

Un sistema di accounting economico per le grid computazionali

Rosario M. Piro

Istituto Nazionale di Fisica Nucleare (INFN) - Sezione di Torino

E-mail: piro@to.infn.it

Luglio 2003

Abstract

La gestione efficace dei numerosi utenti e delle risorse distribuite di una grid (griglia) computazionale richiede un sistema sofisticato che tenga traccia delle risorse utilizzate dai singoli utenti. Un approccio promettente per un tale sistema di accounting riguarda l'implementazione di un modello economico in cui le risorse consumate vengono pagate con crediti virtuali. Nel presente articolo viene presentato il sistema di accounting del progetto European DataGrid.

1 La necessità del grid computing

L'invenzione di Internet, del World Wide Web e soprattutto i protocolli ad essi associati ci hanno permesso di scambiare dati in modo veloce su larga scala. Le possibilità offerte da queste tecnologie hanno dato un grosso impulso ad alcuni settori scientifici ed economici.

Le attuali tecnologie utilizzate su Internet però, mirando principalmente ad un puro *scambio di dati*, potrebbero non essere più sufficienti per soddisfare le future esigenze della ricerca in alcuni particolari campi come la fisica delle alte energie (fisica delle particelle elementari) o la bioinformatica. Il Large Hadron Collider (LHC), l'acceleratore di particelle più potente attualmente in costruzione al CERN di Ginevra, ad esempio, produrrà una tale quantità di dati sperimentali che nuovi sistemi informatici saranno necessari per la loro elaborazione.

Le tecnologie grid, che non mirano soltanto ad uno scambio di informazioni ma anche di *risorse computazionali*, possono essere viste come un'evoluzione di Internet. Ciò permetterà una stretta collaborazione tra le cosiddette organizzazioni virtuali (gruppi di scienziati, centri di ricerca, collaborazioni industriali, ecc.) per mettere a disposizione reciprocamente le proprie risorse [6]. Le possibili applicazioni che potranno trarre

beneficio dalle enormi capacità di calcolo o dall'accesso a database distribuiti riguardano, ad esempio, l'analisi degli esperimenti di fisica delle particelle, la ricerca biomedica, la simulazione meteorologica oppure il design e lo sviluppo di nuovi prodotti industriali.

2 Il progetto European DataGrid

Il progetto European DataGrid (EDG [1, 2]), finanziato in gran parte dall'Unione Europea, è uno dei progetti di grid computing più ambiziosi e coinvolge centinaia di scienziati di tutto il mondo. Anche se le prime applicazioni del progetto sono di tipo scientifico, le possibilità offerte dalla nuova tecnologia avranno sicuramente un impatto su future attività industriali e commerciali. Infatti potrebbe essere più economico per una singola organizzazione affittare risorse di calcolo, se necessario, piuttosto che acquistarne di nuove, soprattutto se è possibile compensare i costi dando in affitto le proprie risorse non utilizzate.

L'implementazione di una grid efficace ed affidabile, di *natura eterogenea e geograficamente distribuita*, richiede lo sviluppo di nuovi approcci di gestione che ottimizzino l'uso globale del sistema senza però compromettere le policy locali (di sicurezza, di accesso autorizzato, ecc.) dei singoli fornitori. L'utente finale dovrà essere in grado di utilizzare la grid in modo trasparente come una singola risorsa virtuale con enormi capacità.

La Figura 1 descrive in modo semplificato i singoli strati dell'implementazione del sistema. La base del modello è l'infrastruttura, ovvero i servizi di rete e le risorse locali interconnesse che spesso vengono chiamate *computational fabric*. I servizi globali necessari per il funzionamento della grid vengono forniti dal cosiddetto *middleware*, il software in mezzo tra le applicazioni e le risorse. Esso è costituito da componenti che distribuiscono il carico di lavoro e i dati necessari tra le singole risorse e le memorie di massa, da componenti

che servono a monitorare lo stato e da componenti che gestiscono l'accesso al computational fabric. Il livello più alto del modello sono le applicazioni vere e proprie che utilizzano i servizi forniti dal middleware per accedere alle funzionalità offerte dalla grid.

