

L'INTEGRAZIONE DELLE ARCHITETTURE CLOUD NEI SISTEMI GIS: LO STANDARD OGC WEB PROCESSING SERVICE (WPS)

G. Calabretta¹

¹ Regione Calabria – Nucleo di Valutazione e Verifica degli Investimenti Pubblici

1 – Elementi introduttivi

Nel corso degli ultimi anni si è assistito all'introduzione di servizi rivolti all'attivazione di nuovi metodi di diffusione ed utilizzo dei dati geografici attraverso Internet. La genesi di questi servizi ha comportato il confrontarsi con nuove sfide per la creazione di applicazioni da utilizzare come servizio per milioni di utenti, piuttosto che applicazioni che girano su singoli computers. Molte delle soluzioni emerse sono risultate estremamente efficaci e la notevole offerta di Software, Piattaforme ed Infrastrutture come servizi (oggi comunemente identificati dalle sigle *SaaS*, *PaaS* e *IaaS*, fig. 1) ha attivato la scoperta di nuovi *pattern* architetturali per l'applicazione ottimale dei paradigmi noti come Cloud Computing (CC) anche ai sistemi informativi geografici. Le caratteristiche principali di tali paradigmi sono basati sull'uso di Internet e di una serie di framework software per gestire dati ed applicazioni. Questo consente all'utente l'utilizzo di applicazioni e l'accesso a dati, da un qualsiasi computer dotato di accesso ad internet.

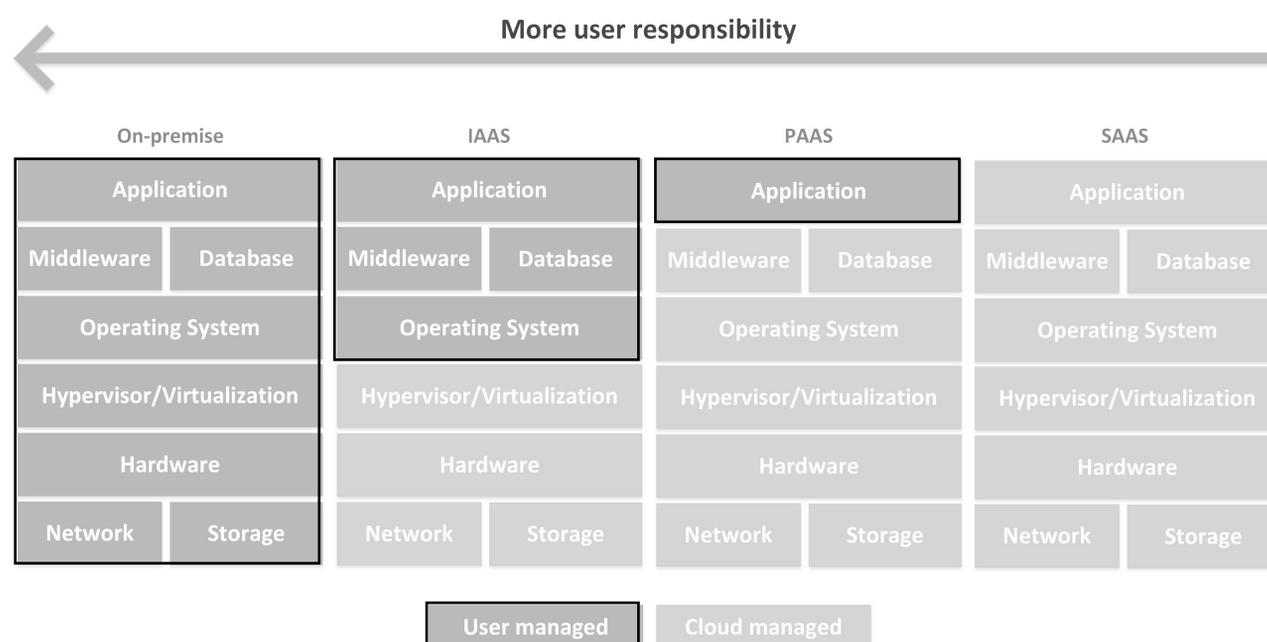


Fig. 1 – Riepilogo delle diverse tipologie di offerte Cloud Computing a partire dal Centro Dati tradizionali (on-premise)

La condivisione e l'accesso ad informazioni geografiche (aspetto chiave di applicazioni come *Google Maps/Google Earth*, *Bing Maps*, *Yahoo maps* tra i più noti) ha poi spinto gli utenti, anche quelli usualmente meno avvezzi, a porsi domande in maniera 'spaziale'.

