On the accuracy of the CELL processor

Rene Alt*, Jean-Luc Lamotte*, Svetoslav Markov**
*Laboratoire Informatique Paris 6
University Pierre et Marie Curie 4 place Jussieu, 75252 Paris cedex 05, France
E-mail : Rene.Alt@lip6.fr
**Institute of Mathematics and Informatics
Bulgarian Academy of Sciences
"G. Bonchev" st., bl. 8, 1113 Sofia, Bulgaria
E-mail : smarkov@bio.bas.bg

SCAN 2008, El Paso, September 29th

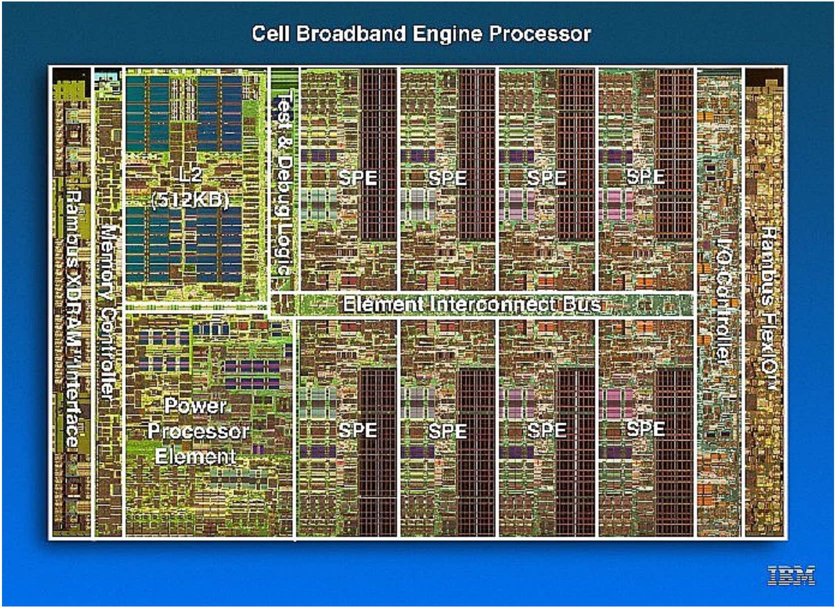
Contents

- Introduction : The CELL processor
- Theory about the effect of rounding modes
- Numerical experiments on a PC
  - Inner products
  - Linear systems
- Numerical experiments on the CELL
  - Inner products
  - Linear systems
- Some effects of the imprecise division of the CELL SPU
- Conclusion: is the CELL well suited for scient. computation ?

The CELL Processor

- Constructed by IBM, Toshiba, Sony for Playstation 3
- Used to build super computers: IBM Roadrunner with 12,960 modified Cell processors and 6948 AMD Opterons has broken the computing record in june 2008 (more than 1 petaflops).
- 64-bit Power processor core
- 8 Synergistic Processing Units (SPU)with Power-based core. top clock speeds exceeds 4 GHz
- 2.5MB of on Chip memory (512KB L2 and 8 * 256KB)
- 234 million transistors, size of 221mm2, 90 nanometer SOI process technology
- SPU: RISC architecture, SIMD organization, local store
- **But: Rounding to zero arithmetic, 12 correct bit division**

The CELL Processor /2



Theory about rounding modes in inner products

- General formula (Alt 1978)
  - $p = \sum_{i=1}^{i=n} x_i y_i \qquad x_i, y_i, p \in \mathbb{R}$
  - $P = \sum_{i=1}^{i=n} X_i Y_i \qquad X_i, Y_i, P \in \mathbb{F}(Float.\ point)$
  - $X_i = x_i(1+\lambda_i),\ \ Y_i = y_i(1+\mu_i)$
- Rounding errors on all F.P. operations
- Set $\rho = P - p$ and $s^2 = \sum_{i=1}^{n} x_i^2 y_i^2$
- $\overline{\alpha}$ and δ: resp mean value and stand dev. of the population of relative errors of the computer.
- $\overline{\rho} = \overline{\alpha}\ p\ \frac{n^2+7n-2}{2n}$
- $\overline{\rho^2} = \overline{\alpha^2}\ \left(\frac{3n^3+41n^2+134n-72}{12n} p^2 + \frac{(n-2)(n^2+3n-6)}{12n} s^2\right)$
  $+\delta^2\ \left(\frac{n+1}{3} p^2 + \frac{n^2+19n-6}{6n} s^2\right)$

