



Newsletter ReCaS-Bari - N.6 - Ottobre 2016

www.recas-bari.it

Questo e i numeri precedenti della Newsletter ReCaS sono disponibili a questo [link](#).

In questo numero:

- Prima release del software sviluppato dal progetto INDIGO-DataCloud
- Il contributo di ReCaS-Bari al processamento dei dati satellitari Sentinel
- Il data center ReCaS-Bari alla Digital Infrastructures for Research 2016
- Crediti

Prima release del software sviluppato dal progetto INDIGO-DataCloud

Il data center ReCaS-Bari è coinvolto attivamente nel progetto europeo (Horizon 2020) INDIGO-DataCloud, che ha come obiettivo quello di sviluppare una piattaforma open source, rivolta alle comunità scientifiche, per l'immagazzinamento e l'analisi dei dati, dispiegabile su hardware eterogeneo.



Il progetto ha rilasciato la prima versione (nome in codice MidnightBlue) del software sviluppato, con licenza Apache 2.0. Il software, scaricabile al link <http://repo.indigo-datacloud.eu> può essere installato su infrastrutture cloud pubbliche e private.

Il rilascio è il punto finale di una fase di indagine dei requisiti che ha coinvolto numerose collaborazioni scientifiche europee, afferenti ad aree tematiche molto diverse fra loro, fra cui biologia strutturale, scienze della terra, fisica, bioinformatica, beni culturali, astrofisica, scienze della vita e climatologia.

Sono state sviluppate numerose componenti software open source, per superare gap tecnologici legati all'ottimizzazione dell'utilizzo di risorse di calcolo e dati distribuiti. Queste componenti sono state quindi aggregate in una suite modulare che, basandosi su interfacce standard basate su TOSCA, contribuisce alla definizione e all'implementazione di una open cloud europea per la scienza, con elevati livelli di automazione. Caratteristiche fondamentali sono:

- incremento dell'efficienza delle risorse e dei servizi dei data center;
- accesso avanzato a dati distribuiti, con federazione tramite protocolli multipli

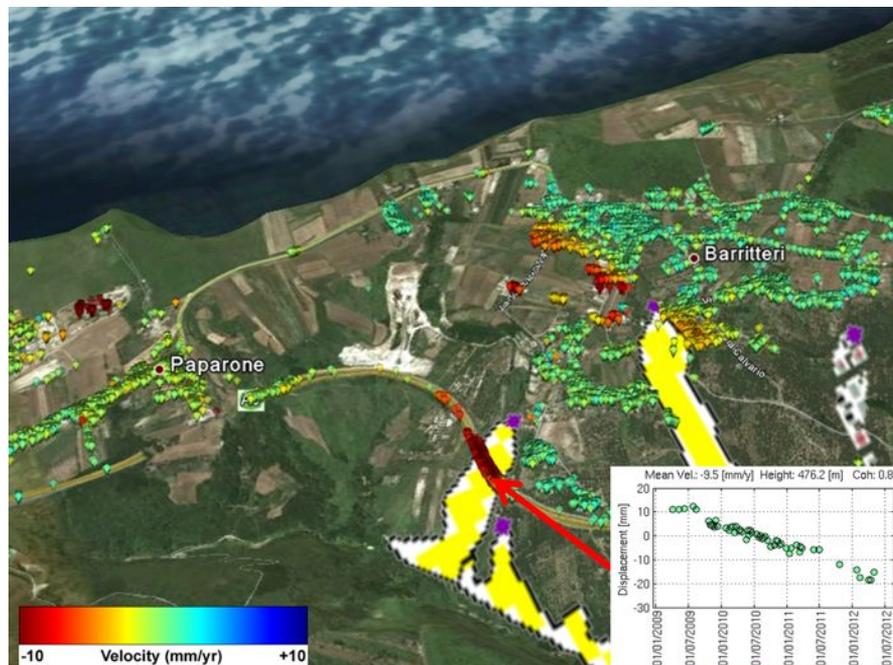
(gli utenti finali non hanno bisogno di conoscere l'esatta localizzazione delle risorse, poiché il layer PaaS si occupa di gestire la complessità dello scheduling e del brokering);

- possibilità di allocare risorse laddove siano immagazzinati i dati, per permettere una più rapida analisi degli stessi;
- automazione del deployment di insiemi complessi di dati e risorse;
- supporto integrato per l'*high-performance big data analytics*;
- autenticazione/autorizzazione tramite SAML, OpenID Connect e X.509;
- integrazione di applicazioni scientifiche in front-end programmabili e in mobile app, per rendere l'esperienza degli utenti *user friendly*.

Il data center ReCaS-Bari si appresta ad includere tra i servizi offerti la maggior parte dei servizi prodotti da INDIGO (vedi il [catalogo](#)).

Il contributo di ReCaS-Bari al processamento dei dati satellitari Sentinel

Il satellite [Sentinel1](#) fornisce dati radar del nostro pianeta a supporto di molte applicazioni ambientali. Il satellite è uno dei risultati del programma europeo [Copernicus](#), finalizzato all'osservazione della terra.



