

## Gestione di applicazioni di calcolo eterogenee con strumenti di cloud computing: l'esperienza di INFN-Torino.

Stefano Bagnasco, Dario Berzano, riccardo Brunetti, Matto Concas, Stefano Lusso

L'obiettivo del progetto di Cloud Computing al Centro di Calcolo della Sezione di Torino dell'INFN era riorganizzare e consolidare le risorse di calcolo di un tipico centro Tier-2 EGI/WLCG come il nostro. L'infrastruttura core, inizialmente pensata per servire il "cliente" principale (nel nostro caso il Tier-2 dell'esperimento ALICE), è stata il nucleo principale attorno al quale l'infrastruttura è progressivamente cresciuta. Dopo una fase preliminare di analisi degli strumenti disponibili è stato realizzato un prototipo attraverso l'uso di strumenti mainstream e open source come OpenNebula, Lustre, GlusterFS e OpenWRT; dalla fase di prototipo si è poi rapidamente passati alla produzione tanto che ormai l'infrastruttura include parecchie decine di host fisici. Siamo così riusciti a integrare in un'ottica di IaaS use case molto diversi: dalla Grid farm tradizionale del Tier-2 EGI/WLCG di ALICE, comprendendo tutti i servizi Grid come CE e BDII, all'Analysis Facility interattiva basata su PROOF, sempre di ALICE, a una mini-farm per l'analisi di immagini tomografiche utilizzata da uno spin-off INFN fino a singole istanze di macchina di calcolo per un utenze spot.

È ora possibile soddisfare dinamicamente anche le esigenze dei gruppi minori o persino di singoli utenti, gradualmente eliminando le farm o i server "proprietary" con conseguenti economie di scala nei costi e nella gestione. La tipologia di macchine virtuali disponibili si divide in "worker" (tipicamente nodi di elaborazione che vengono istanziati e differenziati a partire da una immagine base) e "service" (host virtuali per servizi critici che sono in grado di migrare e poggiano su hardware di classe server); alle due tipologie corrisponde un differente tipo di storage back end (locale nel primo caso, basato su GlusterFS nel secondo).

GlusterFS 3.3 è utilizzato nell'infrastruttura sia per esportare un repository di immagini di macchine virtual (le operazioni di I/O sono rare e nessuna replicazione né aggregazione viene usata) sia per memorizzare le immagini attualmente in esecuzione delle macchine virtuali di classe "service". In questo caso le operazioni di I/O sono frequenti e intensive, e l'affidabilità è garantita da due server che mantengono repliche sincrone. Il sistema rimane in esecuzione anche qualora uno dei due server o dischi non funzioni, e si ripara automaticamente quando esso ritorna disponibile (self-healing). Lo stesso file system è infine usato per aggregare un pool di dischi di grandi dimensioni per servire velocemente dati di HEP ai worker PROOF dell'Analysis Facility. Nel talk saranno brevemente presentati i risultati di benchmark orientati ai tre use case che integrano l'esperienza decisamente positiva grazie alla facilità di gestione e ai tool messi a disposizione.

La connettività dei cluster virtuali è realizzata, anche per ragioni di sicurezza, attraverso l'uso di "virtual router" che permettono di partizionare opportunamente la rete vista dai nodi virtuali. In questo modo ciascuna macchina virtuale o gruppo di macchine che fa riferimento a una certa utenza è completamente isolata da quelle di altri utenti anche se utilizza la stessa macchina fisica.