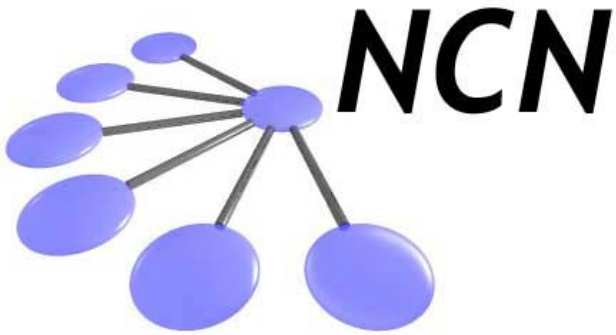
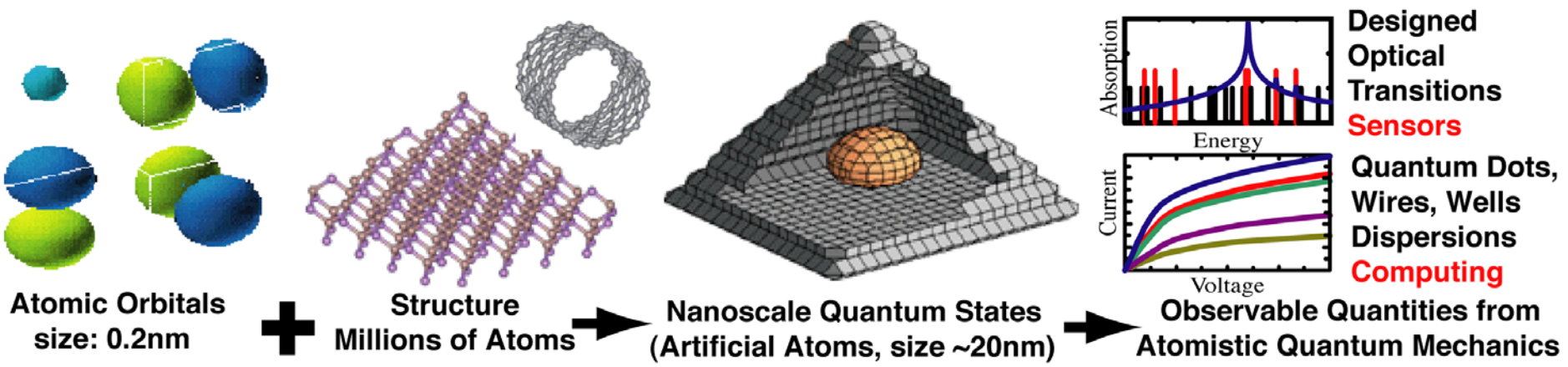
Introduction to the NEMO3D Tool



Gerhard Klimeck

NEMO 3-D Technical Approach



Problem:

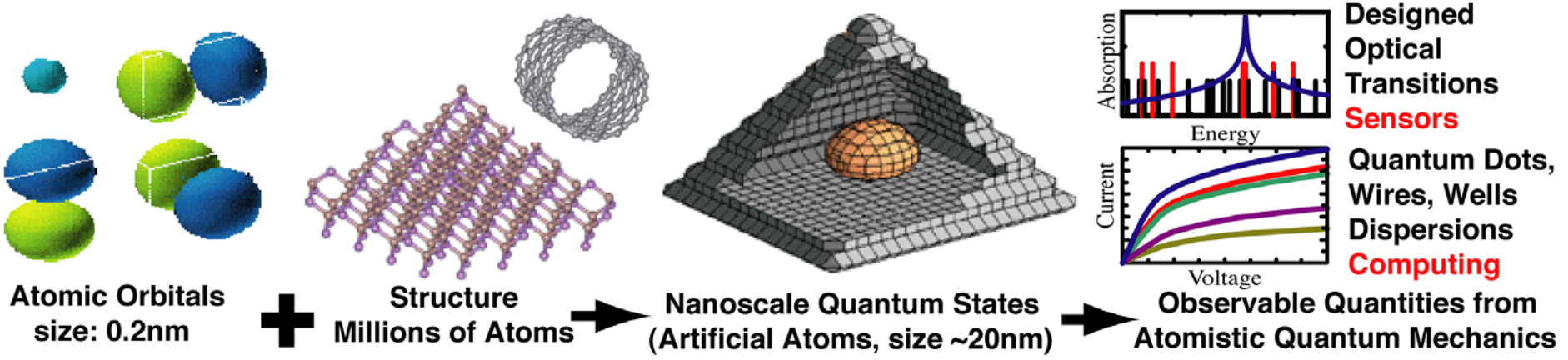
Nanoscale device simulation requirements:

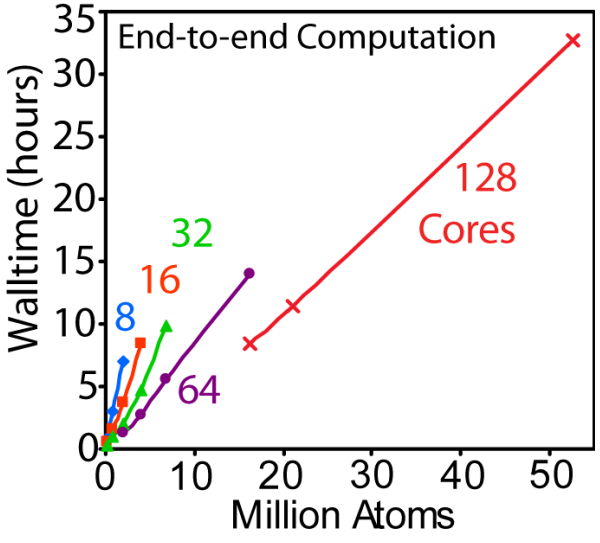
- Cannot use bulk / jellium descriptions, need description atom by atom
  => use pseudo-potential or local orbitals
- Consider finite extend, not periodic
  => local orbital approach
- Bonds in semiconductors are stable!
  => Do not need ab-initio methods!
- Need to include about one million atoms.
  => need massively parallel computers
- The design space is huge:.
  => need a design tool

Approach:

