

 	Titolo	RTI: Leonardo - IBM – Sistemi Informativi – Onit – Prisma – Deas - Reply PS
	Proprietario: Autorità di Bacino del Distretto Idrografico della Sicilia – RTI Autore: RTI Data rilascio: 14/06/2024 Codice: S22028.02.1004PROC Revisione: 04.00	

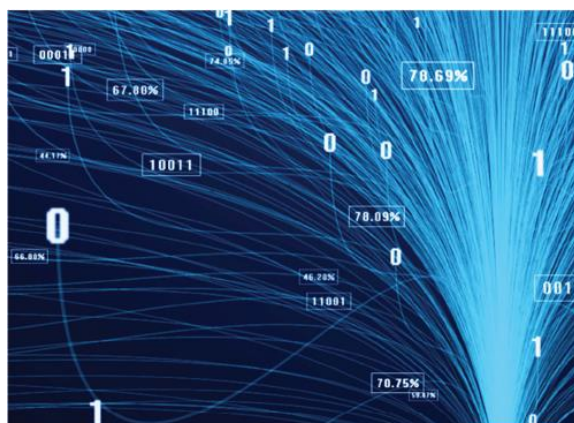
**DIPARTIMENTO REGIONALE DELL'AUTORITA' DI BACINO
DEL DISTRETTO IDROGRAFICO DELLA SICILIA
SERVIZIO 5 - ATTUAZIONE INTERVENTI A TITOLARITA' O A REGIA**

Programma Operativo Ambiente FSC 2014-2020 – Asse 2 – Linea di Azione 2.3.1 Progetto: “Autorità di Bacino del Distretto Idrografico della Sicilia – Interventi per il miglioramento della qualità dei corpi idrici”- Linea d'intervento L3- Aggiornamento quadro conoscitivo in materia di derivazioni - Attività “Servizio informatico per la creazione banca dati documentale dei titoli di derivazione” e “Servizio tecnico di reperimento dati presso uffici competenti al rilascio di titoli concessori e autorizzatori (Genio Civile e Dipartimento dell'acqua e dei rifiuti) verifica e caricamento in banca dati” - CUP: F62G16000000001	Accordo Quadro Consip denominato “Applicativi di Data management e servizi di PMO - Ordine Diretto - Lotto 4 - Servizi Applicativi di Data Management per le PAL Centro/Sud” CIG A.Q. 8184365FA4 CIG DERIVATO 9264301708
--	---

Piattaforma IDROCAP

Servizio LA.DW.1

- Documentazione delle procedure Batch DTS -



Autore del documento: **RTI**

  
Programmazione FSC 2014-2020 Piano Operativo Ambiente

 	Titolo	RTI: Leonardo - IBM – Sistemi Informativi – Onit – Prisma – Deas - Reply PS
	Proprietario: Autorità di Bacino del Distretto Idrografico della Sicilia – RTI Autore: RTI Data rilascio: 14/06/2024 Codice: S22028.02.1004PROC Revisione: 04.00	

1	SEZIONE DI CONTROLLO DEL DOCUMENTO	3
2	DOCUMENTI DI RIFERIMENTO	4
3	ACRONIMI E DEFINIZIONI	4
4	INTRODUZIONE.....	5
5	SINCRONIZZAZIONE TRA DATABASE MASTER E DATABASE SLAVES	6
5.1	METODO DI SINCRONIZZAZIONE.....	6
6	ESECUZIONE PROGRAMMATICA DEL BACKUP DEL DB.....	8
6.1	PROCEDURA DI BACKUP DEL DB PRINCIPALE	8
7	ESECUZIONE PROGRAMMATICA DEL FILE SYSTEM DISTRIBUITO	8

 		Titolo		RTI: Leonardo - IBM – Sistemi Informativi – Onit – Prisma – Deas - Reply PS
		Proprietario: Autorità di Bacino del Distretto Idrografico della Sicilia – RTI Autore: RTI Data rilascio: 14/06/2024 Codice: S22028.02.1004PROC Revisione: 04.00		

1 SEZIONE DI CONTROLLO DEL DOCUMENTO

1.1 LISTA DEI CAMBIAMENTI

Riepilogo delle versioni			
Versione	Data	Nome	Motivo aggiornamento
01.00	31/03/2023	RTI	Prima emissione (Lotto 1)
02.00	26/06/2023	RTI	Estensione funzionalità (Lotto 2)
03.00	26/09/2023	RTI	Estensione funzionalità (Lotto 3)
04.00	14/06/2024	RTI	Estensione funzionalità (Lotto 4)

