

# Temi di ricerca per tesi universitarie sul progetto WNoDeS (Worker Nodes on Demand Service), da svolgersi in collaborazione con l'INFN-CNAF

## I. INTRODUZIONE

L'Istituto Nazionale di Fisica Nucleare (INFN) è attualmente impegnato a livello internazionale in una importante attività legata al calcolo scientifico in cui risorse computazionali e di storage possono essere accedute in modo efficiente tramite interfacce di Grid e Cloud computing.

In questo contesto il Centro Nazionale per la Ricerca e Sviluppo nelle Tecnologie Informatiche e Telematiche (CNAF) dell'INFN, con sede a Bologna, ha prodotto un software chiamato *Worker Nodes on Demand Service* (WNoDeS). Lo scopo di questo software è l'adozione di tecnologie di virtualizzazione in centri di calcolo di medie o grandi dimensioni, la integrazione delle risorse e la loro offerta all'utenza attraverso interfacce standard. Queste interfacce sono attualmente il middleware gLite sviluppato in progetti europei come Enabling Grids for E-science (EGEE) e la Open Cloud Computing Interface (OCCI), in fase di definizione presso l'Open Grid Forum (OGF).

Il software WNoDeS è in produzione al Centro di Calcolo Tier-1 del CNAF, uno degli 11 centri mondiali di calcolo di supporto agli esperimenti di fisica basati sul Large Hadron Collider (LHC) costruito al CERN di Ginevra. Correntemente il CNAF gestisce attraverso WNoDeS circa 1500 macchine virtuali, che si prevede diventino circa 7000 entro la fine del 2010.

L'INFN-CNAF è interessato alla evoluzione del software WNoDeS attraverso contributi da parte di studenti universitari di facoltà scientifiche sotto forma di tesi di laurea. La durata di queste tesi (di primo o di secondo livello) sarà valutata caso per caso.

## II. CONOSCENZE RICHIESTE

Per lavorare sui progetti sotto elencati e ottenere risultati in tempi compatibili con la durata di una tesi è necessario possedere competenze legate alle tecnologie e ai linguaggi di programmazione utilizzati in WNoDeS. Le conoscenze generali richieste sono:

- Familiarità con Python o con Java. Questi sono i due linguaggi utilizzati in WNoDeS.
- Familiarità con il funzionamento della rete Internet e di protocolli di comunicazione.
- Conoscenza non strettamente richieste ma apprezzate: tecnologie di virtualizzazione basate su KVM, Grid middleware, Cloud computing, sistemi di versioning del software come CVS, SVN o Git.

Alcuni dei temi proposti possono richiedere competenze più specifiche di quanto esposto in questo paragrafo.

## III. TEMI PROPOSTI

Per domande sui temi sotto elencati si prega di contattare il responsabile del progetto WNoDeS, dott. Davide Salomoni (Davide.Salomoni@cnaif.infn.it).

### A. Aggiornamento delle immagini virtuali

- **Descrizione:** il progetto si occupa di rendere più automatizzata la distribuzione di immagini virtuali verso i nodi che partecipano al cluster di calcolo in caso di prima configurazione dei nodi o di aggiornamento della versione delle immagini stesse.
- **Conoscenze richieste:** shell programming, Python.
- **Difficoltà:** facile/media.

### B. Distribuzione gerarchica delle immagini virtuali

- **Descrizione:** il progetto si occupa di implementare una gerarchia di caching server per una distribuzione ottimale delle immagini virtuali ai nodi che partecipano al cluster di calcolo.
- **Conoscenze richieste:** Linux system management, shell programming.
- **Difficoltà:** facile/media.

### C. Fault Tolerance nel Name Server di WNoDeS

- **Descrizione:** un componente fondamentale dell'architettura di WNoDeS è un servizio chiamato Name Server, che garantisce la consistenza di vari elementi del cluster di calcolo. È essenziale che questo servizio sia dotato di ridondanza e alta affidabilità. Il progetto si occupa di estendere gli attuali meccanismi di ridondanza e di validare l'affidabilità del sistema.
- **Conoscenze richieste:** fault tolerance di servizi, alta affidabilità di sistemi Linux.
- **Difficoltà:** facile/media.

### D. WNoDeS Management Console

- **Descrizione:** attualmente il software WNoDeS viene gestito con un set di tool di amministrazione molto semplici. Questo progetto si occupa di sviluppare una console di gestione che sia basata su una applicazione web. Una volta completato il progetto di base è possibile considerare l'estensione a device mobili. Vedi <http://aws.amazon.com/console/> per un esempio.
- **Conoscenze richieste:** shell programming, Python, web programming.
- **Difficoltà:** facile/media.