- Use local orbital description for individual atoms in arbitrary crystal / bonding configuration
  - Use s, p, and d orbitals.
  - Use genetic algorithm to determine material parameter fitting
- Compute mechanical strain in the system.
- Develop efficient parallel algorithms to generate eigenvalues/vectors of very large matrices

NEMO 3-D Technical Approach





Demonstration / Capability / Impact:

- 52 million atom electronic structure $(101nm)^3$.
- Quantum dots, nanowires, quantum computing…

Approach:

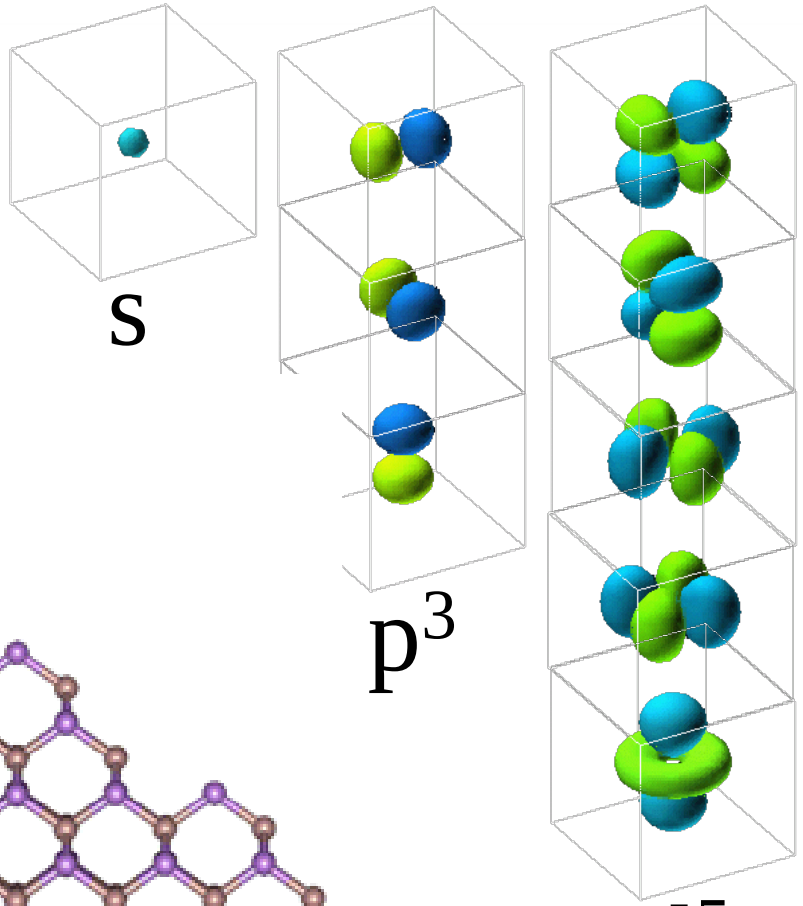
- Use local orbital description for individual atoms in arbitrary crystal / bonding configuration
  - Use s, p, and d orbitals.
  - Use genetic algorithm to determine material parameter fitting
- Compute mechanical strain in the system.
- Develop efficient parallel algorithms to generate eigenvalues/vectors of very large matrices