1.2 LISTA DISTRIBUZIONE DEL DOCUMENTO

Lista di Distribuzione

Società	Nome	Cognome	Ruolo
Regione Siciliana - Dipartimento Regionale dell'Acqua e dei Rifiuti	Vito	Cutrone	RUP
	Raffaele	Di Salvo	DEC
Regione Siciliana – Dipartimento Regionale dell'Autorità di bacino	Giuseppe	Sciacca	Punto Ordinante
	Enzo Marco	Castrogiovanni	Collaboratore RUP/DEC
LEONARDO S.p.A.	Antonio	Rumiz	RUAC
	Ennio	Girimonte	Referente Tecnico
	Pasquale	Fierro	PM

 	Titolo	RTI: Leonardo - IBM – Sistemi Informativi – Onit – Prisma – Deas - Reply PS
	Proprietario: Autorità di Bacino del Distretto Idrografico della Sicilia – RTI Autore: RTI Data rilascio: 14/06/2024 Codice: S22028.02.1004PROC Revisione: 04.00	

2 DOCUMENTI DI RIFERIMENTO

Rif.	Nome Documento	Titolo
DOC1	AQDML3IRCP01SR01 Rev 04.00 Autorità di Bacino del Distretto Idrografico della Sicilia-SdR	LA.DW.1 - Specifica dei Requisiti
DOC2	AQDML3IRCP01SDD01 Rev.04.00 - IDROCAP - Specifiche di progettazione	Piattaforma IDROCAP Servizio LA.DW.1 - Specifiche di progettazione -
DOC3	AQDML3IRCP01SR02 Rev.04.00 - IDROCAP - Specifiche Funzionali	Piattaforma IDROCAP Servizio LA.DW.1 - Specifiche funzionali -

3 ACRONIMI E DEFINIZIONI

Acronimi/Definizione	Descrizione
AdB	Autorità di Bacino del Distretto Idrografico della Sicilia
Amministrazione	Autorità di Bacino del Distretto Idrografico della Sicilia
DRAR	Dipartimento regionale dell'Acqua e dei Rifiuti
DRT	Dipartimento Regionale Tecnico
IDROCAP	Sistema di gestione delle captazioni idrauliche
N/A	Non applicabile
DTS	Data Transformation Services
CPU	Central Processing Unit
SQL	Structured Query Language

 	Titolo	RTI: Leonardo - IBM – Sistemi Informativi – Onit – Prisma – Deas - Reply PS
	Proprietario: Autorità di Bacino del Distretto Idrografico della Sicilia – RTI Autore: RTI Data rilascio: 14/06/2024 Codice: S22028.02.1004PROC Revisione: 04.00	

4 INTRODUZIONE

Il presente documento, ha lo scopo di fornire una panoramica completa e dettagliata delle procedure batch e dei Data Transformation Services (DTS) implementati. Esso copre vari aspetti, tra cui la sincronizzazione tra database master e slaves, l'esecuzione programmatica del backup dei database e del file system distribuito, il deploy dell'infrastruttura mediante script Ansible e il monitoraggio delle performance e delle operazioni.

L'obiettivo principale di questo documento è fornire una guida chiara e comprensibile per gli stakeholder coinvolti, tra cui sviluppatori, amministratori di sistema, responsabili IT e utenti finali. Esso serve come riferimento per comprendere i processi, le tecnologie e le metodologie utilizzate nelle procedure batch e nei DTS, nonché per garantire che queste procedure siano gestite e mantenute in modo efficiente, sicuro e conforme alle best practice del settore.

Le informazioni contenute nel documento sono organizzate in sezioni tematiche, ognuna delle quali affronta un aspetto specifico delle procedure batch e dei DTS. Questa struttura modulare facilita la consultazione e l'aggiornamento del documento, permettendo agli stakeholder di trovare rapidamente le informazioni di cui hanno bisogno e di mantenerle aggiornate nel tempo.

Disclaimer per ciclo di sviluppo completo a lotti

L'approccio è caratterizzato da fasi di realizzazione e prevede l'utilizzo di cicli di: pianificazione e raccolta dei requisiti, progettazione, sviluppo, testing e controllo qualità, rilascio a lotti, monitoraggio e feedback, rilascio completo, sviluppi iterativi e incrementali, in cui il team di sviluppo si concentra sulla realizzazione di un insieme di funzionalità specifiche. Ogni rilascio incrementale contiene un insieme di funzionalità autoconsistenti da sottoporre a validazione/collauda; il team lavora insieme all'utente per definire, progettare, sviluppare, testare e consegnare una porzione del prodotto finale. In questo modo, il prodotto nel suo complesso viene sviluppato in modo graduale ed incrementale; il team ha la possibilità di ricevere feedback costante dal cliente e di adattarsi rapidamente alle esigenze in evoluzione.

