



ISTITUTO NAZIONALE DI FISICA NUCLEARE

Sezione di Padova

INFN/CCR-07/10
November 30, 2007



CCR/17/07/P
Aprile 2007
Versione 1.0

CONFRONTO DI PRESTAZIONI DI APPLICAZIONI HEP CON BENCHMARK SINTETICI

Michele Michelotto¹⁾

¹⁾*INFN-Sezione di Padova, Via F. Marzolo, 8, I-35131 Padova, Italy*

Abstract

I siti candidati ad ospitare i centri di calcolo Tier2 hanno comprato a fine 2006 diversi nodi di calcolo. Tipicamente macchine 1U con doppio processore dual core e 4 GB di memoria. Il costo medio, Iva Compresa, si aggira attorno ai 0.5 Euro/SI2000.

PACS.: 89.20-Ff (Computer Science and Technology)

*Published by **SIS-Pubblicazioni**
Laboratori Nazionali di Frascati*

1 INTRODUZIONE

Questo documento riassume brevemente le tendenze della tecnologia dei nodi di calcolo usati tipicamente negli esperimenti di fisica delle alte energie nella farm di simulazione e ricostruzione dei Tier2/Tier3.

Verranno descritti i processori del momento e le prospettive per i prossimi 12 mesi e i prezzi per i nodi di calcolo ottenuti da alcune sedi INFN a fine 2006.

2 UNITÀ DI MISURA

Dal momento che le richieste degli esperimenti sono in termini di Specint 2000, cioè legati ai valori di CPU Int 2000 della suite SPEC, ho valutato tutte le macchine in base ai valori dichiarati dal sito spec.org per questi processori e moltiplicato gli Specint di un singolo core per il numero dei core. In realtà misurando con la suite SPEC i valori ottenuti sono molto più bassi, sia per l'uso di un diverso compilatore (gcc è il compilatore usato dagli esperimenti) sia per l'uso di ottimizzazioni molto penalizzanti (per esempio il CERN effettua i tender di acquisto misurando gli spec compilato con ottimizzazione *gcc -O2 -fPIC -pthread*).

Per il prezzo invece considero il prezzo IVA Compresa in Euro.

3 IL NODO DI CALCOLO

Il nodo di calcolo più conveniente è il box 1U biprocessore con un alimentatore e un piccolo disco ATA. Alcuni siti sono riusciti ad ottenere buoni prezzi anche per configurazioni blade che in teoria dovrebbero permettere risparmi infrastrutturali, in termini di minori consumi, migliore gestione grazie alla condivisione degli alimentatori, apparati KVM, switch integrati etc.

4 I PROCESSORI

I processori interessati dall'indagine sono tutti dual core, permettendo quindi di avere quattro core in ogni macchina biprocessore. I produttori sono AMD e Intel.

AMD aveva la leadership tecnologica ad inizio 2006 con i processori della serie 2xx (260, 265, 270, 280) costruiti in tecnologia 90 nm con prestazioni doppie dei single core e consumi identici, quindi raddoppiando gli Specint/Watt.

A metà anno però Intel ha reso disponibile il primo processore veramente dual core (Woodcrest, con nome in codi 51x0, es 5110 o 5160) in tecnologia 65 nm con clock fino a 3.0 GHz e consumi paragonabili a quelli AMD.

La risposta di AMD è stata la serie di processori 22xx, in pratica un restyling della serie precedente con un leggero aumento di prestazioni ma senza passare alla tecnologia 65nm.

A fine anno Intel ha prodotto i primi processori quasi quad-core. Si tratta infatti di due chip dual core affiancati montati nello stesso processore. Il primo vero quad-core dovrebbe essere il Barcelona K10 di AMD o una nuova versione veramente quad-core, cioè con un singolo chip con i quattro core. Questi chip sono attesi in produzione per la fine del 2007.

A fine 2007, inizio 2008 sono previsti anche le versioni a 45 nm di questi chip, questo

dovrebbe portare ad un aumento di clock a parità di consumo oppure ad una diminuzione di consumi e quindi maggiori densità di integrazione (macchine a 4 vie, macchine con due board dual processor nello stesso case).

5 MEMORIA

Tutte le macchine analizzate con quattro core avevano anche quattro Gigabyte di memoria. La tendenza è quindi quella di avere un Gigabyte per ogni core. Tuttavia l'esperimento Alice ha la particolare necessità di avere 2 GB per core. Ovviamente questo implica un costo per SI2k maggiore dal momento che gli SI2K sono basati in pratica unicamente sul processore.

Il passaggio del codice degli esperimenti da 32bit a 64 bit dovrebbe permettere un aumento di prestazioni tra il 15% e il 40% ma anche un aumento delle necessità di memoria che non riesco a quantificare (tra 10% e 100%).

