

## STUDIO DI FATTIBILITA'

### "Tecnologie di Grid-computing per le P.M.I. del distretto di elettronica e tecnologie avanzate di Genova"

Progetto finanziato dal P.S.T.L. - Parco Scientifico Tecnologico della Regione Liguria nell'ambito del "Docup Obiettivo 2 (2000-2006) Misura 3.7 Sottomisura D"

Il presente studio di fattibilità è finalizzato a valutare la possibilità di trasferire tecnologie ed applicazioni GRID all'interno di sistemi e distretti industriali, con particolare riferimento alle PMI del distretto high tech di Genova.

#### Premessa

Lo studio si inserisce all'interno di un contesto industriale e di ricerca in fase di profonda evoluzione, sia a livello internazionale che territoriale. Come può una PMI tecnologica italiana competere con concorrenti di Taiwan, Singapore, Malaysia, Cina che hanno manodopera a basso costo riuscendo a praticare prezzi più competitivi? Puntando sull'efficienza, sullo sviluppo di prodotti e processi innovativi, adottando tecnologie di produzione d'avanguardia. Un simile salto qualitativo è spesso difficile per una singola azienda, ma può diventare meno oneroso se affrontato nell'ambito di una crescita comune di un distretto industriale.

E' in questa ottica che entra in gioco la tecnologia GRID che consente di usufruire di servizi comuni, pur mantenendo le individualità. Il motto "piccolo è bello" è ancora parzialmente vero: piccolo è ancora bello purché sia condiviso in rete.

#### La tecnologia GRID – Cos'è?

Esistono due principali filoni di sviluppo della tecnologia:

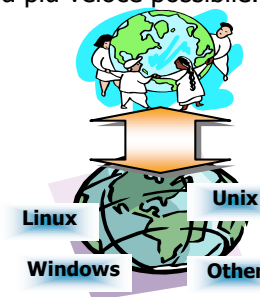


##### *Computational Grid.*

Siamo nel campo di applicazioni di calcolo ad elevate prestazioni, l'utente necessita di più "teraflops" possibili: progettazione e testing, processi batch, computer grafica, modelli statistici, previsioni meteo, etc... La tecnologia GRID combina ampie e differenti tipologie di risorse per affrontare problemi che non potrebbero essere risolti su di un singolo sistema, o almeno per farlo nella maniera più veloce possibile.

##### *Data Grid.*

Siamo nel campo di applicazioni "data-intensive". Nei prossimi anni enormi quantità di dati scientifici e non giungeranno da differenti siti del pianeta: non solo esperimenti astrofisici, ma anche dati provenienti dalle più svariate applicazioni in ambito scientifico ed industriale. La tecnologia Grid consente di memorizzare, organizzare ed analizzare dati mantenendoli in repository, librerie digitali, database geograficamente distribuiti.



Così come il Web è uno strumento per *condividere informazioni* attraverso Internet, GRID è uno strumento per *condividere risorse di calcolo e risorse di storage* attraverso Internet.

Simulare un enorme computer virtuale che, in maniera intelligente, metta a disposizione le proprie risorse partendo da sistemi eterogenei connessi tra loro: le imprese hanno a disposizione "on demand" tutta la potenza necessaria ad un costo precedentemente inaccessibile.

#### Le potenzialità delle reti GRID per le PMI

Le reti GRID rappresentano l'evoluzione nell'utilizzo della rete Internet. Oggi Internet viene utilizzata prevalentemente per "comunicare", mediante scambio di informazioni e di dati. Tramite le GRID, Internet servirà per "condividere" potenza di calcolo e dati. In analogia alle rete elettrica (o a quella idrica), che da più centrali (o sorgenti) trasporta l'energia, tramite una infrastruttura complessa, fino agli utenti finali, così le GRID, tramite l'infrastruttura Internet, renderanno disponibili agli utenti "centrali di calcolo" e "centrali di immagazzinamento dati", indipendentemente dalla loro collocazione geografica.

## La tecnologia GRID – Elementi innovativi

La tecnologia introduce diversi elementi innovativi che possono risultare di sicuro interesse.

### Gestione

Migliore gestione delle risorse di rete, che spesso risultano sotto utilizzate. Mediamente le macchine desktop sono sfruttate meno del 5% ed i server presentano picchi di inattività.

### Affidabilità

Ad oggi spesso una maggiore affidabilità si ottiene attraverso una duplicazione dei componenti: l'hardware risulta molto costoso. In un sistema Grid ben organizzato se un componente della rete fallisce, le altre parti sono in grado di sopprimere.