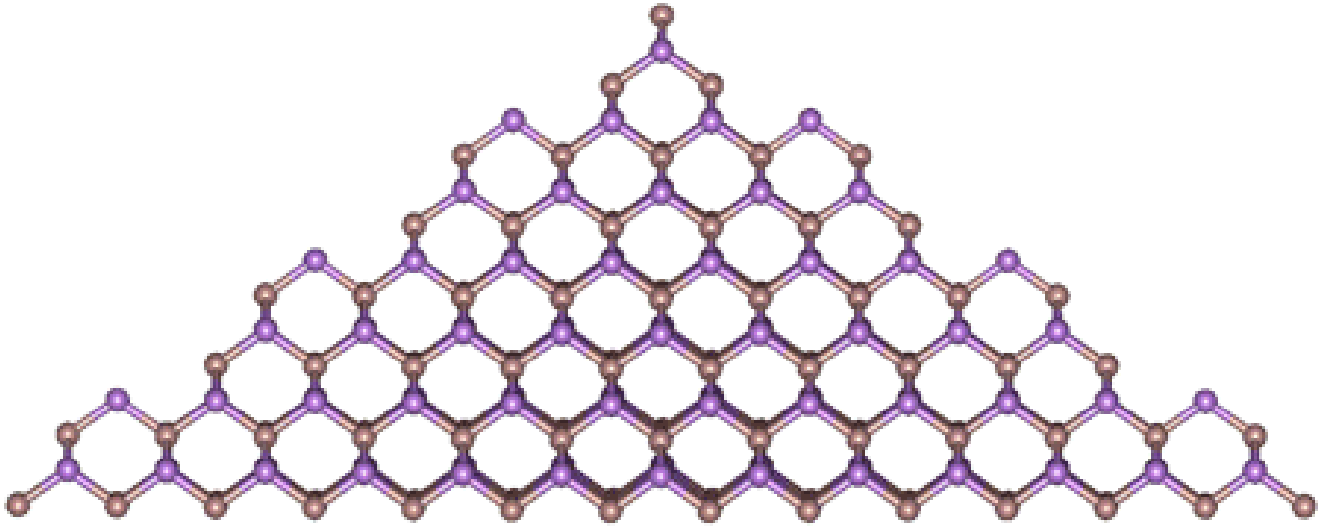
Atomistic Tight-Binding Hamiltonian

$$\psi_{i,j,k} = \sum_{l}^{orbitals} C^{l}_{i,j,k}\phi_{l}$$

$$\phi_{l} = s, p^3, d^5$$

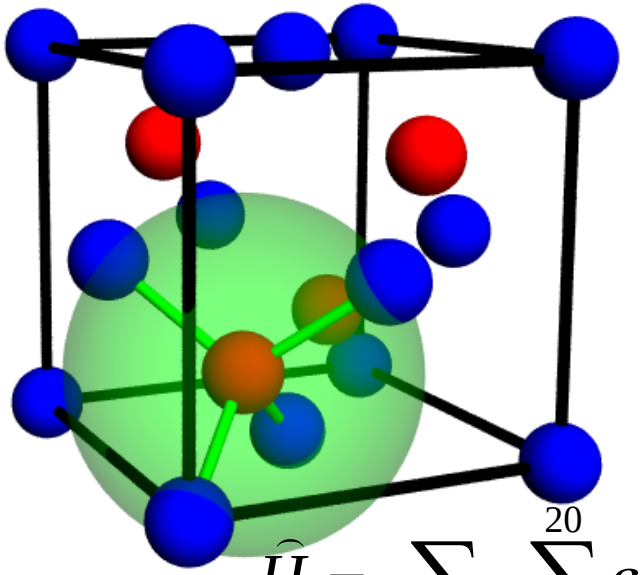
Wavefunction Basis





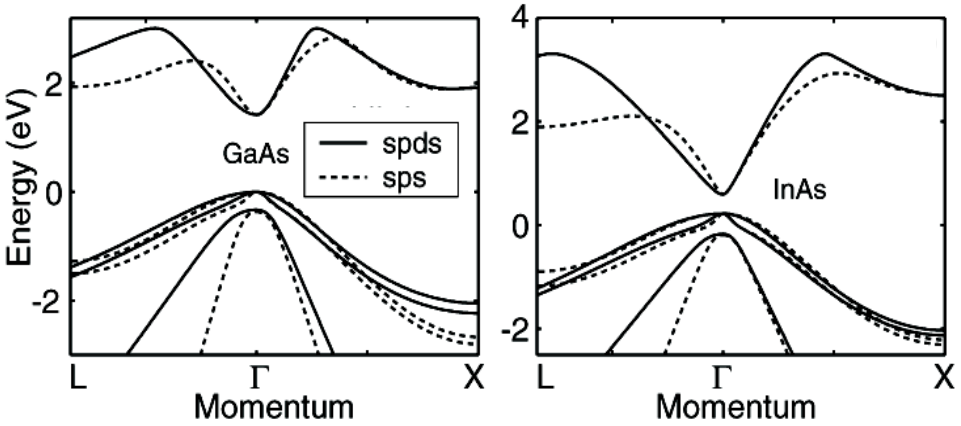
Crystalline Structure

Nearest-Neighbor sp3d5s* Model



- Realistic zinc-blende lattice
- Each atom has 10 spin-degenerate orbitals ($sp^3d^5s^*$)
- Each atom has 4 nearest neighbors
- TB parameters found by fitting of bulk band structure to experiment

$$\hat{H} = \sum_{atoms\ R} \sum_{\substack{\alpha=1 \\ bands}}^{20} \varepsilon_{R\alpha} c^{+}_{R\alpha} c_{R\alpha} + \sum_{atoms\ R} \sum_{atoms\ R'=1}^{4nn} \sum_{\substack{\alpha=1 \\ bands}}^{20} \sum_{\substack{\alpha'=1 \\ bands}}^{20} t_{R\alpha,R'\alpha'} c^{+}_{R\alpha} c_{R'\alpha'}$$