Questa tipologia di servizi sono organizzate in Infrastrutture di Dati Spaziale (SDIs) ed i servizi offerti sono basati sulla disponibilità attraverso la rete Internet. In particolare nel corso degli ultimi decennio, sono state oltre trenta le specifiche di webservice standard per l'accesso e la condivisione dei dati geografici introdotte dall'Open Geospatial Consortium (OGC). Molte di queste sono già diventati degli standard ISO, ad esempio la specifica Web Map Service (WMS) che dal 2005 è diventata la norma ISO 19128, o il Geography Markup Language (GML), diventata la norma ISO 19136 dal 2007. A puro titolo esemplificativo, un insieme tipico di servizi offerti oggi da una infrastruttura SDI, può comprendere un webmap service (WMS), un webmapping service (WMTS), un webfeature service (WFS), a webcoverage service (WCS), un webcatalogue service (CS-W), ed altri.

Gli standard OGC risultano oramai largamente diffusi. Un decisivo impulso a tale sviluppo è dato dalla crescente disponibilità di reti a banda larga e dall'offerta sempre più a buon mercato dei fornitori di servizi di Cloud Computing (Amazon EC, RackSpace ed altri), che ha consentito l'aumento delle performance e della scalabilità dei prodotti basati sulle specifiche OGC.

Ad ogni modo data la notevole accessibilità di dati attraverso webservice, ha reso il processamento degli stessi in modalità remota, come la naturale evoluzione dei tradizionali paradigmi di GIS computing. A questo fine, fin dal 2005, OGC ha iniziato a lavorare sul Web Processing Service (WPS) Interface Standard -una interfaccia che consente ai processi computazionali di essere esposti come Webservice - aggiungendo in tal modo un focus sull'analisi ed il geoprocessing online dei dati, insieme alle modalità di accesso ed interrogazione più 'tradizionali'. L'obiettivo del servizio, è infatti fornire una modalità di fruizione personalizzabile, attraverso i processi di analisi e trasformazione esposti, dei dati presenti nelle varie SDIs.

2 - Il Web Processing Service

Le specifiche OGC del servizio WPS (OGC, 2007) descrivono una serie standardizzata di operazioni per rendere fruibile ed eseguire ogni tipo di funzionalità di geoprocessing sul web. In accordo alle specifiche OGC, un processo è definito come un algoritmo di calcolo o un modello che opera su dei dati georiferiti. Nel dettaglio, le specifiche WPS, descrivono tre tipi di operazioni che vengono gestite in modalità *stateless*: *GetCapabilities*, *DescribeProcess* e *Execute*.

L'operazione *GetCapabilities* è comune ad ogni tipo di OGC Web Service e restituisce i metadata del servizio stesso. Nel caso del WPS restituisce una breve descrizione del processo offerto dalla specifica istanza WPS. Per ottenere maggiori informazioni del servizio offerto, il WPS è in grado di restituire i metadata del processo attraverso l'operazione *DescribeProcess*. Questa fornisce la descrizione di tutti i parametri che sono richiesti per attivare il processo stesso. Noti tali parametri l'utente può infine lanciare l'operazione *Execute*. Analogamente agli altri OGC Web Service, il WPS comunica attraverso HTTP-GET and HTTP-POST, utilizzando una messagistica basata su OGC-specific XMLencoding. WPS supporta l'esposizione simultanea di processi anche attraverso SOAP, lasciando tuttavia al client la possibilità di scegliere il meccanismo di interfaccia più appropriato.

La risposta ad una richiesta *Execute* è un documento XML che include i metadata sulla richiesta, oltre agli outputs del processo, codificati ed inglobati nella risposta XML. Questa tipologia di risposta è, in genere, consigliata quando l'output non supera le dimensioni di qualche MBytes, e l'utente necessita dei metadata nella risposta.

