Quantum Monte Carlo simulations in chemistry at the petascale level and beyond

A. Scemama[1], M. Caffarel[1], E. Oseret[2], W. Jalby[2]

[1]Laboratoire de Chimie et Physique Quantiques / IRSAMC, Toulouse, France
[2]Exascale Computing Research / Intel, CEA, GENCI, UVSQ Versailles, France

28 June 2012


Quantum Monte Carlo methods

- Solve the Schrödinger equation with random walks
- State-of-the-art and routine approaches in physics : nuclear physics, condensed-matter, spin systems, quantum liquids, infrared spectroscopy . . .
- Still of confidential use for the electronic structure problem of quantum chemistry (as opposed to post-HF and DFT)
- Reason : Very high computational cost for small/medium systems

But :

- Very favorable scaling with system size compared to standard methods
- Ideally suited to extreme parallelism

Quantum Monte Carlo methods

- Solve the Schrödinger equation with random walks
- State-of-the-art and routine approaches in physics : nuclear physics, condensed-matter, spin systems, quantum liquids, infrared spectroscopy ...
- Still of confidential use for the electronic structure problem of quantum chemistry (as opposed to post-HF and DFT)
- Reason : Very high computational cost for small/medium systems

But :

- Very favorable scaling with system size compared to standard methods
- Ideally suited to extreme parallelism

Quantum Monte Carlo methods

- Solve the Schrödinger equation with random walks
- State-of-the-art and routine approaches in physics : nuclear physics, condensed-matter, spin systems, quantum liquids, infrared spectroscopy ...
- Still of confidential use for the electronic structure problem of quantum chemistry (as opposed to post-HF and DFT)
- Reason : Very high computational cost for small/medium systems

But :

- Very favorable scaling with system size compared to standard methods
- Ideally suited to extreme parallelism

Quantum Monte Carlo methods

- Solve the Schrödinger equation with random walks
- State-of-the-art and routine approaches in physics : nuclear physics, condensed-matter, spin systems, quantum liquids, infrared spectroscopy ...
- Still of confidential use for the electronic structure problem of quantum chemistry (as opposed to post-HF and DFT)
- Reason : Very high computational cost for small/medium systems

But :

- Very favorable scaling with system size compared to standard methods
- Ideally suited to extreme parallelism

Quantum Monte Carlo methods

- Solve the Schrödinger equation with random walks
- State-of-the-art and routine approaches in physics : nuclear physics, condensed-matter, spin systems, quantum liquids, infrared spectroscopy ...
- Still of confidential use for the electronic structure problem of quantum chemistry (as opposed to post-HF and DFT)
- Reason : Very high computational cost for small/medium systems

But :

- Very favorable scaling with system size compared to standard methods
- Ideally suited to extreme parallelism

Quantum Monte Carlo for molecular systems

Problem : Solve stochastically the Schrödinger equation for N electrons in a molecule

$$
\begin{aligned}
E &= \frac{\int \mathrm{d}\mathbf{r}_1 \ldots \mathrm{d}\mathbf{r}_N \Phi(\mathbf{r}_1, \ldots, \mathbf{r}_N) \mathcal{H} \Phi(\mathbf{r}_1, \ldots, \mathbf{r}_N)}{\int \mathrm{d}\mathbf{r}_1 \ldots \mathrm{d}\mathbf{r}_N \Phi(\mathbf{r}_1, \ldots, \mathbf{r}_N) \Phi(\mathbf{r}_1, \ldots, \mathbf{r}_N)} \\
&\sim \sum \frac{\mathcal{H}\Psi(\mathbf{r}_1, \ldots, \mathbf{r}_N)}{\Psi(\mathbf{r}_1, \ldots, \mathbf{r}_N)}, \text{ sampled with } (\Psi \times \Phi)
\end{aligned}
$$

$\mathcal{H}$: Hamiltonian operator

E : Energy

$\mathbf{r}_1, \ldots, \mathbf{r}_N$: Electron coordinates

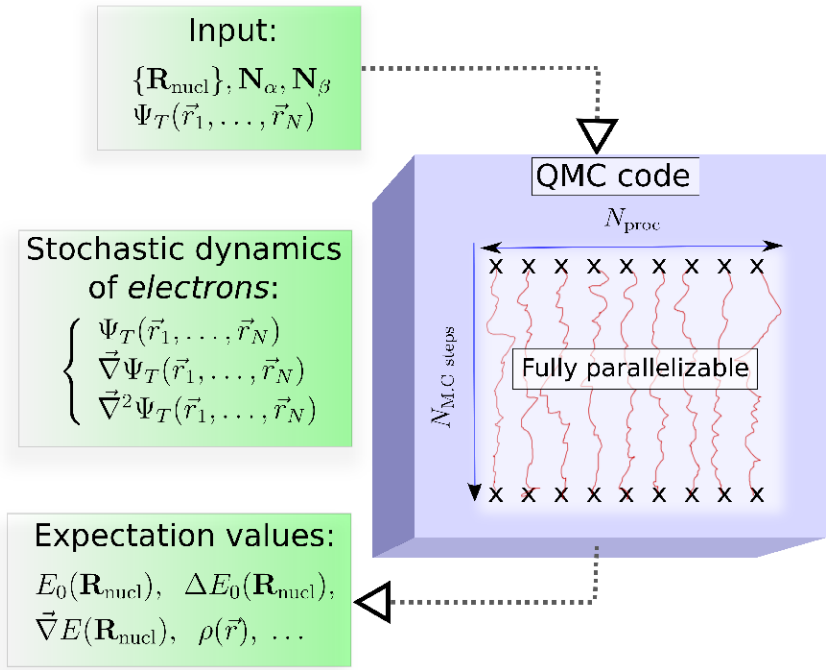
Φ : Exact wave function

Ψ : Trial wave function

QMC Algorithm

- Walker : vector of $\mathbb{R}^{3N}$ containing the electron positions
- Drifted diffusion of walkers with birth/death process to generate the $3N$-density $(\Psi \times \Phi)$ (needs Ψ, $\nabla\Psi$, $\Delta\Psi$)