6 INDAGINE

Ho chiesto a tutte le sedi Tier2 e a Babar i costi e le prestazioni delle macchine relative agli ultimi acquisti. Ho avuto risposte da LNL, Catania, Roma1, Torino, Pisa, Napoli e Babar-Padova.. Vi allego la tabella:

data	cpu	formato	clock	core	SI2K/core	mem	€/box	specint	€/€/Si2k
Nov-06	dual 2214	1U	2200	4	1352	4GB	2478	5408	0.458
Nov-06	dual 2214	1U	2200	4	1352	8GB	3228	5408	0.597
Sep-06	dual 280	1U	2400	4	1492	4GB	3270	5968	0.548
Dec-06	dual 280	1U	2400	4	1492	4GB	2916	5968	0.489
Nov-06	dual 5160 xeon 32	1U	3000	4	2900	4GB	3960	11600	0.341
May-06	bit	blade	2000	4	1663	4GB	4121	6652	0.620
	dual 5110	blade	1600	4	1719	4GB	3165	6876	0.460
Feb-07	dual 2216	blade	2400	4	1474	4GB	3193	5896	0.542
Nov-06	dual 2218	blade	2600	4	1648	4GB	3636	6592	0.552

average	0.512
st dev	0.085

Le prime macchine sono di tipo 1U mentre le ultime 3 sono di tipo blade. Notate per le prime due colonne la differenza di prezzo tra macchine con 4GB e con 8GB.

I blade di xeon hanno dei processori a 32 bit e consumano molto poco mentre tutte le altre sono architetture a 64 bit.

L'ultima colonna riporta i prezzi per Si2K (quello derivato dal sito spec.org). Sotto sono riportati la media e la deviazione standard.

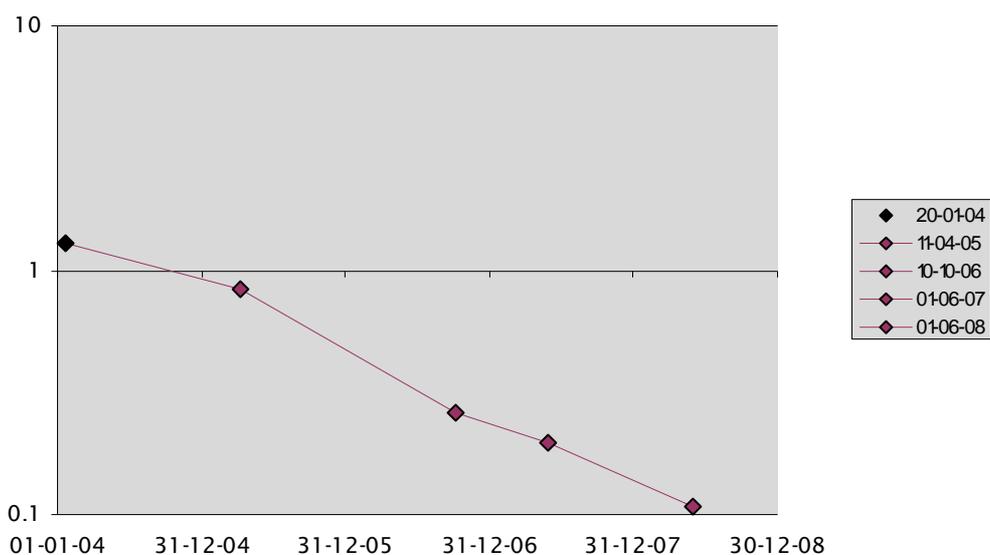
Per confronto vediamo i risultati delle gare presso il Tier1 negli ultimi anni. Vediamo che i prezzi sono decisamente più bassi di quelli che riescono ad ottenere i Tier2, dati i volumi

interessati.

Acquisto	Processore	SI2000	Euro/SI2K
	Xeon 2 x 2,4 GHz 2 GB	1600	1.312
Jan-04	2 x Xeon 3 GHz 4 GB	2000	1.295
Apr-05	2xOpteron 252, 4 GB	3138	0.841
Oct-06	2x5140, 6 GB	9436	0.262
Jun-07	2x 2core o 4core		0.195
Jun-08	2 x 4core -16GB		0.108

Le ultime due righe in grassetto, sono ottenute estrapolando le ultime tre gare, ipotizzando che continui l'andamento esponenziale in figura:

Tier1



7 FUTURO

In questo momento i prezzi dei processori soprattutto quelli con clock più elevato sono scesi per lasciare i prezzi più alti ai processori quad-core. Anche i prezzi delle memoria sono scesi. Posso stimare prezzi tra il 10% e il 30% inferiori sulle stesse macchine durante il 2007 con il solito trend in discesa. Se si dovessero raddoppiare le memorie invece i prezzi per Specint rimarrebbero gli stessi.