### Sicurezza

Autenticazione a chiave pubblica e chiave privata, certificati digitali e single sign-on consentono uno scambio sicuro di informazioni sensibili tra gli utenti dell'Organizzazione Virtuale (VO).

### Condivisione

Condivisione di risorse (cpu, file system, sw, servizi, licenze, larghezza di banda,...) tra sistemi eterogenei attraverso standard precisi.



### Calcolo parallelo e distribuito

La tecnologia consente la gestione di applicazioni e servizi distribuiti e l'esecuzione di sotto-processi su risorse differenti, ottenendo spesso notevoli guadagni in termini di performance.

## Il Progetto

Il progetto è stato sviluppato dalle aziende esecutrici seguendo fasi successive:

- Indagine sullo stato dell'arte
- Stima investimenti necessari all'implementazione di reti Grid
- Progettazione e implementazione di rete Grid pilota
- Progettazione portale Grid per l'accesso alle risorse
- Implementazione servizi di prova



In seguito ad un'approfondita analisi dello stato dell'arte della tecnologia è stato possibile individuare i software ed i prodotti potenzialmente in grado di soddisfare le problematiche evidenziate nell'ambito delle PMI ed effettuare una stima degli investimenti necessari all'implementazione di un'infrastruttura Grid. Di conseguenza le scelte di progettazione sono ricadute sulle soluzioni che presentano al momento il miglior rapporto fra investimento e valore aggiunto aziendale contestualizzato alla realtà delle PMI.



Una generica architettura Grid presenta diversi moduli ognuno dei quali è dedicato a specifiche funzionalità all'interno dell'organizzazione virtuale:

- Gestione dei servizi
- Gestione delle policy
- Comunicazione
- Sicurezza

Una volta individuate le componenti core di una generica infrastruttura Grid, si è scelto di non restare vincolati ad un'unica soluzione architettonica specifica di una tipologia di servizio per non perdere di vista il processo di standardizzazione tracciato dalla comunità scientifica internazionale. La continua e rapida evoluzione della tecnologia e dei suoi standard di riferimento impone una continua attenzione ai software open-source in fase di sviluppo, con tutte le problematiche e le difficoltà che ne conseguono.

Si è quindi proceduto con l'implementazione parallela di diversi ambienti in modo da mantenere un'apertura verso un ampio range di possibili applicazioni.

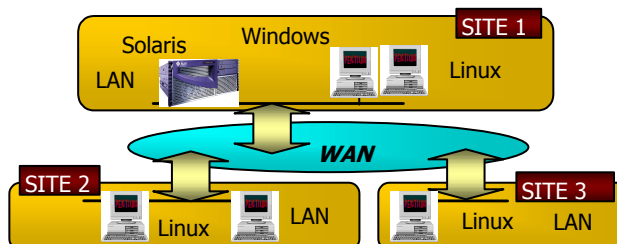
Il prototipo presenta tre differenti ambienti, uno orientato all'integrazione dei servizi, uno per il Computational GRID ed uno per il Data GRID:

### GRID Globus

Ambiente destinato a diventare lo standard de facto della tecnologia. Un'architettura in grado di contenere differenti tipologie di servizio in un'unica soluzione integrata. Utilizzando standard aperti, consente di implementare soluzioni per la gestione di organizzazioni estremamente complesse con elevati standard di qualità e sicurezza, coordinando risorse distribuite ed eterogenee.

### GRID LSF

Architettura per il calcolo distribuito e parallelo orientata all'erogazione di servizi verso soggetti che necessitano di aumentare la potenza di calcolo aziendale. L'idea di base è fornire potenza di elaborazione on-demand agli utenti in modo semplice e sicuro. E' compito della Grid gestire il bilanciamento del carico computazionale tra i diversi nodi della rete in base a politiche di scheduling e tecniche di ottimizzazione delle risorse.



### GRID SRB

Architettura per la virtualizzazione dello storage ottenuta attraverso un middleware client-server che fornisce un'interfaccia di accesso uniforme a dati e metadati geograficamente distribuiti e di tipo eterogeneo. Questa soluzione è già ampiamente utilizzata per gestire grandi moli di dati in campo scientifico in discipline come l'astronomia, le neuroscienze, la fisica e la chimica. Grazie alle sue capacità di condividere dati eterogenei e distribuiti, sta prendendo piede anche in campo industriale con diverse soluzioni di business. In particolare data sharing per gestire la documentazione e l'archivio dei dati aziendali, l'integrazione e la collaborazione orientata al settore biomedico, gestionale, normativo.