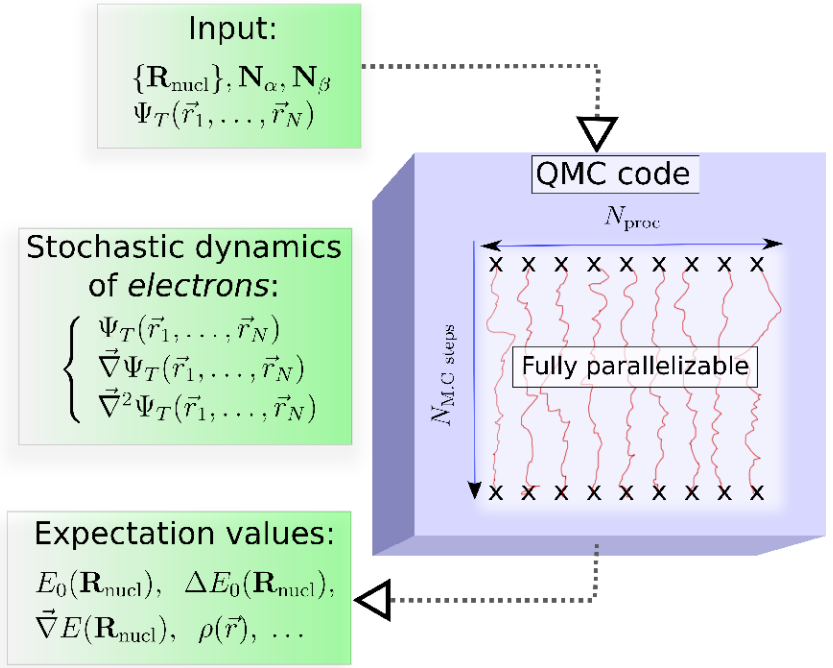
- Compute $\frac{\mathcal{H}\Psi(\mathbf{r}_1,\ldots,\mathbf{r}_N)}{\Psi(\mathbf{r}_1,\ldots,\mathbf{r}_N)}$ for all positions
- The energy is the average of all computed $\frac{\mathcal{H}\Psi(\mathbf{r}_1,\ldots,\mathbf{r}_N)}{\Psi(\mathbf{r}_1,\ldots,\mathbf{r}_N)}$
- Extreme parallelism : Independent populations of walkers can be distributed on different CPUs



QMC Algorithm

- Walker : vector of $\mathbb{R}^{3N}$ containing the electron positions
- Drifted diffusion of walkers with birth/death process to generate the $3N$-density $(\Psi \times \Phi)$ (needs Ψ, $\nabla\Psi$, $\Delta\Psi$)

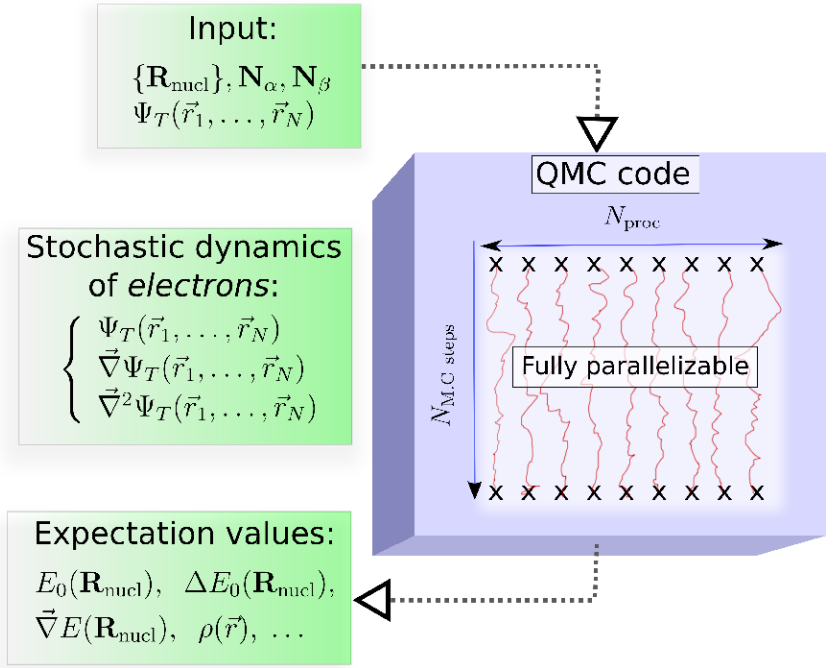
- Compute $\frac{\mathcal{H}\Psi(\mathbf{r}_1,\ldots,\mathbf{r}_N)}{\Psi(\mathbf{r}_1,\ldots,\mathbf{r}_N)}$ for all positions
- The energy is the average of all computed $\frac{\mathcal{H}\Psi(\mathbf{r}_1,\ldots,\mathbf{r}_N)}{\Psi(\mathbf{r}_1,\ldots,\mathbf{r}_N)}$
- Extreme parallelism : Independent populations of walkers can be distributed on different CPUs