Per ogni lotto, il team di sviluppo si concentra sulla identificazione dei requisiti e delle specifiche di progettazione che sono stati definiti per quel particolare lotto. Se necessario, il team può apportare modifiche ad elementi dei lotti precedenti per riflettere le nuove esigenze raccolte dal cliente od integrare nuove informazioni acquisite durante la fase di monitoraggio e feedback. Tuttavia, queste modifiche vengono gestite in modo strutturato e controllato, in modo da minimizzare l'impatto sullo sviluppo complessivo del prodotto.

In sintesi, dopo aver eseguito diversi cicli di rilascio a lotti ed aver affrontato eventuali problemi significativi, l'applicazione sarà pronta per un rilascio completo a tutti gli utenti, ciò consente di sviluppare il prodotto in modo graduale, adattarsi ai feedback ricevuti dal cliente e gestire le modifiche alle specifiche di progettazione in modo efficiente.

Per ulteriori informazioni sulla metodologia di progettazione fare riferimento al documento "Appendice 2 al Capitolato tecnico speciale Lotti 1-2-3 - Cicli e Prodotti".

Per avere un ulteriore approfondimento sulle tematiche qui trattate prendere come principale riferimento il documento "Manuale di gestione applicazione".

Programmazione FSC 2014-2020 Piano Operativo Ambiente		  	MINISTERO DELLA TRANSIZIONE ECOLOGICA
QUALSIASI COPIA CARTACEA DI QUESTO DOCUMENTO È SOLO PER RIFERIMENTO Riservato RTI - Amministrazione			Pag. 5 di 9 Stato: Rev.04.00

 	Titolo	RTI: Leonardo - IBM – Sistemi Informativi – Onit – Prisma – Deas - Reply PS
	Proprietario: Autorità di Bacino del Distretto Idrografico della Sicilia – RTI Autore: RTI Data rilascio: 14/06/2024 Codice: S22028.02.1004PROC Revisione: 04.00	

5 SINCRONIZZAZIONE TRA DATABASE MASTER E DATABASE SLAVES

La sincronizzazione tra il database master e i database slaves è un componente fondamentale dell'infrastruttura di IDROCAP, che garantisce la consistenza dei dati, la ridondanza e la disponibilità elevata del servizio. La sincronizzazione dei dati tra il database master e i database slaves avviene attraverso un processo di replicazione, nel quale le modifiche apportate al database master vengono copiate e applicate automaticamente ai database slaves.

L'obiettivo principale della sincronizzazione è assicurare che il database slave contenga sempre una copia aggiornata dei dati presenti nel database master. Questo permette di distribuire il carico di lavoro tra i diversi nodi, migliorando le performance del sistema e fornendo una maggiore tolleranza ai guasti. In caso di malfunzionamento del database master, uno dei database slaves può essere promosso a master per garantire la continuità delle operazioni e ridurre i tempi di inattività.

Il ruolo del database master è quello di gestire e coordinare le operazioni di lettura e scrittura dei dati, mentre il database slave è responsabile di ricevere e applicare le modifiche inviate dal master. Inoltre può essere utilizzato anche per eseguire operazioni di lettura, alleggerendo il carico sul database master e migliorando la scalabilità e le performance complessive del sistema.

5.1 METODO DI SINCRONIZZAZIONE

Nell'infrastruttura di riferimento, si utilizza una versione specifica di MySQL (≥ 8.0), compatibile con il framework utilizzato dalla Web Application, con gli altri microservizi e con gli altri componenti del sistema. È importante assicurarsi che tutti i nodi del database utilizzino la stessa versione di MySQL e che eventuali aggiornamenti o modifiche siano attentamente pianificati e testati per evitare problemi di compatibilità o malfunzionamenti.

Replicazione

Il metodo di sincronizzazione utilizzato nell'infrastruttura per mantenere i dati coerenti tra il database master e il database slave è la replicazione. La replicazione in MySQL è un processo asincrono basato su un'architettura master-slave, in cui il database master è responsabile delle operazioni di scrittura e il database slave si occupa di ricevere e applicare le modifiche provenienti dal master.