Se dovessero diffondersi le macchine con quattro core le previsioni sono difficili. La macchina con 8 core avrebbe bisogno di 8 o 16 GB di memoria, processori più costosi ma non

il doppio più costosi, mentre il resto dei componenti avrebbe circa lo stesso costo. Le prestazioni dovrebbero raddoppiare a parte una leggera diminuzione dovuta al clock inferiore. Se dovessi stimare un numero direi 3.5 Euro/Si2k.

8 SPECINT PUBBLICATI, SPECINT MISURATI E EVENTI/SECONDO

Vorrei finire questo documento con alcune considerazioni sull'uso degli Specint misurati.

Come ho già detto gli Specint pubblicati sono ottenuti solitamente con compilatori Intel o Microsoft con sistemi operativi spesso MS Windows.

Misurando il codice prodotto con il compilatore gcc le prestazioni scendono notevolmente. Imponendo poi le ottimizzazioni usate dagli esperimenti CERN (-O2 per non spingere, -fPIC per generare codice Position Independent per lineare con le shared librerie, -pthread per generare codice multithreaded) le prestazioni scendono ulteriormente.

		spec	gcc -O3	gcc CERN	drop	drop CERN
Intel	PIII	1222	834	701	68%	57%
Intel	Xeon 2.2	1554	1062	881	68%	57%
Intel	Xeon 2.4	1826	1193	982	65%	54%
Intel	Xeon 2.6	1976	1264	1055	64%	53%
	Xeon					
Intel	5148LV	9060	5818	4770	64%	53%
Intel	Xeon 5150	11044	6732	5032	61%	46%
Intel	Xeon 5160	12244		5564		45%
AMD	AMD 275	5556	4464	2998	80%	54%
AMD	AMD 270	5128	4120	2672	80%	52%

Nella tabella qui sopra vedete per alcuni processori il calo di prestazioni passando a gcc. Si ottengono solo il 61-68% delle prestazioni pubblicate per processori Intel e l'80% per AMD. Usando il "gcc CERN" c'è un calo ulteriore.

Per questo motivo ho cercato di misurare il rapporto tra produttività con codice di HEP (es. eventi/secondo) e Specint di questi processori.

9 ROOTMARK

Per esempio nella prossima tabella potete vedere il rapporto tra Rootmark e SI2K pubblicati per alcuni processori. Si nota l'aumento di prestazioni passando da codice a 64 bit a codice 32 bit, soprattutto per i processori AMD. Secondo questa tabella con codice a 32 bit i processori AMD a parità di Specint hanno un resa del 22% più alta e addirittura del 40% con codice a 64 bit. Tuttavia in valore assoluto la macchina più performante è di gran lunga quella con processore Intel.

La parte alta delle tabelle riporta i valori assoluti, la parte bassa in rapporto al Woodcrest 5160 preso come processore di riferimento.

	clock GHz			SI2K base	rootmarks 64/64	rootmark 32/64	64 bit 32 bit ratio rootmark	rootmarks 64/64 clock ratio	rootmark 32/64 clock ratio	rootmarks 64/64 SI2K ratio	rootmark 32/64 SI2K ratio
Intel 5160	3.00	3G	4 cores	3061	2005.80	1584.60	127%	668.6000	528.2000	655.276	517.674
AMD 265	1.80	4G	4 cores	1168	1083.40	743.30	146%	601.8889	412.9444	927.568	636.387
AMD 248	2.20	2G	2 cores	1411	1273.70	884.60	144%	578.9545	402.0909	902.693	626.931
AMD 248	2.20	8G	2 cores	1411	1299.50	884.20	147%	590.6818	401.9091	920.978	626.648
AMD 2218	2.60	8G	4 cores	1648	1546.40	1079.60	143%	594.7692	415.2308	938.350	655.097
Intel 5160	3.00	3G	4 cores	100.00%	100.00%	100.00%	100%	100.00%	100.00%	100.00%	100.00%
AMD 265	1.80	4G	4 cores	38.16%	54.01%	46.91%	115%	90.02%	78.18%	141.55%	122.93%
AMD 248	2.20	2G	2 cores	46.10%	63.50%	55.82%	114%	86.59%	76.12%	137.76%	121.11%
AMD 248	2.20	8G	2 cores	46.10%	64.79%	55.80%	116%	88.35%	76.09%	140.55%	121.05%
AMD 2218	2.60	8G	4 cores	53.84%	77.10%	68.13%	113%	88.96%	78.61%	143.20%	126.55%

10 PYTHIA

Di nuovo la stessa tabella usando Pythia per generare 300 eventi Susy.