Il WPS prevede due modalità operative (fig. 2):

- *Synchronous (o sincrona)*, consigliata per le operazioni che si prevedono rapide
- *Asynchronous (asincrona)*, per computazioni più lunghe temporalmente; in tal caso il servizio fornisce un token che può essere utilizzato per controllare lo stato.

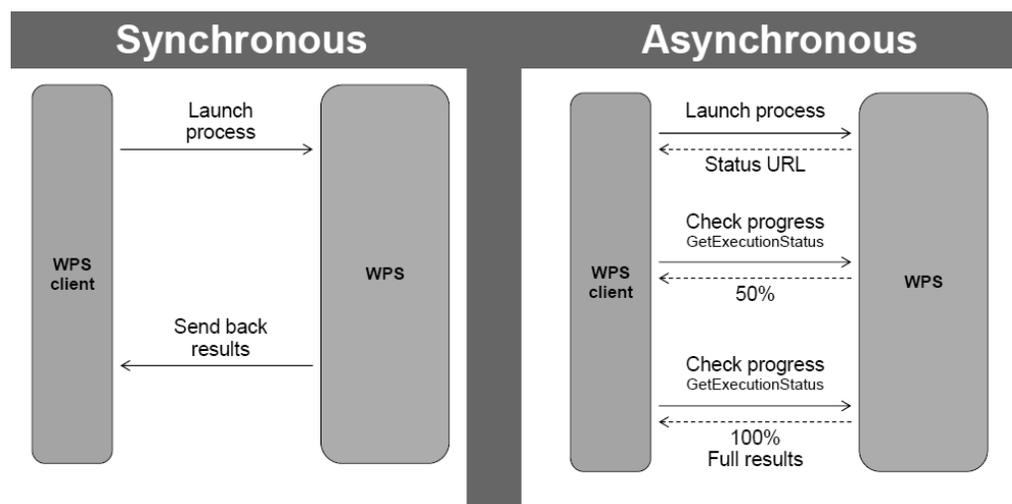


Fig. 2 – Modalità operative del WPS

In aggiunta a ciò, la specifica OGC descrive il processamento di dati URL-referenziati e l'immagazzinamento dei risultati. Quest'ultimo aspetto appare particolarmente interessante, in quanto consente di accedere ai risultati lato server, utilizzando una semplice URL, senza dover rilanciare il processo stesso. E' possibile comunque richiedere i risultati in formato raw (ad esempio una richiesta di buffer, può ritornare l'immagine dell'elemento con il buffer aggiunto, codificato in un file png).

Recentemente dei membri dell'OGC members hanno intrapreso una ipotesi di definire un insieme finito di algoritmi di analisi che possa caratterizzare il profilo di un WPS (WPS profile). Le principali categorie di analisi sono risultate: Geometria di base; Aggregazioni: Prossimità; Misure statistiche; Misure di distribuzione; misure di similarità.

Il WPS definisce tre tipologie di dati:

- *Complex Data* che include le immagini, i file XML, CSV, e strutture dati personalizzati o proprietarie;
- *Literal Data* che include le stringhe di testo o i singoli valori numerici;
- *Bounding Box Data* che includono le coordinate geografiche di un'area rettangolare.

Nel 2014 è stata proposta la revisione dello standard WPS 1.0, con la nuova specifica WPS 2.0 (diventa effettiva a partire da Marzo 2015). Questa incorpora le modifiche richieste alla precedente versione ed aggiunge alcune migliorie. A differenza della versione 1.0 fornisce un nuovo modello concettuale che può essere utilizzato per l'utilizzo del WPS in tipologie di architetture differenti come REST o SOAP. Sono state introdotte alcune nuove funzionalità importanti come la possibilità di cancellare un processo, o il supporto per lo standard OGC SensorML.

3 - Principali framework per l'implementazione di servizi OGC WPS

Di seguito sono riportati e descritti sinteticamente i principali progetti *opensource* che supportano la creazione e pubblicazione di servizi WPS (fig. 3). I progetti Degree e Geoserver sono già progetti ufficiali del Open Source Geospatial Foundation (www.osgeo.org), mentre i progetti Zoo e PyWPS sono al momento in periodo di incubazione. Tutti i software descritti sono presenti nella distribuzione OSGeo Live (live.osgeo.org).