QMC Algorithm

- Walker : vector of $\mathbb{R}^{3N}$ containing the electron positions
- Drifted diffusion of walkers with birth/death process to generate the $3N$-density $(\Psi \times \Phi)$ (needs Ψ, $\nabla\Psi$, $\Delta\Psi$)

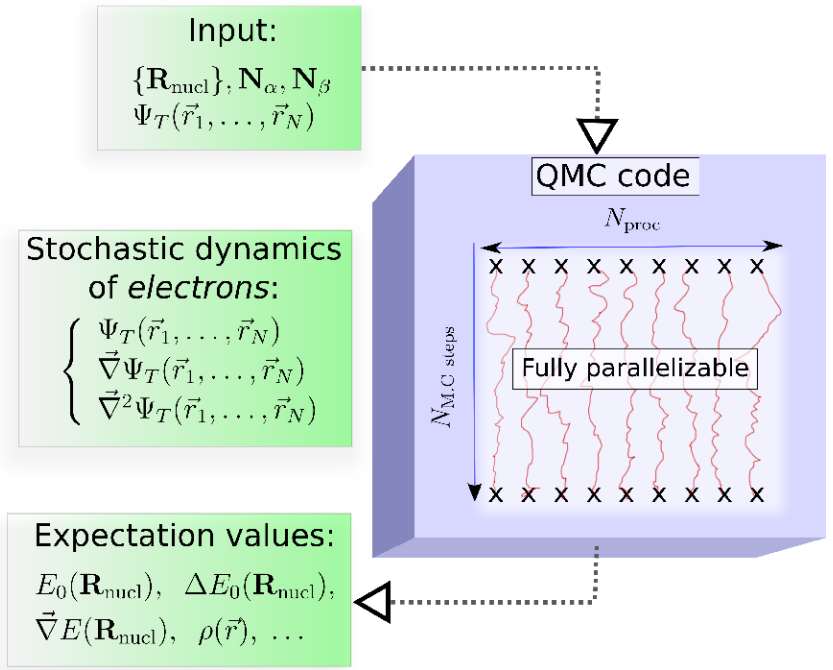
- Compute $\frac{\mathcal{H}\Psi(\mathbf{r}_1,\ldots,\mathbf{r}_N)}{\Psi(\mathbf{r}_1,\ldots,\mathbf{r}_N)}$ for all positions
- The energy is the average of all computed $\frac{\mathcal{H}\Psi(\mathbf{r}_1,\ldots,\mathbf{r}_N)}{\Psi(\mathbf{r}_1,\ldots,\mathbf{r}_N)}$
- Extreme parallelism : Independent populations of walkers can be distributed on different CPUs



QMC Algorithm

- Walker : vector of $\mathbb{R}^{3N}$ containing the electron positions
- Drifted diffusion of walkers with birth/death process to generate the $3N$-density $(\Psi \times \Phi)$ (needs Ψ, $\nabla\Psi$, $\Delta\Psi$)

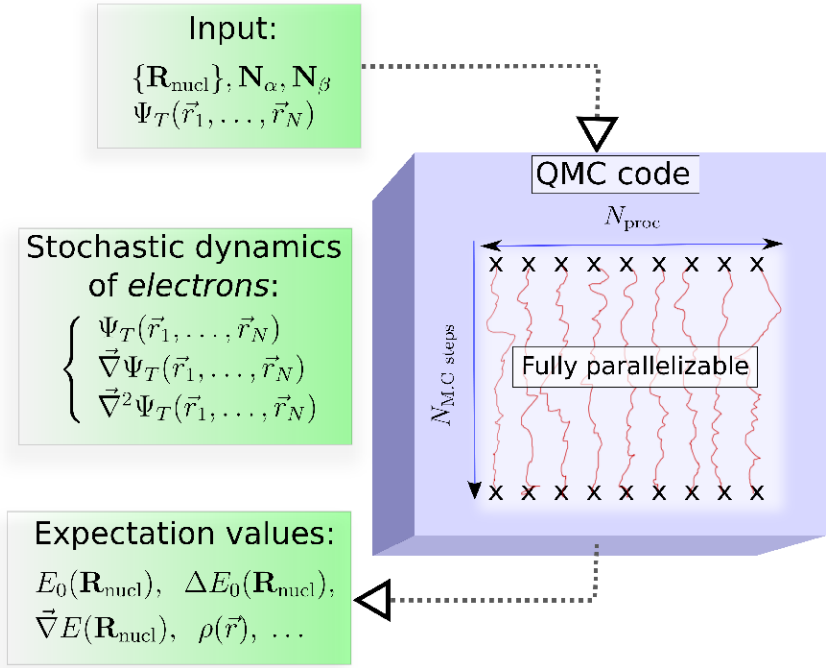
- Compute $\frac{\mathcal{H}\Psi(\mathbf{r}_1,\ldots,\mathbf{r}_N)}{\Psi(\mathbf{r}_1,\ldots,\mathbf{r}_N)}$ for all positions
- The energy is the average of all computed $\frac{\mathcal{H}\Psi(\mathbf{r}_1,\ldots,\mathbf{r}_N)}{\Psi(\mathbf{r}_1,\ldots,\mathbf{r}_N)}$
- Extreme parallelism : Independent populations of walkers can be distributed on different CPUs



