

Il progetto DebianGis (2005-2009)

Francesco P. Lovergine (*)

(*) CNR -- Istituto di Studi sui Sistemi Intelligenti per l'Automazione
tel: 080 5929431, fax: 080 5929460, e-mail: f.lovergine@ba.issia.cnr.it

Sommario

Il progetto DebianGis nasce alla fine del 2004 con l'idea di supportare e mantenere la prima distribuzione GNU/Linux multi-architettura orientata ai GIS e fondata sulla nota distribuzione generalista Debian. I software GIS open source possono validamente essere integrati per implementare sistemi informativi geografici orientati ad applicazioni desktop e server, e annoverano alcuni importanti tool riconosciuti dalla comunità GIS come prodotti di primaria importanza (es. GDAL/OGR, Mapserver, Grass, ecc.). L'integrazione dello *stack* di programmi FOSS per il GIS richiede però una notevole esperienza di system integration da parte dell'*end-user* e una piattaforma omogenea di supporto e test. Il team di sviluppo ha lavorato negli anni sul supporto dei software GIS di frequente utilizzo e garantendo un alto standard qualitativo per i prodotti integrati nell'archivio Debian. Ad oggi DebianGis rappresenta un importante *test bed* per la maggior parte dei progetti open source GIS, e distribuzione di riferimento per la *OSGeo Foundation*.

Abstract

The DebianGis project was born at the end of 2004 with the goal of supporting and maintaining the first multi-architecture and GIS oriented GNU/Linux distribution, as founded on the general purpose Debian distribution. Free and open source GIS softwares can be effectively integrated to create geo information systems for desktop and server, by means of some important tools which are well-known in the GIS community as being primary interest programs (e.g. GDAL/OGR, Mapserver, Grass, etc.). Integrating a complete stack of FOSS GIS programs requires that the end-user has a deep experience of system integration and also a complete and homogeneous support and testing platform. The development team worked in the latest years on supporting frequently used GIS software and ensuring an high-level quality standard for all products added to the Debian main archive. DebianGis is today an important test bed for many open source GIS and a reference distribution for OSGeo Foundation.

Introduzione.

Il progetto di sviluppo DebianGis nasce nell'ambito peculiare della integrazione e supporto di software specializzati (nello specifico GIS e programmi affini) su piattaforma GNU/Linux. Il sistema operativo generalmente e brevemente denominato come "Linux" (con una tipica sineddoche che riduce il sistema al solo *kernel*) è diventato nel corso dei suoi diciotto anni di esistenza la principale alternativa open source - e non solo tale - alla *suite* di sistemi operativi proprietari sviluppati dalla Microsoft Corporation. Rispetto a questi ultimi, *Linux* ha "predato" gli standard e l'approccio integrato dei tradizionali sistemi UNIX, favorendo la creazione di innumerevoli "distribuzioni" del sistema stesso, caratterizzate da diverse tecnologie di pacchettizzazione del software e collezioni di programmi, tutte variamente orientate ai più differenti ambiti applicativi. Sebbene le distribuzioni più generaliste (*general-purpose*) siano anche le più diffuse e in genere utilizzabili ugualmente bene in ambito desktop e server, quando si considerano aree applicative molto specialistiche - quali i sistemi informativi geografici - si scopre che esse sono anche scarsamente adeguate come prodotti *off-the-shelf*. Allo stesso tempo, l'offerta di prodotti GIS