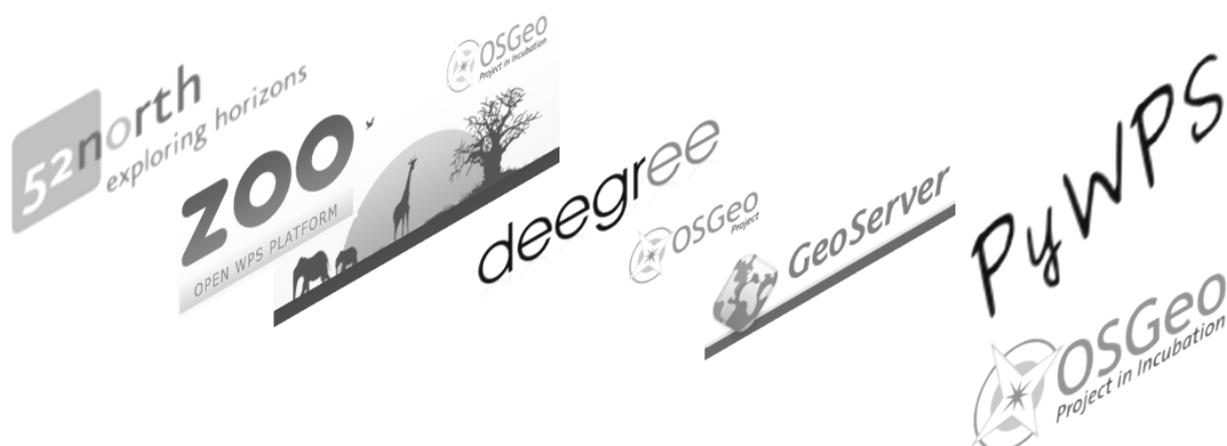


Fig. 3 – Principali framework opensource per la pubblicazione di servizi WPS

52°North WPS - <http://52north.org>

L'iniziativa 52° North per il free e opensource software nel campo geospaziale è stata fondata nel 2004 ed è una rete di sviluppo e ricerca internazionale con la missione di tradurre i risultati della ricerca in applicativi che possano essere utilizzati produttivamente nel dominio geospaziale. Per rafforzare tale processo 52° North è strutturata come una iniziativa ad adesioni libere che include partner dell'industria, della ricerca ed utenti professionali. I prodotti sono resi disponibili sotto un doppio modello di licenza, ossia la GNU General Public License 2 (GPL2) e parallelamente sotto una licenza commerciale.

Il software North52 WPS possiede funzionalità opensource OGC compliant basate su codice java che agiscono come plugin per il Java Tomcat servlet ed come plugin per clients opensource molto diffusi come uDig, Jump and OpenLayers.

E' pienamente compatibile con la specifica OGC WPS 1.0.0 (*document 05-007r7*), e può essere facilmente configurato attraverso una interfaccia accessibile da browser. Sono integrate al suo funzionalità di geoprocessing da applicativi come Sextante e Grass. Sono state aggiunte inoltre delle funzionalità backend per gli script in linguaggio R, ed un connettore per Esri Arcgis Server.