### Il portale

Per ognuno dei precedenti ambienti è stato implementato un portale che consente un accesso ai servizi semplice e intuitivo, nascondendo i dettagli implementativi legati all'utilizzo ottimale delle risorse di calcolo della Grid sottostante. Attraverso un'ulteriore integrazione si è ottenuto un unico portale di accesso alla GRID in grado di fornire al tempo stesso funzionalità di base e servizi atti a soddisfare le esigenze di sfruttamento delle risorse da parte delle aziende, integrando nel framework le funzionalità peculiari di Computational e Data Grid.

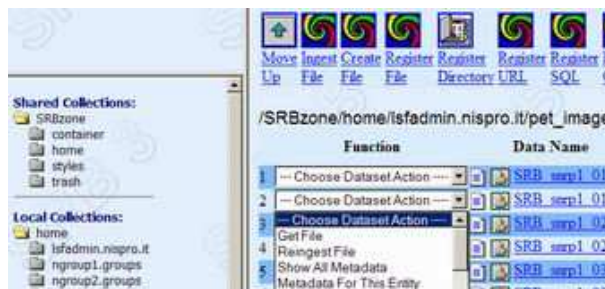


Grazie ad una completa integrazione l'utente ha di fronte un unico punto di accesso ad una GRID polifunzionale che mette a disposizione servizi di base e applicazioni specifiche.

Sia la gestione dei processi e dei file di dati, sia i risultati dell'elaborazione, sia le risorse disponibili e la configurazione generale della Grid, possono essere controllati da un'unica interfaccia web.

L'utente, una volta autenticato, in base alle proprie autorizzazioni ha accesso a determinati servizi o applicativi che gli forniscono, in maniera del tutto trasparente, la potenza di calcolo della Grid.

L'interfaccia di gestione dello storage virtuale mette a disposizione le risorse attraverso una comune connessione ad internet offrendo una vasta pletora di operazioni sui dati (inserimento, organizzazione dei dati, replicazione, etc...).

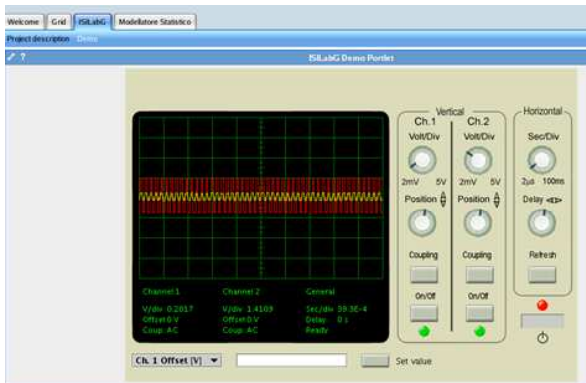
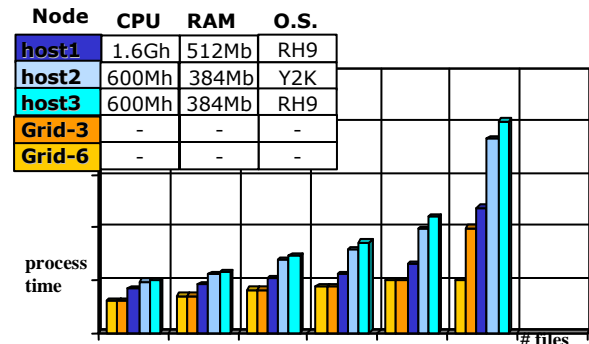


## Le applicazioni

Il costante confronto con le PMI e la consapevolezza della continua evoluzione tecnologica hanno portato a focalizzare l'attenzione su alcune tipologie di applicativi:

### Post-elaborazione del Burn-in

Nel contesto di RGM S.p.a., un'azienda produttrice di alimentatori ad alta tecnologia, trasformatori, circuiti ibridi di controllo e di potenza, l'obiettivo è distribuire la post-elaborazione dell'output prodotto dal software di gestione del burn-in mediante fieldpoint su più nodi, e controllare il ciclo di vita dei job attraverso una gestione centralizzata dei processi. L'architettura Grid fornisce come ulteriore valore aggiunto la possibilità di un controllo remoto via web ed un'ampia scalabilità della soluzione.



### Controllo Remoto di Sistemi di Misura

Per la gestione e il monitoraggio remoto di impianti e strumentazioni, per il monitoraggio di processi di misura, può risultare utile utilizzare tecniche per il controllo basate su Grid che agevolano l'integrazione delle fasi di acquisizione ed elaborazione e la cooperazione fra aziende.