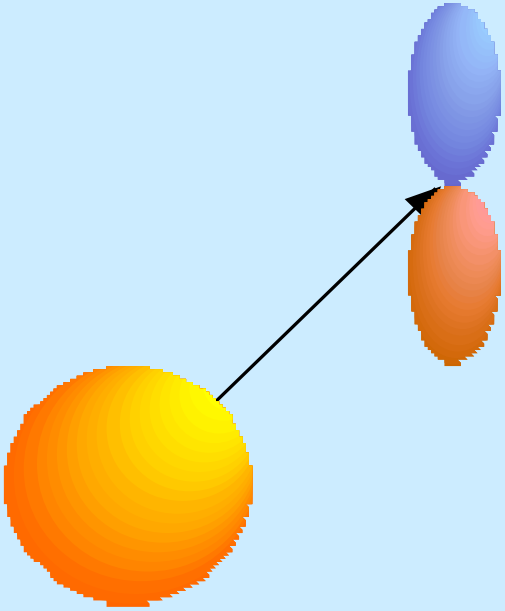


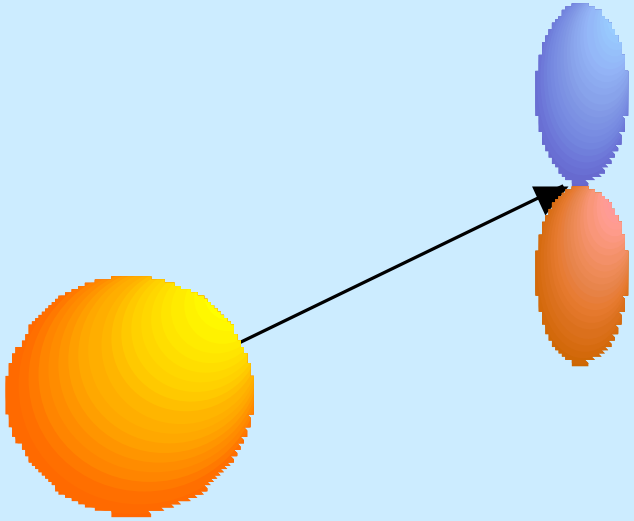
Systems with up to 52 million atoms

Matrix order: $1x10^9$
Lanczos Diagonalization

Modeling Strain

Bond Lengths and Angles Change





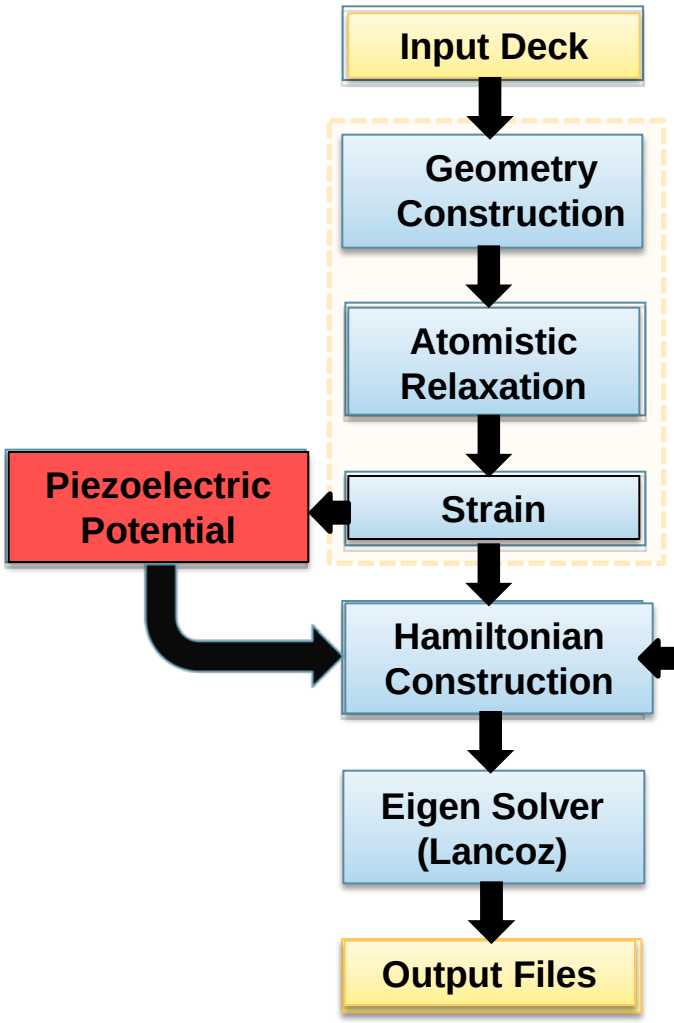
$$\mathbf{d} = d\left(l\mathbf{e}_x + m\mathbf{e}_y + n\mathbf{e}_z\right)$$

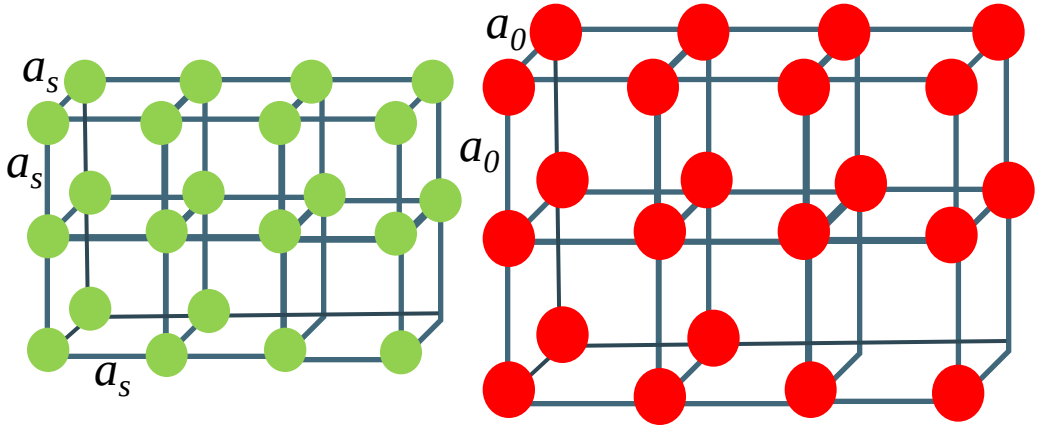
$$\mathbf{d}' = d'\left(l'\mathbf{e}_x + m'\mathbf{e}_y + n'\mathbf{e}_z\right)$$

$$E_{sa,\,zc} = n\,V_{sa,\,pc\sigma}$$

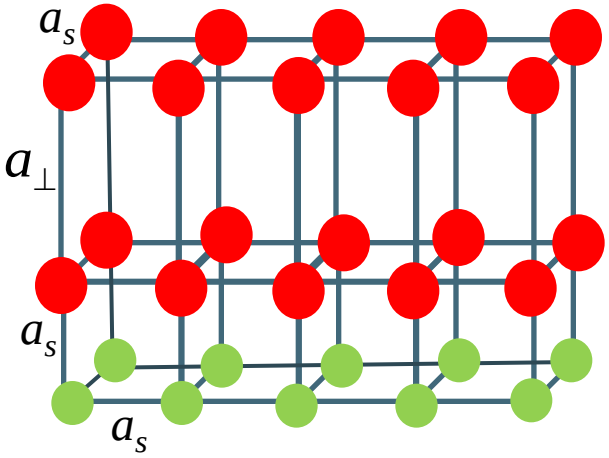
$$E_{sa,\,zc} = n'\,V_{sa\,,\,pc\sigma}\left(\frac{d}{d}\right)^{\eta}$$

NEMO 3-D Calculation Flow – VFF Strain Model





An epitaxial layer is grown, on a substrate with different lattice constant, the epilayer deforms (strain) :



VFF Model

$$a_\perp = a_0 - 2\frac{C_{12}}{C_{11}}[a_s - a_0] \Longleftrightarrow \mathbf{R}' = (1+\varepsilon)\mathbf{R}$$