QMC Algorithm

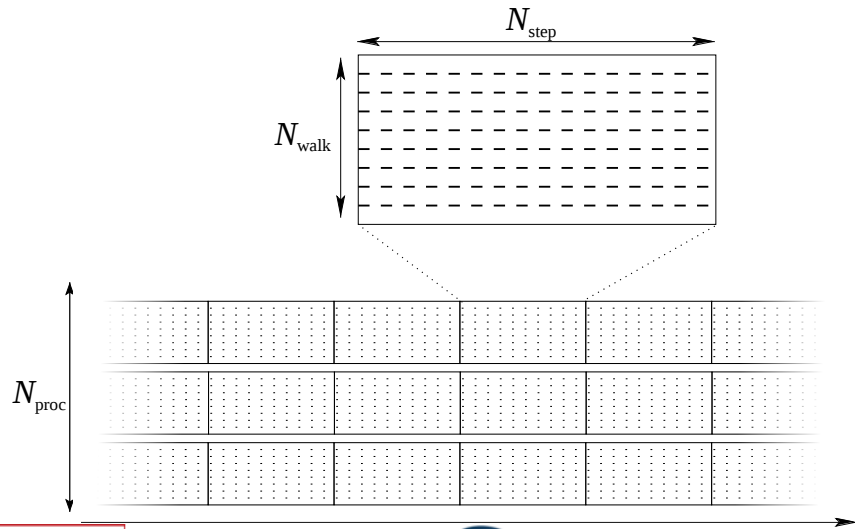
- Walker : vector of $\mathbb{R}^{3N}$ containing the electron positions
- Drifted diffusion of walkers with birth/death process to generate the $3N$-density $(\Psi \times \Phi)$ (needs Ψ, $\nabla\Psi$, $\Delta\Psi$)

- Compute $\frac{\mathcal{H}\Psi(\mathbf{r}_1,\ldots,\mathbf{r}_N)}{\Psi(\mathbf{r}_1,\ldots,\mathbf{r}_N)}$ for all positions
- The energy is the average of all computed $\frac{\mathcal{H}\Psi(\mathbf{r}_1,\ldots,\mathbf{r}_N)}{\Psi(\mathbf{r}_1,\ldots,\mathbf{r}_N)}$
- Extreme parallelism : Independent populations of walkers can be distributed on different CPUs



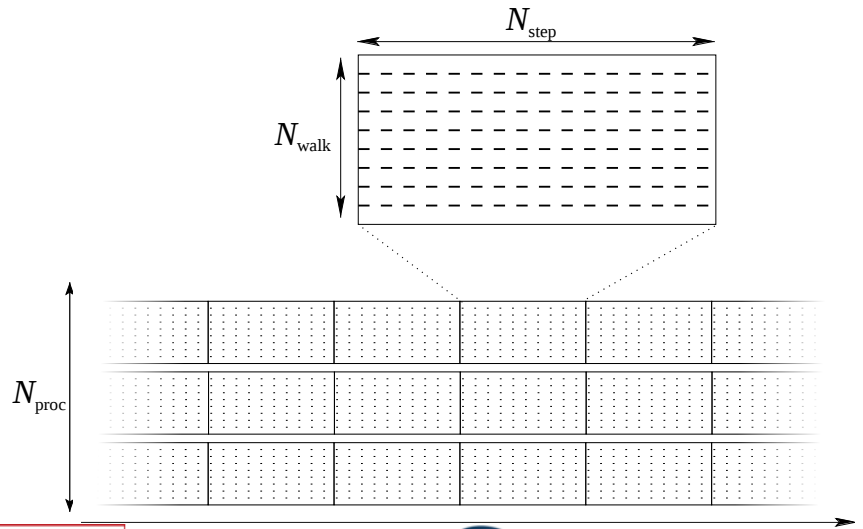
Implementation in QMC=Chem

- Block : N_{walk} walkers executing N_{step} steps
- Compute as many blocks as possible, as quickly as possible
- Block averages have a Gaussian distribution



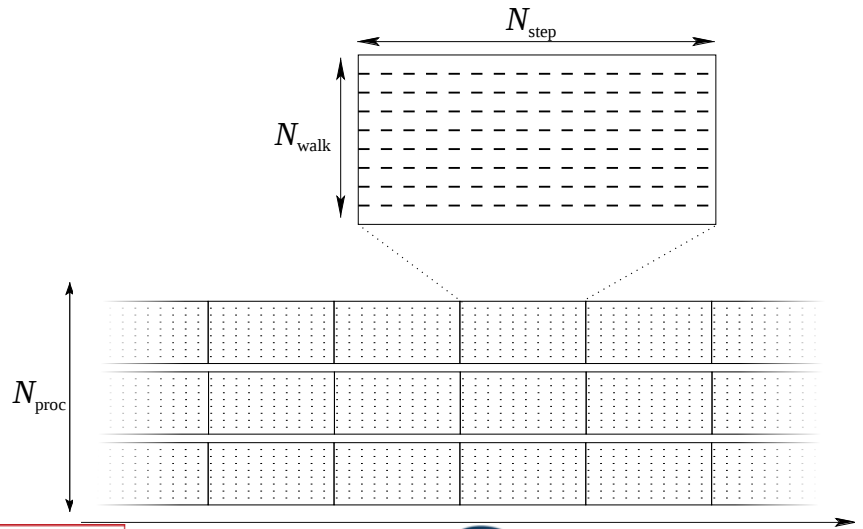
Implementation in QMC=Chem

- Block : N_{walk} walkers executing N_{step} steps
- Compute as many blocks as possible, as quickly as possible
- Block averages have a Gaussian distribution