Theory about rounding modes in inner products(2)

- Rounding to nearest (S.P.)

$$\overline{\alpha} = 0, \quad \delta = 0.408\ 2^{-t} \Longrightarrow \begin{cases} \overline{\rho} = 0 \\ \boxed{\sqrt{\overline{\rho^2}} \simeq \dfrac{\delta\ s}{\sqrt{6}}\ \sqrt{n}} \end{cases}$$

- Rounding to zero

$$\overline{\alpha} = -0.693\ 2^{-t}, \quad \delta = 0.431\ 2^{-t} \Longrightarrow \begin{cases} \boxed{\overline{\rho} \simeq \overline{\alpha}\ p\ \dfrac{n}{2}} \\ \overline{\rho^2} \simeq \overline{\alpha^2}\ n^2(\frac{3p^2+s^2}{12}) \end{cases}$$

- t : nb of bits of the mantissa (24 or 53)
- Hyp : The relative rounding errors are equidistributed and all errors are independant
- Theory very well verified by experimentation

Theory about imprecise data

(Markov, Lamotte, Alt, 2004)

- Each data X is represented by a gaussian funct. with known mean val. and std. dev. : $X = (m, \sigma)$. Such a number is called a stochastic number. (Vignes, Chesneaux)
- Operations on stoch. numb. are those on Gaussian functions: $\mathrm{s}+, \mathrm{s}-, \mathrm{s}*, \mathrm{s}/$
- Addition and multiplication by scalars :

$$X_1 \;{}_{\mathrm{s}}{+} X_2 \overset{def}{=} \left(m_1 + m_2;\ \sqrt{\sigma_1^2 + \sigma_2^2}\right),$$

$$\gamma * X \overset{def}{=} \left(\gamma m;\ |\gamma|\sigma\right),\ \gamma \in \mathbb{R}.$$

- Monoid w r t addition $\longrightarrow$ extended to group structure
- Multiplication by scalar $\longrightarrow$ s-space similar to vect. sp.

Theory about imprecise data /2

- From the formula on the sum :
  $X_1 \ {}_{\mathrm{s}}\!+ X_2 \ {}_{\mathrm{s}}\!+ \ldots {}_{\mathrm{s}}\!+ X_n = \left(\sum_{i=1}^{n} x_i, \sqrt{\sigma_1^2 + \sigma_2^2 + \ldots + \sigma_n^2} \right)$
  replacing all σ_i by $\overline{\sigma}$
  Error on the sum $\simeq \overline{\sigma}\sqrt{n}$
- Also proved for inner products of real $\times$ stochastic
- Experimentally true but not proved for inner products of stochastic $\times$ stochastic
- Remark: The CESTAC method and the CADNA software takes into account both data errors and round-off errors.

Inner products, rounding to nearest

- Compute : $\sum_{i=1}^{n} x_i y_i$ in SP and DP with rounding to nearest. x_i, y_i randomly generated between a and b.
- Theory: relative error in $O(\sqrt{n})$
- Compute $\beta = 2^{24}|relat.error|/\sqrt{n}$
- Results with $a = -100$, $b = 100$.

n	$Result$	$relat.error$	$ndig.$	β
10	$0.50807393E+04$	$0.53093494E-07$	7.00	0.282
100	$0.21252789E+05$	$0.33788535E-07$	7.00	0.057
1000	$0.11471347E+06$	$0.19676749E-06$	6.71	0.104
10000	$-0.12076502E+06$	$0.10804285E-05$	5.97	0.181
100000	$-0.61547312E+06$	$0.59440595E-05$	5.23	0.315