	clock GHz			SI2K base	pythia 64/64 evt/sec	pythia 32/64 evt/sec	64 bit 32 bit ratio pythia	pythia 64/64 clock ratio	pythia 32/64 clock ratio	pythia 64/64 SI2K ratio	pythia 32/64 SI2K ratio
Intel 5160	3.00	3G	4 cores	3061	189.15	152.86	124%	63.0500	50.9533	61.794	49.938
AMD 265	1.80	4G	4 cores	1168	110.78	92.40	120%	61.5444	51.3333	94.846	79.110
AMD 248	2.20	2G	2 cores	1411	135.29	116.70	116%	61.4955	53.0455	95.882	82.707
AMD 248	2.20	8G	2 cores	1411	135.48	116.69	116%	61.5818	53.0409	96.017	82.700
AMD 2218	2.60	8G	4 cores	1648	162.00	134.41	121%	62.3077	51.6962	98.301	81.559
Intel 5160	3.00	3G	4 cores	100.00%	100.00%	100.00%	100%	100.00%	100.00%	100.00%	100.00%
AMD 265	1.80	4G	4 cores	38.16%	58.57%	60.45%	97%	97.61%	100.75%	153.49%	158.42%
AMD 248	2.20	2G	2 cores	46.10%	71.53%	76.34%	94%	97.53%	104.11%	155.17%	165.62%
AMD 248	2.20	8G	2 cores	46.10%	71.63%	76.34%	94%	97.67%	104.10%	155.38%	165.61%
AMD 2218	2.60	8G	4 cores	53.84%	85.65%	87.93%	97%	98.82%	101.46%	159.08%	163.32%

In questo caso di nuovo il processore più potente è il 5160 di intel ma il processore AMD produce più il 50% di eventi in più per ogni Specint.

11 CMS OSCAR

Il codice di CMS c'è solo a 32 bit. Qui ho usato OSCAR per produrre gli eventi e poi ne ho effettuato la fase di Digits e la ricostruzione:

	clock GHz			SI2K base	CMS DST evt/sec	CMS DIGIS evt/sec	CMS OSCAR evt/sec	CMS DST clock ratio	CMS DIGIS clock ratio	CMS OSCAR clock ratio	CMS DST SI2K ratio	CMS DIGIS SI2K ratio	CMS OSCAR SI2K ratio
Intel 5160	3.00	3G	4 cores	3061	0.5997	0.0703	0.1713	0.1999	0.0234	0.0571	0.1959	0.0230	0.0560
AMD 265	1.80	4G	4 cores	1168	0.2850	0.0340	0.0909	0.1583	0.0189	0.0505	0.2440	0.0291	0.0778
AMD 248	2.20	2G	2 cores	1411	0.3829	0.0503	0.1083	0.1740	0.0229	0.0492	0.2714	0.0356	0.0768
AMD 248	2.20	8G	2 cores	1411	0.3865	0.0503	0.1084	0.1757	0.0229	0.0493	0.2739	0.0356	0.0768
AMD 2218	2.60	8G	4 cores	1648	0.4258	0.0486	0.1311	0.1638	0.0187	0.0504	0.2584	0.0295	0.0796
Intel 5160	3.00	3G	4 cores	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
AMD 265	1.80	4G	4 cores	38.16%	47.52%	48.36%	53.06%	79.21%	80.61%	88.44%	124.55%	126.75%	139.07%
AMD 248	2.20	2G	2 cores	46.10%	63.85%	71.55%	63.22%	87.07%	97.57%	86.21%	138.51%	155.22%	137.15%
AMD 248	2.20	8G	2 cores	46.10%	64.45%	71.55%	63.28%	87.88%	97.57%	86.29%	139.81%	155.22%	137.28%
AMD 2218	2.60	8G	4 cores	53.84%	71.00%	69.13%	76.53%	81.93%	79.77%	88.31%	131.88%	128.41%	142.15%

12 CONCLUSIONI

Il panorama dei processori per worker node nel 2007 è molto promettente grazie ai nuovi processori quad-core, alla piena transizione alla tecnologia a 65nm e all'apparizione della tecnologia a 45 nm. Tuttavia rimane difficile fare previsioni in termini di prezzo prestazione dal momento che le macchine con quad core non sono ancora sul mercato, sebbene siano nei listini. Dai prezzi ottenuti a fine 2006 di 0.5 Euro/Specint2000 si può prevedere per il 2007 prezzi attorno a 0.3 – 0.4 Euro/Si2K ma bisogna fare molta attenzione alle prestazioni reali delle macchine a parità di specint teorici.

Nel Tier1 si ottengono solitamente rapporto costo/prestazioni migliori grazie agli elevati volumi di macchine acquistati che potrebbero permettere di arrivare ai 0.2 Euro/Si2K. E' probabile nel corso dell'anno uno spostamento dall'uso dei Spec CPU 2000 Int ad altri benchmark della famiglia CPU 20.