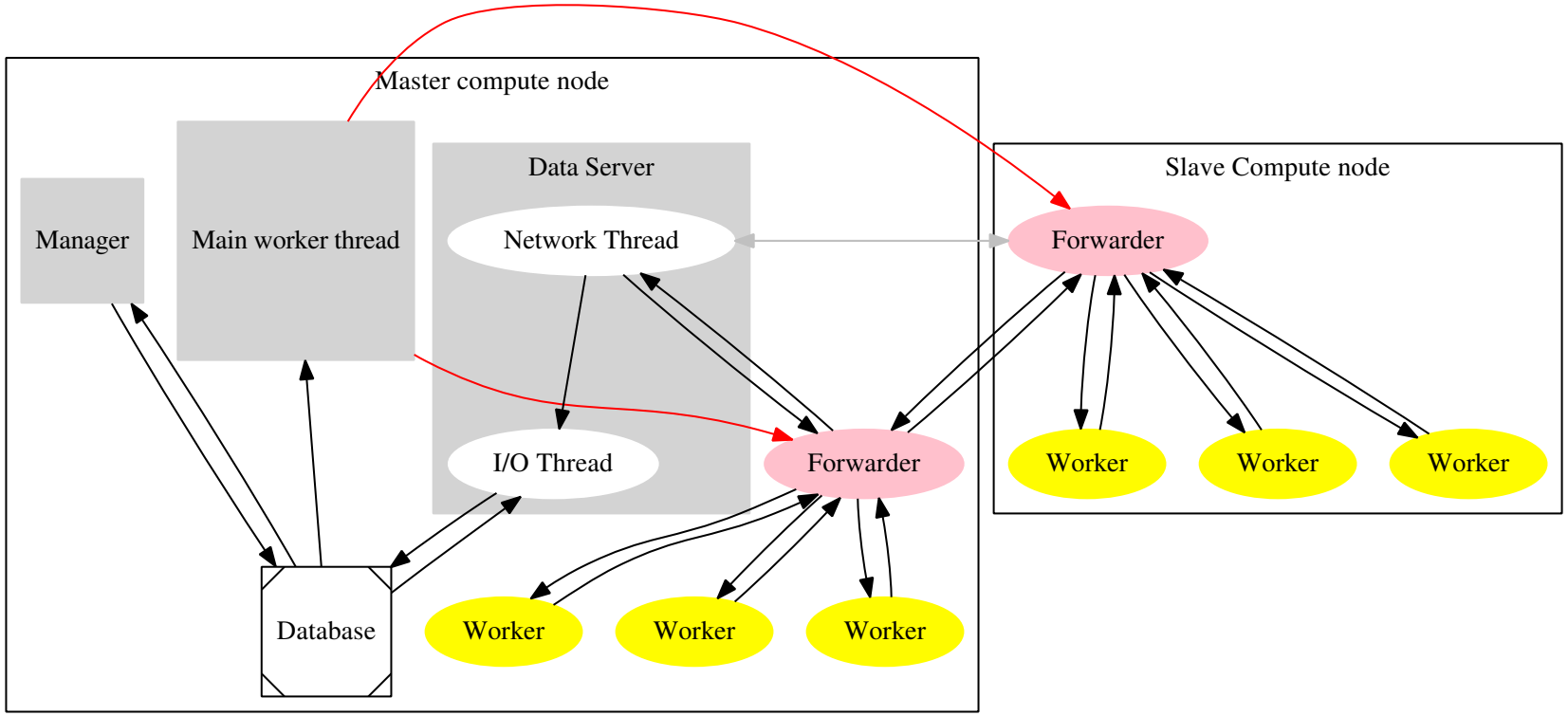
Implementation in QMC=Chem

- Block : N_{walk} walkers executing N_{step} steps
- Compute as many blocks as possible, as quickly as possible
- Block averages have a Gaussian distribution