Mappa delle velocità di spostamento superficiale nei pressi del Comune di Barritteri, in Calabria, sovrapposta al catalogo dei fenomeni franosi del progetto IFFI. Target al suolo di colorazione verde sono stabili, mentre quelli in movimento sono evidenziati con colorazioni differenti a seconda della direzione di movimento del target stesso. In basso a destra è mostrato il trend degli spostamenti di un target in area instabile.

[GAP](#) (spin-off dell'università di Bari) e [Planetek Italia](#), nell'ambito di un progetto di ricerca chiamato FAST4MAP, cofinanziato dalla Agenzia Spaziale Italiana, stanno sperimentando l'utilizzo di algoritmi sviluppati per monitorare spostamenti millimetrici della superficie terrestre utilizzando i dati provenienti da Sentinel 1. Un esempio di applicazione di tali metodologie di indagine è mostrato nella figura, in relazione all'area peri-urbana del Comune di Barritteri, in Calabria.

Questi algoritmi, che si basano su tecniche avanzate di interferometria differenziale multi-temporale, richiedono il processamento contestuale di decine di immagini satellitari. I volumi di dati sono pari a centinaia di gigabyte per ogni area oggetto di analisi (qualche centinaio di chilometri quadri) e la elaborazione necessita di ingenti risorse di calcolo. Utilizzando alcune centinaia di core il processo di elaborazione può impegnare alcune decine di ore.

La collaborazione con ReCaS-Bari su questo tema aveva già avuto una prima fase quando ReCaS, GAP e Planetek avevano collaborato per realizzare una sfida tra ricercatori

(http://www.planetek.it/news_eventi/tutti_gli_eventi/2015/04/gpu4eo_challenge_2015), con un premio messo in palio da Planetek, per l'ottimizzazione di algoritmi per il processamento SAR su GPU.

A valle di tale prima collaborazione di successo, si sta prototipando l'integrazione e l'automazione della intera catena di processamento sulla architettura HTC.

Il prossimo passaggio sarà la migrazione di alcuni step di processamento utilizzando le GPU, sull'ambiente HPC di ReCaS.

L'opportunità maggiore offerta da ReCaS a queste attività è la disponibilità, in un unico ambiente facilmente integrabile delle risorse HTC e HPC, unitamente alle classiche macchine virtuali dell'ambiente cloud, offre delle possibilità uniche nel realizzare un'architettura mista, che privilegia l'uso del giusto strumento per ogni singolo task di processamento.

Il data center ReCaS-Bari alla [Digital Infrastructures for Research 2016](#)

ReCaS-Bari ha esposto un suo poster alla *Digital Infrastructures for Research 2016*, che si è svolta a Cracovia, in Polonia, dal 28 al 30 settembre 2016. La conferenza, organizzata da [EGI](#), [EUDAT](#), [GÉANT](#), [OpenAIRE](#) and [RDA Europe](#) prosegue il percorso tracciato nella precedente edizione tenuta un anno fa a Bari. Nel [poster](#) è stato presentato il sistema di monitoring utilizzato per monitorare l'attività del data center ReCaS-Bari, del quale viene



riportata una anteprima.

In particolare il poster illustra l'attività appena avviata di monitoraggio su richiesta delle applicazioni utente, al fine di fornire la possibilità di acquisire metriche, realizzare grafici, impostare allarmi e notifiche personalizzate per ciascun utente di ReCaS-Bari, e per ciascun servizio dell'utente. Di fatto il monitoring di ReCaS-Bari al momento è utilizzato dagli amministratori



dell'infrastruttura per aggregare in un'applicazione web lo stato dei sensori dei servizi, in modo che ciascun aspetto della farm di calcolo sia costantemente sotto osservazione, e, in caso di malfunzionamenti, effettui una segnalazione agli amministratori stessi. Ogni servizio è associato a un gruppo di amministratori, che sono gli unici ad essere notificati in caso di malfunzionamenti.

Il nuovo servizio offerto dal gruppo di monitoring prevede invece l'estensione agli utenti delle capacità di monitoraggio: ogni utente può richiedere quali fra i propri servizi, o fra quelli del proprio gruppo di ricerca, poter monitorare, oppure quali grafici generare, e le modalità di notifica (messaggistica, email...), al fine di essere avvisato in caso di malfunzionamenti, oppure per misurare le performance della propria applicazione. In caso di servizi ben noti, quali web services, la creazione del pacchetto di monitoring è immediata; nel caso di applicazioni con parametri meno tipici, il gruppo di monitoring interagisce con l'utente, analizzando il servizio e definendo le linee guida per la creazione del pacchetto di opzioni per il monitoring.

Questo approccio, che va nella direzione del *Monitoring as a Service* tipico dei servizi cloud evoluti, permette l'autodiagnosi e il profiling delle proprie applicazioni e della piattaforma sottostante, fornendo all'utente un controllo sempre più potente sulla parte di infrastruttura di calcolo a lui dedicata.

Crediti

Hanno collaborato a questo numero della newsletter:

Marica Antonacci, Giacinto Donvito, Rosanna Gallitelli, Giorgio Maggi, Stefano Nicotri, Vincenzo Spinoso.