open source risulta ricca e strutturata, sia in ambito *server* che *desktop*, ma senza offrire una accessibilità immediata per l'utente non esperto. Esperienza normale sino a qualche anno fa, per chi volesse attingere alla considerevole offerta di prodotti FOSS GIS, erano le lunghe sessioni di lavoro per collezionare le giuste versioni di sorgenti e librerie di programmi, studiarne la documentazione di installazione, compilare e installare i prodotti in totale autonomia sulla propria macchina, con le relative difficoltà legate alla scarsa familiarità dell'utente con l'ambiente di sviluppo, la pletora di linguaggi di programmazione impiegata, i problemi legati alla piattaforma sottostante e le relative *toolchain*. Tali esperienze si traducevano spesso per i non “iniziati” in intere giornate impegnate a risolvere problemi funzionali, consultare mailing list e documenti disparati, e infine ottenere risultati magari lontani dalle aspettative o soluzioni incomplete. A tali difficoltà iniziali si aggiungeva la necessità di reiterare il processo ad ogni maggiore aggiornamento del sistema operativo o dei programmi, ogni volta scontrandosi con nuovi problemi. L'approccio “autarchico” non favoriva peraltro la replicabilità delle implementazioni su più macchine e il supporto per la sicurezza, cruciale per soluzioni orientate ai *server*.

Il progetto DebianGis nasce alla fine del 2004 come risposta all'esigenza di semplificare il supporto delle soluzioni FOSS GIS in ambiente GNU/Linux, consentendo all'end-user di costruire o usare applicazioni basate su un ampio *stack* di programmi e librerie perfettamente integrati, ben supportati, omogenei e installabili facilmente, unitamente agli altri programmi del sistema operativo. L'obiettivo dichiarato in quello che è oggi considerato il DebianGis Manifesto (Lovergine, 2004) è la creazione di una distribuzione orientata ai GIS, completamente integrata alla distribuzione general purpose Debian, in quel tempo disponibile nella versione 3.0, *codename* Woody. A distanza di cinque anni un team di più di venti sviluppatori e *maintainer* hanno partecipato e si sono coordinati per il rilascio di tre versioni successive della suite di programmi e librerie che costituiscono DebianGis, con un numero di pacchetti e uno standard qualitativo via via crescenti.

Il progetto Debian

Il progetto Debian è costituito da un gruppo di volontari, sviluppatori e utenti, che si dedicano alla produzione di un sistema operativo che si compone di software libero. Formalmente, il progetto è incorporato in Software in the Public Interest Inc., una organizzazione no-profit con sede a New York. I volontari sono però sparsi su tutto il pianeta e lavorano quasi completamente via rete, salvo per meeting occasionali e una conferenza annuale.

Il prodotto principale del progetto è Debian GNU/Linux, una distribuzione di sistema operativo che si basa sul *kernel* Linux ed è costituito ad oggi da più di trentamila pacchetti software integrati. Versioni non ufficiali del sistema si basano su *kernel* GNU/Hurd e BSD. Il sistema è supportato ufficialmente su undici diverse architetture a 32 e 64bit, con rilasci maggiori in media ogni 2-3 anni e *point release* minori più volte nel corso dell'anno. Un *security team* assicura aggiornamenti *just in time* per i pacchetti software per i quali si riscontrino problemi di sicurezza dopo il rilascio.

Il progetto è stato fondato ufficialmente nell'agosto del 1993 da Ian Murdock, quando lo stesso concetto di “distribuzione” non esisteva, e conta nel 2009 più di mille *developer* ufficiali e un numero altrettanto consistente di collaboratori, che si occupano di tutti gli aspetti del supporto, dallo sviluppo vero e proprio, alla documentazione e mantenimento degli archivi e *mirroring* degli stessi in tutto il mondo. Dall'inizio Debian GNU/Linux è stata la prima distribuzione *community driven*, aperta alla collaborazione di tutti gli sviluppatori e utenti. Rimane ad oggi la più significativa distribuzione non basata su una entità commerciale, e dotata di una Costituzione, Contratto Sociale e documenti di *policy* che governano lo sviluppo e l'evoluzione del progetto su base tecnica e democratica. La distribuzione Debian è stata la prima a introdurre il concetto di micro-pacchettizzazione del software, unitamente al formato *.deb* che presenta informazioni di dipendenza per esprimere le inter-relazioni tra i pacchetti software e assicurare la consistenza del sistema. L'intero processo di sviluppo e la struttura di creazione e distribuzione del software sono aperti e documentati. Dal 2005 una compagnia con sede in Sudafrica, Canonical Ltd, produce la

distribuzione Ubuntu, orientata ad ambiente desktop e derivata da Debian GNU/Linux. Per tale distribuzione viene fornito anche supporto commerciale.