Nel contesto dell'infrastruttura, con un solo master e un solo slave, si utilizza la replicazione master-slave unidirezionale. Di seguito vengono forniti alcuni dettagli tecnici su come è stato architettato l'intero processo:

- **Registrazione delle modifiche:** ogni volta che viene eseguita un'operazione di scrittura sul database master, MySQL registra le modifiche in un file chiamato "binary log" (binlog). Il binlog contiene tutte le informazioni necessarie per replicare le modifiche sui database slaves.

 	Titolo	
	Proprietario: Autorità di Bacino del Distretto Idrografico della Sicilia – RTI Autore: RTI Data rilascio: 14/06/2024 Codice: S22028.02.1004PROC Revisione: 04.00	RTI: Leonardo - IBM – Sistemi Informativi – Onit – Prisma – Deas - Reply PS

- Connessione tra master e slave: il database slave stabilisce una connessione con il database master utilizzando un account con i privilegi appropriati. Il database slave richiede quindi al master le informazioni sulle modifiche da applicare.
- Trasferimento delle modifiche: il database master invia al database slave gli eventi registrati nel binlog che corrispondono alle modifiche da replicare. Il database slave riceve gli eventi e li memorizza in un file chiamato "relay log".
- Applicazione delle modifiche: un processo chiamato "SQL thread" sul database slave legge gli eventi dal relay log e li applica al database slave in modo sequenziale. Questo garantisce che le modifiche vengano applicate nello stesso ordine in cui sono state eseguite sul database master, mantenendo la coerenza dei dati.

La replicazione asincrona utilizzata in MySQL ha il vantaggio di non bloccare il database master durante il processo di sincronizzazione, permettendo di mantenere alte prestazioni anche in presenza di un elevato numero di operazioni. Tuttavia, è importante monitorare attentamente lo stato della replicazione per assicurarsi che eventuali ritardi o problemi vengano prontamente individuati e risolti, evitando discrepanze tra il database master e il database slave.

Per garantire la corretta configurazione e il funzionamento della replicazione, è necessario prestare attenzione a diversi aspetti, tra cui la configurazione dei file my.cnf sia sul master che sullo slave, l'impostazione del server ID univoco per ciascun nodo e la creazione di un account dedicato con i giusti privilegi per la comunicazione tra master e slave.

Monitoraggio della sincronizzazione

Per il monitoraggio della sincronizzazione così come in tutti gli altri servizi, si utilizza Prometheus:

- Raccolta delle metriche: Prometheus raccoglie una serie di metriche dai nodi del database, tra cui lo stato della replicazione, il ritardo di replicazione, il numero di query per secondo e altre informazioni relative alle performance e alla salute del sistema. Queste metriche vengono raccolte utilizzando il node_exporter per MySQL per Prometheus, che espone le metriche in un formato compatibile con Prometheus.
- Configurazione degli allarmi: sulla base delle metriche raccolte, sono configurati allarmi per essere avvisati tempestivamente per il blocco del processo di replicazione o per l'elevato utilizzo delle risorse, come CPU e memoria.
- Notifiche: quando un allarme viene attivato, Prometheus invia una notifica tramite il componente Alertmanager.
- Visualizzazione e analisi delle metriche: le metriche raccolte da Prometheus possono essere visualizzate e analizzate utilizzando la consueta dashboard di grafana.

 	Titolo	
	Proprietario: Autorità di Bacino del Distretto Idrografico della Sicilia – RTI Autore: RTI Data rilascio: 14/06/2024 Codice: S22028.02.1004PROC Revisione: 04.00	RTI: Leonardo - IBM – Sistemi Informativi – Onit – Prisma – Deas - Reply PS

6 ESECUZIONE PROGRAMMATICA DEL BACKUP DEL DB

6.1 PROCEDURA DI BACKUP DEL DB PRINCIPALE

Per garantire la sicurezza e l'integrità dei dati, viene eseguito un backup programmato del database ogni notte. Il processo di backup viene gestito utilizzando una combinazione di strumenti e tecnologie specifici che assicurano l'efficienza e l'affidabilità dell'operazione.

Il tool principale utilizzato per eseguire il dump del database è mysqldump, uno strumento fornito da MySQL per esportare i dati in un formato compatibile con SQL.

La temporizzazione del processo di backup viene gestita attraverso un cronjob, che è configurato per avviare il processo di backup ogni notte, assicurando che i dati vengano salvati regolarmente senza intervento manuale.