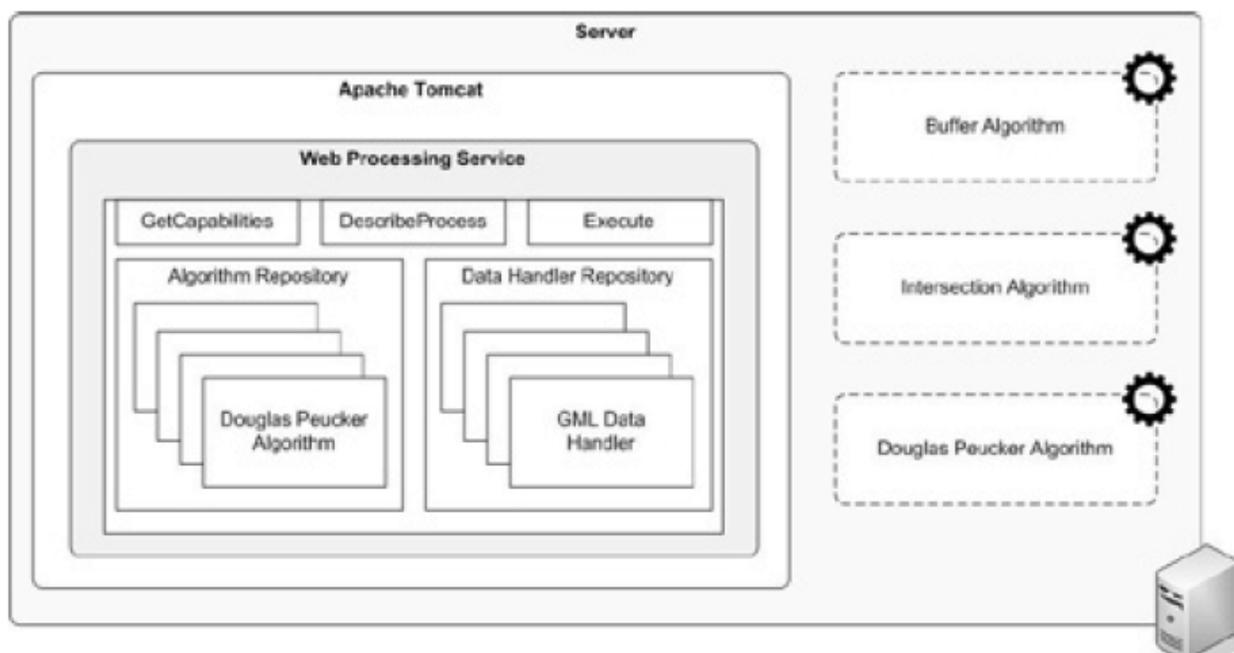


Fig. 4 – Esempio architettura North 52° WPS, per la generalizzazione di elementi geografici

Zoo-Project

ZOO-Project è una implementazione opensource di OGC WPS, rilasciata sotto una licenza di tipo MIT/X-11 e fa parte dei progetti in incubazione OSGeo. Il software mette a disposizione un framework che facilita lo sviluppo dei servizi e permette la creazione di servizi WPS in cascata (chain WPS Web services), principalmente grazie ad una crescente suite di esempi basate su librerie opensource e JavaScript API lato server, in grado di attivare e connetter i vari servizi del framework stesso. Lo ZOO-Project ha una architettura aperta sviluppata con codice modulare principalmente in C, Python and JavaScript.

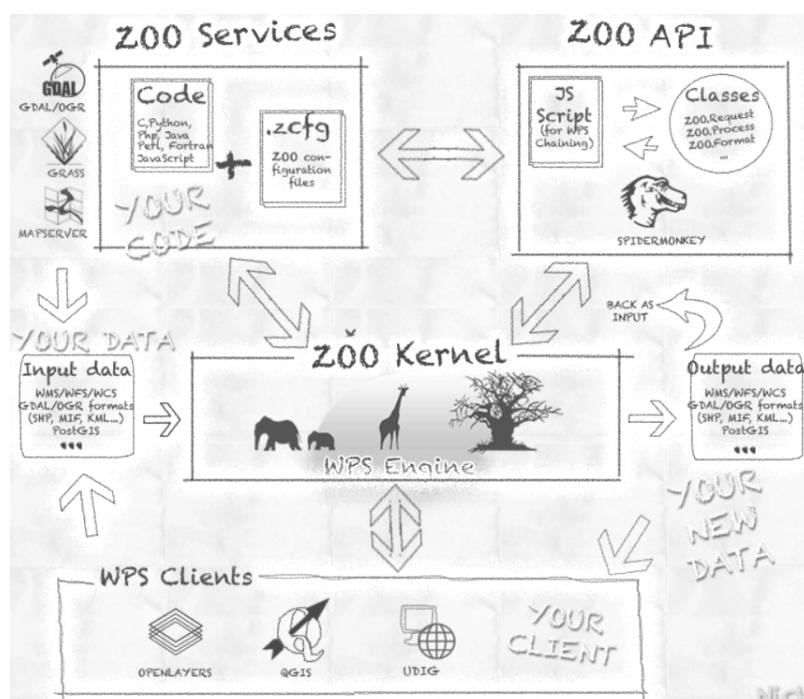


Fig. 5 – Un riepilogo grafico delle possibili architetture per il framework ZOO dal sito di progetto