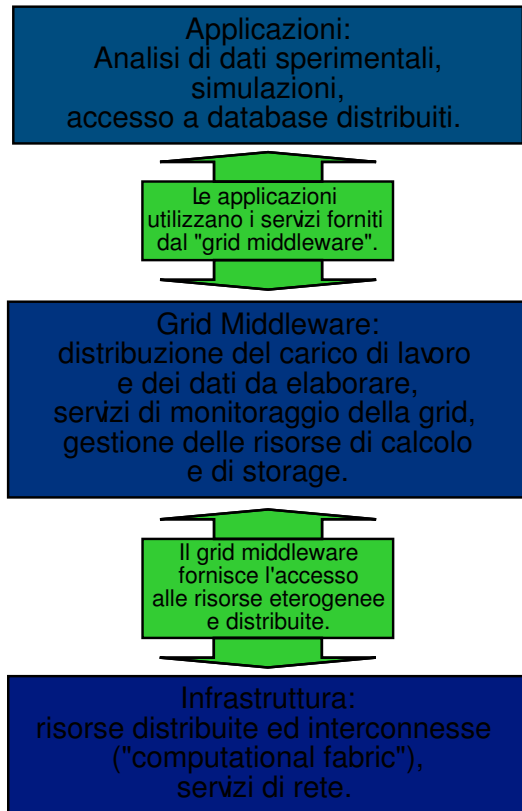


Figura 1. Strati dell'implementazione di una grid

Il middleware di EDG è tra l'altro basato su un pacchetto software denominato Globus Toolkit [7] che fornisce servizi di livello base per la sicurezza, la scoperta e la gestione di risorse, l'accesso ai dati, la comunicazione all'interno della grid e così via. Il Grid Information Service (GIS) di Globus, per esempio, permette di accedere tramite un'interfaccia semplice a informazioni sulle risorse connesse e sul loro stato. L'autenticazione ed autorizzazione di utenti e risorse si basa sulla Grid Security Infrastructure (GSI) che utilizza certificati di identità in formato X.509 e la crittografia asimmetrica (tramite chiave pubblica e privata).

2.1 La gestione del carico di lavoro

Uno dei compiti più importanti del middleware è la *distribuzione del carico di lavoro*, ovvero la distribuzio-

ne delle applicazioni utente (spesso chiamate job) sulle singole risorse a disposizione. Siccome la natura distribuita della grid e il numero elevato di utenti causano un carico di lavoro imprevedibile, le prestazioni e quindi la qualità del servizio (QoS) dipendono molto da una distribuzione (scheduling) opportuna dei singoli job.

La Figura 2 mostra un diagramma semplificato del sistema di gestione del carico (Workload Management System) di EDG [3], che viene sviluppato sotto la responsabilità dell'Istituto Nazionale di Fisica Nucleare (INFN).

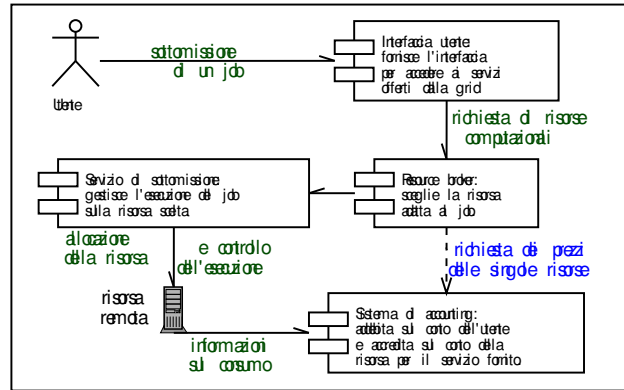


Figura 2. Sistema di gestione del carico di lavoro (Workload Management System)

L'accesso alla grid avviene attraverso un'interfaccia utente (User Interface). L'utente sottomette un job composto dall'applicazione da eseguire e da una descrizione dell'applicazione, espressa in un linguaggio semplice che consente di notificare al sistema le esigenze dell'eseguibile (in termini di sistema operativo, memoria minima, eventuali pacchetti software installati, ecc.). In modo simile l'utente può effettuare la cancellazione di un job già sottomesso, la richiesta di informazioni sullo stato di esecuzione e la richiesta di recuperare l'eventuale output del job.