#### E. Miglioramenti all'interfaccia Web di OCCI

- **Descrizione:** l'interazione tra gli utenti e il layer Cloud di WNoDeS (che implementa la Open Cloud Computing Interface) può avvenire attraverso una applicazione web. Questa interfaccia web è attualmente in fase di sviluppo; il progetto si occupa di estenderla. Una volta completato il progetto di base è possibile considerare l'estensione a device mobili. Vedi <http://aws.amazon.com/console/> per un esempio.
- **Conoscenze richieste:** Java, web programming.
- **Difficoltà:** media.

#### F. Ottimizzazione delle prestazioni di I/O

- **Descrizione:** il progetto si occupa di valutare metodi per aumentare le prestazioni di I/O su Virtual Machines e di studiare ed implementare meccanismi di Quality of Service (QoS).
- **Conoscenze richieste:** Linux, shell programming, network QoS.
- **Difficoltà:** media.

#### G. Scalabilità del Name Server di WNoDeS

- **Descrizione:** Per garantire la scalabilità di WNoDeS occorre gestire molte connessioni contemporaneamente, evitando un numero eccessivo di thread attivi. Il progetto si occupa del miglioramento della scalabilità del servizio introducendo l'utilizzo di notifiche asincrone congiuntamente all'utilizzo di una queue per la gestione delle richieste.
- **Conoscenze richieste:** comunicazione tra processi, programmazione multi-threading in Python.
- **Difficoltà:** media.

#### H. Test di Cloud Storage

- **Descrizione:** la fornitura di capacità significative di storage virtuale in WNoDeS è attualmente in fase preliminare. Il progetto si occupa di testing e di sviluppo della parte di Cloud Storage.
- **Conoscenze richieste:** shell programming, file system distribuiti.
- **Difficoltà:** media.

#### I. Integrazione di meccanismi di autenticazione

- **Descrizione:** l'autenticazione degli utenti a WNoDeS è basata internamente su certificati digitali X.509. Gli utenti, tuttavia, possono essere autenticati anche con altri metodi, come Kerberos, username/password o federazioni Shibboleth. Il progetto intende partecipare allo sviluppo e al test di un gateway di autenticazione tra questi metodi e i certificati X.509, che possono anche essere di breve durata (short-lived certificates).
- **Conoscenze richieste:** sicurezza informatica, certificati digitali.
- **Difficoltà:** media.

#### J. Integrazione di meccanismi di autorizzazione

- **Descrizione:** la validazione delle richieste d'uso delle risorse in WNoDeS avviene attraverso un framework di autorizzazione esterno e generico. Il progetto si occupa del test e dello sviluppo delle possibilità offerte da questo framework.
- **Conoscenze richieste:** uso di API, wrapping di API da C a Python.
- **Difficoltà:** sono possibili diversi sviluppi, di varia difficoltà.

#### K. Revisione del Messaging Protocol di WNoDeS

- **Descrizione:** i processi di WNoDeS comunicano tra di loro attraverso un protocollo di comunicazione molto semplice sviluppato ad-hoc. Il progetto si occupa di rivedere questo protocollo e di passare a un approccio basato su standard.
- **Conoscenze richieste:** comunicazione tra processi, programmazione asincrona.
- **Difficoltà:** media/difficile.

#### L. Notifiche asincrone tra processi WNoDeS

- **Descrizione:** alcuni processi di WNoDeS interagiscono tra loro con un meccanismo a polling. Il progetto si occupa di verificare la possibilità di implementare notifiche asincrone relativi a passaggi di stato tra processi.
- **Conoscenze richieste:** comunicazione tra processi, programmazione asincrona.
- **Difficoltà:** media/difficile.

#### M. Studio di fattibilità di procedure per rendere le immagini virtuali indipendenti dall'architettura di calcolo di un sito

- **Descrizione:** si vogliono utilizzare le immagini virtuali indipendentemente dalle architetture di calcolo, del sistema di storage utilizzato e del modello computazionale adottato in un sito. Il progetto deve quindi identificare e sviluppare le procedure necessarie per rendere le immagini virtuali indipendenti dal sito, utilizzare standard come l'Open Virtualization Format (OVF) per le specifiche e prevedere la possibilità anche di una distribuzione di immagini a livello geografico.
- **Conoscenze richieste:** shell programming, Python.
- **Difficoltà:** media/difficile.

## IV. ALTRO

Oltre ai progetti sopra menzionati, la roadmap di WNoDeS prevede numerosi altri sviluppi. In caso di interesse si prega di contattare il dott. Davide Salomoni ([Davide.Salomoni@cnaif.infn.it](mailto:Davide.Salomoni@cnaif.infn.it)).