Versione	Data iniziale – finale di rilascio
Debian GNU/Linux 3.1 (<i>sarge</i>)	June 6 th , 2005.
Debian GNU/Linux 4.0 r0-r8 (<i>etch</i>)	April 8 th , 2007 – April 8 th 2009...
Debian GNU/Linux 5.0.0-3 (<i>lenny</i>)	April 14 th – September 5 th , 2009...

Tabella 1. Versioni e rilasci di interesse per DebianGis

Lo distribuzione di Debian GNU/Linux si articola in diversi *flavor* che vengono aggiornati in base allo stadio di sviluppo e al rilascio della successiva distribuzione stabile:

- *oldstable*: la precedente versione stabile
- *stable*: la versione stabile corrente, suggerita per la totalità degli utenti
- *testing*: la successiva versione stabile, in corso di sviluppo e test per bug bloccanti
- *unstable*: la versione di sviluppo, in perenne modifica
- *experimental* un *test bed* di compilazione per versioni dei programmi radicalmente nuove

Lo sviluppo vero e proprio avviene nel flavor *unstable* e appositi programmi provvedono a migrare in *testing* i software privi di bug critici e dipendenti solo da software a loro volta privi di bug critici. Quando *testing* raggiunge un livello complessivo di qualità sufficiente, la distribuzione viene rilasciata come stabile e la nuova *stable* viene aggiornata da quel momento solo per problemi gravi o di sicurezza, riscontrati successivamente al rilascio. Altri software provvedono a compilare automaticamente tutti i pacchetti sulle varie architetture ad ogni nuovo rilascio di ciascuno di essi. Ogni pacchetto software, libreria o programma che sia, viene curato da uno o più maintainer che provvedono ad assicurare la correttezza del programma, l'aderenza agli standard (*policy*) e aggiungono la sovrastruttura (*maintainer scripts*) necessaria alla installazione, configurazione ed eventuale rimozione o aggiornamento del pacchetto stesso, oltre a modifiche (*patch*) per all'adattamento all'ambiente Debian GNU/Linux multi-architettura e alla correzione di bug funzionali più o meno significativi riscontrati sul software come rilasciato all'origine.

L'intero processo di sviluppo si avvale di sistemi software aperti realizzati *in-house* di ausilio allo sviluppatore e agli utenti, che comprendono il *build* automatico, la sottomissione e gestione di *bug report*, il monitoraggio dello stato dei pacchetti, i *framework* di compilazione e installazione in ambiente controllato, la gestione degli archivi, lo sviluppo condiviso e un gran numero di strumenti di analisi del software ad hoc. Il processo *new maintainer*, che permette a nuovi sviluppatori e utenti di partecipare allo sviluppo della distribuzione richiede da uno a due anni di esperienza sul campo per raggiungere un livello di conoscenza dell'intero processo di sviluppo tale da consentire una cooperazione efficace. Per maggiori informazioni sul progetto e la sua evoluzione si rimanda necessariamente alla copiosa documentazione per gli sviluppatori (Debian Project, 2009) e alle *mailing list* di sviluppo del progetto Debian.

FOSS per i GIS

I *free and open source software* per i sistemi informativi geografici comprendono una serie di librerie e programmi orientati sia alla realizzazione di servizi su server che all'uso desktop. Nella maggior parte dei casi tali software hanno uno standard qualitativo elevato, tanto da essere impiegati comunemente anche all'interno di prodotti proprietari (per es. la libreria GDAL/OGR è comunemente adottata all'interno dei prodotti ESRI) e sistemi applicativi critici. La maggioranza dei formati *raster* e vettoriali e dei protocolli oggi utilizzati nei GIS proprietari sono comunque aperti e documentati, e/o standard dell'*Open Gis Consortium* e questo ha consentito la realizzazione di *driver* alternativi in codice aperto per la maggior parte di essi. Non esistono nel mondo FOSS

