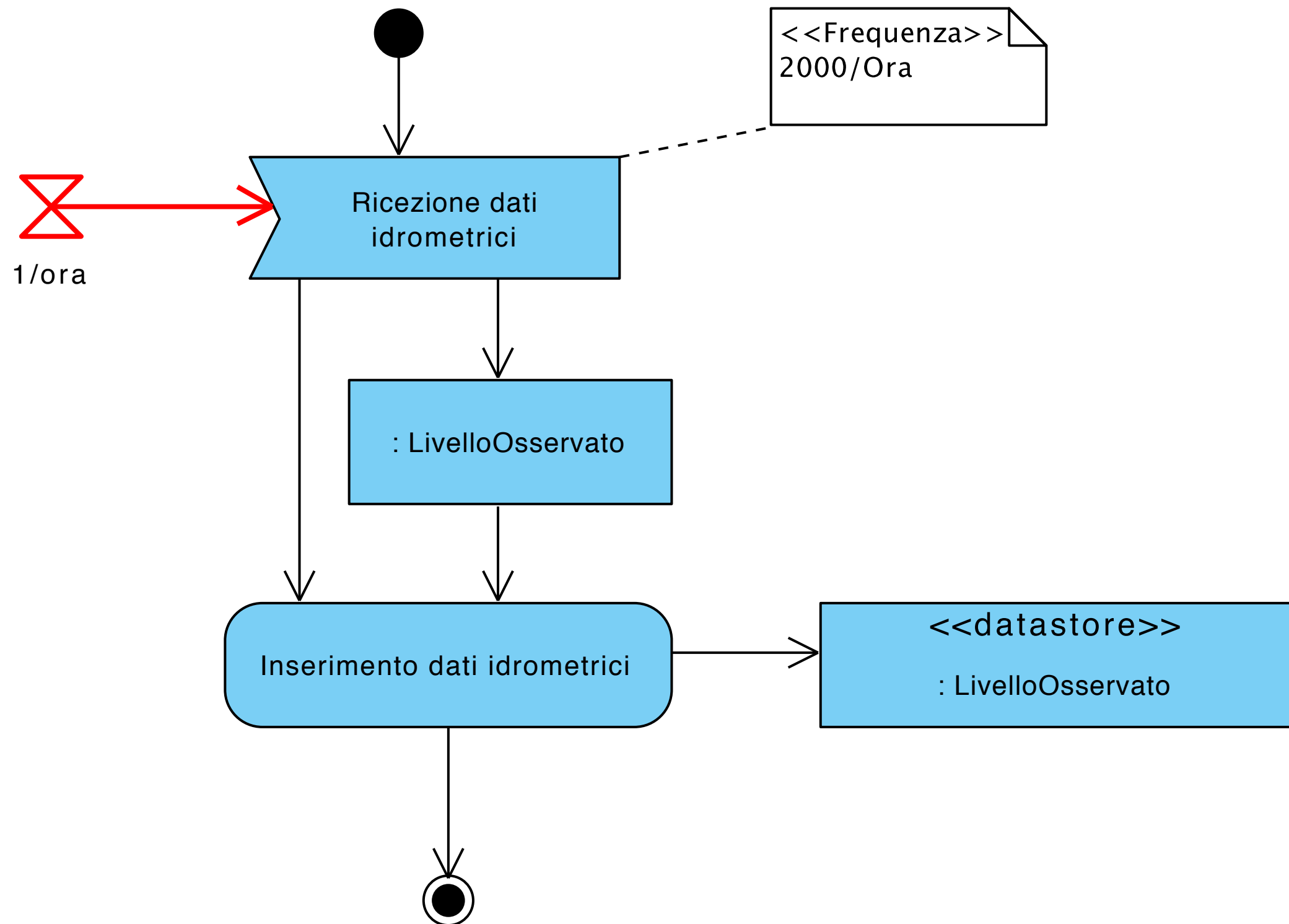


# Acquisizione Dati Idrometrici





# Archiviazione Dati Idrometrici



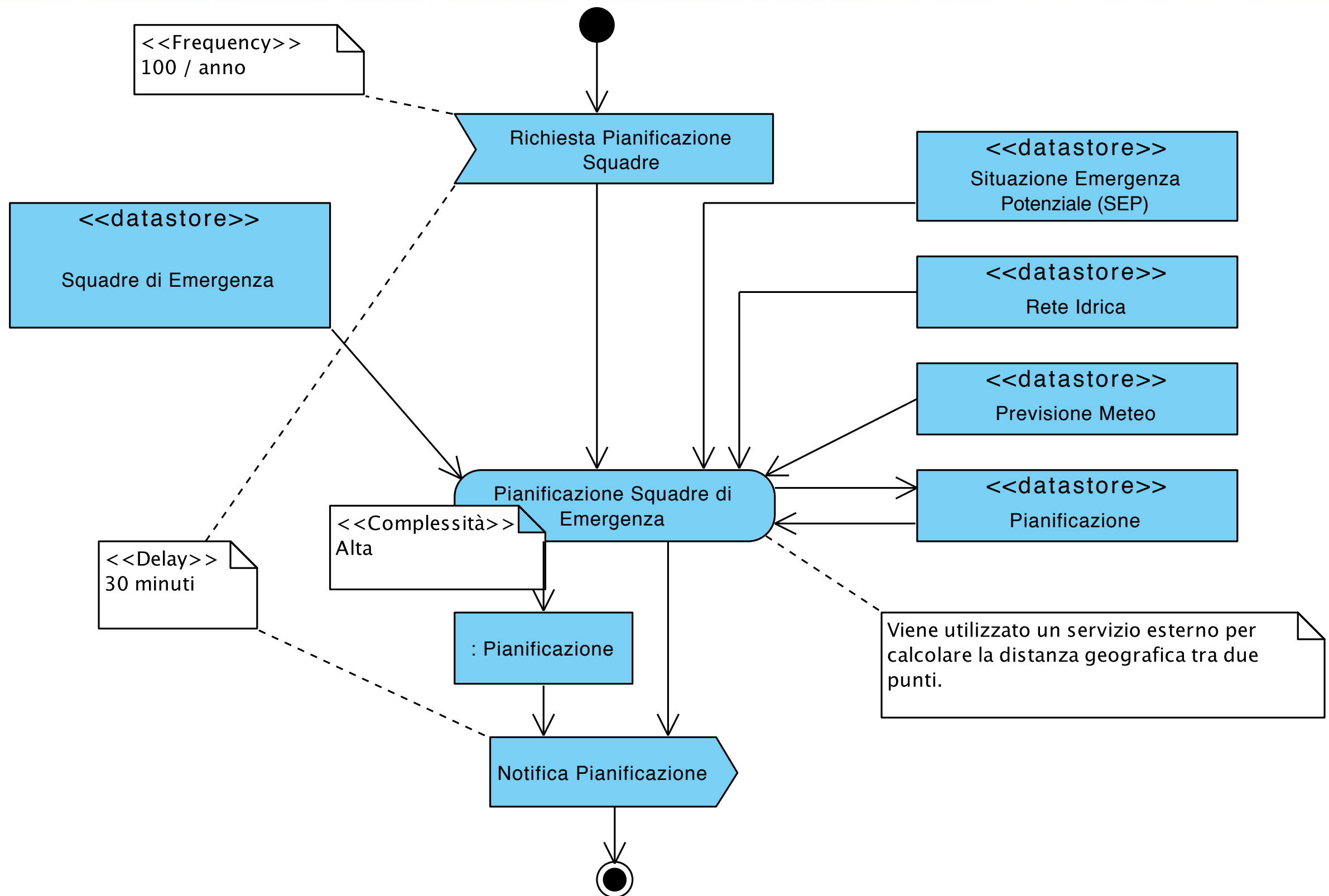
# Pianificazione Squadra di Emergenza

Una volta che il sistema ha riconosciuto le Situazioni di Emergenza Potenziali, l'operatore del Centro di Supervisione valuta la necessità di gestire tali emergenze, richiedendo al sistema di pianificare e/o ripianificare le squadre di emergenza.

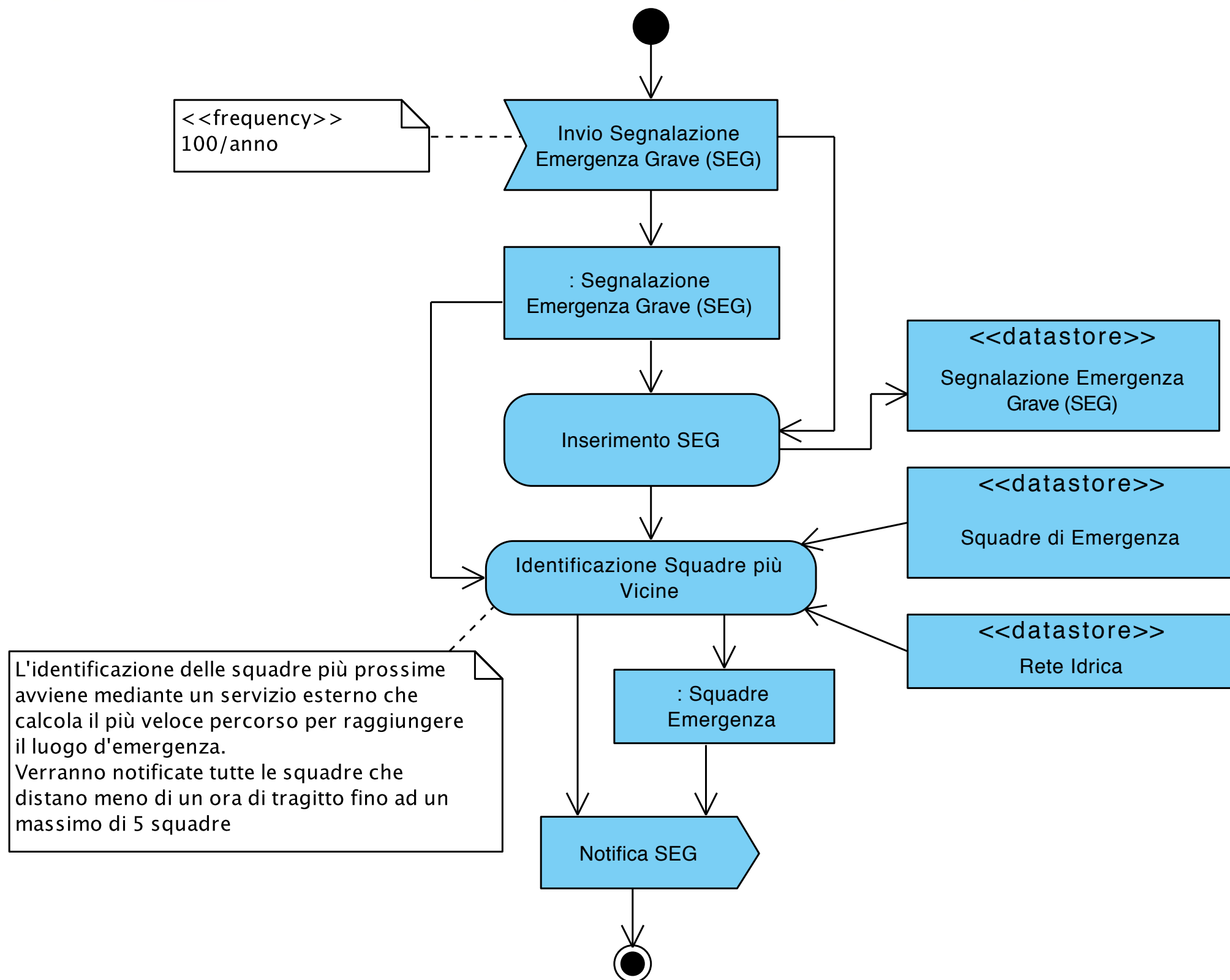
Durante tale operazione, l'algoritmo di pianificazione utilizza dei servizi esterni che calcolano la distanza stradale tra due punti geolocalizzati. Per questo motivo il sistema non prevede una mappa viaria della nazione al suo interno.

A seconda di queste informazioni, l'algoritmo sceglie in che modo dislocare le squadre di emergenza sul territorio.