Il Resource Broker è il componente responsabile dello *scheduling* che può essere suddiviso in tre fasi: la prima fase è la scoperta delle risorse disponibili su cui l'utente (o l'organizzazione virtuale a cui appartiene) è autorizzato, il che richiede l'interrogazione del sistema di informazione (GIS). La seconda fase viene denominata *matchmaking* e corrisponde alla valutazione in termini di sistema operativo, architettura (processori, dispositivi), ecc. per individuare l'insieme di risorse su cui un'esecuzione del job sia possibile. L'ultima fase è la decisione di scheduling, ovvero la scelta di una (o più) particolari risorse su cui eseguire il job dell'utente. I criteri che influiscono sulla scelta possono essere

vare a seconda delle esigenze dell'intero sistema: la disponibilità locale di dati a cui il job deve accedere per minimizzare il traffico di rete, l'occupazione corrente delle singole risorse per minimizzare il tempo di attesa, ecc.

La sottomissione del job sulla risorsa scelta non viene effettuata dal Resource Broker stesso, ma da un servizio che serve anche per monitorare lo stato di esecuzione per poter notificare all'utente la fine dell'esecuzione, o un eventuale fallimento.

La maggior parte dei passi necessari fino al recupero dell'output viene registrata in un servizio denominato Logging & Bookkeeping (non mostrato in Figura 2). Ciò permette agli amministratori e agli utenti del sistema di controllarne il corretto funzionamento.

2.2 Il sistema di accounting

Il sistema di accounting di EDG (DataGrid Accounting System, DGAS [4, 5]) viene sviluppato dalla Sezione di Torino dell'INFN per fornire un metodo per tener traccia della potenza computazionale utilizzata dai singoli utenti. Inoltre favorisce algoritmi di scheduling basati su *economie computazionali* per raggiungere un bilanciamento del carico di lavoro e quindi un'ottimizzazione delle prestazioni dell'intera grid.

I prezzi per l'uso vengono assegnati da una Price Authority (PA) a seconda della qualità delle risorse e del loro corrente carico. Un bilanciamento globale può essere raggiunto assegnando i prezzi in modo da incentivare gli utenti (o il Resource Broker in loro nome) a sottomettere nuovi job su risorse libere. L'architettura del sistema di accounting prevede che il Resource Broker richieda informazioni sui prezzi dalle PA prima di effettuare la decisione di scheduling.

L'accounting del consumo (mostrato in Figura 3) viene effettuato tramite un sistema di server distribuiti, denominati Home Location Register (HLR) ed ispirati alla rete di telefonia mobile GSM. La risorsa che ha eseguito l'applicazione di un utente comunica l'uso, in termini di tempo CPU, spazio disco occupato, memoria utilizzata ecc., all'HLR responsabile per il conto dell'utente. L'HLR richiede il prezzo per unità della risorsa e calcola il costo complessivo dell'esecuzione. Una volta conosciuto il costo totale versa il pagamento (in termini di GridCredits ovvero crediti virtuali) all'HLR della risorsa. Quest'ultimo accredita sul conto della risorsa e conferma la ricevuta del versamento all'HLR dell'utente che finalmente addebita sul conto dell'utente.

Un tale sistema permette di implementare un'economia virtuale (economia computazionale) in cui l'autorizzazione per accedere ai servizi della grid dipende