GIS prodotti generalisti come nella controparte proprietaria. Le soluzioni FOSS, in questo come altri contesti, si ottengono facendo coesistere ed interoperare programmi e librerie differenti, prodotti da team di sviluppo indipendenti e con linguaggi diversi. La conseguenza più immediata di tale approccio è l'esistenza di lunghe catene di dipendenze fra software differenti e vincoli sulle versioni interoperabili e le configurazioni. Se da un lato questo consente di implementare soluzioni raffinate di processing e fornire servizi anche avanzati impilando strati software di terze parti efficacemente, l'utente casuale può risultare disorientato dalla necessità di imparare a dominare prodotti non omogenei per interfacce e modalità di impiego.

L'apprendimento delle componenti utili alla realizzazione di soluzioni FOSS GIS viene peraltro complicata dalla necessità di integrare autonomamente le varie componenti software sulla piattaforma di lavoro prescelta, operazione che si dimostra ostica per la maggior parte degli utenti anche su sistemi operativi proprietari più "convenzionali".

A livello di linguaggi di programmazione le soluzioni FOSS GIS spaziano da linguaggi compilati come C/C++ e Java, a linguaggi interpretati come Python, Perl, Tcl, Ruby o PHP. Ove presenti, le relative interfacce grafiche sono disponibili in ambienti X toolkit, GTK+, Qt, WxWidgets e Tk. Le soluzioni web utilizzano in modo consistente l'approccio AJAX del web 2.0 e librerie di classi Javascript più o meno estese. In alcuni casi esistono soluzioni basate su ambienti di programmazione specializzati per il calcolo scientifico (per es. R). La grande varietà di ambienti di programmazione utilizzati rende il supporto dei più comuni toolkit GIS particolarmente impegnativo per l'utente "casuale" e risulta imprescindibile un consistente ausilio di strumenti avanzati di pacchettizzazione e uno sforzo di mantenimento dei pacchetti software pre-compilati sulle comuni piattaforme GNU/Linux e Microsoft Windows. Tale esigenza è solo parzialmente soddisfatta dai team principali di sviluppo dei vari progetti FOSS GIS e nel caso dei sistemi GNU/Linux è resa particolarmente problematica dall'assenza di una piattaforma unificata di sviluppo e *port* che sia indipendente dalla distribuzione.

Il progetto DebianGis

Il progetto DebianGis nasce nell'ottobre 2004 con la pubblicazione del DebianGis Manifesto (Lovergine, 2004) sulla mailing list degli annunci per gli sviluppatori del progetto Debian. Alcuni pacchetti di interesse GIS, in particolare Grass e Gdal, erano già in qualche modo presenti nell'archivio Debian, se pure con diversi problemi funzionali. In quel periodo lo sviluppo nel progetto Debian stava passando globalmente dalla micro-pacchettizzazione individuale allo sviluppo in team, anche con l'ausilio di collaboratori "esterni" al progetto.

La creazione di un team per il mantenimento coordinato dei pacchetti GIS già presenti e l'integrazione di nuovi software indicati come di massimo interesse dalla comunità GIS, ha motivato la nascita di DebianGis. La scelta di una distribuzione *community-driven* come Debian GNU/Linux per l'ambiente di sviluppo risultava naturale sia per l'esperienza pregressa degli sviluppatori fondatori che partecipavano alle prime attività, sia perchè in quegli anni era una scelta sostanzialmente obbligata dall'assenza di alternative altrettanto aperte e ben strutturate.

Nel corso già dei primi mesi sono stati implementati i nuovi tool di sviluppo coordinato che il progetto Debian stava realizzando in quel periodo: il *repository* comune CVS prima e SVN dalla fine del 2006; il *content management system* Gforge di "Alioth"; un wiki di documentazione del progetto, più tardi ospitato nel wiki principale di Debian. Il progetto DebianGis ha dimostrato subito di attirare l'attenzione della comunità FOSS GIS (e soprattutto gli sviluppatori "upstream" dei progetti GIS principali) come primo e unico sforzo strutturato di mantenimento dell'intero *stack* di librerie e programmi in ambiente GNU/Linux. È risultato di fondamentale efficacia anche l'inteso feed-back fra l'attività di mantenimento in DebianGis e i singoli progetti FOSS GIS per la qualità e quantità di bug tracciati e proposte di *patch*.