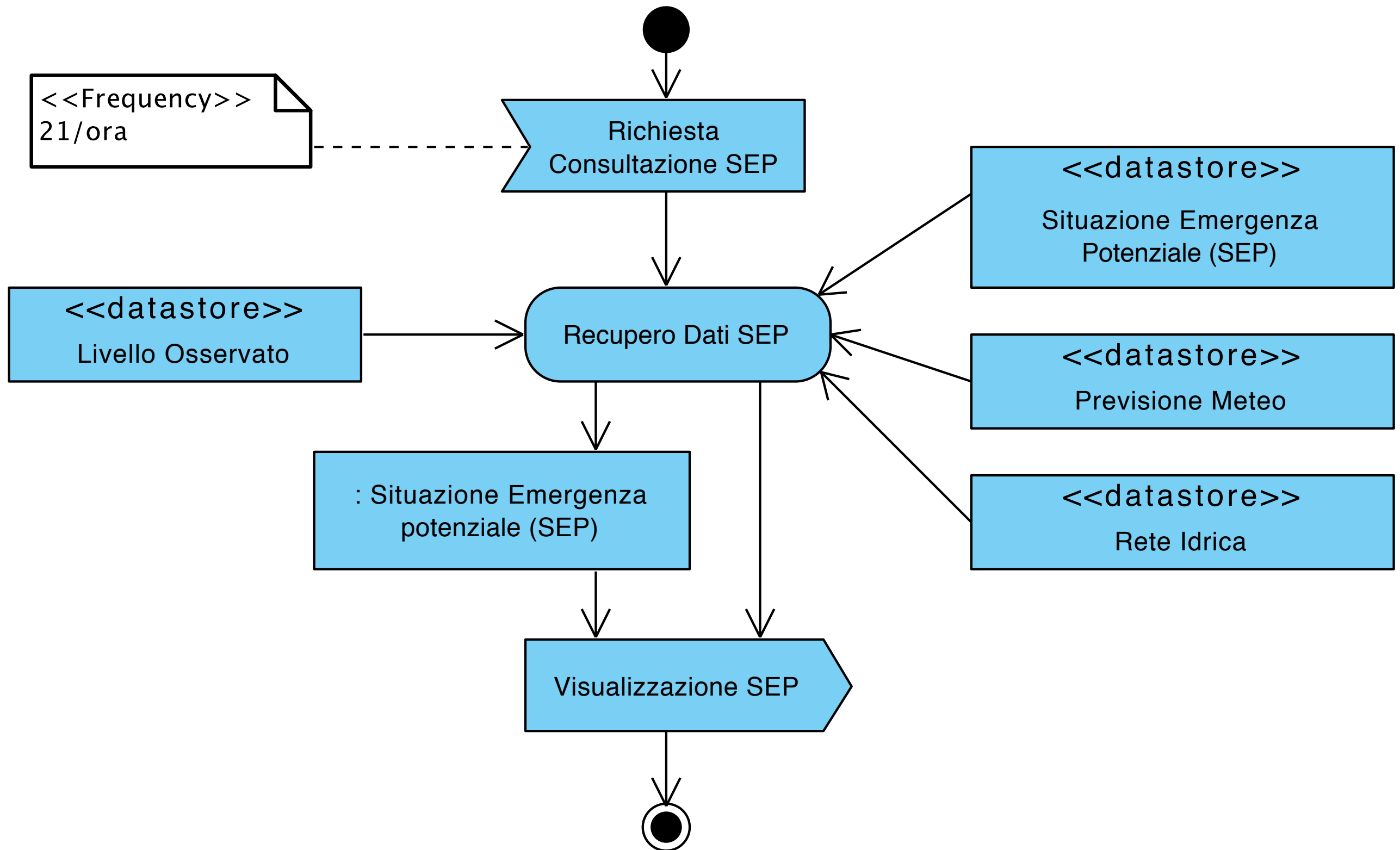
# Pianificazione Squadre di Emergenza



# Acquisizione SEG

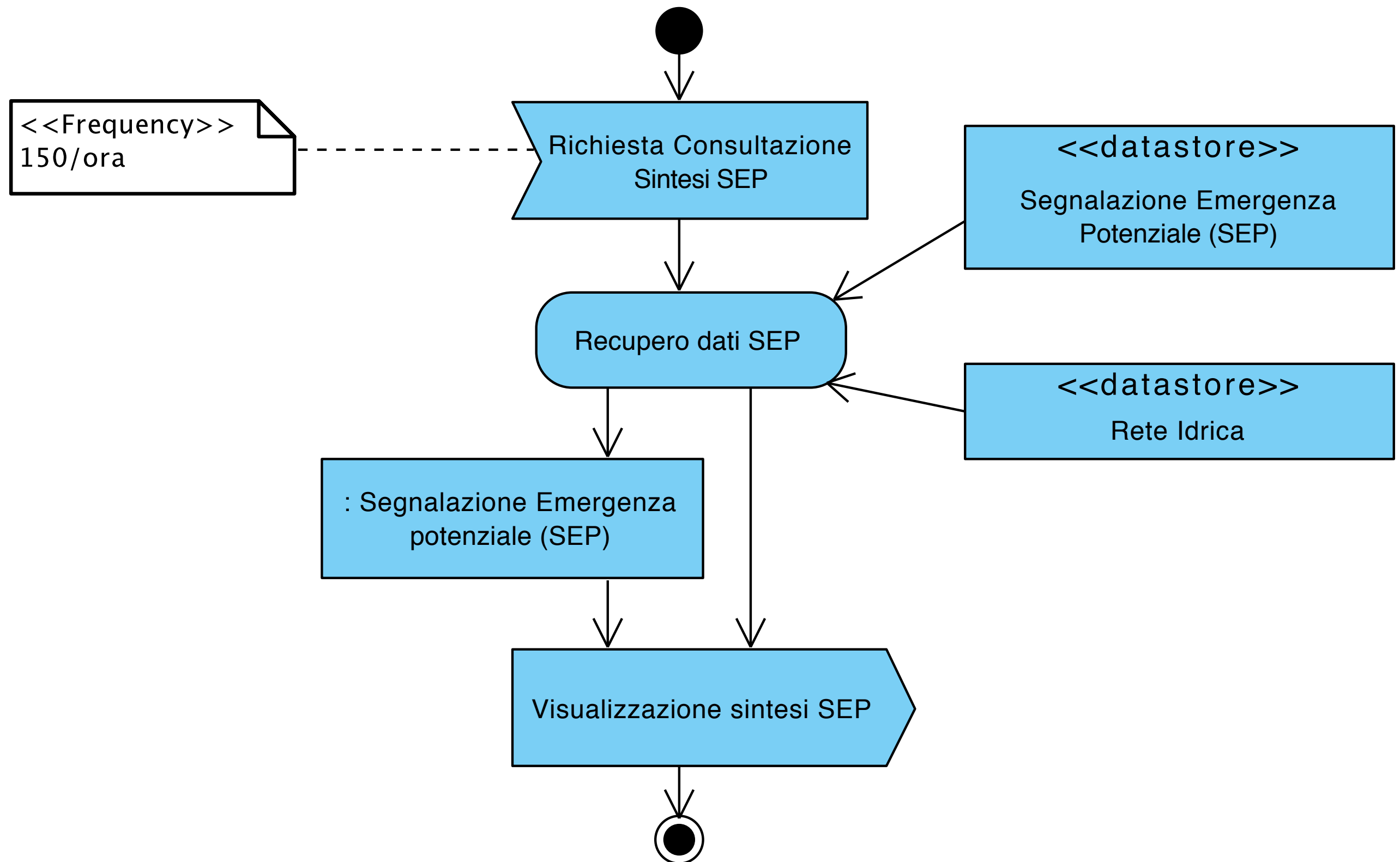


# Consulta SEP





# Consulta Sintesi SEP



# Architettura Logica

# Soluzione 1

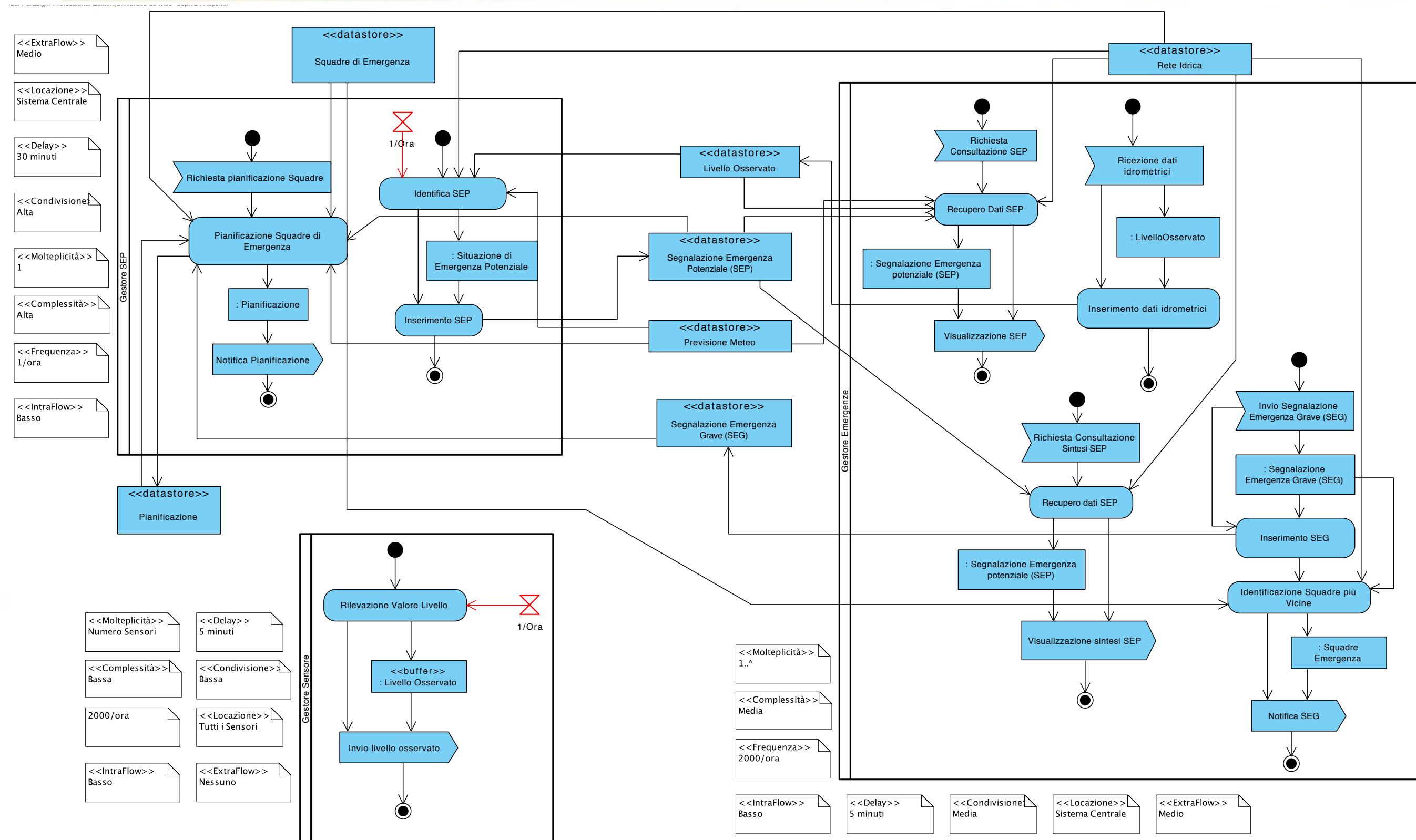
# Descrizione - Soluzione 1

L'architettura separa le attività di inserimento/ricezione dati dalle attività di identificazione e pianificazione delle Emergenze.

- ❑ Un componente che si occupa dell'invio dei dati idrometrici
- ❑ Un componente si occupa dell'identificazione e successivamente della pianificazione delle SEP.
- ❑ Un componente si occupa della ricezione e dell'invio dei dati; inserisce il valori idrometrici, fornisce le informazioni relative alle SEP sia in forma completa, che in forma sintetica e gestisce le Segnalazioni di Emergenza Grave.

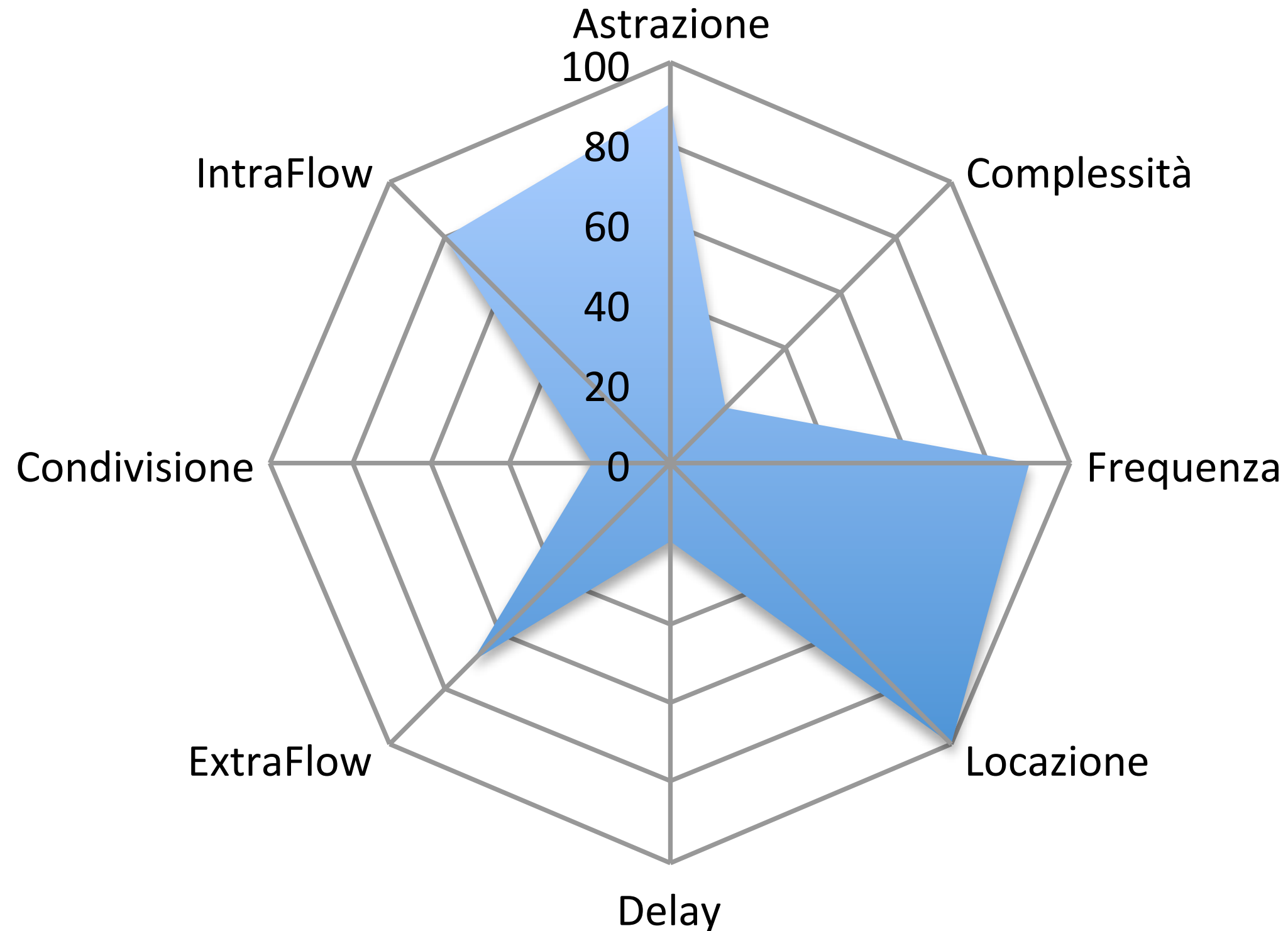
Suddivisione in base alla complessità.

# Soluzione 1





# Footprint Overlap - Soluzione 1



# Soluzione 2

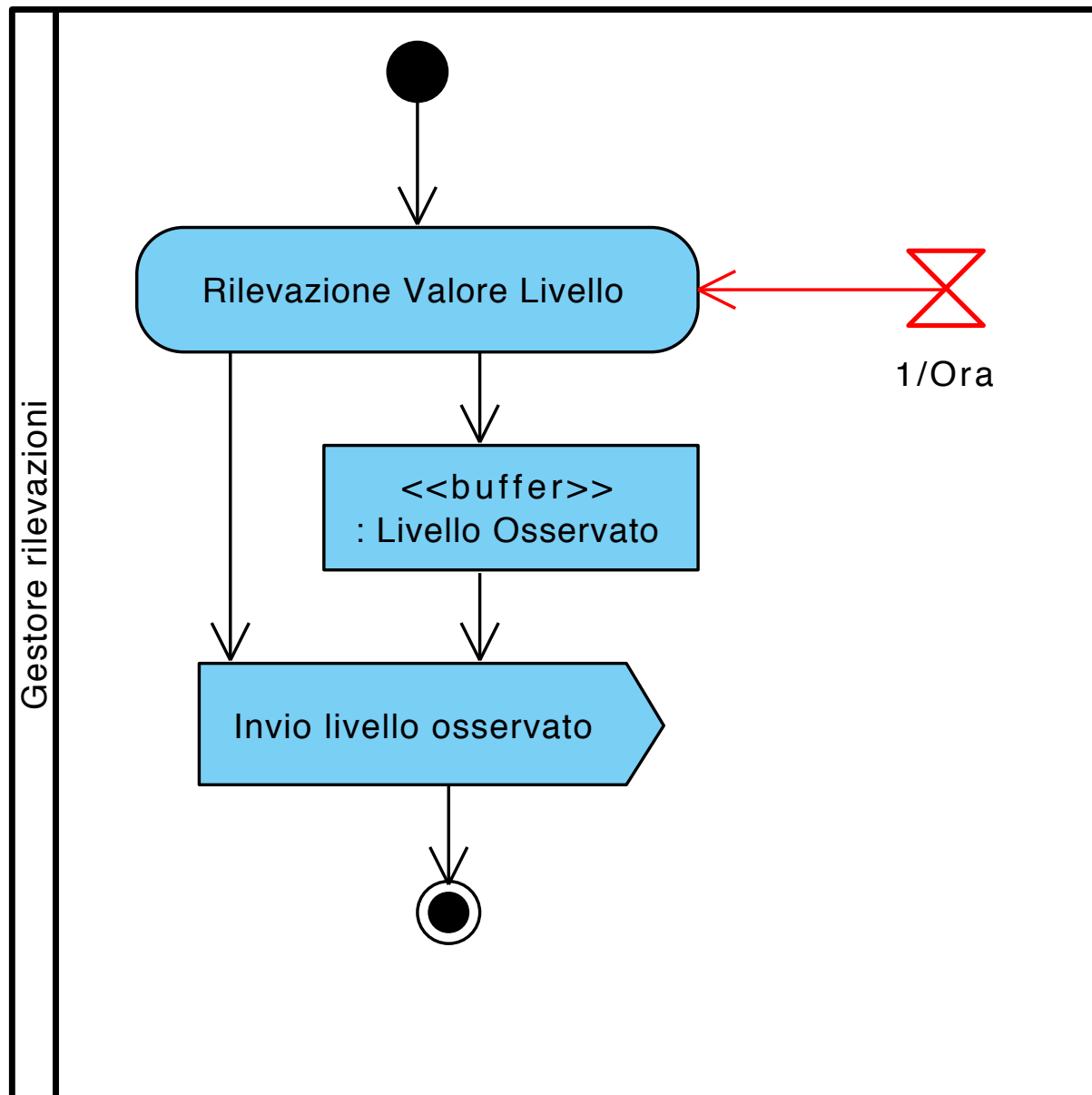
# Descrizione - Soluzione 2

Architettura è divisa in componenti logici:

- ❑ Gestore Rilevazioni DI: contiene le attività legate all'acquisizione e l'invio di informazioni da parte dei sensori.
- ❑ Gestore Archiviazioni DI: contiene le attività legate alla ricezione e alla memorizzazione delle informazioni rilevate dai sensori nel database.
- ❑ Gestore SEP: contiene le attività collegate all'identificazione, alla memorizzazione e alla pianificazione delle SEP.
- ❑ Gestore Consultazione SEP: contiene le attività legate alla visualizzazione delle SEP.
- ❑ Gestore SEG: contiene le attività legate all'acquisizione e alla notifica delle SEG.