Relative error in $O(\sqrt{n})$

Inner products, rounding to zero

- Compute : $\sum_{i=1}^{n} x_i y_i$ in SP and DP with rounding to zero. x_i, y_i randomly generated between a and b.
- Theory: relative error in $O(n)$
- Compute $\beta = 2^{24}|relat.error|/n$
- Results with $a = -100$, $b = 100$.

n	$Result$	$relat.\ error$	$ndig.$	β
10	$0.50807378E+04$	$0.23521963E-06$	6.63	0.395
100	$0.21252777E+05$	$0.58518668E-06$	6.23	0.098
1000	$0.11471073E+06$	$0.23639801E-04$	4.63	0.397
10000	$-0.12072600E+06$	$0.32415066E-03$	3.49	0.544
100000	$-0.61404406E+06$	$0.23159622E-02$	2.64	0.389

Relative error in $O(n)$

Inner products, uncertain data

- Computation of sums of n positive numbers. Uncertain data in a known range.
- Several samples (30) generated in the range of uncertainty.
- Mean val. and std. dev. on the result computed from the samples.
- Theory: The standard deviation increases as $\sqrt{n}$ (Alt Lamotte Markov 2006)
- Example, relat err =0.001, numbers generated in $[0, 100]$

n	$Result$	$relat.\ error$	$ndig.$	$Std.dev/\sqrt{n}$
10	$0.56501328D+03$	$0.186E-04$	4.73	$0.455E-01$
100	$0.49115716D+04$	$0.184E-04$	4.74	$0.500E-01$
1000	$0.51429866D+05$	$0.687E-05$	5.16	$0.472E-01$
10000	$0.49883813D+06$	$0.147E-05$	5.83	$0.541E-01$
100000	$0.49863373D+07$	$0.437E-06$	6.36	$0.532E-01$

Relative error in $O(\sqrt{n})$

Inner products, CELL processor

- Compute : $\sum_{i=1}^{n} x_i y_i$ in SP and DP on the CELL.
  x_i, y_i randomly generated between two bounds a and b, Gaussian distribution.
- Compute $\beta = 2^{24} |relat.error|/n$
- Results with $a = 0$, $b = 10$.

n	$Result$	$relat.\ error$	$ndig.$	β
4	$+5.835503e+01$	$+1.29e-07$	6.89	0.539
16	$+3.127193e+02$	$+3.62e-07$	6.44	0.379
64	$+1.624892e+03$	$+1.34e-06$	5.87	0.352
256	$+5.865698e+03$	$+5.70e-06$	5.24	0.373
1024	$+2.485485e+04$	$+2.28e-05$	4.64	0.374
4096	$+9.967466e+04$	$+8.91e-05$	4.05	0.365
8192	$+1.994063e+05$	$+1.82e-04$	3.74	0.372

Relative error in $O(n)$

Inner products, CELL processor /2

- Number of decimal significant digits using:
  - Standard algorithm
  - Special function sdot
  - Special operation FMA (Fused Multiply and Add).

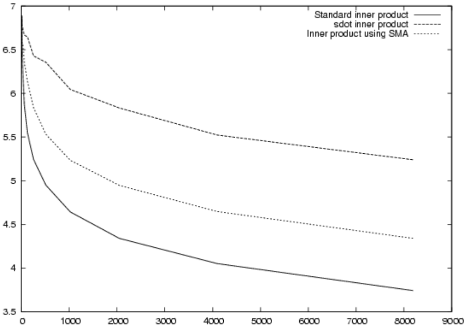


Figure: Nb significant digits on the CELL

Inner products, CELL processor /3

- Computation of $\beta = 2^{24} * |relat\_error|/n$ using:
  - Standard algorithm
  - Special function sdot
  - Special operation FMA