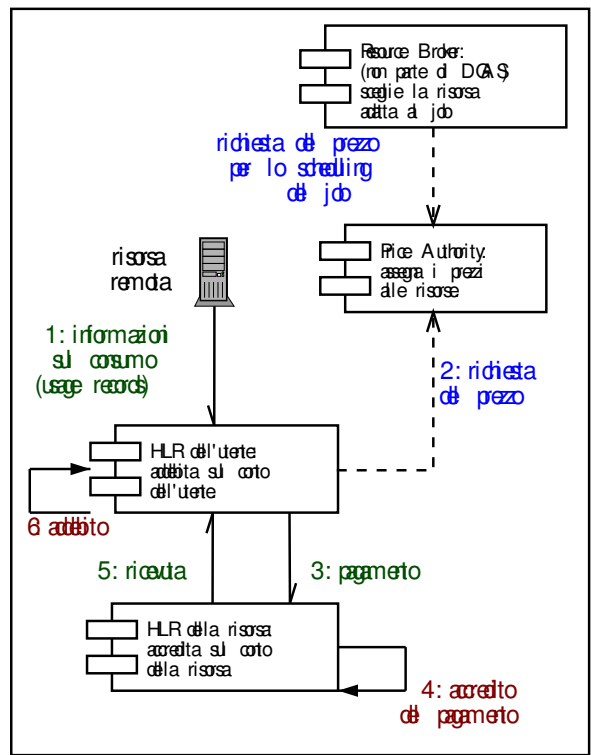


Figura 3. Sistema di accounting di EDG

anche dall'autorizzazione economica dell'HLR che potrebbe bloccare l'esecuzione in mancanza di crediti sul conto dell'utente. In modo naturale viene creato un mercato di scambio in cui le organizzazioni virtuali partecipanti guadagnano crediti mettendo a disposizione le loro risorse libere e, in caso di necessità, spendono crediti per poter utilizzare risorse altrui.

Il modello economico da applicare deve garantire una relativa stabilità del prezzo e l'esistenza di un punto d'equilibrio tra domanda e offerta per permettere un bilanciamento del carico di lavoro e quindi per raggiungere un'ottimizzazione della qualità del servizio. Il modello da implementare è un obiettivo della ricerca attuale, sia all'INFN di Torino che ad altri centri di ricerca in tutto il mondo. Occorre definire opportune funzioni di utilità per consumatori e fornitori di risorse che impediscano guerre di prezzi cicliche e che permettano di calcolare i prezzi delle risorse senza un aumento ingiustificato del carico di lavoro.

2.3 Conclusione

Come già accennato, le prime applicazioni di EDG sono di natura scientifica. Attualmente esiste una grid di test, denominata testbed, di cui fanno parte siti di

istribuiti in vari paesi europei, che viene utilizzata per testare non soltanto il middleware sviluppato dal progetto europeo, ma anche il funzionamento di software utilizzato per la ricerca nei settori di fisica delle alte energie, earth observation e biomedicina.

Le grid computazionali, una volta raggiunta una maturità sufficiente, potrebbero anche essere sfruttate per scopi commerciali, soprattutto in ambiti che richiedono la collaborazione di più imprese. I partner di un consorzio per lo sviluppo di un nuovo aereo civile, ad esempio, potrebbero condividere risorse computazionali ma anche l'accesso a software altamente specializzato.

Una prima ricaduta dell'applicazione delle nuove tecnologie probabilmente sarà la riduzione degli investimenti hardware delle grandi aziende con sedi distribuite in tutto il mondo. In futuro le grid potranno diminuire la necessità di costruire grandi centri di calcolo indipendenti.

Questi esempi sottolineano l'importanza di un'opportuna distribuzione del carico di lavoro tra le risorse condivise. Un accounting affidabile, in combinazione con un modello economico opportuno, potrebbe portare ad una distribuzione più bilanciata, aumentando così le prestazioni del sistema.

In collaborazione con l'Eurix Group Torino.
(www.eurix.it)

Riferimenti bibliografici

- [1] Sito ufficiale del progetto European DataGrid: <http://www.eu-datagrid.org/>
- [2] B. Segal, "Grid Computing: The European Data Grid Project", IEEE Nuclear Science Symposium and Medical Imaging Conference, Lione, Francia, 15-20 Ottobre 2000.
- [3] Sito ufficiale dell'EDG Workload Management Work Package: <http://www.infn.it/workload-grid/>
- [4] Sito ufficiale del DataGrid Accounting System: <http://www.to.infn.it/grid/accounting/>
- [5] A. Guarise, A. Werbrouk e R. Piro, "DataGrid Accounting System Architecture - v1.0", Technical Report DataGrid-01-TED-0126-1.0, INFN di Torino, Febbraio 2003.
(www.to.infn.it/grid/accounting/curr_arch.pdf)
- [6] I. Foster, C. Kesselman e S. Tuecke, "The Anatomy of the Grid - Enabling Scalable Virtual Organizations", The International Journal of High Performance Computing Applications, Volume 15, Number 3, Fall 2001.
- [7] Sito ufficiale del progetto Globus: www.globus.org