Degree - <http://www.degree.org/>

Il progetto Degree è un framework per infrastrutture dati spaziali basato su Javache che implementi i principali standard geospaziali, fornendo tra gli altri i servizi di geoprocessing. Ha avuto inizio nel 2002, ed è distribuito sotto la GNU Lesser General Public License (LGPL) e dopo vari aggiornamenti è divenuto un progetto OSGeo. Il Degree WPS è un'implementazione che supporta lo standard lo standard WPS 1.0.0, e le richieste di tipo KVP (Key Value Pair, ovvero del tipo <http://qualcheurl.com?key1=value1&key2=value2>), XML and SOAP.

GeoServer - <http://geoserver.org/>

GeoServer è un software lato server basato su Java, costruito su Geotools un Java GIS toolkit, che permette la visualizzazione e la modifica e la condivisione di dati geospaziali, anche di grandi dimensioni. Geoserver è un progetto ufficiale della fondazione OSGEO e possiede una community tra le più attive del settore. Il software implementa la maggior parte degli standard OGC (WMS, WFS, WCS, WPS) e possiede diverse librerie di visualizzazione integrate, come ad esempio OpenLayers, consentendo una generazione di mappe molto veloce. La licenza è di tipo free software abbassando notevolmente le barriere di ingresso in rapporto ai prodotti GIS commerciali. Inoltre è disponibile open source. E' uno dei software OSGeo con un notevole disponibilità supporto di tipo commerciale, a riprova dell'affidabilità del progetto e degli investimenti in corso sullo stesso. All'interno di GeoServer è presente un client WPS, che permette di costruire facilmente la richieste da inviare al server. In particolare l'estensione WPS lato server di GeoServer supporta dati Raster e Vector, consente l'accesso diretto alle sorgente dati, la pubblicazione dei risultati come nuovi layer, e di grande utilità supporta le trasformazioni per il rendering in documenti di stile XML SLD. Il metodo asincrono per i processamenti temporalmente più

intensivi, è pienamente supportato. Altra notevole funzionalità è data dalla possibilità di implementazione e lancio di scripting di geoprocessing in diversi linguaggi (Python, Ruby), senza dover necessariamente riavviare il server. Allo stato attuale esiste il pieno supporto alle specifiche WPS 1.0.0, e parziale per le nuove introdotte dallo standard WPS OGC 2.0.0.

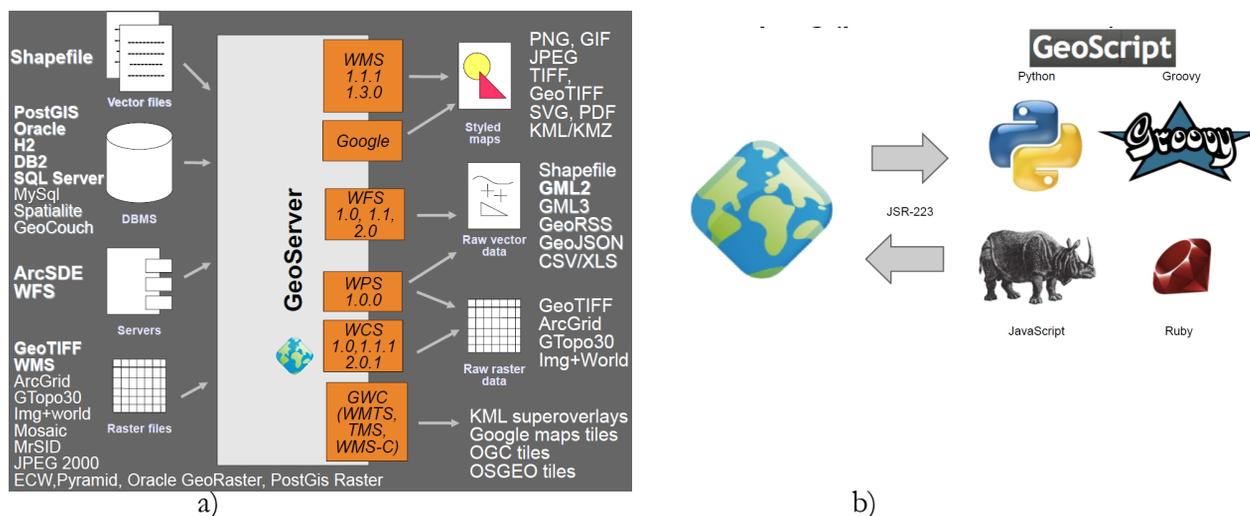


Fig. 6 – Riepilogo degli standard supportati da Geoserver (a), e dei linguaggi di script per il geoprocessing (b)