Per gestire la conservazione e la rotazione dei backup, viene utilizzato un sistema di log rotate. Questo sistema permette di mantenere solo il backup dell'ultimo mese nel server locale, eliminando automaticamente i backup più vecchi e risparmiando spazio di archiviazione.

Inoltre, l'infrastruttura Cloud fornisce ulteriori misure di protezione dei dati attraverso l'utilizzo di snapshot del disco. Questi snapshot creano ogni giorno delle copie istantanee dell'intero disco e permettono di ripristinare rapidamente lo stato del sistema in caso di malfunzionamenti o perdite di dati.

7 ESECUZIONE PROGRAMMATICA DEL FILE SYSTEM DISTRIBUITO

La protezione delle pratiche digitalizzate è di fondamentale importanza per garantire la continuità operativa e la sicurezza di IDROCAP. Le pratiche digitalizzate costituiscono lo storico e quindi deve essere trattato come risorsa critica, che deve essere preservata e resa disponibile in ogni momento. La perdita di queste pratiche potrebbe causare gravi conseguenze, quali interruzioni nelle operazioni, perdite di dati e potenziali danni amministrativi.

Per far fronte a questa esigenza e garantire un elevato livello di protezione delle pratiche, è stata adottata un'architettura a 4 nodi per il file system distribuito MinIO. Questa soluzione consente di offrire una ridondanza dei dati e una disponibilità costante, assicurando che le pratiche digitalizzate siano sempre accessibili anche in caso di guasti o malfunzionamenti di singoli nodi.

L'implementazione di un cluster con 4 nodi permette di ottenere una fault tolerance elevata e di ridurre al minimo, se non annullare, i tempi di fermo del servizio nella maggior parte dei casi.

Backup infrastrutturale

Il backup infrastrutturale viene effettuato quotidianamente attraverso la creazione di snapshot dei volumi delle macchine virtuali che ospitano il file system distribuito MinIO.

Programmazione FSC 2014-2020 Piano Operativo Ambiente	   MINISTERO DELLA TRANSIZIONE ECOLOGICA
QUALSIASI COPIA CARTACEA DI QUESTO DOCUMENTO È SOLO PER RIFERIMENTO Riservato RTI - Amministrazione	
Pag. 8 di 9 Stato: Rev.04.00	

 	Titolo	RTI: Leonardo - IBM – Sistemi Informativi – Onit – Prisma – Deas - Reply PS
	Proprietario: Autorità di Bacino del Distretto Idrografico della Sicilia – RTI Autore: RTI Data rilascio: 14/06/2024 Codice: S22028.02.1004PROC Revisione: 04.00	

Il ripristino degli snapshot può essere necessario in situazioni eccezionali, come ad esempio il fallimento di più di due nodi del cluster MinIO.

Backup settimanale tramite dump dei volumi di MinIO

Per garantire un ulteriore livello di protezione dei dati, oltre ai backup infrastrutturali quotidiani, vengono effettuati backup settimanali tramite dump dei volumi di MinIO. Il processo di dump consiste nella creazione di una copia completa dei dati contenuti nei volumi, che viene archiviata all'interno della stessa macchina, in un volume separato e cifrato. La cifratura del volume di backup garantisce la sicurezza e la riservatezza delle informazioni, proteggendo i dati da eventuali accessi non autorizzati o attacchi informatici.

Il sistema di backup settimanale prevede l'utilizzo del log rotate per gestire le copie dei dump nel tempo. Grazie a questa tecnica, i backup vengono ruotati automaticamente, mantenendo sempre a disposizione un mese di copie di backup.

La programmazione dei backup settimanali tramite dump dei volumi di MinIO viene effettuata mediante l'utilizzo di cronjob. Il cronjob è configurato per avviare il processo di dump dei volumi di MinIO una volta alla settimana, garantendo così la regolarità e l'affidabilità del sistema di backup.

In sintesi, i backup settimanali tramite dump dei volumi di MinIO rappresentano una misura aggiuntiva di protezione dei dati, che, insieme ai backup infrastrutturali quotidiani, garantiscono la sicurezza e la disponibilità delle pratiche digitalizzate e delle informazioni all'interno dell'organizzazione.

Monitoraggio e allarmi del processo di backup settimanale

Tutte le procedure di backup del file system distribuito sono anch'esse soggette a monitoring costante per mezzo di Prometheus, in quanto tali macchine hanno l'agent node_exporter installato.