Parallelism in QMC=Chem

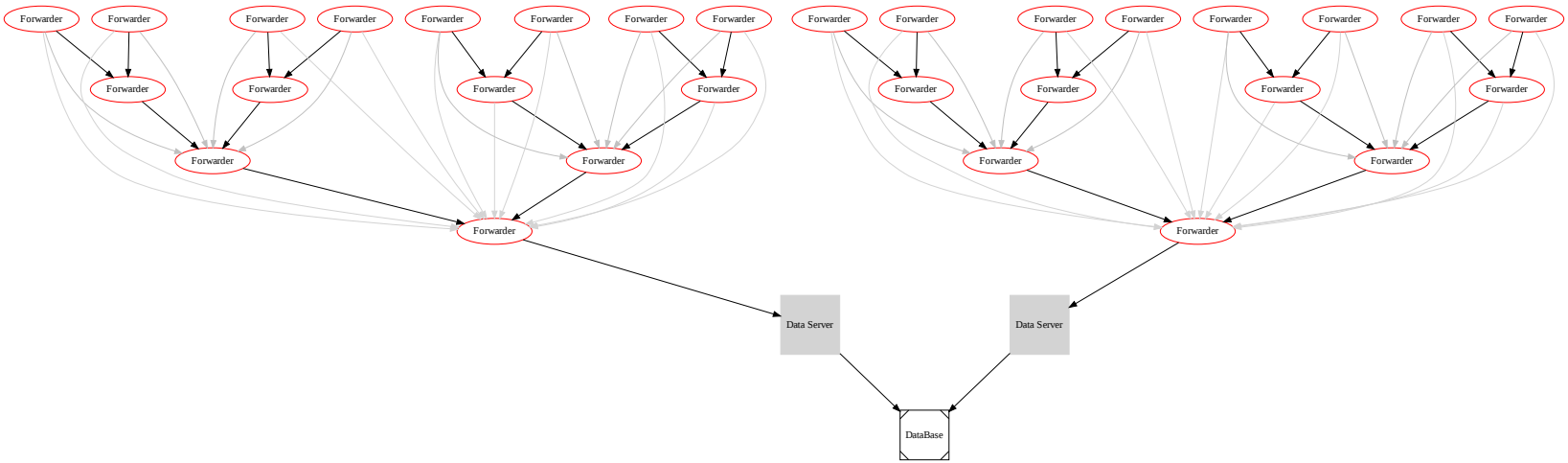
All I/O and network communications are asynchronous



Fault-tolerance

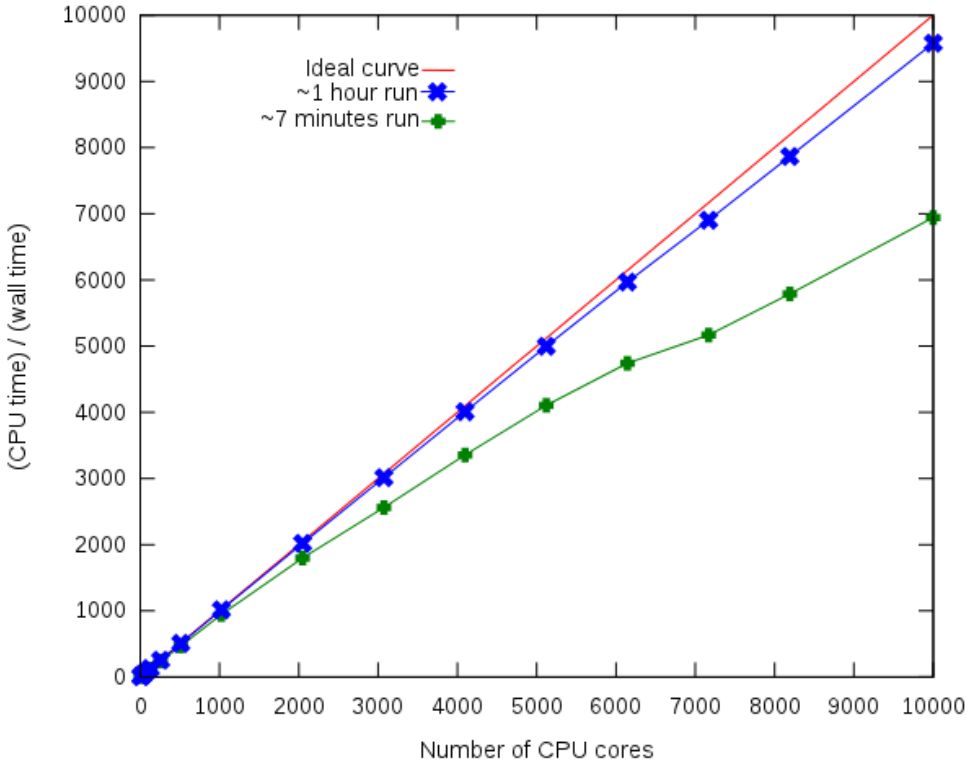
Extreme parallelism $\longrightarrow$ possible system failures

- Blocks are Gaussian $\rightarrow$ losing blocks doesn't change the average
- Simulation survives to removal of any node
- Restart always possible from data base



QMC=Chem scaling

Almost ideal scaling $\longrightarrow$ single-core optimization is crucial.



Hot-spots in a Monte Carlo step

- Matrix inversion $\mathcal{O}(N^3)$ (DP,Intel MKL)
- Sparse$\times$dense matrix products $\mathcal{O}(N^2)$ (SP,our implementation)

Efficiency of the matrix products :

- Static analysis (MAQAO) : Full-AVX (no scalar operations), inner-most loops perform 16 flops/cycle
- Decremental analysis (DECAN) : good balance between flops and memory operations
- Up to 64% of the peak measured on Xeon E5

Hot-spots in a Monte Carlo step

- Matrix inversion $\mathcal{O}(N^3)$ (DP,Intel MKL)
- Sparse$\times$dense matrix products $\mathcal{O}(N^2)$ (SP,our implementation)

Efficiency of the matrix products :

- Static analysis (MAQAO) : Full-AVX (no scalar operations), inner-most loops perform 16 flops/cycle
- Decremental analysis (DECAN) : good balance between flops and memory operations
- Up to 64% of the peak measured on Xeon E5

Hot-spots in a Monte Carlo step

- Matrix inversion $\mathcal{O}(N^3)$ (DP,Intel MKL)
- Sparse$\times$dense matrix products $\mathcal{O}(N^2)$ (SP,our implementation)

Efficiency of the matrix products :

- Static analysis (MAQAO) : Full-AVX (no scalar operations), inner-most loops perform 16 flops/cycle
- Decremental analysis (DECAN) : good balance between flops and memory operations
- Up to 64% of the peak measured on Xeon E5

Hot-spots in a Monte Carlo step

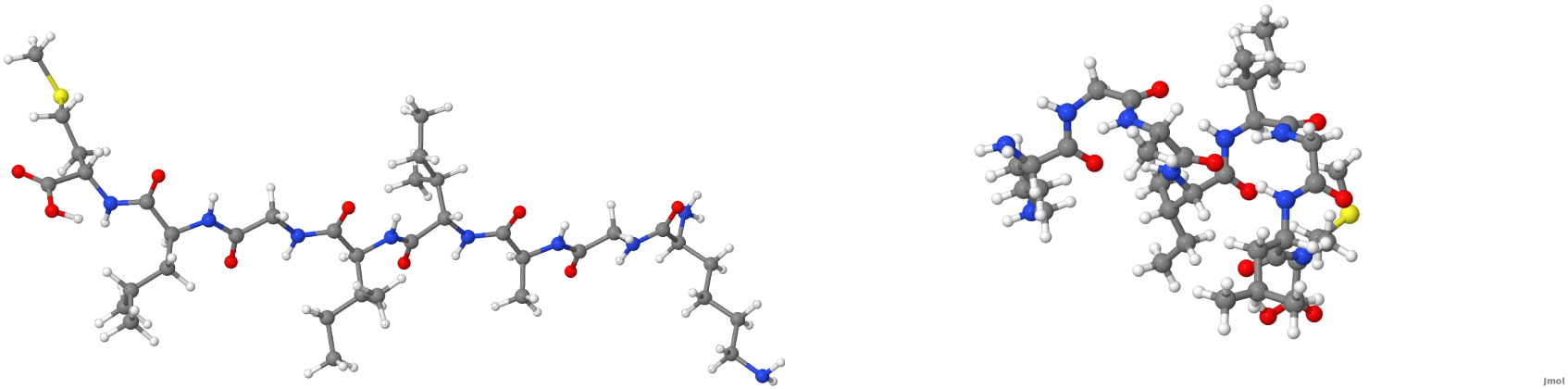
- Matrix inversion $\mathcal{O}(N^3)$ (DP,Intel MKL)
- Sparse$\times$dense matrix products $\mathcal{O}(N^2)$ (SP,our implementation)

Efficiency of the matrix products :

- Static analysis (MAQAO) : Full-AVX (no scalar operations), inner-most loops perform 16 flops/cycle
- Decremental analysis (DECAN) : good balance between flops and memory operations
- Up to 64% of the peak measured on Xeon E5

Amyloid β peptide simulation on Curie

First step in our scientific project : All-electron calculation of the energy difference between the β-strand and the α-helix conformations of amyloid peptide Aβ(28-35)



122 atoms, 434 electrons, cc-pVTZ basis set (2960 basis functions)

Amyloid β peptide simulation on Curie

Scientific results (cc-pVTZ basis set) :

- Standard DFT (B3LYP) : 10.7 kcal/mol
- DFT with empirical corrections (SSB-D) : 35.8 kcal/mol
- All-electron MP2 : 39.3 kcal/mol
- CCSD(T) would require at least 100 million CPU hours
- QMC in $<$ 2 million CPU hours (1 day) : 39.7 $\pm$ 2. kcal/mol
- QMC calculations can be made on these systems $\longrightarrow$ study of the interaction of Copper ions with β-amyloids

Technological results :

- Sustained 960 TFlops/s (Mixed SP/DP) on 76 800 cores of Curie
- $\sim$ 80% parallel speed-up. (Today, it would be $>$ 95 % : run termination was optimized)

Amyloid β peptide simulation on Curie

Scientific results (cc-pVTZ basis set) :

- Standard DFT (B3LYP) : 10.7 kcal/mol
- DFT with empirical corrections (SSB-D) : 35.8 kcal/mol
- All-electron MP2 : 39.3 kcal/mol
- CCSD(T) would require at least 100 million CPU hours
- QMC in < 2 million CPU hours (1 day) : 39.7 $\pm$ 2. kcal/mol
- QMC calculations can be made on these systems $\longrightarrow$ study of the interaction of Copper ions with β-amyloids

Technological results :

- Sustained 960 TFlops/s (Mixed SP/DP) on 76 800 cores of Curie
- $\sim$ 80% parallel speed-up. (Today, it would be > 95 % : run termination was optimized)

Amyloid β peptide simulation on Curie

Scientific results (cc-pVTZ basis set) :

- Standard DFT (B3LYP) : 10.7 kcal/mol
- DFT with empirical corrections (SSB-D) : 35.8 kcal/mol
- All-electron MP2 : 39.3 kcal/mol
- CCSD(T) would require at least 100 million CPU hours
- QMC in < 2 million CPU hours (1 day) : $39.7 \pm 2.$ kcal/mol
- QMC calculations can be made on these systems $\longrightarrow$ study of the interaction of Copper ions with β-amyloids