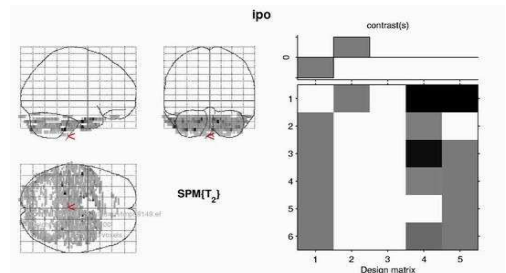
### Modellatore Statistico

Per la diagnosi e il controllo di qualità è stato realizzato un servizio di classificazione, previsione o analisi di correlazioni su dati provenienti da processi industriali e aziendali. Distribuisce algoritmi computazionalmente pesanti sulla Grid per fornire una risposta più veloce all'utente, garantisce affidabilità e sicurezza.

### Diagnosi precoce del morbo di Alzheimer

E' sempre più attuale la necessità di condividere informazioni residenti all'interno di diverse organizzazioni (sanitarie e non) in maniera sicura, scalabile e performante.

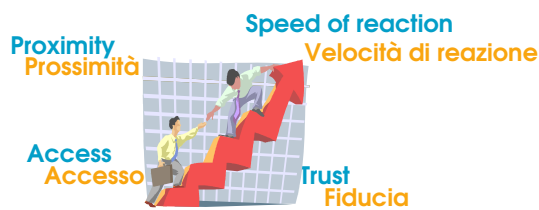
L'introduzione della tecnologia Grid consente un accesso uniforme, affidabile e sicuro a dati e metadati geograficamente distribuiti su piattaforme eterogenee, un trasferimento dei dati più sicuro e veloce e la possibilità di elaborazioni distribuite.



## Conclusioni

Questo studio ha evidenziato gli aspetti delle tecnologie GRID di potenziale utilità per le PMI, rilevando tuttavia anche le barriere che si frappongono ad una loro rapida diffusione: le tecnologie GRID, sviluppate in ambito scientifico, sono ancora relativamente immature per un diffuso uso a livello industriale, soprattutto nel segmento PMI. Inoltre la cultura informatica delle imprese di medio-piccole dimensioni rappresenta spesso un ostacolo all'introduzione di sistemi IT avanzati nei cicli produttivi aziendali.

Nonostante siano emersi alcuni aspetti critici, le tecnologie GRID, potenzialmente, rappresentano una significativa opportunità per le PMI.



- La possibilità di avere a disposizione importanti risorse di calcolo e gestione dei dati, modulabili a seconda delle esigenze temporanee dell'azienda.
- La creazione di Virtual Organization (VO), caratterizzate da diverse competenze, pone le basi per una cooperazione maggiore di quella finora consentita da Internet su aspetti di gestione delle risorse finora poco "sentiti" dalla realtà del mondo imprenditoriale medio-piccolo

**Partners dello studio:**

**NIS - Network Integration and Solutions S.r.l. -**  
Via Benedetto Marcello, 38 - 20124 Milano  
tel: +39 02 54117529 - fax: +39 02 5456367  
Corso Torino, 14/3 - 16129 Genova  
tel: +39 010 5954946 - fax: +39 010 8680159  
web: [www.nispro.it](http://www.nispro.it)  
email: [info@nispro.it](mailto:info@nispro.it)

**Techfin S.r.l.**

Via XXV Aprile, 4-4 16123 Genova  
tel. +39 010 2516672 fax +39 010 2462616

e-mail: [miracoli@techfin.it](mailto:miracoli@techfin.it)



**D.I.B.E. - Biophysical and Electronic Engineering Department**  
via Opera Pia 11a - 16145 Genoa – Italy  
web: <http://www.dibe.unige.it>  
email: [arianna.poggi@dibe.unige.it](mailto:arianna.poggi@dibe.unige.it)  
Tel. +39 010 353 2271  
Fax. +39 010 353 2175



**Biolab – D.I.S.T.** (Dipartimento di Informatica, Sistemistica e Telematica)  
Viale Causa 13, Villa Bonino (piano -1) 16145 - Genova  
tel: +39 010 353 2789 / fax: +39 010 353 2948  
web: <http://www.bio.dist.unige.it/>  
email: [ministaff@bio.dist.unige.it](mailto:ministaff@bio.dist.unige.it)



**IBM Italia**  
Circonvallazione Idroscalo  
20090 Segrate - Milano  
tel: +39 800-017-001  
web: [www.ibm.it](http://www.ibm.it)  
email: [info@ibm.it](mailto:info@ibm.it)



**EBIT AET S.p.a.**  
Via Siffredi 58 - 16153 Genova  
Tel. +39-010-6547464  
Fax +39-010-6547465  
web: [www.ebit-aet.com](http://www.ebit-aet.com)  
email: