

Sistema Informatico SISSA

Situazione e previsioni di sviluppo della Computer Room SISSA

Dott. Davide Brunato

Introduzione

Come richiesto, in previsione della disponibilità dei nuovi fondi per i progetti di calcolo approvati, è stata verificata la situazione della Computer Room della Scuola e analizzata la possibile evoluzione per il prossimo periodo. Le informazioni sono state raccolte con verifiche in Computer Room e mediante colloqui con il personale tecnico del SIS, di SISSA Medialab e Democritos. L'analisi è suddivisa secondo l'utilizzo dei sistemi, con un conteggio delle apparecchiature operanti nel locale e dello spazio occupato attualmente. Si conclude con un'analisi della situazione e proposte operative.

Servizi Centrali e Amministrativi

Vi sono attualmente 45 server (tipologie miste, prevalentemente sistemi *rack* biprocessore) e due sistemi di storage (EMC CX-700 e Compaq MA8000). E' pianificato l'accorpamento dei fileserver di settore a quello centrale, con una conseguente riduzione del numero totale dei server. I futuri ex-server di settore, dei Compaq/HP Proliant G2/G3, verranno riutilizzati per il rinnovo di servizi ormai datati o mancanti (DNS principale, spooler, server updates), per ridondare servizi importanti (webmail) o per implementazioni di test (VOIP/Asterisk, XEN). Lo spostamento dei servizi sui Proliant porterà alla dismissione dei vecchi server in formato tower/miditower, permettendo un miglior utilizzo dello spazio in sala macchine e una migliore affidabilità dei servizi.

Saranno aggiunti i nuovi server arrivati per il multimediale (3 server *rack* biprocessore) e i 2 server SUN V40z quadriprocessore, che sostituiranno le Alpha 4000 nel servizio di fileserver centralizzato. E' prevista anche la sostituzione di alcuni server per l'Amministrazione (3-4 server nuovi).

Per quanto riguarda i sistemi di storage è già stato ordinato un incremento dello spazio disco per il CX-700, con l'aggiunta di due cassette e 15 dischi da 300GB. E' possibile che il vecchio sistema di storage Compaq MA8000 venga riutilizzato per creare un'area *scratch* in rete.

Nel totale il numero di server per i servizi centrali e amministrativi dovrebbe ridursi entro il prossimo anno a 34-35, fatto salvo alcune aggiunte provenienti dalla sala dati del Bellavista. Lo spazio occupato dai sistemi verrà razionalizzato passando da 7 a 5 *rack*, recuperando inoltre la superficie occupata attualmente dalle due Alpha 4000.

Servizi di Calcolo (HPC)

Attualmente vi sono circa 110 server di calcolo:

- ***cerbero*** 72 server biprocessore Opteron
- ***helium*** 8 server Xeon biprocessore
- ***briareo*** 29 nodi biprocessore Pentium III

si prevede un aumento a 145 nodi di calcolo, con l'aggiunta un nuovo cluster di calcolo, comprensivo di nodo di I/O e relativo storage. Il cluster "briareo" dovrebbe

essere spostato a Basovizza. Lo spazio occupato dai nodi di calcolo è di 4 rack e dovrebbe passare a 6 rack con le nuove realizzazioni.

SISSA Medialab

Attualmente hanno 16 server operativi e 2 da dismettere/sostituire. Non hanno ancora deciso cosa sostituiranno, ma in ogni caso il numero di server dovrebbe variare di poco (due server al massimo).

Le loro macchine occupano un rack e circa 2 mq di spazio. Si è chiesto di eliminare un vecchio server con venti dischi IDE, peraltro guasto, per diminuire il consumo e liberare spazio.

Altri Server

Vi sono altri 4 server in formato tower:

- PC biprocessore P4 *run.sissa.it* (Settore di CM, utilizzato per il calcolo);
- server Proliant G2 del gruppo Democritos;
- PC *eschilo.sissa.it* (Percacci, per misurazioni sulle reti dati);
- Server *tt71.sissa.it* (Settore di Neurobiologia, Prof. Torre, DB per progetto di ricerca)

Il server di Democritos rimarrà nei rack dei servizi con gli altri Proliant, mentre le altre macchine verranno spostate piccolo rack 60x60, accanto al rack di Medialab, per ridurre lo spazio occupato e far posto alle nuove macchine di calcolo.

Analisi situazione

Si stima che nei prossimi mesi in Computer Room si passerà da 177 a 203 sistemi. Il consumo attuale della parte in continuità della computer room si attesta intorno ai 40 KW, con equivalente produzione di calore.

Con l'inserimento dei nuovi nodi il consumo della parte calcolo dovrebbe aumentare perlomeno di 15-17 KW (nell'ipotesi migliore, ovvero nodi con processori Opteron), per cui è necessario pensare ad un incremento della potenza refrigerante fornita dall'impianto di condizionamento della Computer Room. E' necessario inserire un nuovo condizionatore che copra l'aumento di potenza previsto. Si auspica che venga richiesto anche un secondo condizionatore di ridondanza, per evitare che un guasto ad una delle apparecchiature possa portare ad una situazione critica, come già successo in passato. Vista l'affidabilità, la qualità ed il supporto tecnico fornito si chiede espressamente che tali apparecchiature siano di marca Emerson Network Power, come peraltro lo è già l'impianto attuale.

Sarà necessario inserire ulteriori punti di distribuzione elettrica, per alimentare il nuovo cluster di calcolo. Vi è però il problema della disponibilità della continuità assoluta. Attualmente le due macchine UPS sopportano un carico pari al 60% della loro potenza. Con questa situazione il singolo UPS non è in grado di supportare l'intero carico in continuità: in caso di rottura di uno dei due UPS, il gruppo

rimanente commuterebbe in by-pass su rete elettrica, lasciando la computer room senza continuità.

La soluzione alternativa di spostare parte del carico su linee prive di continuità assoluta non è praticabile, se non in minima parte, e non risolverebbe l'aggravio del carico sugli UPS. Infatti il numero di server dotati di doppio alimentatore switched è molto ridotto (23) e quasi tutti sono sistemi dei servizi centrali, diversi di questi essenziali.

Inoltre non si ritiene plausibile alimentare i server, neppure nodi di calcolo diskless, completamente sotto la non continuità, visti i continui sbalzi di tensione che si verificano, soprattutto nei mesi estivi. Nel caso di nodi di calcolo i danni ammonterebbero alle ore macchina spese per elaborare i job. Visti il costi di questi sistemi, anche senza contare i costi di gestione e di funzionamento e il danno per la ricerca arrecato all'utenza, si stima che non sia economicamente opportuno lasciare tali server privi della continuità assoluta.

Sebbene quindi sia pensabile inserire punti elettrici in computer room privi della continuità assoluta, allo scopo di distribuire su questa il distribuibile, ma il problema reale sarebbe quello di avere una maggiore affidabilità dell'impianto UPS esistente. Questo può essere fatto in due modi:

1. Aggiungendo un altro UPS da 40 KW, in parallelo a quelli esistenti
2. Con un contratto di manutenzione sugli UPS che preveda tempi di intervento prestabiliti

A sostegno della seconda ipotesi si sottolinea che il produttore dei gruppi (ditta SIEL) prevede contratti di assistenza con intervento in 4 ore dalla chiamata, con copertura sui giorni lavorativi. Sono previsti anche contratti a copertura completa di 365 giorni all'anno. Chiaramente un contratto di manutenzione per questo tipo di apparati ha dei costi elevati, quindi sarà da valutare, nella prospettiva Santorio, qual è la soluzione meno costosa nell'ottica dei tempi previsti per lo spostamento presso la nuova sede.

Trieste, 3 Agosto 2006

Davide Brunato

Coordinatore dei Sistemi Centrali e Reti della SISSA