Suddivisione basata sulla frequenza e sulla locazione dei componenti.

# Gestore Rilevazioni DI



<<Molteplicità>>  
2000

<<Delay>>  
2 minuti

<<IntraFlow>>  
Basso

<<Complessità>>  
Molto bassa

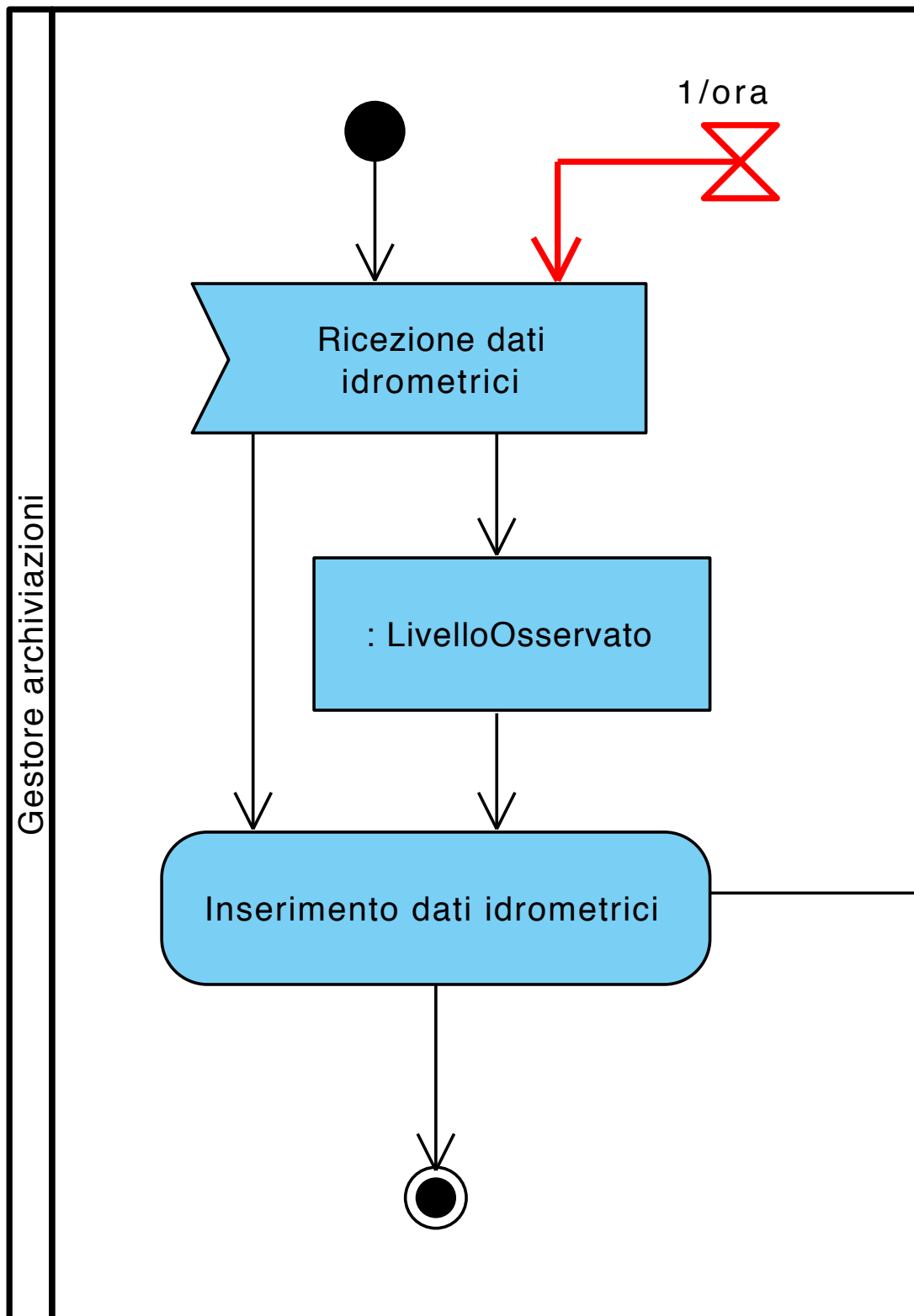
<<Condivisione>>  
Nessuna

<<ExtraFlow>>  
Basso

<<Frequenza>>  
2000/ora

<<Locazione>>  
Tutti i Sensori

# Gestore Archiviazione DI



<<Molteplicità>>  
2000

<<Delay>>  
2 minuti

<<IntraFlow>>  
Medio

<<Complessità>>  
Molto bassa

<<Condivisione>>  
Bassa

<<ExtraFlow>>  
Nessuno

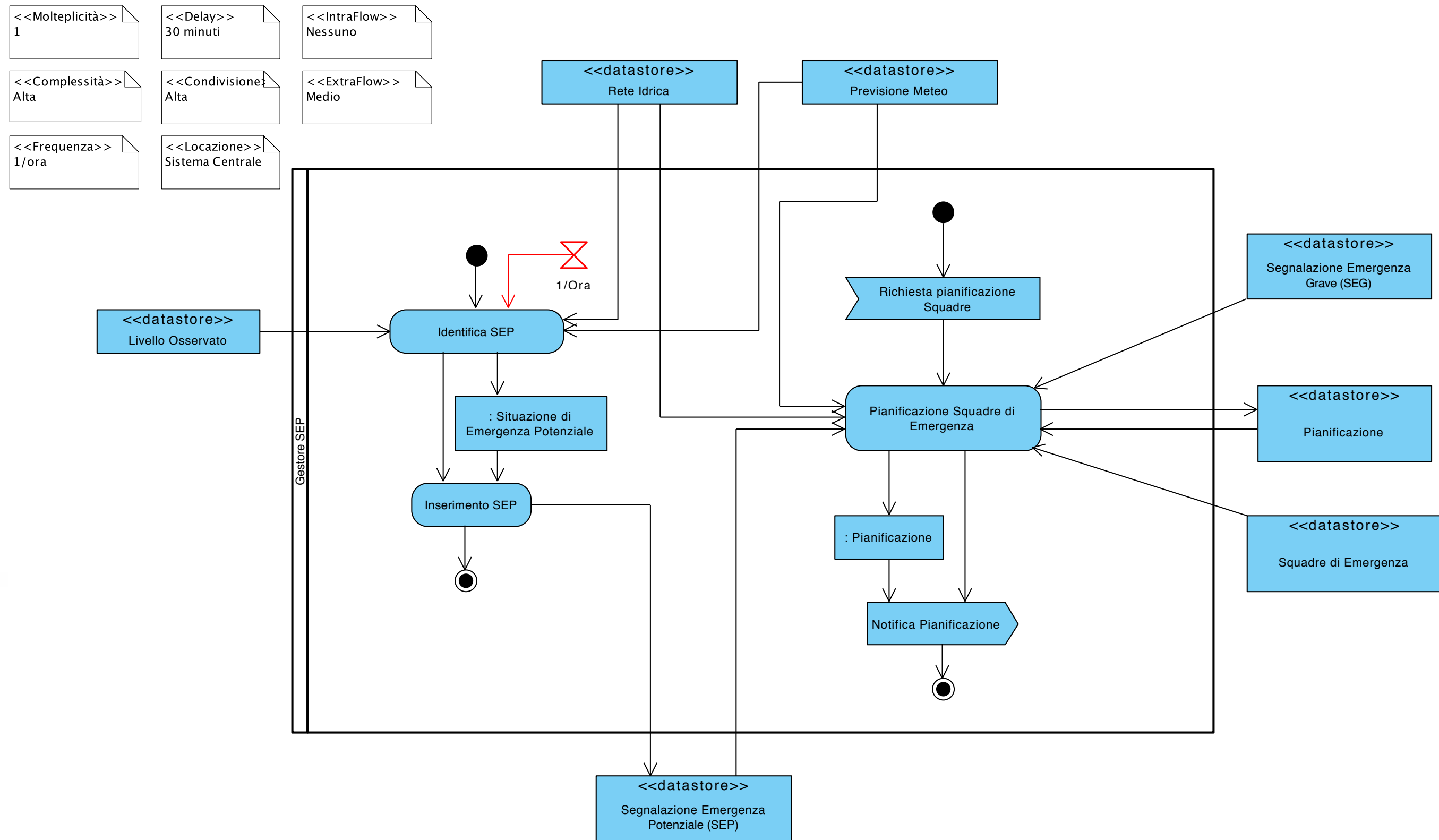
<<Frequenza>>  
2000/ora

<<Locazione>>  
Tutti i Sensori

<<datastore>>  
: LivelloOsservato



# Gestore SEP



# Gestore Consultazione SEP

<<Molteplicità>>  
0..\*

<<Delay>>  
20 secondi

<<IntraFlow>>  
Nessuno

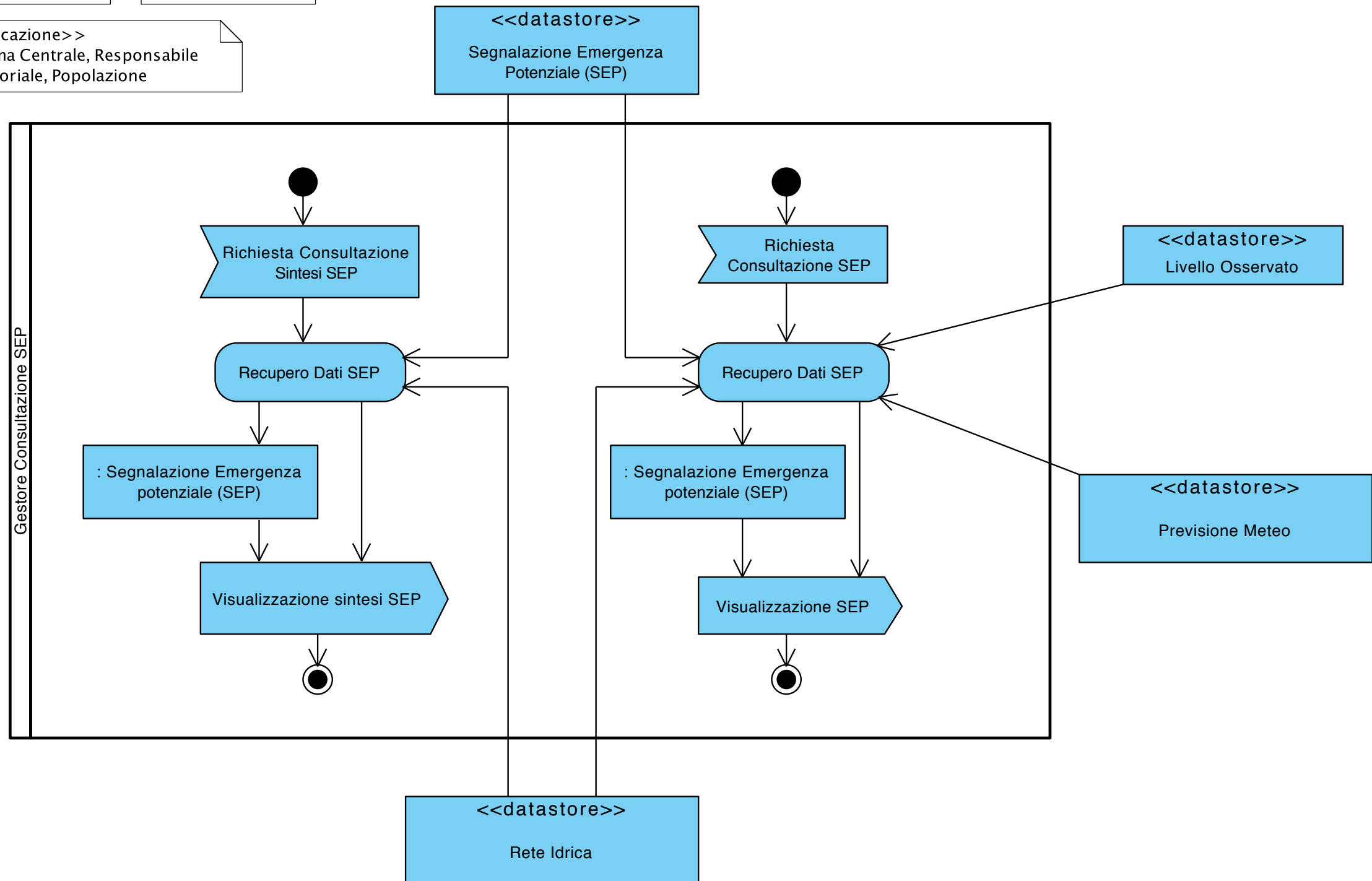
<<Complessità>>  
Bassa

<<Condivisione>>  