Svariate librerie "di base" in precedenza mantenute singolarmente (per es. GMT, HDF4/5), sono passate nel corso del tempo sotto il mantenimento del team di progetto, consentendo un migliore

controllo delle configurazioni e delle relazioni fra i pacchetti e migliorando la qualità globale dei software di interesse GIS. Il team di sviluppo, pur con un inevitabile *turn-over* dei partecipanti nel corso degli anni, ha attirato fino a venti persone, variamente impegnate nella attività di mantenimento e sviluppo di nuovi pacchetti.

Lo stato corrente

Attualmente 66 pacchetti sorgenti sono complessivamente in mantenimento del team DebianGis e almeno altri 40 di medio interesse potrebbero essere pacchettizzati in prospettiva per le prossime versioni di rilascio di Debian GNU/Linux (Debian GIS Project, 2009). Oltre alla attività di sviluppo del ramo principale *unstable*, vengono mantenuti cosiddetti “*backports*” delle nuove versioni di alcuni programmi (ritenuti di interesse primario) per la distribuzione *stable*. Questo consente agli utenti anche l'impiego di versioni aggiornate per singole applicazioni, senza sacrificare supporto e consistenza rispetto agli aggiornamenti di distribuzione successivi. È ragionevole pensare che tale attività *a latere* si espanda in un prossimo futuro. L'attività di management si sta espandendo con l'integrazione dei nuovi strumenti disponibili per i *Debian Blends*, fra i quali il progetto DebianGis ricade come tipologia di prodotto, ai fini soprattutto di un sempre migliore livello di standard qualitativo.

Librerie	Programmi desktop	Programmi server
Shapelib, Gdal/Ogr, Spatialite, Spatialindex, Geos, HDF4/5, Jts, Libgeotiff, Terralib, Netcdf, Proj, Ogd, OpenCV, Openscenegraph, Hdf-Eos4/5, Shapely	AVCE00, E00Compr, Gdal binaries, Earth3d, Grass, Gmt, Gpsbabel, Gpsd, Josm, Gpx2shp, Gts, h5utils, Ncview, Paraview, Thuban, Worldwind, Viking, gpstrans, R	Mapserver, Mapnik, Postgis, gpsdrive, Grass

Tabella 2. Principali pacchetti presenti in DebianGis

I test funzionali comprendono attualmente i sistemi di test di regressione associati ai programmi -- ove disponibili -- e vengono eseguiti periodicamente e ad ogni nuovo rilascio. A questi si accompagna la normale attività di test degli utenti avanzati che utilizzano il *flavor testing* di Debian GNU/Linux. Un obiettivo a medio termine comprende l'integrazione dei relativi framework di test e debug direttamente nei pacchetti sorgenti per consentire una ottimale automatizzazione dei processi di test del codice.

Riferimenti bibliografici

- Lovergine F. P. (2004), “Introducing the DebianGis subproject”, <http://lists.debian.org/debian-devel-announce/2004/10/msg00007.html>
- Lovergine F. P. (2006), “The DebianGis Project”, FOSS4GIS 2006, Lousanne, Sept.
- Lovergine F.P. et al (2009), “DebianGis Wiki”, <http://wiki.debian.org/DebianGis/>
- Debian Project (2009), “Debian Developer's Corner”, <http://www.debian.org/devel/>
- Debian GIS Project (2009), “The DebianGis Package Thermometer”, <http://pkg-grass.alioth.debian.org/debiangis-status.html>
- Debian Project (2009), “Debian Pure Blends”, <http://wiki.debian.org/DebianPureBlends>
- OSGeo Foundation, <http://www.osgeo.org/>