PyWps - <http://pynps.wald.intevation.org/>

PyWPS è una implementazione open source, molto leggera dello standard OGC WPS. Permette agli utenti di realizzare processi di analisi geospaziale attraverso il linguaggio Python e pubblicarli su Internet attraverso un server CGI come ad esempio Apache. PyWPS è degno di nota soprattutto per l'opportunità che offre di rendere accessibili librerie di geoprocessing di alta affidabilità come quelle GRASS, attraverso il linguaggio Python, semplificandone notevolmente l'utilizzo e l'apprendimento. E' anch'esso un progetto in incubazione OSGeo.

Infine, tra i prodotti commerciali sono da citare la suite *ESRI ArcGis Server* e la suite *Intergraph Erdas Apollo*, nella versione Professional, prevedono funzioni e metodi per pubblicazione di webservice conformi allo standard OGC WPS.

Conclusioni

Nel presente lavoro, è stata descritta una panoramica delle possibilità offerte dallo standard OGC WPS. La crescente disponibilità di reti a banda larga, connesse all'incremento di offerte di servizi dei diversi paradigmi di Cloud Computing, apre promettenti scenari per una larga diffusione dello standard WPS, all'interno di applicazioni destinati anche all'utenza non specialistica. Si è ritenuto utile dare risalto ai framework applicativi opensource, principalmente legati alla fondazione OSGeo, che consentono la pubblicazioni di tali servizi con procedure affidabili e con curve di apprendimento meno ostiche.

In definitiva, è opportuno precisare che lo standard WPS non è detto che debba rappresentare in ogni ambito la soluzione ottimale. Basti pensare, ad esempio, alle di analisi ed alle funzioni di geoprocessing che possono essere portate a termine più facilmente nei progetti GIS la cui architettura prevede l'accesso alle funzionalità di geodatabase come Postgres/PostGIS. Lo standard WPS presenta tuttavia delle peculiarità notevoli come ad esempio la possibilità di attivare modelli di processamento complessi resi disponibili via Internet (ad esempio modelli meteorologici), anche attraverso dati resi disponibili da terzi, o come la standardizzazione on-the-fly del rendering delle mappe. La diffusione dei servizi basati sugli standard WPS amplierà notevolmente gli scenari delle architetture GIS di nuova concezione, e si ritiene possa rappresentare un punto di riferimento importante per l'utilizzo dei geodati negli anni venire.

Riferimenti Bibliografici

OGC, Open Geospatial Consortium, 2007, OGC OpenGIS® Web Processing Service.

OGC, Open Geospatial Consortium, 2015, OGC WPS 2.0 Interface Standard.

52° North Initiative for Geospatial OpenSource SoftwareGmbH, Münster, Germany.
<http://52north.org/>

Aime A. Status of GeoServer WPS, FOSS4G-NA 2015, San Francisco, March 9th-12th 2015

Evangelidis K., *Geospatial services in the Cloud*, Computers & Geosciences N.63 (2014)116–122

Foerster T., 2009, *Integrating Web-based Sensor Information into Geospatial Mass-market Applications through OGC Web Processing Services*, International Journal on Advances in Intelligent Systems, vol 2 no 2&3.

Geomatics Workbooks, volume 12 FOSS4G Europe, Como 2015, Editors Brovelli, Minghini, Negretti.

Nebert, D. (Ed.), 2004, *Developing Spatial Data Infrastructures: The SDI Cookbook v.2.0, Global Spatial Data Infrastructure*, GSDI. Open Geospatial Consortium.

Wehrmann T., 2011, *Data processing using Web Processing Service orchestration within a Spatial Data Infrastructure*, 34th International Symposium for remote Sensing of the Environment, Sydney April 10-15 April 2011