Alta

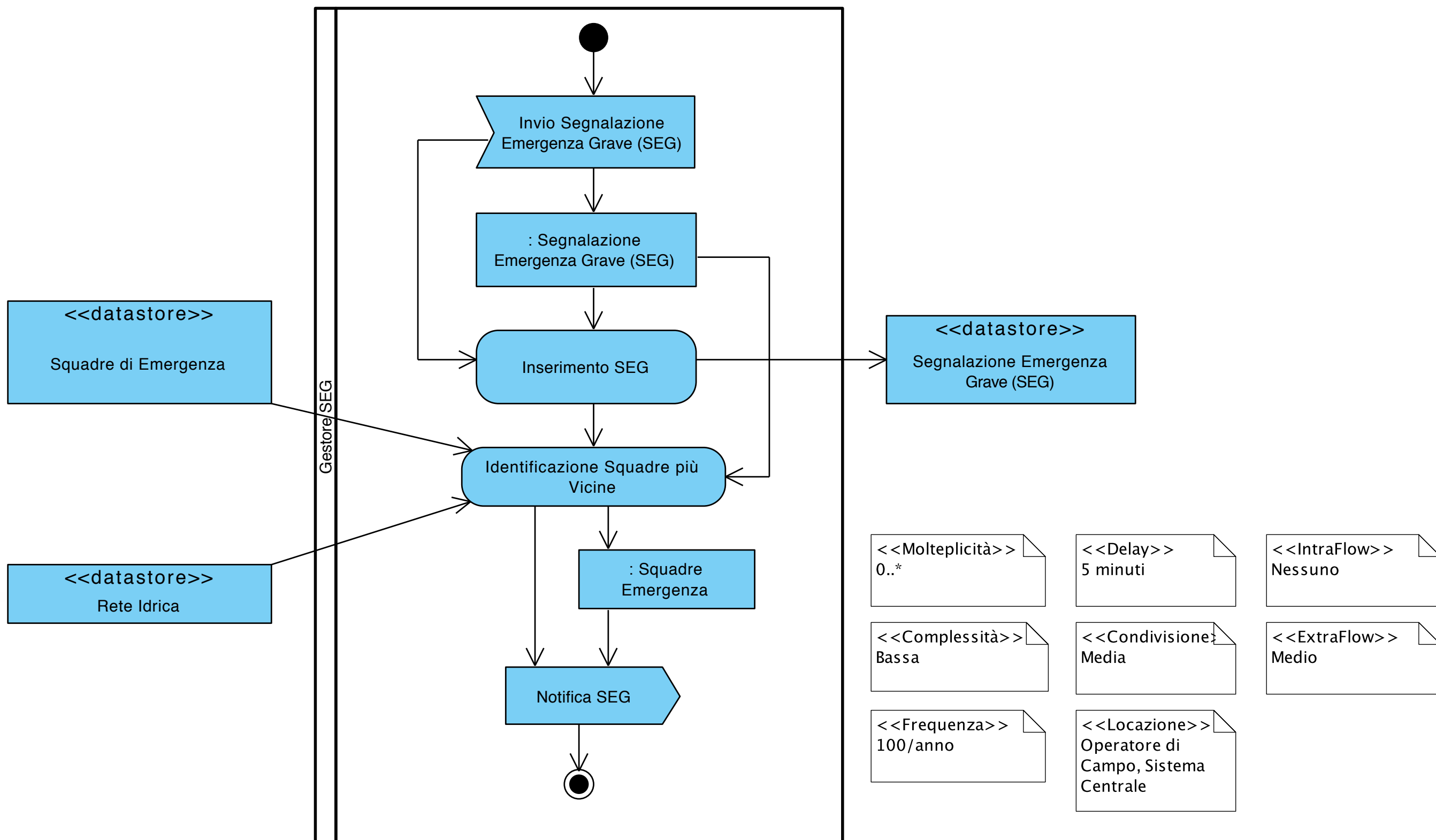
<<ExtraFlow>>  
Alto

<<Frequenza>>  
150/ora

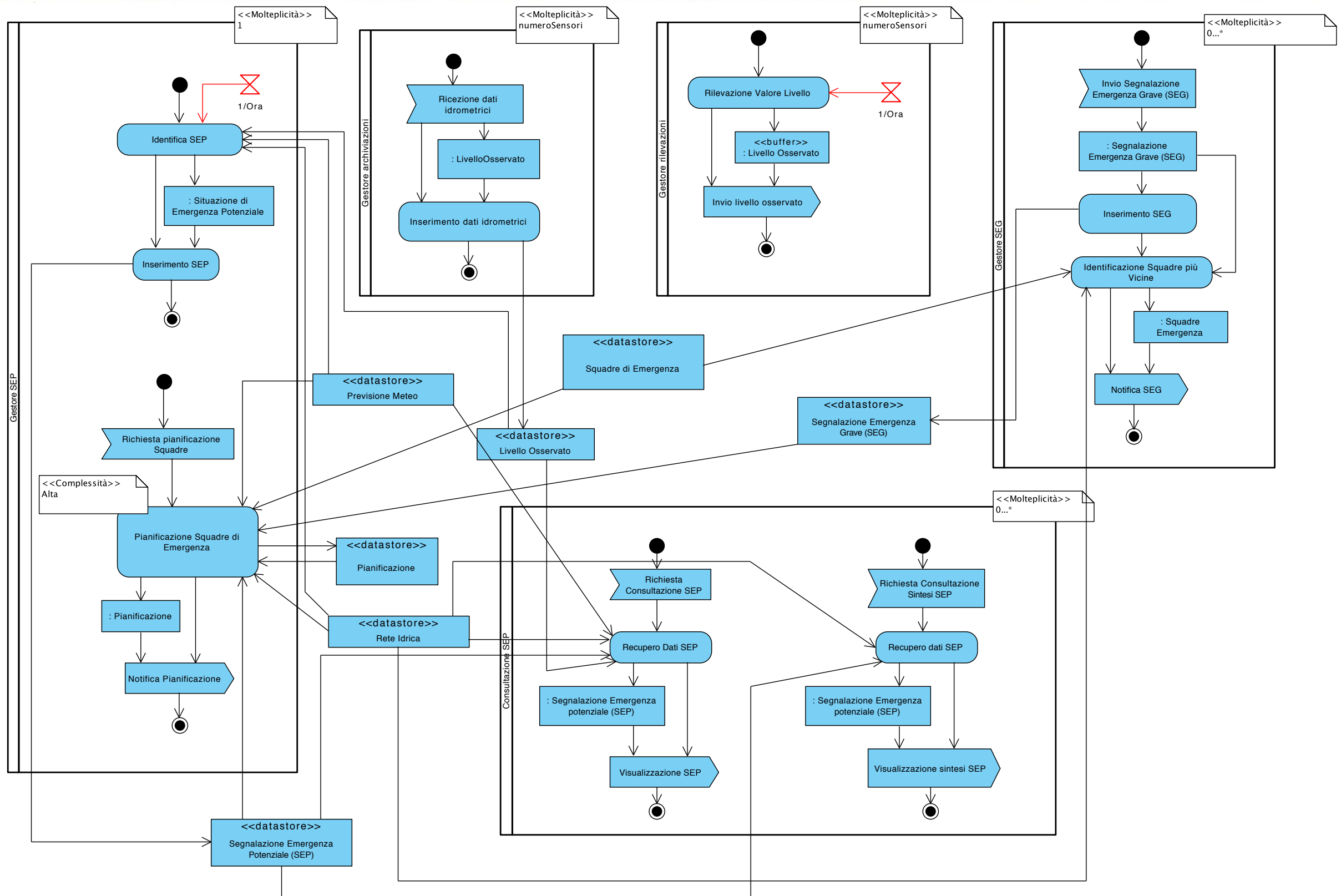
<<Locazione>>  
Sistema Centrale, Responsabile  
Territoriale, Popolazione



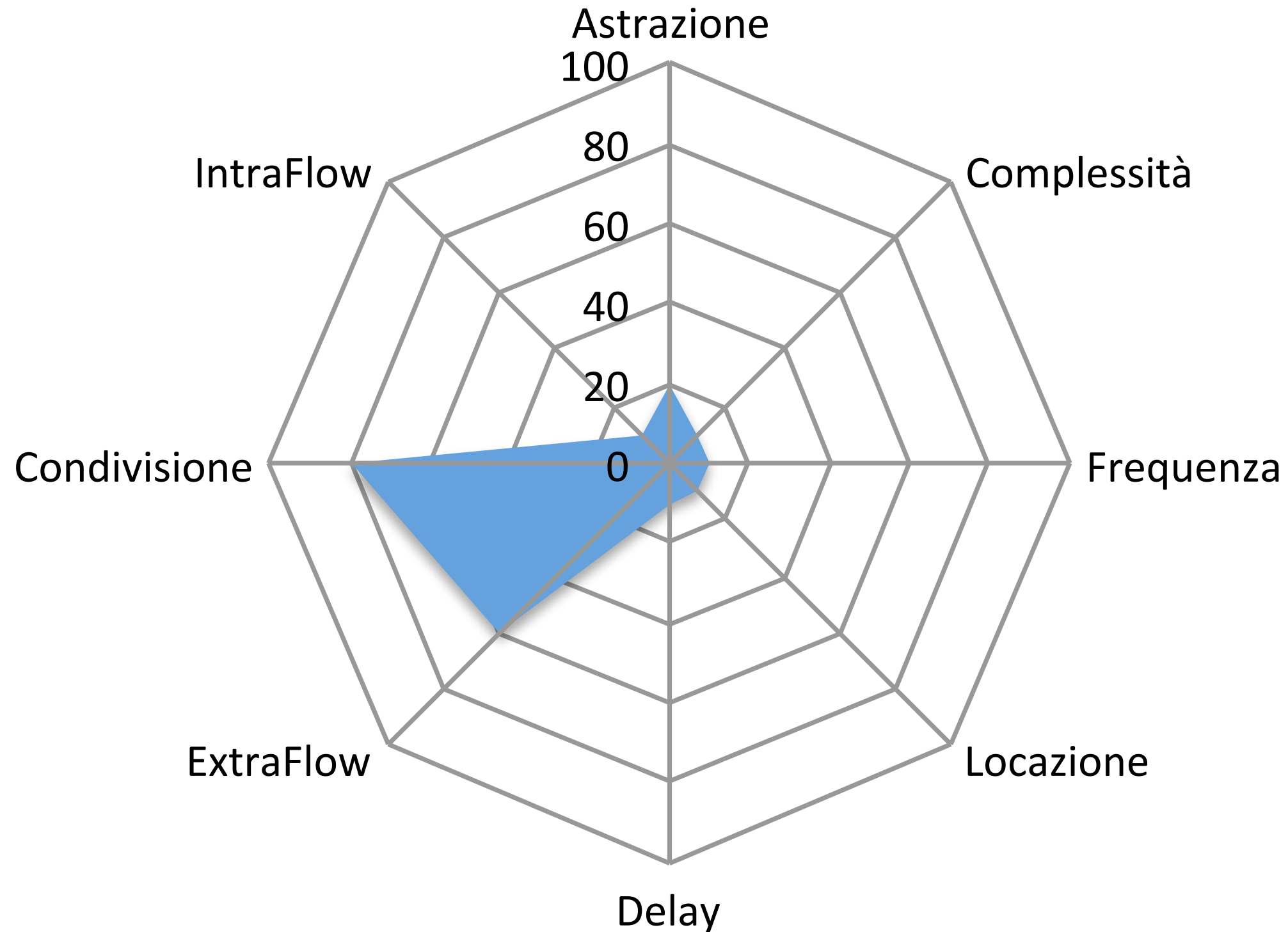
# Gestore SEG



# Soluzione 2



# Footprint Overlap - Soluzione 2





# Scelta - Soluzione 2

Abbiamo scelto la seconda soluzione proposta perché minimizza la complessità e la locazione.

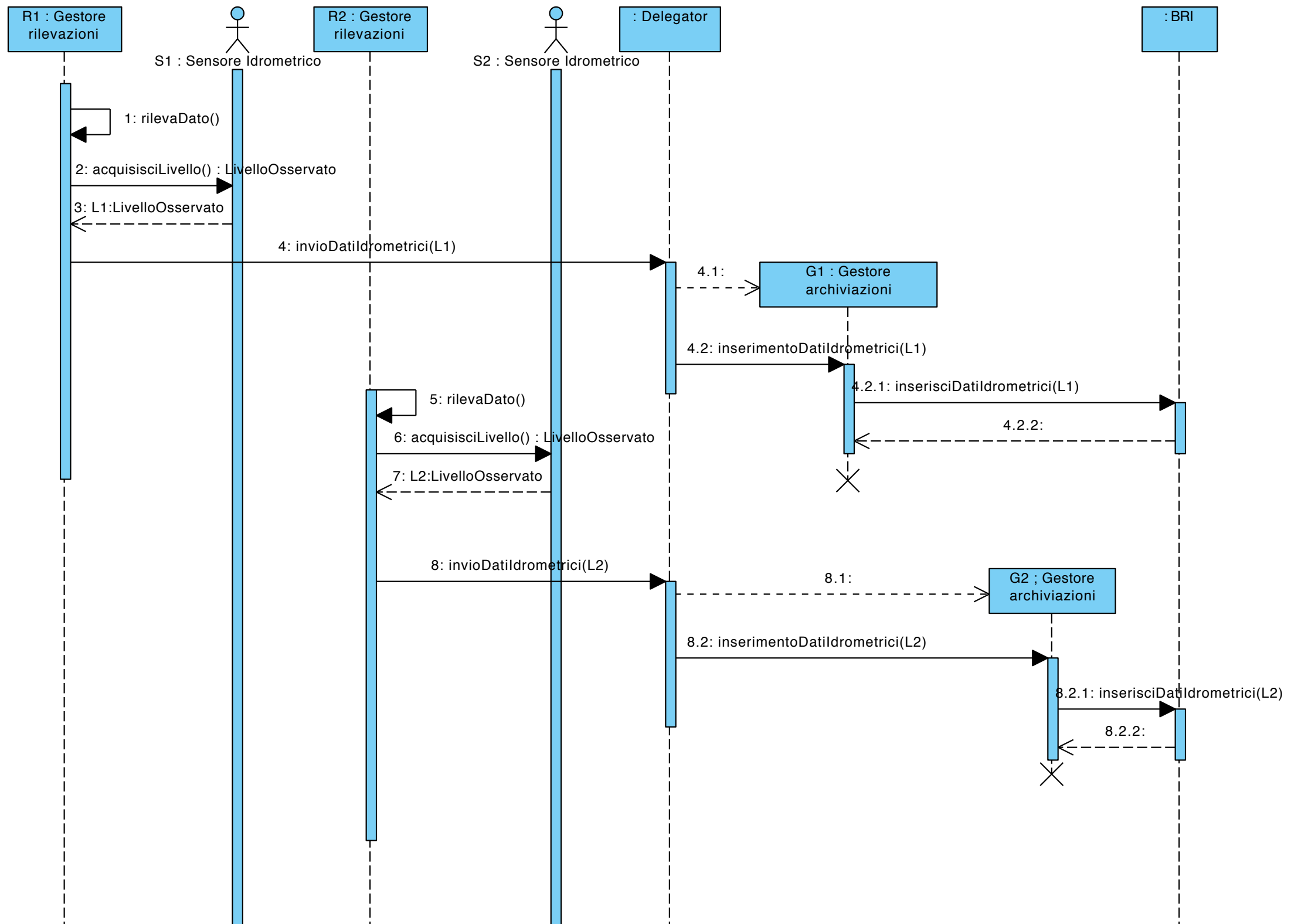
## **Pro:**

- ☐ Complessità, gli algoritmi di complessità alti sono eseguiti dallo stesso componente logico.
- ☐ Locazione, la divisione dei componenti fisici è mantenuta nei componenti logici.
- ☐ Frequenza, i componenti logici eseguono le azioni con frequenza simile.

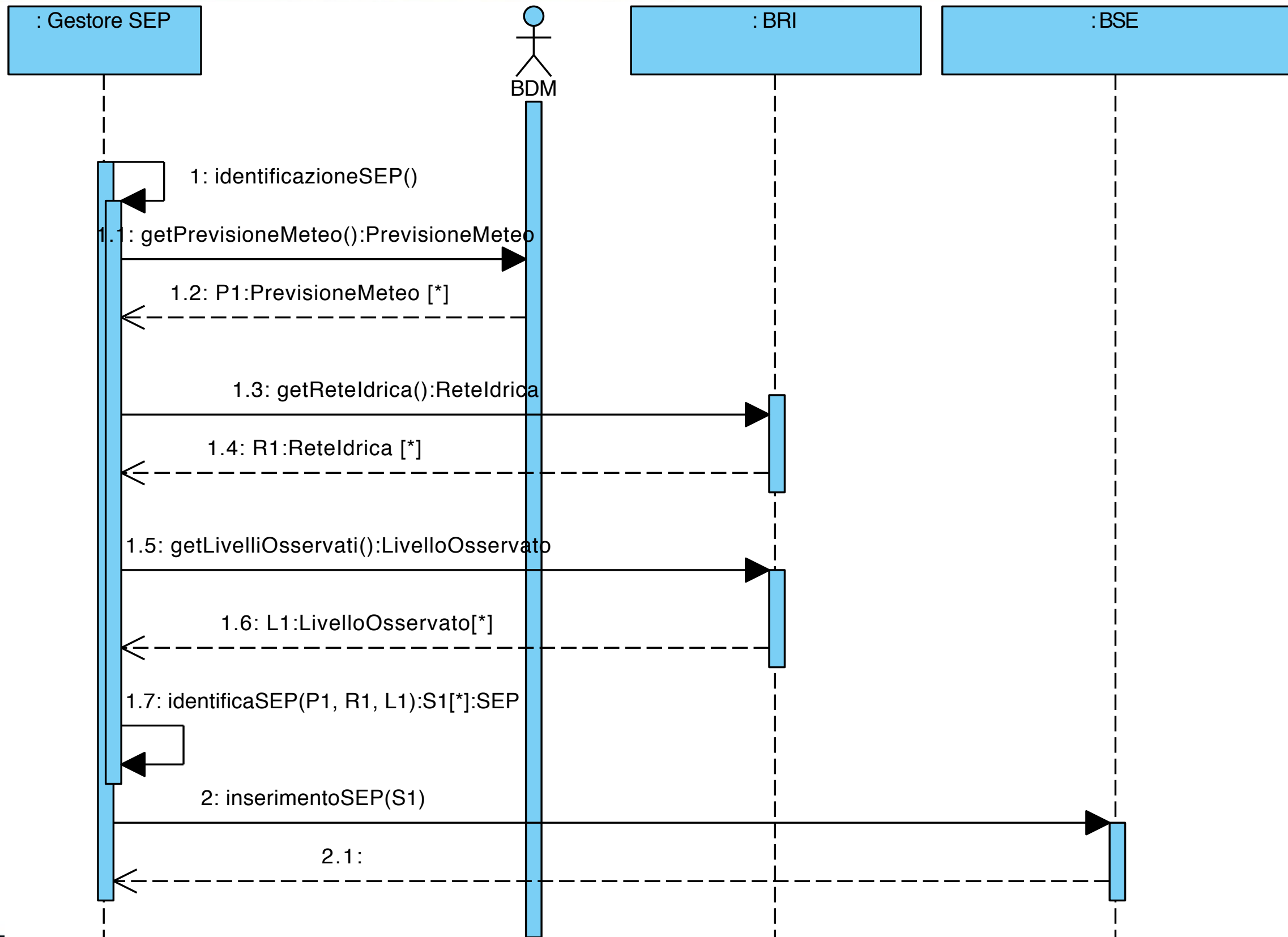
## **Contro:**

- ☐ Condivisione dati elevata tra le componenti logiche.