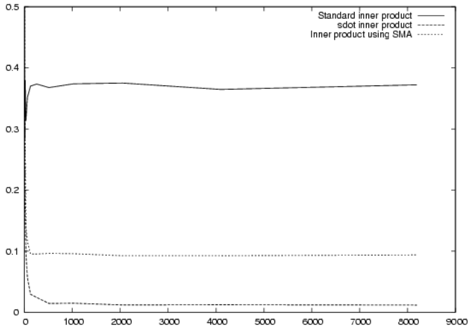


Figure: Sum of numbers, coeff β on the CELL

Linear system, rnd to nearest and to zero

- Solve : $AX = B$ (Std gauss. elimination) with rnd to nearest and rnd to zero arithmetic and on the CELL

$$\text{with } A = \begin{pmatrix} 5 & 7 & 6 & 5 \\ 7 & 10 & 8 & 7 \\ 6 & 8 & 10 & 9 \\ 5 & 7 & 9 & 10 \end{pmatrix} \text{ and } B = \begin{pmatrix} 23 \\ 32 \\ 33 \\ 31 \end{pmatrix} \text{ (Wilson)}$$

Computed solution

Exact	rnd to nearest	rnd to zero	*CELL*
1	$0.99999970E+00$	$0.10000181E+01$	$1.0000135e+00$
1	$0.10000004E+01$	$0.99998879E+00$	$9.9999154e-01$
1	$0.99999982E+00$	$0.99999547E+00$	$9.9999678e-01$
1	$0.10000001E+01$	$0.10000027E+01$	$1.0000020e+00$

Rounding to nearest more accurate)

Linear system on the CELL

- Solve the same system on the CELL SPU with direct standard Gaussian elimination and CELL SIMD intructions.

Computed solution

Exact	Standard Gauss	CELL SIMD instructions
1	$1.0000135e+00$	$1.0407186e+00$
1	$9.9999154e-01$	$9.7517973e-01$
1	$9.9999678e-01$	$9.9008989e-01$
1	$1.0000020e+00$	$1.0059216e+00$

Pb due to division, ex:The J.-M. Muller sequence

$$U_n = 111 - \frac{1130}{U_{n-1}} + \frac{3000}{U_{n-1}.U_{n-2}}$$

$$U_0 = 5.5 \quad U_1 = \frac{61}{11}$$

	Std Division	SMID instructions
$\frac{61}{11}$	5.5454545e+00	5.5451584e+00
3	+5.5901684e+00	+5.5858231e+00
4	+5.6335177e+00	+5.5364304e+00
5	+5.6761970e+00	+3.8882599e+00
6	+5.7406115e+00	-4.0252731e+01
7	+6.2242178e+00	+1.1990636e+02
8	+1.3412224e+01	+1.0095535e+02
9	+6.2684993e+01	+1.0005419e+02

Conclusion

Is the CELL optimal for scientific computation ?

- Very fast processor allowing very fast super computers
- Not designed for accurate floating point computation
- Rounding to zero mode and division replaced by imprecise multiplication by inverse.
- Error on the computation of inner products proportional to n. BLAS more efficient but still with an error in $O(n)$.
- $\Longrightarrow$ Be careful to the problem, chose a nice algorithm and control the solution.
- In the tested version double precision arithmetic was a 15 Gigaflops one. Next version of CELL sould be with 100Gflops double precision.

References

• IBM: The Cell project at IBM Research, Michael Gschwind, www.research.ibm.com/cell/

• Alt, R., Error Propagation in Fourier Transforms., *Math. Comp. in Sim*, 20, 1978 , 37-43.

• Markov, S., Alt, R., Lamotte, J.L., Stochastic Arithmetic: S-spaces and Some Applications, *Numer. Algorithms* 37 (1-4), 2004, 275-284.

• Alt, R., Lamotte, J.L., Markov, S., Abstract structures in stochastic arithmetic, Proc. of (IPMU 2006), Paris, France, July 2–7, 2006, 794-801.

• J.-M. Chesneaux and J. Vignes, Les fondements de l'arithmtique stochastique, C.R. Acad. Sci., Paris, sr.1, 315, 1992, 1435-1440.

• Vignes, J., A stochastic arithmetic for reliable scientific computation, *Math. and Comp. in Sim.* 35, 1993, 233-261.

• Vignes, J., Discrete stochastic arithmetic for validating results of numerical software, *Numer. Algo*, 37, 2004, 377-390.