Strain tensor

Strain Impact on Electronic Structure

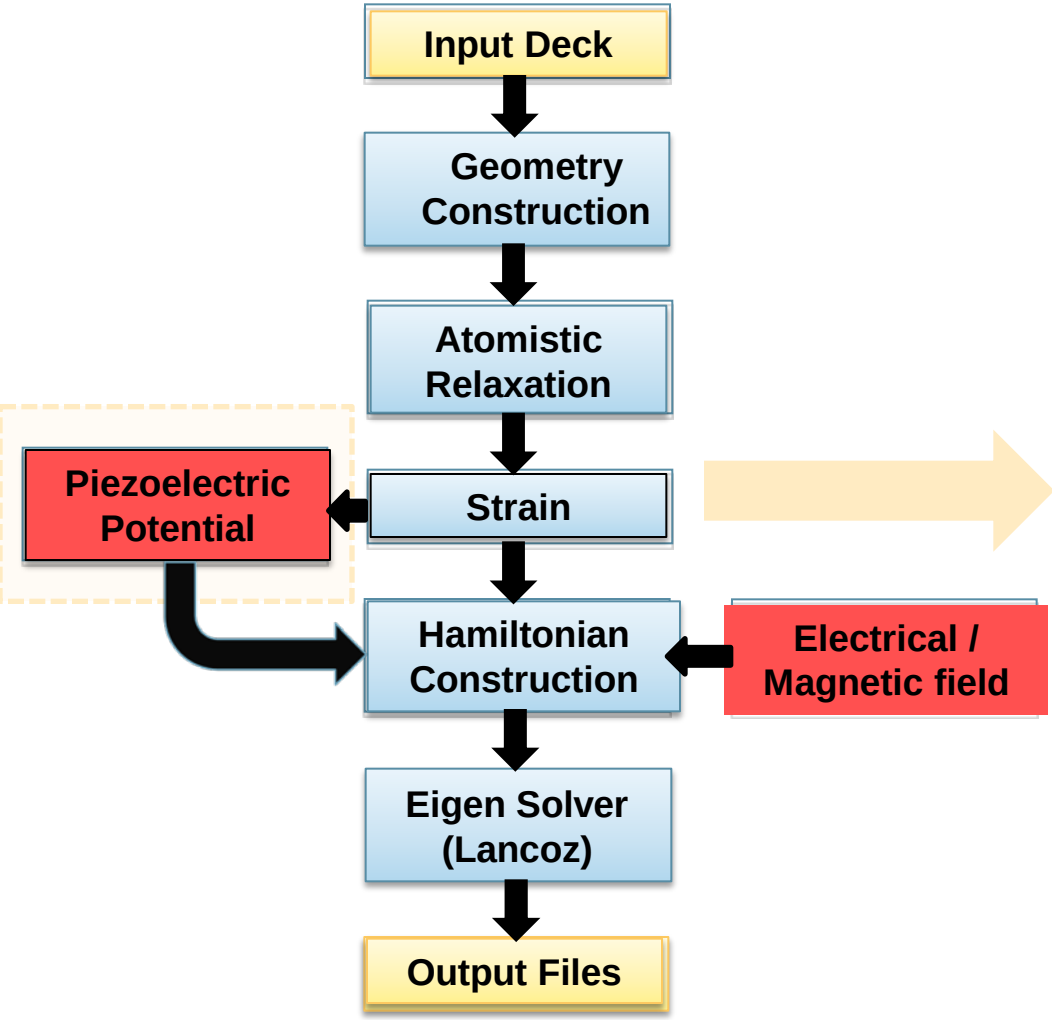
- Strain ➔ Changes bond lengths/angles ➔ deforms crystals
- Band edges are shifted and degeneracy of HH-LH is lifted
- Strain impact on electronic structure:

Conduction Band ➔ $E_c = E_{c0} + a_c(\varepsilon_{xx}+\varepsilon_{yy}+\varepsilon_{zz})$

Valence Band ➔ $E_{HH/LH} = E_{HH0/LH0} + a_v(\varepsilon_{xx}+\varepsilon_{yy}+\varepsilon_{zz}) \pm 0.5b_v\,(\varepsilon_{xx}+\varepsilon_{yy}-2\varepsilon_{zz})$

where a_c = -5.08eV, a_v=1.0eV and b_v = -1.8eV for InAs

NEMO 3-D Calculation Flow -Piezo-Electric Model



Strain tensor ➔ Polarization

$$P_1 = 2e_{14}\begin{pmatrix}\varepsilon_{yz}\\ \varepsilon_{xz}\\ \varepsilon_{xy}\end{pmatrix}$$

$$P_2 = 2B_{114}\begin{pmatrix}\varepsilon_{xx}\varepsilon_{yz}\\ \varepsilon_{yy}\varepsilon_{xz}\\ \varepsilon_{zz}\varepsilon_{xy}\end{pmatrix} + 2B_{124}\begin{pmatrix}\varepsilon_{yz}(\varepsilon_{yy}+\varepsilon_{zz})\\ \varepsilon_{xz}(\varepsilon_{zz}+\varepsilon_{xx})\\ \varepsilon_{xy}(\varepsilon_{xx}+\varepsilon_{yy})\end{pmatrix} + 4B_{156}\begin{pmatrix}\varepsilon_{xz}\varepsilon_{xy}\\ \varepsilon_{yz}\varepsilon_{xy}\\ \varepsilon_{yz}\varepsilon_{xz}\end{pmatrix}$$