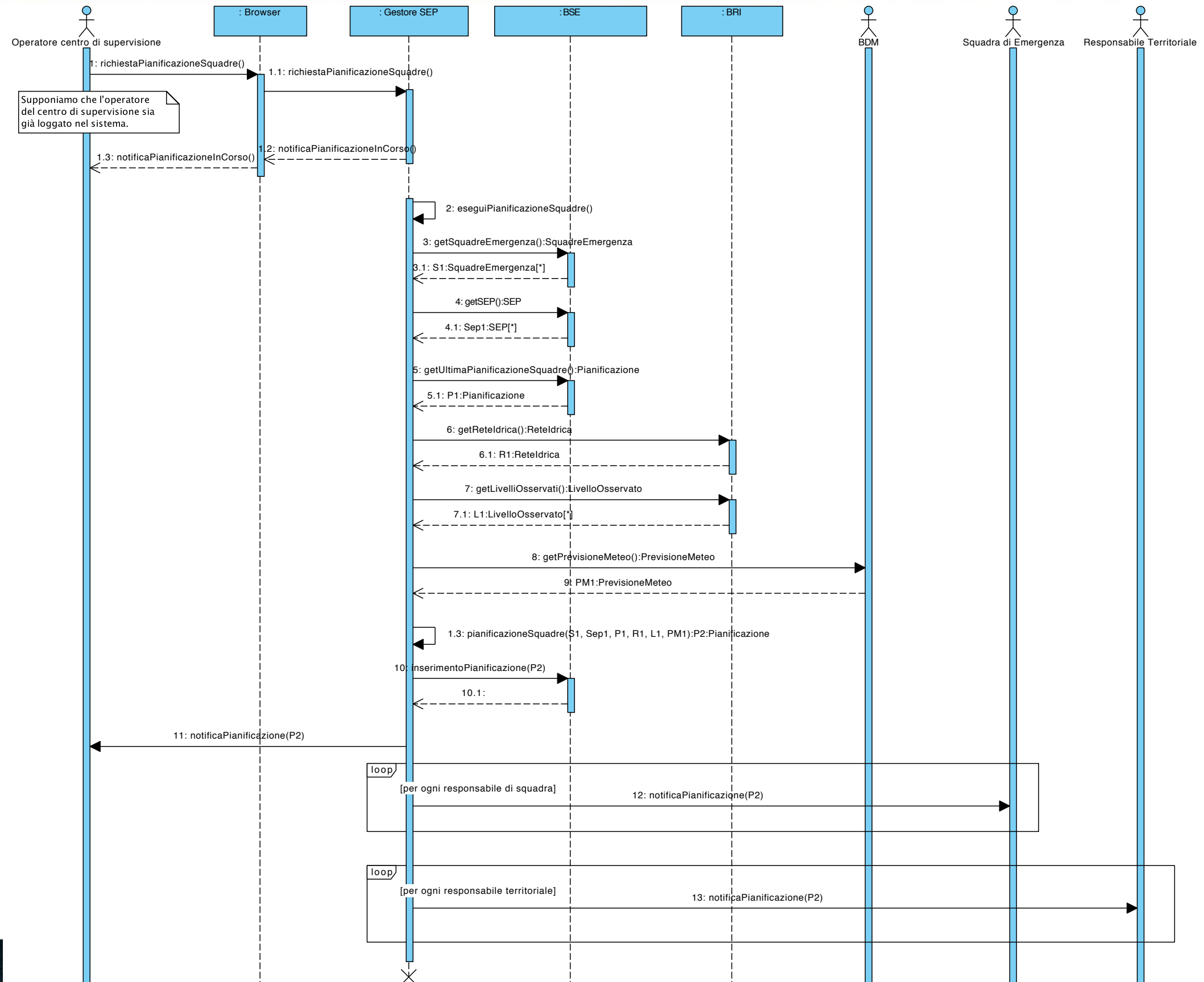
# Acquisizione Dati Idrometrici



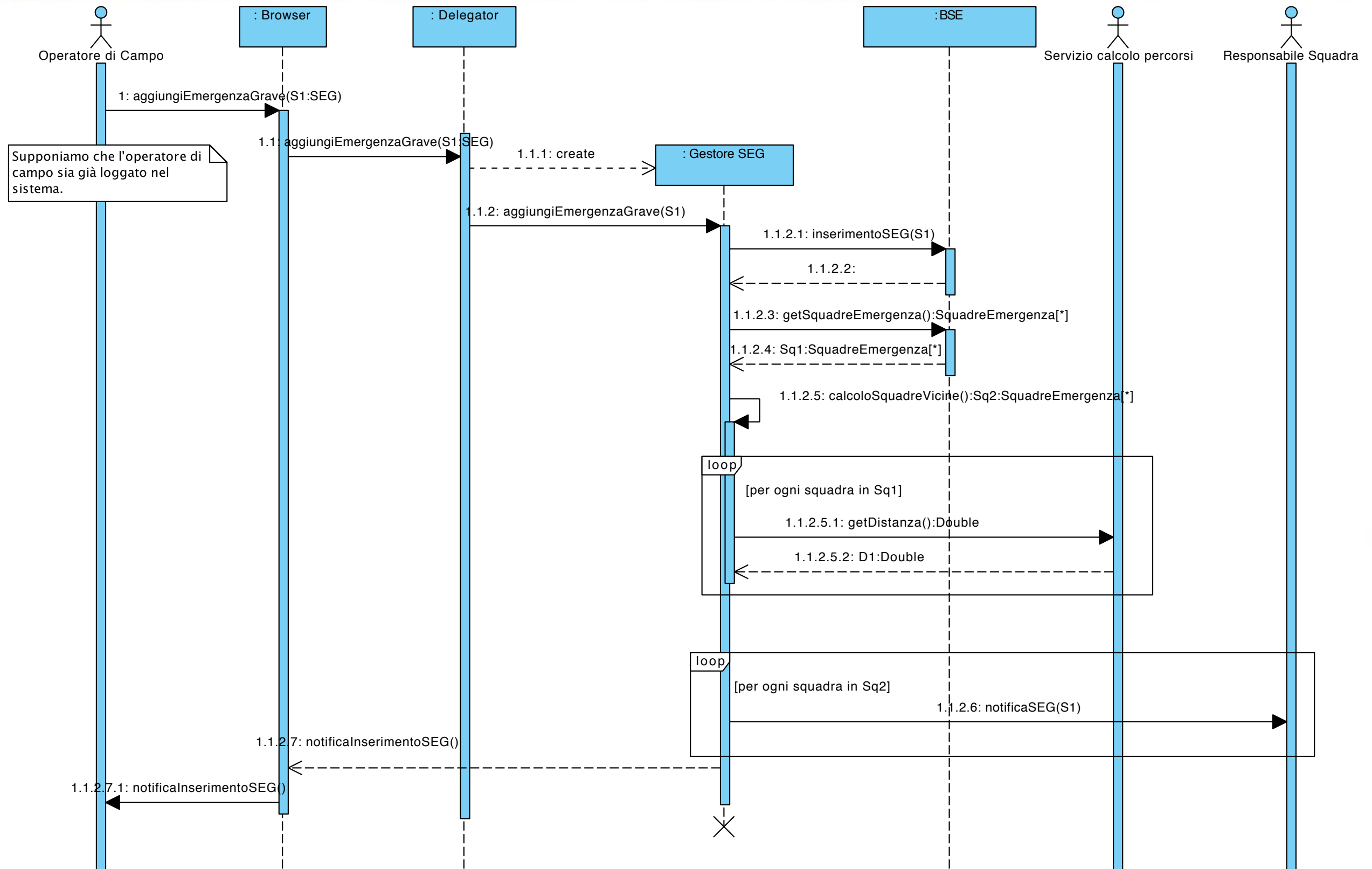
# Identificazione SEP



# Pianificazione SEP

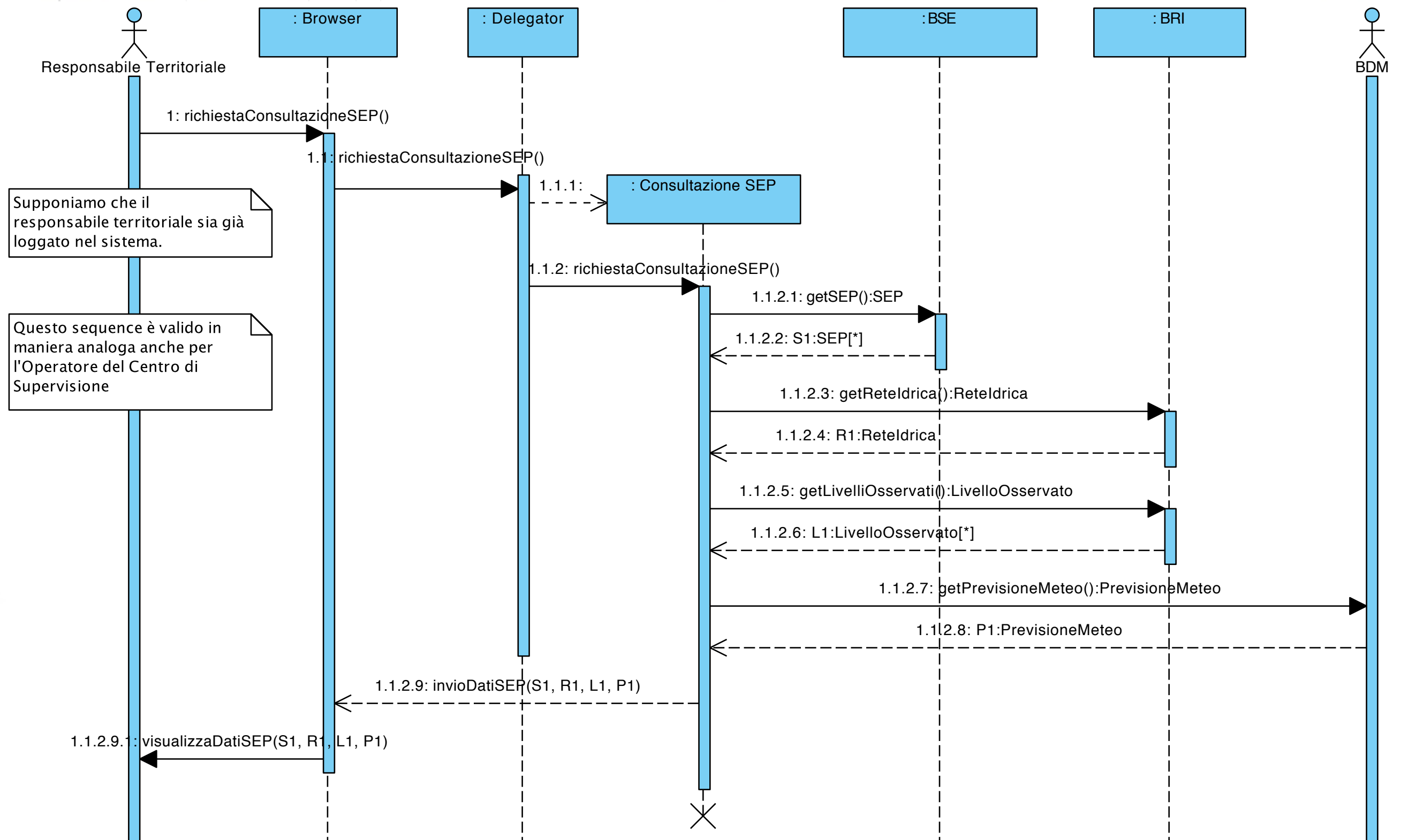


# Acquisizione SEG

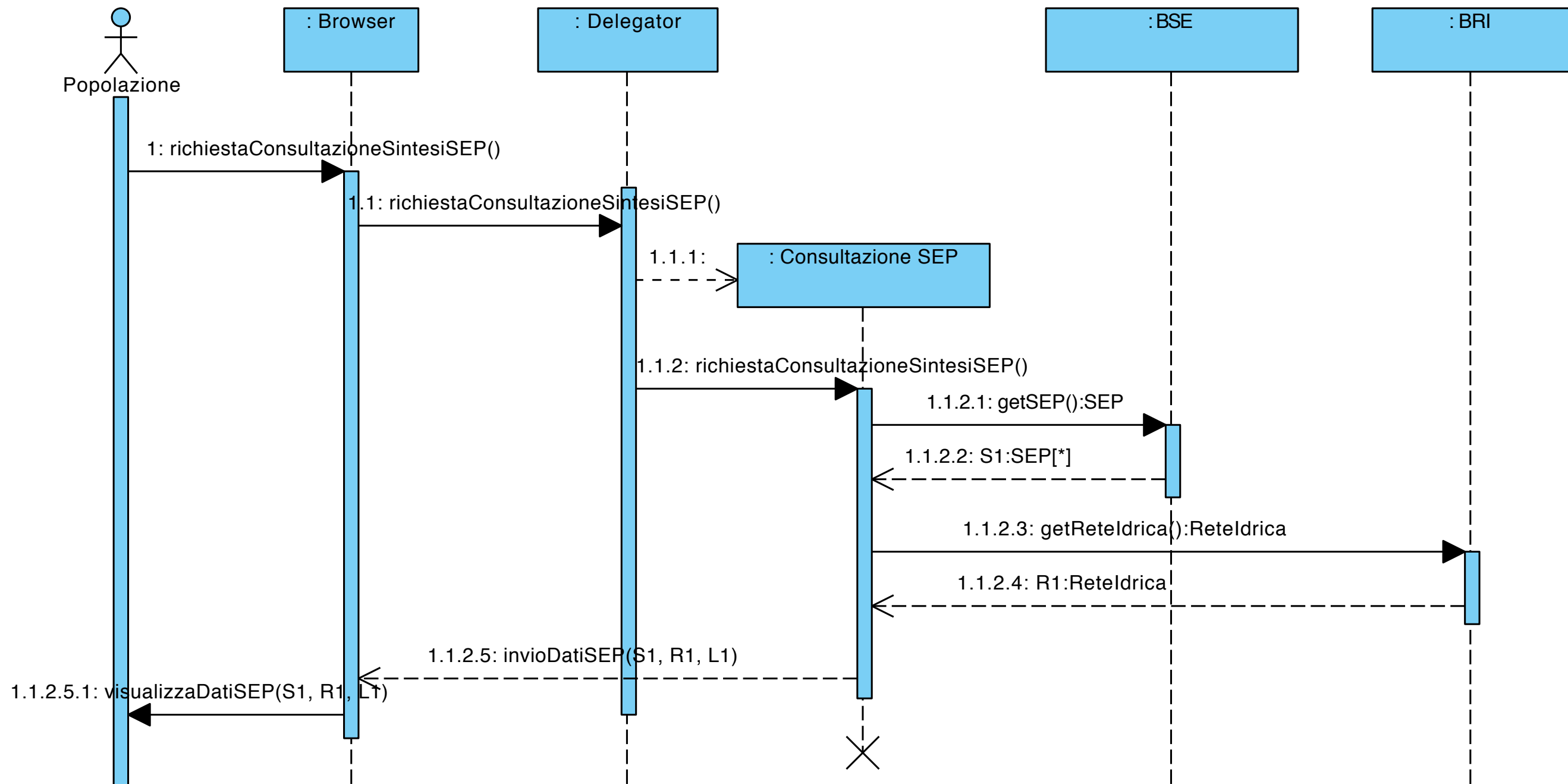




# Consulta SEP



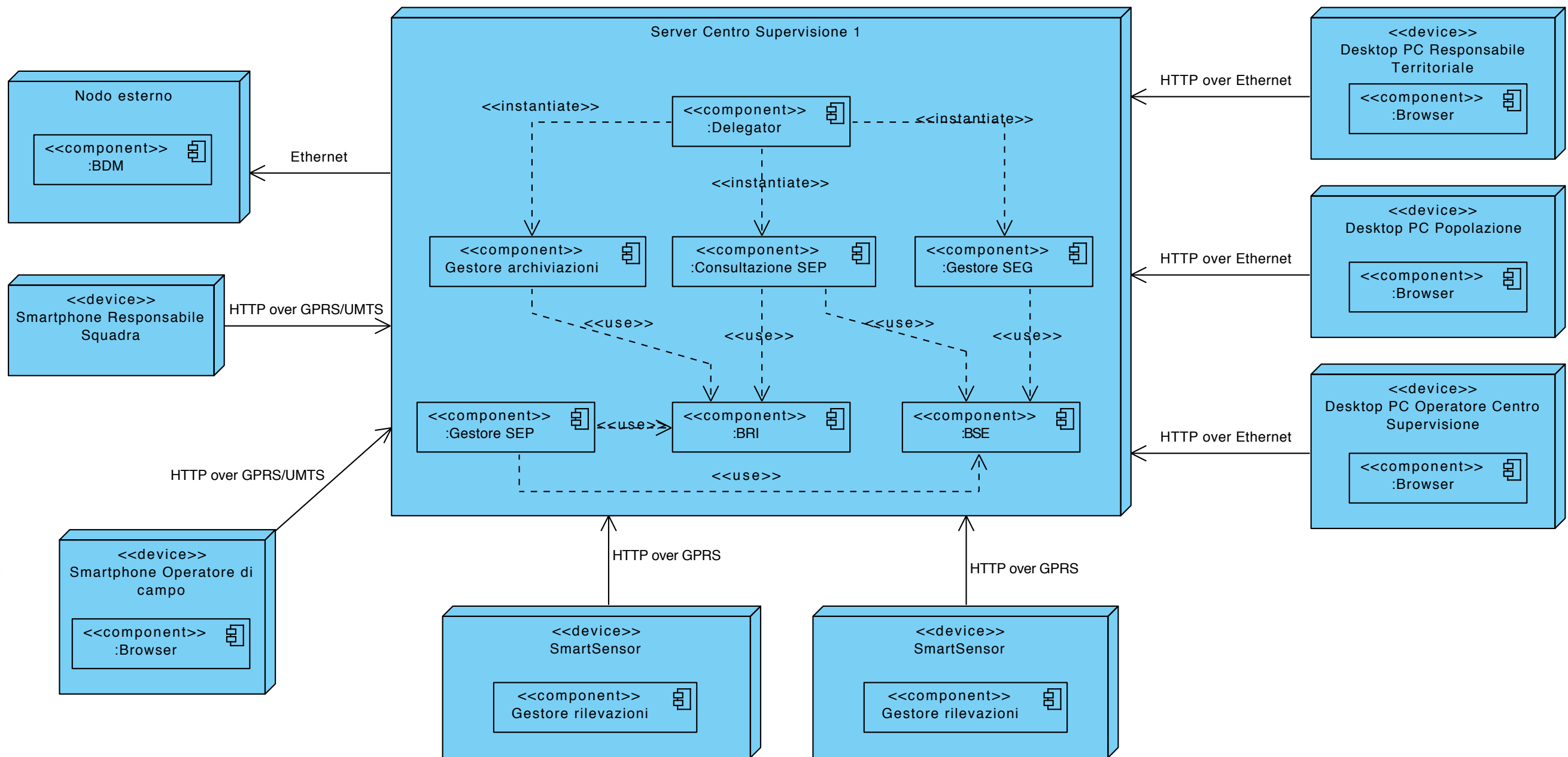
# Consulta Sintesi SEP



# Architettura Concreta

# Soluzione 1

# Soluzione 1





# Soluzione 1 - Valutazione

## **Pro:**

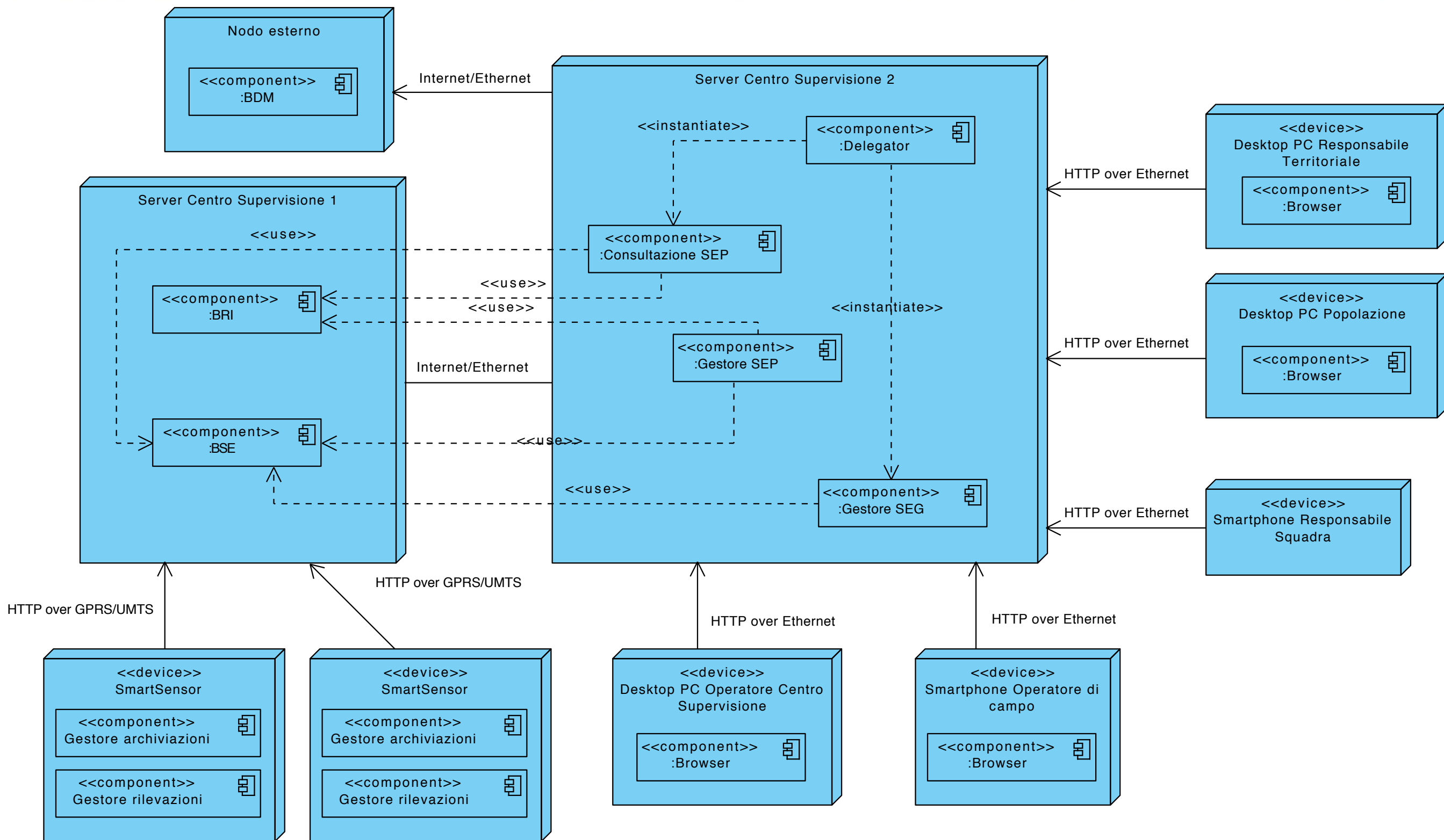
- ☐ Bassi costi di gestione e sviluppo.
- ☐ Gestione centralizzata del sistema.
- ☐ Gli SmartSensor e gli Smartphone degli operatori non necessitano di potenza di calcolo.
- ☐ Moderata larghezza di banda necessaria.

## **Contro:**

- ☐ Ritardi (il server potrebbe diventare un collo di bottiglia).
- ☐ Bassa resistenza ai guasti (un malfunzionamento del server blocca tutto il sistema).

# Soluzione 2

# Soluzione 2



# Soluzione 2 - Valutazione

## **Pro:**

- ☐ Separazione tra nodo di elaborazione e nodo di archiviazione.
- ☐ Le operazioni di acquisizione e memorizzazione dei dati idrometrici (alta frequenza) vengono eseguite in modo distribuito.
- ☐ Possibilità di specializzazione dell'hardware necessario (server per elaborazione e server per archiviazione).

## **Contro:**

- ☐ Elevato flusso dati tra unità di elaborazione e unità di memorizzazione
- ☐ Problemi di sicurezza legati all'accesso al database BRI da parte degli SmartSensor.

# Soluzione Scelta - Soluzione 1

Si è scelto di utilizzare la prima soluzione proposta perché ottimizza il flusso dati tra database e unità di calcolo; in questo modo possiamo assicurare una maggiore velocità di esecuzione degli algoritmi di pianificazione che dipendono fortemente dall'accesso ai dati.

Inoltre la soluzione evita i problemi di sicurezza legati all'accesso distribuito al database da parte degli SmartSensor.

Per migliorare l'Availability del sistema si potrebbe implementare un Disaster Recovery Plan che consenta, in caso di guasto al sistema principale, di avere una postazione esistente pronta all'uso. Inoltre è opportuno eseguire un backup periodico delle basi di dati in un luogo geograficamente distante.

# Tecnologie - Comunicazione

Si è scelto di utilizzare la normale rete dati GPRS/UMTS offerta dagli operatori telefonici per la comunicazione tra i Sensori, gli Smartphone ed il nodo di elaborazione centrale.

## **Pro:**

- ☐ Tutte le infrastrutture di rete sono preesistenti.
- ☐ Bassi costi.

## **Contro:**

- ☐ Le prestazioni dipendono dallo stato della rete GPRS/UMTS.



# Panoramica dei Costi

- ❑ Sensore Idrometrico (Sensore a ultrasuoni + Arduino + Modulo GPRS/UMTS + Scocca protettiva)
- ❑ Assemblaggio e messa in posizione dei Sensori Idrometrici
- ❑ Rete dati tra Sensori e Server (GPRS/UMTS)
- ❑ Smartphone (Modello Economico)
- ❑ Formazione degli Operatori
- ❑ Rete dati tra Smartphone e Server (GPRS/UMTS)
- ❑ Server
- ❑ Sviluppo e progettazione

# Stime dei Costi - Hardware

Nome	Prezzo singolo	Quantità necessaria	Prezzo complessivo
Sensore ULM20 - CAE	300,00 €	2000	600.000,00 €
Arduino + Modulo GPRS/GSM + Case blindato	250,00 €	2000	500.000,00 €
Assemblaggio - Installazione	250,00 €	2000	500.000 €
Aruba Private Cloud (16 core, 32GB RAM, 1TB, 5 IP, 4 Vlan)	471,60 €/mese	1	471,60 €/mese
Servizio Previsione Meteo	30.000 €/Anno	1	2.500 €/mese
Smatphone	200,00 €	400	80.000 €
Piano tariffario SIM dati	10 €/mese	2400	24.000 €/mese
Manutenzione Ordinaria	12/anno	1	350,00 €
Manutenzione Straordinaria (riparazione guasti)	~12/anno	1	450,00 €
<b>Investimento Hardware</b>			<b>1.680.000,00 €</b>
<b>Canone Mensile</b>			<b>27.771,60 €/mese</b>

# Stime dei Costi - Software

Ruolo	Durata (Giorni lavorativi)	Numero	Costo (Giornaliero)
Project Manager	250	1	1.500,00 €
Sviluppatori	250	5	600,00 €
Formazione Operatore	5	5	80,00 €
Progettazione e Sviluppo			1.125.000,00 €
Totale			1.127.000,00 €

**Totale complessivo:**

2.807.000,00 € + 27.771,60 € Mensili

# Architettura dei Dati

# Metodologia

Sono stati prodotti gli schemi logici relazionali delle basi di dati BRI (Base dati Rete Idrica), BSE (Base Dati Segnalazione Emergenza) e BDM (Base Dati Meteo).

Partendo dagli schemi logici relazionali, mediante un processo di reverse engineering, sono stati prodotti gli schemi concettuali delle basi dati BRI, BSE e BDM.

# BDM - Schema logico

**Sensore**(codiceSensore, marca, modello, tipo\*, idCentralina)

**RilevazioneSensore**(codiceSensore, data, valore)

**Centralina**(idCentralina, long, lat, ASLM, idCellaGeografica)

**CellaGeografica**(idCella, longitudine, latitudine)

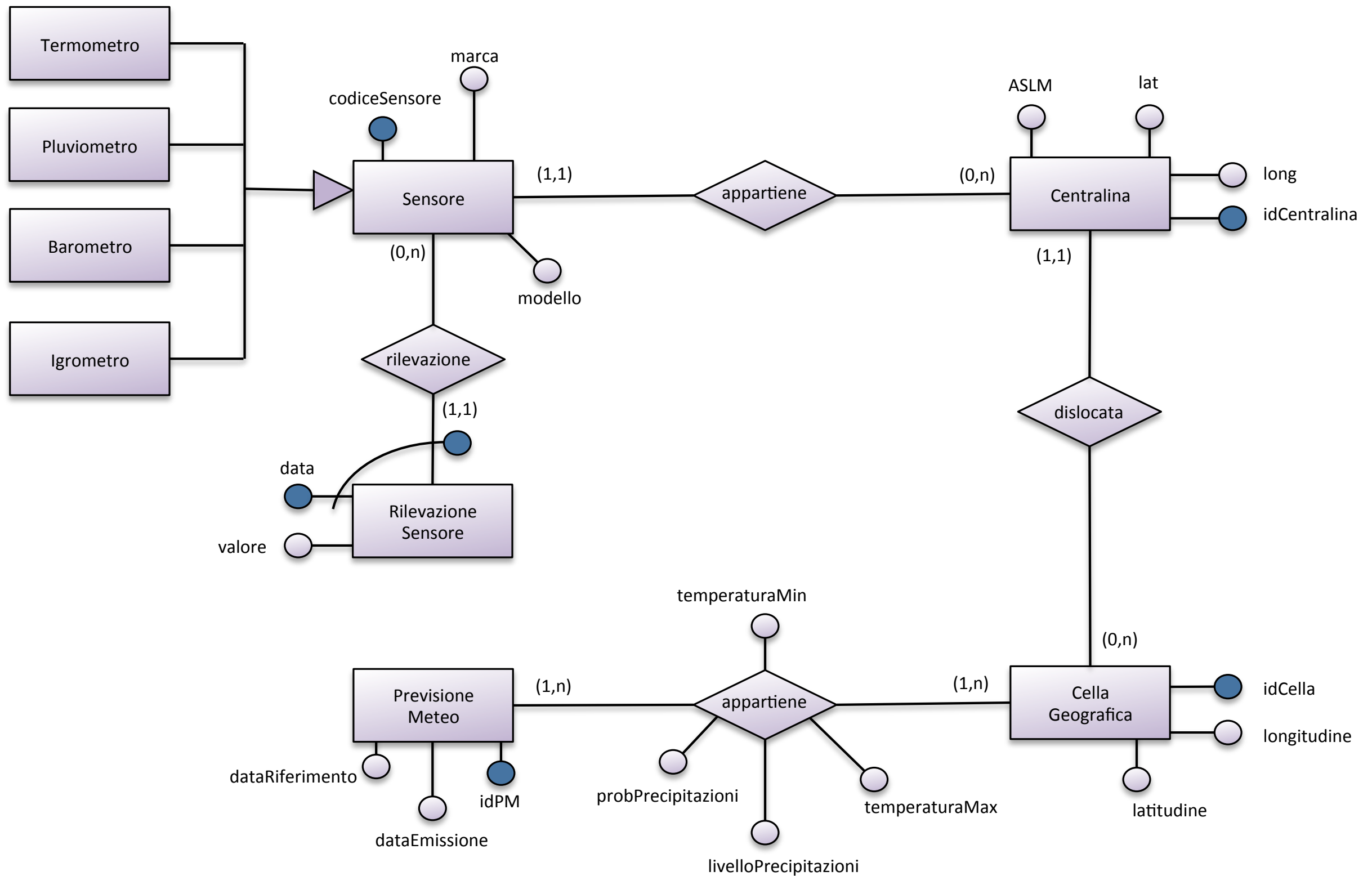
**PrevisioneMeteo**(idPM, dataEmissione, dataRiferimento)