Technological results :

- Sustained 960 TFlops/s (Mixed SP/DP) on 76 800 cores of Curie
- $\sim$ 80% parallel speed-up. (Today, it would be > 95 % : run termination was optimized)

Amyloid β peptide simulation on Curie

Scientific results (cc-pVTZ basis set) :

- Standard DFT (B3LYP) : 10.7 kcal/mol
- DFT with empirical corrections (SSB-D) : 35.8 kcal/mol
- All-electron MP2 : 39.3 kcal/mol
- CCSD(T) would require at least 100 million CPU hours
- QMC in < 2 million CPU hours (1 day) : $39.7 \pm 2.$ kcal/mol
- QMC calculations can be made on these systems $\longrightarrow$ study of the interaction of Copper ions with β-amyloids

Technological results :

- Sustained 960 TFlops/s (Mixed SP/DP) on 76 800 cores of Curie
- $\sim$ 80% parallel speed-up. (Today, it would be > 95 % : run termination was optimized)

Amyloid β peptide simulation on Curie

Scientific results (cc-pVTZ basis set) :

- Standard DFT (B3LYP) : 10.7 kcal/mol
- DFT with empirical corrections (SSB-D) : 35.8 kcal/mol
- All-electron MP2 : 39.3 kcal/mol
- CCSD(T) would require at least 100 million CPU hours
- QMC in < 2 million CPU hours (1 day) : 39.7 $\pm$ 2. kcal/mol
- QMC calculations can be made on these systems $\longrightarrow$ study of the interaction of Copper ions with β-amyloids

Technological results :

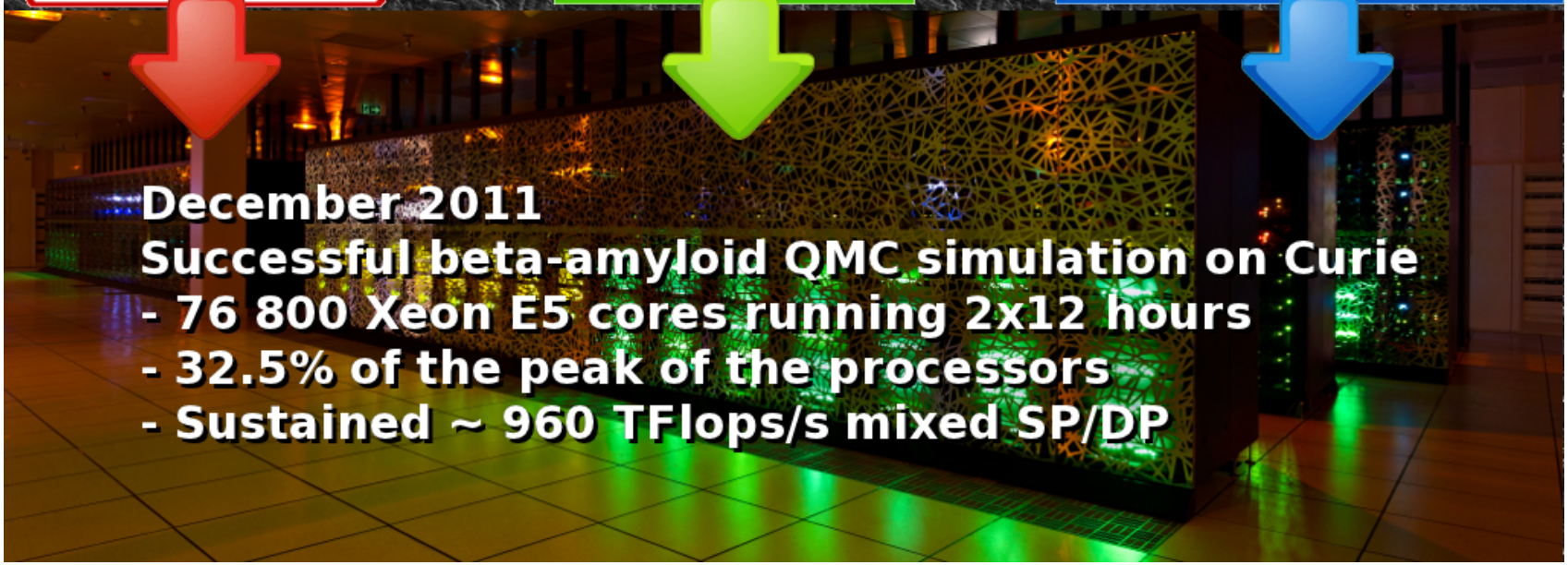
- Sustained 960 TFlops/s (Mixed SP/DP) on 76 800 cores of Curie
- $\sim$ 80% parallel speed-up. (Today, it would be > 95 % : run termination was optimized)

Partners

M. Caffarel, A. Scemama
* QMC Methods
* QMC=Chem code

GENCI: Leader of national HPC resources
* Provides Curie

W. Jalby, E. Oseret
* Performance analysis tools
* Xeon E5 optimization



December 2011
Successful beta-amyloid QMC simulation on Curie
- 76 800 Xeon E5 cores running 2x12 hours
- 32.5% of the peak of the processors
- Sustained ~ 960 TFlops/s mixed SP/DP