Polarization ➔ Charge density

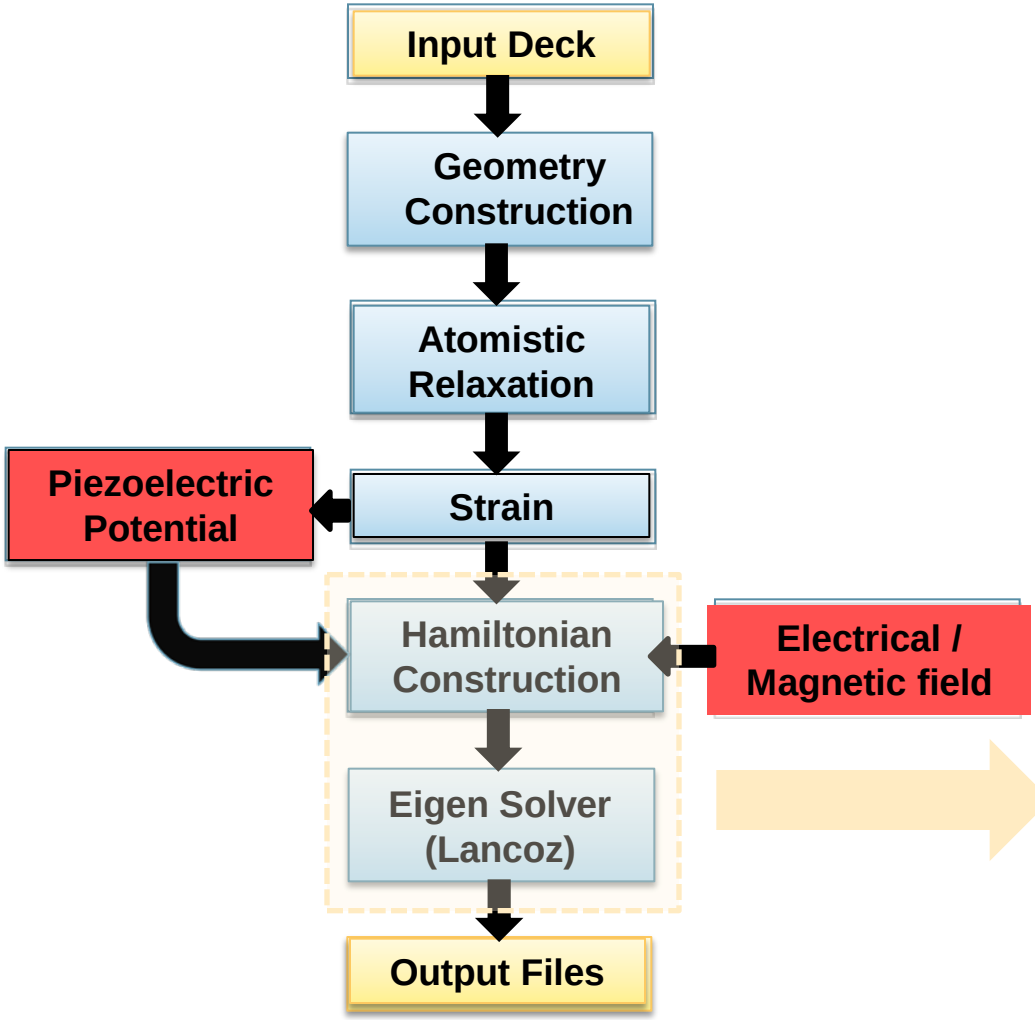
$$\rho_{piezo}(r) = -\nabla.P$$

$$P = P_1 + P_2$$

Charge density ➔ Potential

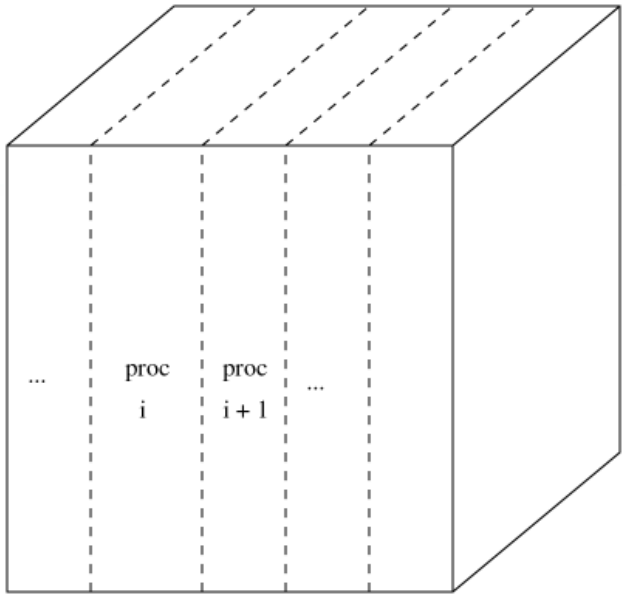
$$\rho_p(r) = \varepsilon_0 \nabla.[\varepsilon_s(r)\nabla V_p(r)]$$

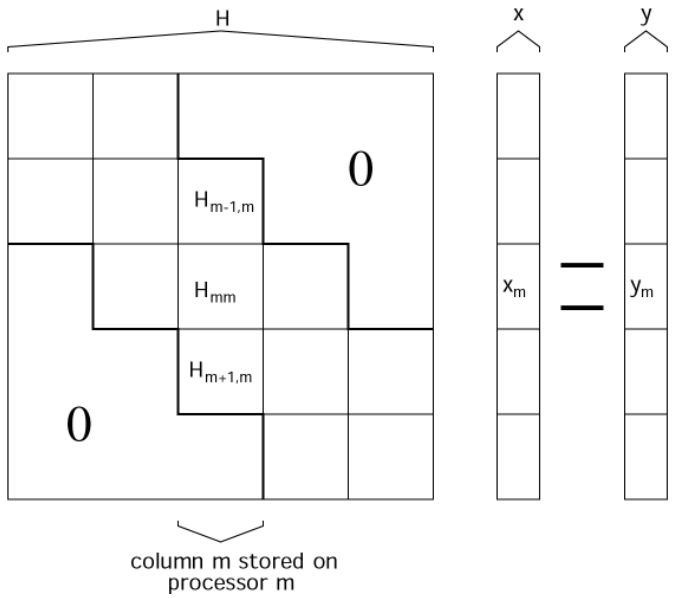
NEMO 3-D Calculation Flow – Electronic Structure Model



- Real-size QD typically involves multi-million atoms.
- $sp^3d^5s^*$ TB requires 20 bases per each atom with consideration of spin-orbit couplings → (Natomx20) by (Natomx20) Hamiltonian.
- Need to obtain interior spectrum eigenstates
- NEMO 3-D: Designed from the bottom up to be parallel: MPI.
- NEMO 3-D code has been tested to run on (almost) any HPC cluster

Parallelization and Methods





- Divide Simulation domain into slices.
- Communication only from one slice to the next (nearest neighbor)
- Communication overhead across the surfaces of the slices.

- Limiting operation:
  complex sparse matrix-vector multiplication
- Enable Hamiltonian storage or
  re-computation on the fly.

- Electronic structure needs eigenvalues and eigenvectors. Matrix is Hermitian
- NEMO 3-D methods:
  - Standard 2-pass Lanczos
  - PARPACK about 10x slower
  - Folded Spectrum Method (Zunger), also typically slower than Lanczos

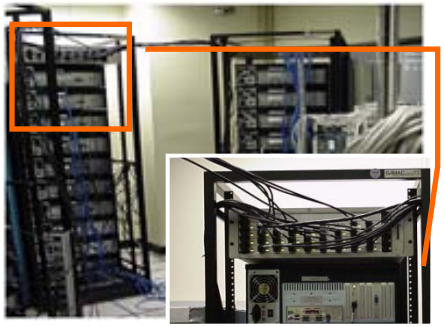
From Beowulf Concept (1998 JPL, Tom Sterling) to Commodity Products in 4 Generations

Hyglac (1997)
16 Pentium Pros 200MHz
128 MB RAM per node
2 GB total
5GB Disc per node
80 GB total
100 Mb/s ethernet crossbar
Linux, MPI
3.2GFlops