**Previsione-Cella**(idPM, idCella, temperaturaMin, temperaturaMax, probPrecipitazioni, livelloPrecipitazioni)

\*Il tipo del sensore è Termometro, Pluviometro, Barometro o Igrometro



# BDM - ER



# BRI - Schema logico

**Fiume**(idFiume, nome)

**TrattoFiume**(idTrattoFiume, lunghezzaTratto, livelloMassimo, livelloAllerta, portata, idFiume, nodoInizio, nodoFine)

**Nodoldrico**(idNodoldrico, altitudine, latitudine, longitudine)

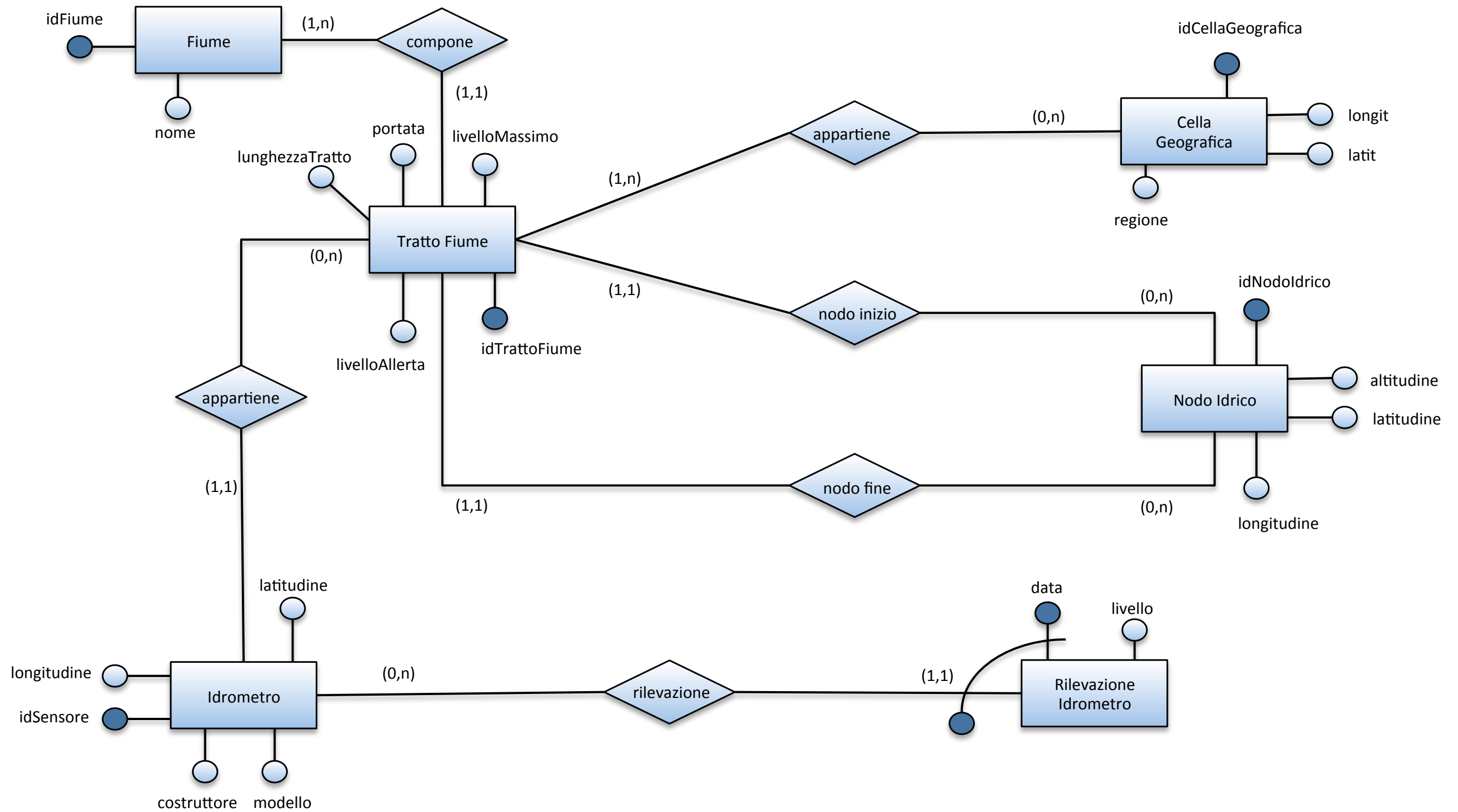
**CellaGeografica**(idCellaGeografica, longit, latit, regione)

**TrattoFiume-Cella**(idCella, idTratto)

**Idrometro**(idSensore, costruttore, modello, latitudine, longitudine, idTrattoFiume)

**RilevazioneIdrometro**(idSensore, data, livello)

# BRI - ER



# BSE - Schema logico

**SEG**(idSeg, tipo, latitudine, longitudine, altitudine, data, idTrattoFiume, CFOperatore)

**TrattoFiume**(idTrattoFiume, latitudineInizio, latitudineFine, longitudineInizio, longitudineFine)

**SEP**(idSEP, data, livelloPericolo)

**Pianificazione**(idPianificazione, data)

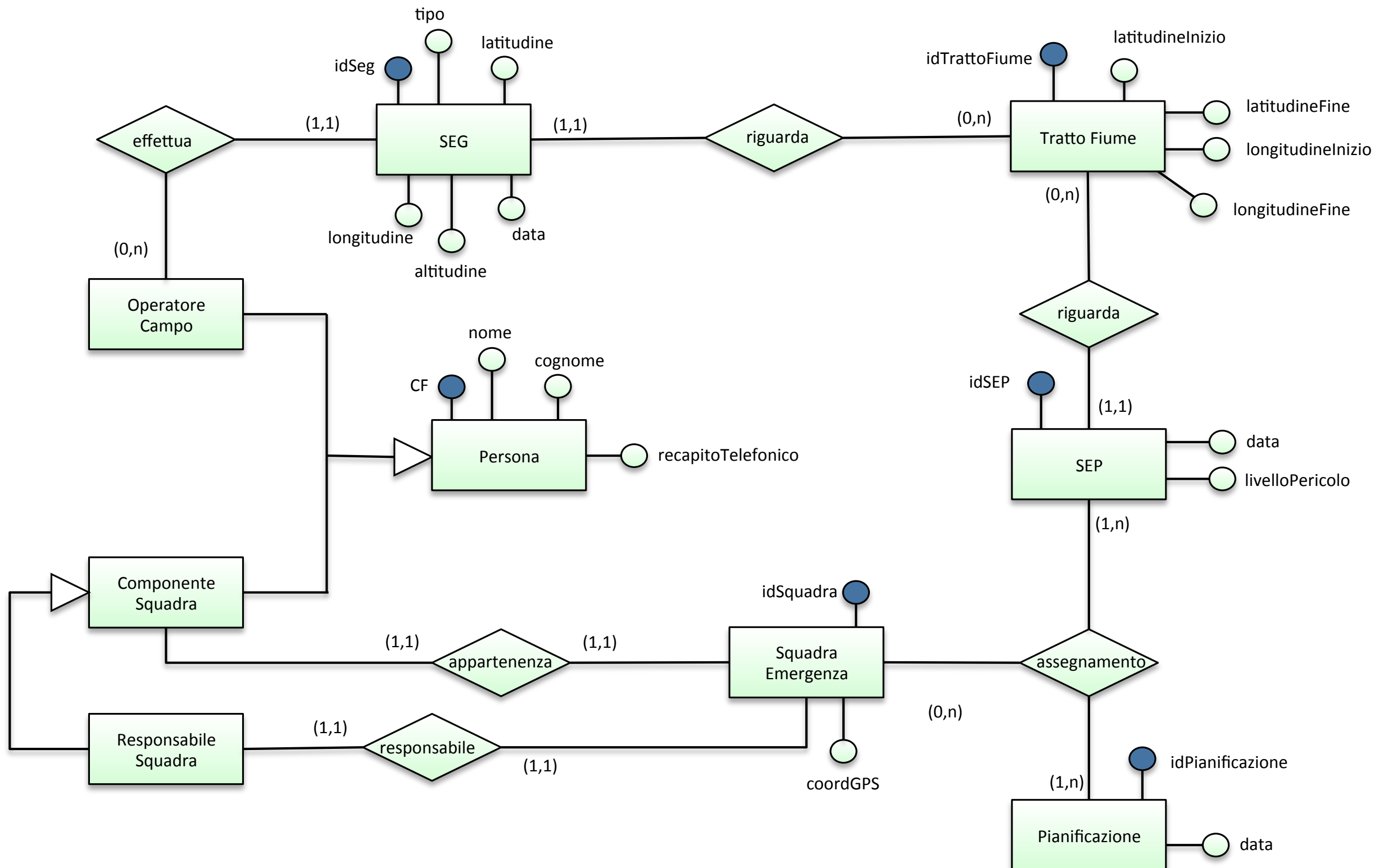
**SquadraEmergenza**(idSquadra, CFResponsabile, coordGPS)

**Assegnamento**(idPianificazione, idSEP, idSquadra)

**OperatoreCampo**(CF, nome, cognome, recapitoTelefonico)

**ComponenteSquadra**(CF, nome, cognome, recapitoTelefonico, idSquadra)

# BSE - ER





# Corrispondenze/Eterogeneità (1)

## Corrispondenze tra BRI e BDM

BRI	BDM	Tipo	Scelta
Idrometro	Sensore	Iperonimia (IS-A)	Si mantiene Sensore
Idrometro.costruttore	Sensore.marca	Sinonimia	Si mantiene Sensore.marca
Idrometro.idSensore	Sensore.codiceSensore	Sinonimia	Si mantiene Sensore.codiceSensore
RilevazioneIdrometro.idSensore	RilevazioneSensore.codiceSensore	Sinonimia	Si mantiene RilevazioneSensore.codiceSensore
RilevazioneIdrometro.livello	RilevazioneSensore.valore	Sinonimia	Si mantiene RilevazioneSensore.valore
CellaGeografica.idCellaGeografica	CellaGeografica.idCella	Sinonimia	Si mantiene CellaGeografica.idCella
CellaGeografica.longit	CellaGeografica.longitudine	Sinonimia	Si mantiene CellaGeografica.longitudine
CellaGeografica.latit	CellaGeografica.latitudine	Sinonimia	Si mantiene CellaGeografica.latitudine



# Corrispondenze/Eterogeneità (2)

Corrispondenze tra BRI e BSE			
BRI	BSE	Tipo	Scelta
TrattoFiume.nodoInizio	TrattoFiume.latitudineInizio	Differenti tipi	Si sceglie nodoldrico come tipo di dato rappresentato
	TrattoFiume.longitudineInizio	Differenti tipi	Si sceglie nodoldrico come tipo di dato rappresentato
TrattoFiume.nodoFine	TrattoFiume.latitudineFine	Differenti tipi	Si sceglie nodoldrico come tipo di dato rappresentato
	TrattoFiume.longitudineFine	Differenti tipi	Si sceglie nodoldrico come tipo di dato rappresentato

# Integrazione

# Scelte Architetture (1)

Si è scelto di unificare gli schemi logici utilizzando il processo

## **EII (Enterprise Information Integration)**

che ci permette di fornire un'unica interfaccia per la visualizzazione di tutti i dati presenti nei singoli database.

Questo può essere implementato in due modi:

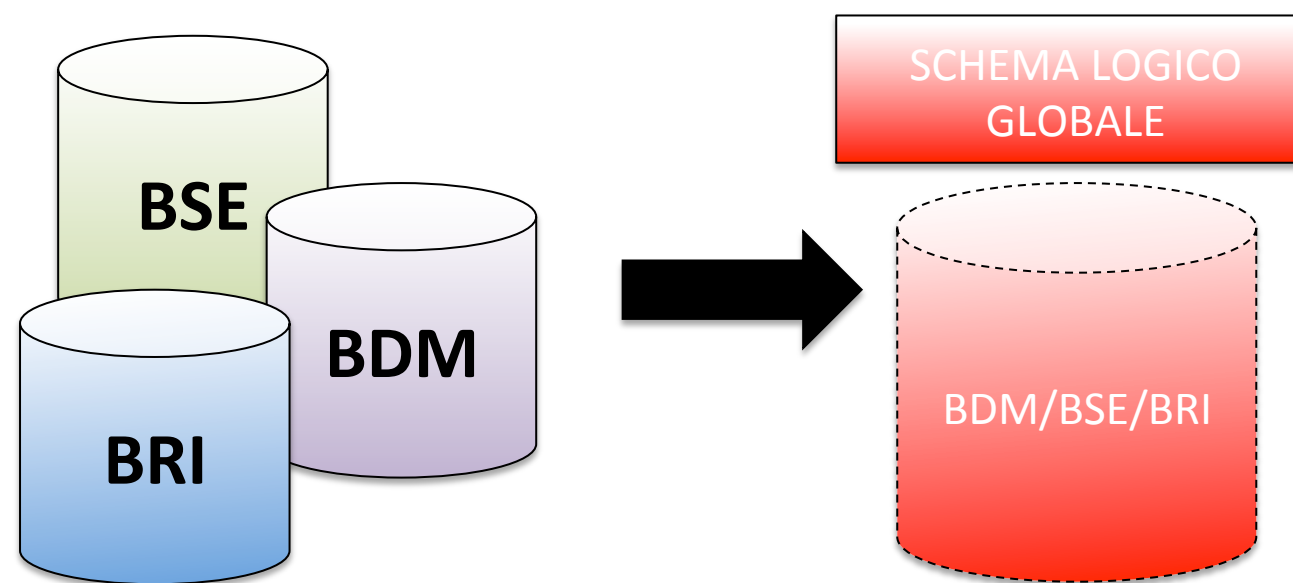
- ☐ **Data Warehousing:** viene creato un nuovo database fisico in cui vengono copiati tutti i dati dei singoli DB secondo lo schema logico globale (durante questa operazione possono essere effettuate operazioni per il miglioramento della qualità dei dati, e.g. *Record Linkage*). Tutte le interrogazioni vengono eseguite sul nuovo database. Periodicamente è necessario effettuare un riallineamento dei dati del database globale con quelli dei singoli database locali.
- ☐ **Virtual Data Integration:** viene creato uno schema logico globale che fornisce un'interfaccia unificata di visualizzazione dei dati. Tutte le interrogazioni che vengono effettuate su questa interfaccia vengono riadattate, attraverso un mediatore, agli schemi logici dei singoli database.

# Scelte Architettureali (2)

La soluzione proposta si basa su **Virtual Data Integration**:

L'architettura prevede che i dati rimangano nei database attuali e siano acceduti tramite un mediatore che riadatta le interrogazioni sullo schema logico virtuale a interrogazioni locali sui singoli database.