Gordon Bell Prize 1997

Nimrod (1999)
32 Pentium IIIs 450MHz
512 MB RAM per node
16 GB total
8GB Disc per node
256 GB total
100 Mb/s ethernet crossbar
Linux, MPI
14.4 GFlops

Pluto (2001)
64 Pentium IIIs 800MHz
dual CPUs
2 GB RAM per node
64 GB total
10 GB Disc per node
320 GB total
2 Gb/s Myricom crossbar
Linux, MPI
51.2 GFlops



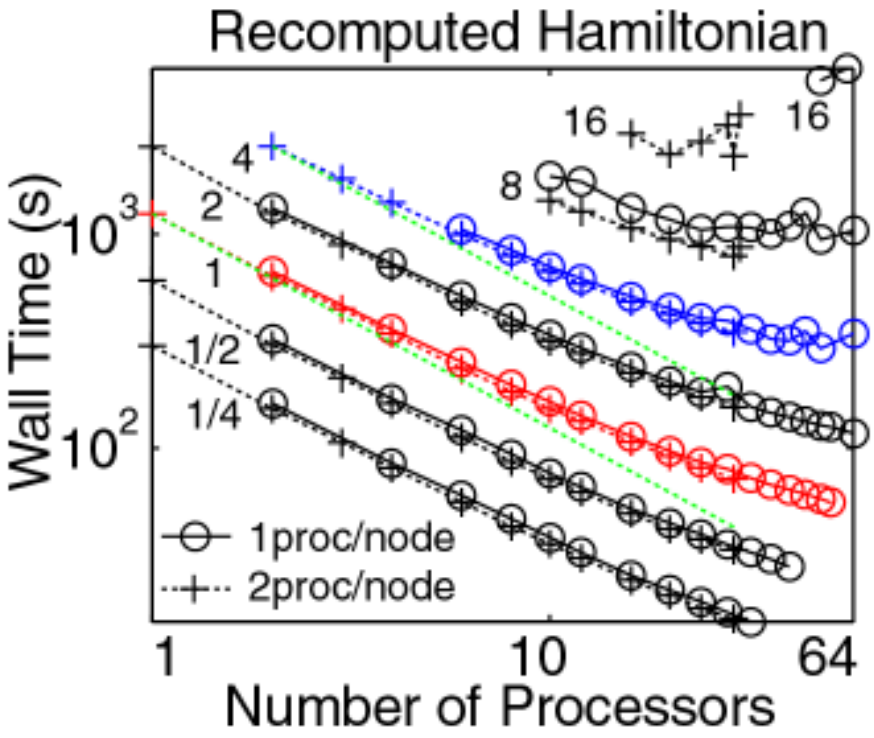
NewYork (2002)
66 Xserve G4 1GHz
1GB RAM per node
33 GB total
60 GB Disc per node
2 TB total
100 Mb/s ethernet crossbar
MAC OS X, MPI
495GFlops