I mapping utilizzati dal mediatore sono stati definiti con la modalità **Global As View (GAV)**.

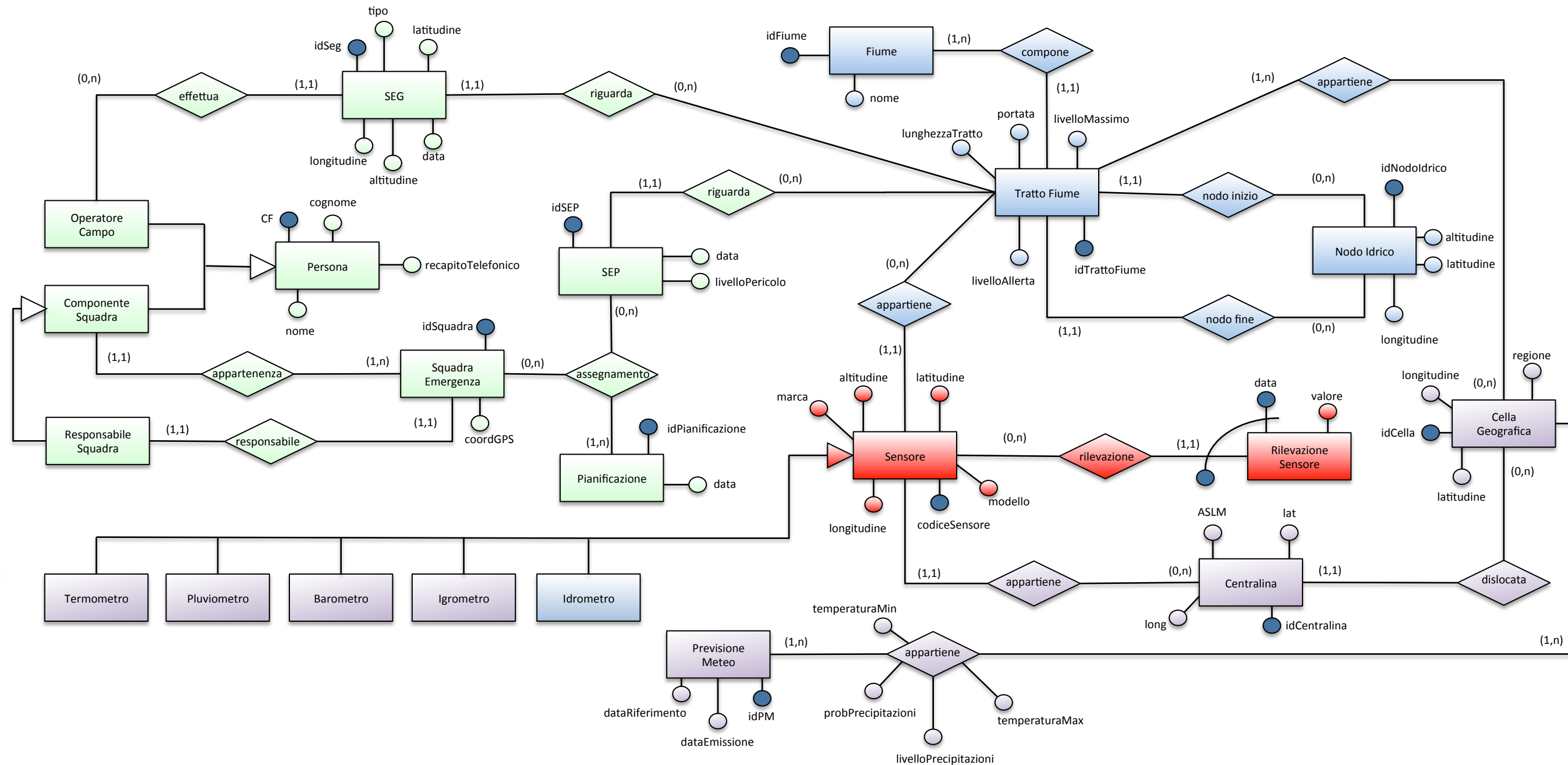


# Assunzioni

- ❑ I dati della tabella BDM.Sensore e BRI.Idrometro sono disgiunti, e non ci sono conflitti tra le chiavi primarie delle due tabelle. Allo stesso modo anche tra BDM.RilevazioneSensore e BRI.RilevazioneIdrometro non esistono conflitti.
- ❑ le tuple di BSE.TrattoFiume sono un sottoinsieme delle tuple di BRI.TrattoFiume (vengono copiate in BSE solo le tuple interessate da una SEP).
- ❑ Le tuple di BRI.CellaGeografica sono un sottoinsieme delle tuple di BDM.CellaGeografica. In BRI sono contenute le sole celle geografiche che sono attraversate da almeno un tratto di fiume; nel momento in cui vengono inserite in BRI, viene aggiunta l'informazione relativa alla regione di appartenenza della cella geografica (attributo "regione").



# Modello ER Integrato





# Schema logico Integrato (1)

**SEG**(idSeg, tipo, latitudine, longitudine, altitudine, data, idTrattoFiume, CFOperatore)

**SEP**(idSEP, data, livelloPericolo)

**Pianificazione**(idPianificazione, data)

**SquadraEmergenza**(idSquadra, CFResponsabile, coordGPS)

**Assegnamento**(idPianificazione, idSEP, idSquadra)

**OperatoreCampo**(CF, nome, cognome, recapitoTelefonico)

**ComponenteSquadra**(CF, nome, cognome, recapitoTelefonico, idSquadra)

**Sensore**(codiceSensore, marca, modello, tipo, altitudine, latitudine, longitudine, idCentralina, idTrattoFiume)

# Schema logico Integrato (2)

**RilevazioneSensore**(codiceSensore, data, valore)

**Centralina**(idCentralina, long, lat, ASLM, idCellaGeografica)

**CellaGeografica**(idCella, longitudine, latitudine, regione)

**PrevisioneMeteo**(idPM, dataEmissione, dataRiferimento)

**Previsione-Cella**(idPM, idCella, temperaturaMin, temperaturaMax, probPrecipitazioni, livelloPrecipitazioni)

**Fiume**(idFiume, nome)

**TrattoFiume**(idTrattoFiume, lunghezzaTratto, livelloMassimo, livelloAllerta, portata, idFiume, nodoInizio, nodoFine)

**Nodoldrico**(idNodoldrico, altitudine, latitudine, longitudine)

**TrattoFiume-Cella**(idCella, idTratto)

# Mapping (1)

```
CREATE VIEW SEG(idSeg, tipo, latitudine, longitudine, altitudine, data,  
idTrattoFiume, CFOperatore) AS  
  SELECT idSeg, tipo, latitudine, longitudine, altitudine, data, idTrattoFiume,  
  CFOperatore  
FROM BSE.SEG
```

```
CREATE VIEW SEP(idSEP, data, livelloPericolo) AS  
  SELECT idSEP, data, livelloPericolo FROM BSE.SEP
```

```
CREATE VIEW Pianificazione(idPianificazione, data) AS  
  SELECT idPianificazione, data FROM BSE.Pianificazione
```

```
CREATE VIEW SquadraEmergenza(idSquadra, CFResponsabile, coordGPS) AS  
  SELECT idSquadra, CFResponsabile, coordGPS  
FROM BSE.SquadraEmergenza
```

```
CREATE VIEW Assegnamento(idPianificazione, idSEP, idSquadra) AS  
  SELECT idPianificazione, idSEP, idSquadra  
FROM BSE.Assegnamento
```

# Mapping (2)

**CREATE VIEW** OperatoreCampo(CF, nome, cognome, recapitoTelefonico) **AS**

**SELECT** CF, nome, cognome, recapitoTelefonico  
**FROM** BSE.OperatoreCampo

**CREATE VIEW** ComponenteSquadra(CF, nome, cognome, recapitoTelefonico, idSquadra) **AS**

**SELECT** CF, nome, cognome, recapitoTelefonico, idSquadra  
**FROM** BSE.ComponenteSquadra

**CREATE VIEW** RilevazioneSensore(codiceSensore, data, valore) **AS**

**SELECT** codiceSensore, data, valore  
**FROM** BDM.RilevazioneSensore  
**UNION**

**SELECT** idSensore **AS** codiceSensore, data, livello **AS** valore  
**FROM** BRI.RilevazioneIdrometro

# Mapping (3)

```
CREATE VIEW Sensore(codiceSensore, marca, modello, tipo,  
altitudine, latitudine, longitudine, idCentralina, idTrattoFiume) AS  
SELECT S.codiceSensore AS codiceSensore, S.marca AS  
marca, S.modello AS modello, S.tipo AS tipo, C.ASLM AS  
altitudine, C.lat AS latitudine, C.long AS longitudine,  
S.idCentralina AS idCentralina, null AS idTrattoFiume  
FROM BDM.Sensore AS S JOIN BDM.Centralina AS C ON  
S.idCentralina  
UNION  
SELECT idSensore AS codiceSensore, costruttore AS marca,  
modello AS modello, "Idrometro" AS tipo, null AS altitudine,  
latitudine AS latitudine, longitudine AS longitudine, null AS  
idCentralina, idTrattoFiume AS idTrattoFiume  
FROM BRI.Idrometro
```



# Mapping (4)

```
CREATE VIEW Centralina(idCentralina, long, lat, ASLM, idCellaGeografica) AS  
SELECT idCentralina, long, lat, ASLM, idCellaGeografica  
FROM BDM.Centralina
```

```
CREATE VIEW CellaGeografica(idCella, longitudine, latitudine, regione) AS  
SELECT M.idCella AS idCella, M.longitudine AS longitudine, M.latitudine  
AS latitudine, I.regione AS regione  
FROM BDM.CellaGeografica M LEFT OUTER JOIN BRI.CellaGeografica I  
ON M.idCella = I.idCella
```

```
CREATE VIEW PrevisioneMeteo(idPM, dataEmissione, dataRiferimento) AS  
SELECT idPM, dataEmissione, dataRiferimento  
FROM BDM.PrevisioneMeteo
```

```
CREATE VIEW Previsione-Cella(idPM, idCella, temperaturaMin,  
temperaturaMax, probPrecipitazioni, livelloPrecipitazioni) AS  
SELECT idPM, idCella, temperaturaMin, temperaturaMax,  
probPrecipitazioni, livelloPrecipitazioni  
FROM BDM.Previsione-Cella
```



# Mapping (5)

```
CREATE VIEW Fiume(idFiume, nome) AS  
  SELECT idFiume, nome  
  FROM BRI.Fiume
```

```
CREATE VIEW TrattoFiume(idTrattoFiume, lunghezzaTratto,  
  livelloMassimo, livelloAllerta, portata, idFiume, nodoInizio, nodoFine) AS  
  SELECT idTrattoFiume, lunghezzaTratto, livelloMassimo, livelloAllerta,  
  portata, idFiume, nodoInizio, nodoFine  
  FROM BRI.TrattoFiume
```

```
CREATE VIEW Nodoldrico(idNodoldrico, altitudine, latitudine,  
  longitudine) AS  
  SELECT idNodoldrico, altitudine, latitudine, longitudine  
  FROM BRI.Nodoldrico
```

```
CREATE VIEW TrattoFiume-Cella(idCella, idTratto) AS  
  SELECT idCella, idTratto  
  FROM BRI.TrattoFiume-Cella
```

# Query (1)

Esempio(1) di interrogazione sullo schema globale che visita due schemi locali:

*“Selezionare tutte le previsioni meteo relative alle celle nelle quali passa il fiume Adige nella regione Veneto”.*

```
SELECT CG. idCella, PC.probPrecipitazioni, PC.livelloPrecipitazioni
FROM Fiume F
    JOIN TrattoFiume TF ON F.idFiume = TF.idFiume
    JOIN TrattoFiume-Cella CT ON CT.idTratto = TF.idTrattoFiume
    JOIN CellaGeografica CG ON CG.idCella = CT.idCella
    JOIN Previsione-Cella PC ON PC.idCella = CG.idCella
WHERE F.nome = “Adige” AND CG.Regione = “Veneto”
```

# Unfolding (1)

Unfolding dell'interrogazione precedente:

```
SELECT CGM.idCella, PC.probPrecipitazioni, PC.livelloPrecipitazioni
FROM BRI.Fiume F
    JOIN BRI.TrattoFiume TF ON F.idFiume = TF.idFiume
    JOIN BRI.TrattoFiume-Cella CT ON
        CT.idTratto = TF.idTrattoFiume
    JOIN BRI.CellaGeografica CGI ON CGI.idCella = CT.idCella
    LEFT OUTER JOIN BDM.CellaGeografica CGM ON
        CGM.idCella = CGI.idCella
    JOIN BDM.Previsione-Cella PC ON PC.idCella = CGM.idCella
WHERE F.nome = "Adda" AND CGI.Regione = "Veneto"
```

# Query (2)

Esempio(2) di interrogazione sullo schema globale che visita due schemi locali:

*“Selezionare l’elenco di tutti i sensori idrometrici e pluviometrici”.*

**SELECT** codiceSensore, marca, modello

**FROM** Sensore S

**WHERE** S.tipo = “Idrometro” OR S.tipo = “Pluviometro”

# Unfolding (2)

Unfolding dell'interrogazione precedente:

```
SELECT codiceSensore, marca, modello  
FROM BDM.Sensore S  
WHERE S.tipo = "Pluviometro"  
UNION  
SELECT idSensore AS codiceSensore, costruttore AS marca,  
modello  
FROM BRI.Idrometro
```

# Open Data (1)

Gli Open Data sono caratterizzati da:

- ☐ Essere accessibili gratuitamente
- ☐ Essere pubblicati in formato aperto
- ☐ Avere una licenza aperta

Questi dati vengono pubblicati in via del tutto gratuita sul sito web del committente del progetto, sotto forma di uno o più file disponibili per il download.

I costi di produzione dei file sono compresi in quelli di manutenzione server e i costi di pubblicazione sul sito sono trascurabili.


Per scaricare il file non è richiesta alcuna iscrizione.



# Open Data (2)

Si è scelto di distribuire liberamente i dati relativi alle serie storiche di rilevazione dei livelli dei corsi d'acqua, i dati relativi alle SEP e quelli relativi alle SEG.

Per quanto riguarda i primi due dati, le informazioni non causano alcun problema di privacy in quanto non riguardano persone fisiche. Al contrario, per le informazioni relative alle SEG è necessario fornire i dati privi dell'identificativo dell'operatore di campo che ha effettuato la segnalazione.

Sono state analizzate diversi tipi di licenze per la pubblicazione di dati aperti; dopo aver analizzato i possibili impieghi è stato scelto di rilasciare i dati con una licenza di tipo **CC BY** (Creative Commons con obbligo di citare la fonte dei dati). 

Esistono altre tipologie di licenze come:

ODbL (Open Database License) che vincola gli utilizzatori a citare la fonte ed a condividere i dati senza alterarli e CC0 (Creative Commons 0) dove i dati sono di pubblico dominio.

# Open Data (3)

Si è deciso di pubblicare i dati in un formato accessibile a tutti, sia privati sia aziende, consultabile in modo semplice e che non richieda una conoscenza di linguaggi di programmazione particolari.

Per adempiere a questo scopo è stato scelto il formato **CSV** (Comma separated values).

Il formato dati è il seguente:

`SegnalazioneEmergenzaGrave.csv`

`tipo, latitudine, longitudine, altitudine, data, nomeFiume`

`SituazioniEmergenzaPotenziale.csv`

`data, livello_di_pericolo, latitudineInizio, latitudineFine,  
longitudineInizio, longitudineFine`

`RilevazioneIdrometro.csv`

`marca, modello, tipo, altitudine, latitudine, longitudine,  
data, valore`

Si noti che le informazioni sensibili, protette da copyright o di competenza altrui sono state omesse.

# Open Data (4)

I dati Open forniti dal sistema possono essere utilizzati per studi scientifici a cui interessa incrociare i dati fluviali con altre tipologie di dati (e.g. Morfologia del territorio) o per effettuare statistiche sulla piovosità in determinate zone.