Affordable Supercomputer for ~$100k

Parallelization Benchmarks

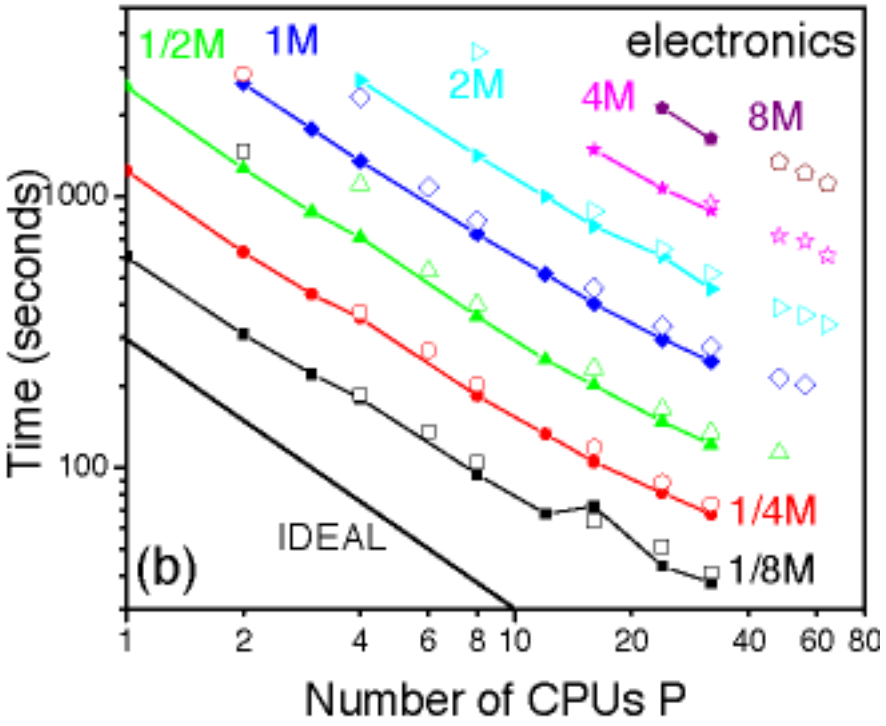
2000

32 nodes Dual Intel P3, 800MHz,
1GB RAM, 100Mbs ethernet



2003

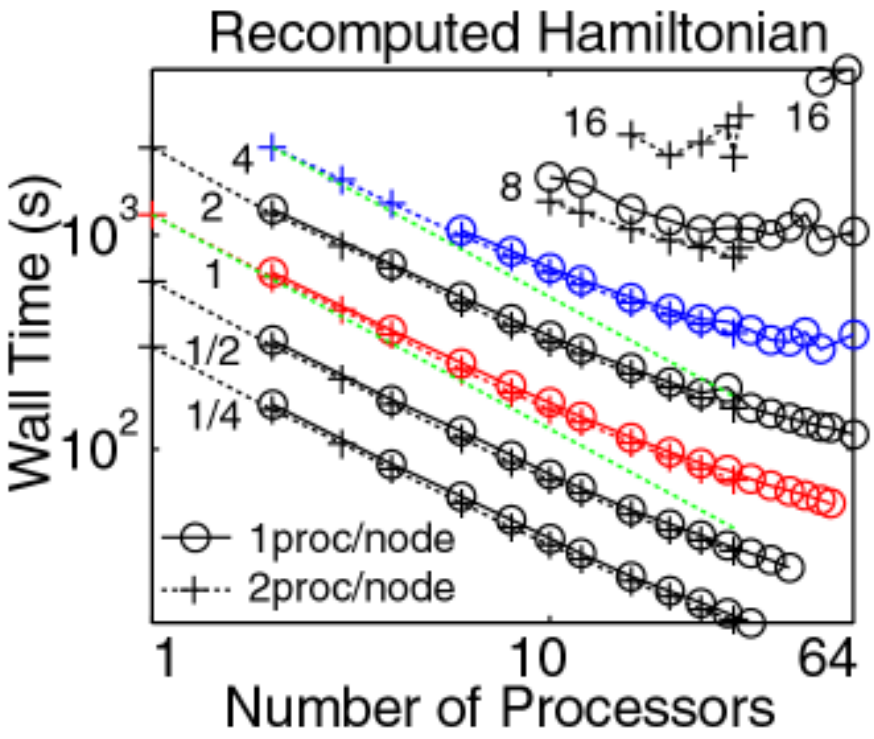
32 node, G5 Apple cluster
2GB RAM, Dual Gigabit Ethernet.



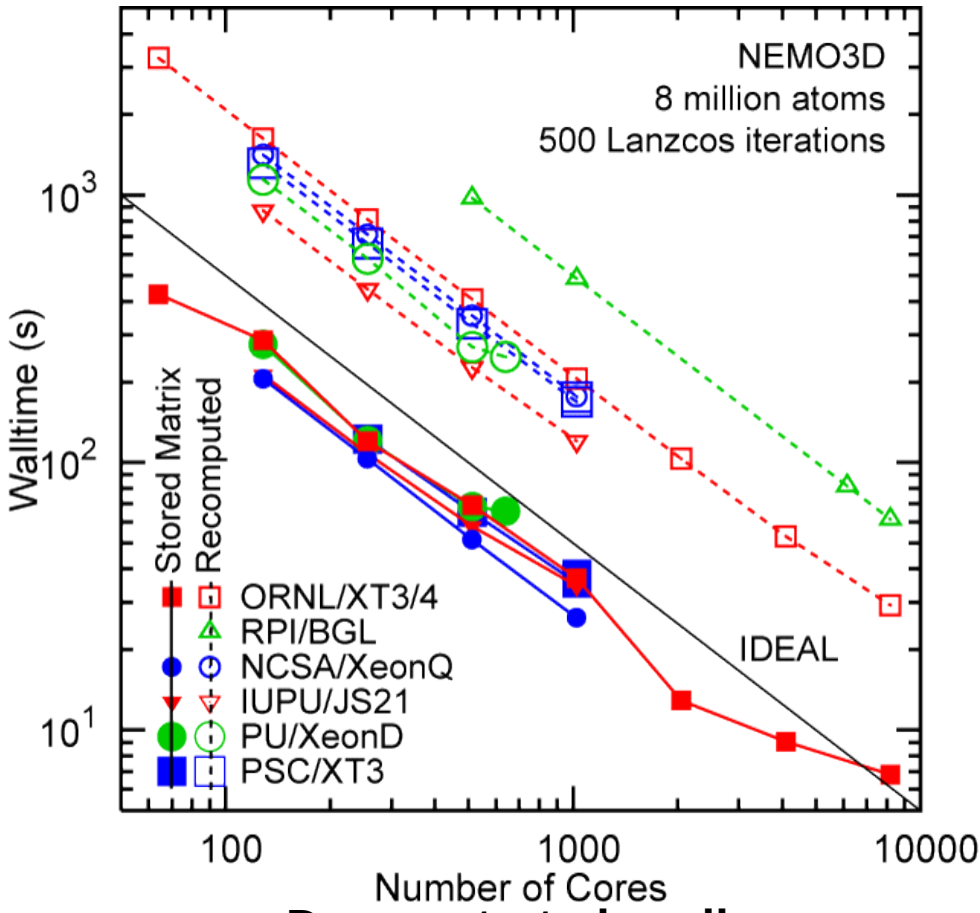
Parallelization Benchmarks

2000

32 nodes Dual Intel P3, 800MHz,
1GB RAM, 100Mbs ethernet



2007
IBM BGL, Cray XT3,
Intel Clusters, IBM JS21



Demonstrated scaling to 8,192 cores

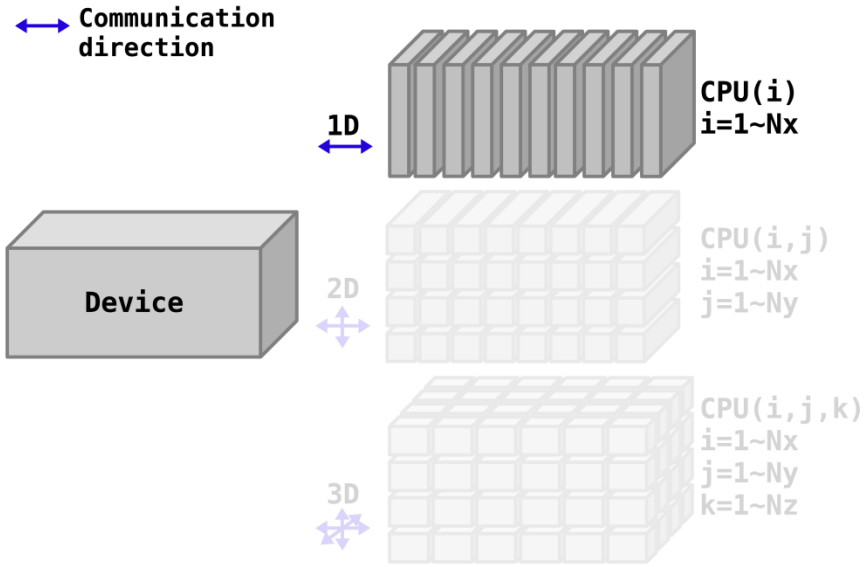
Parallelization Benchmarks

Scaling to 8,192 processors in NEMO3D

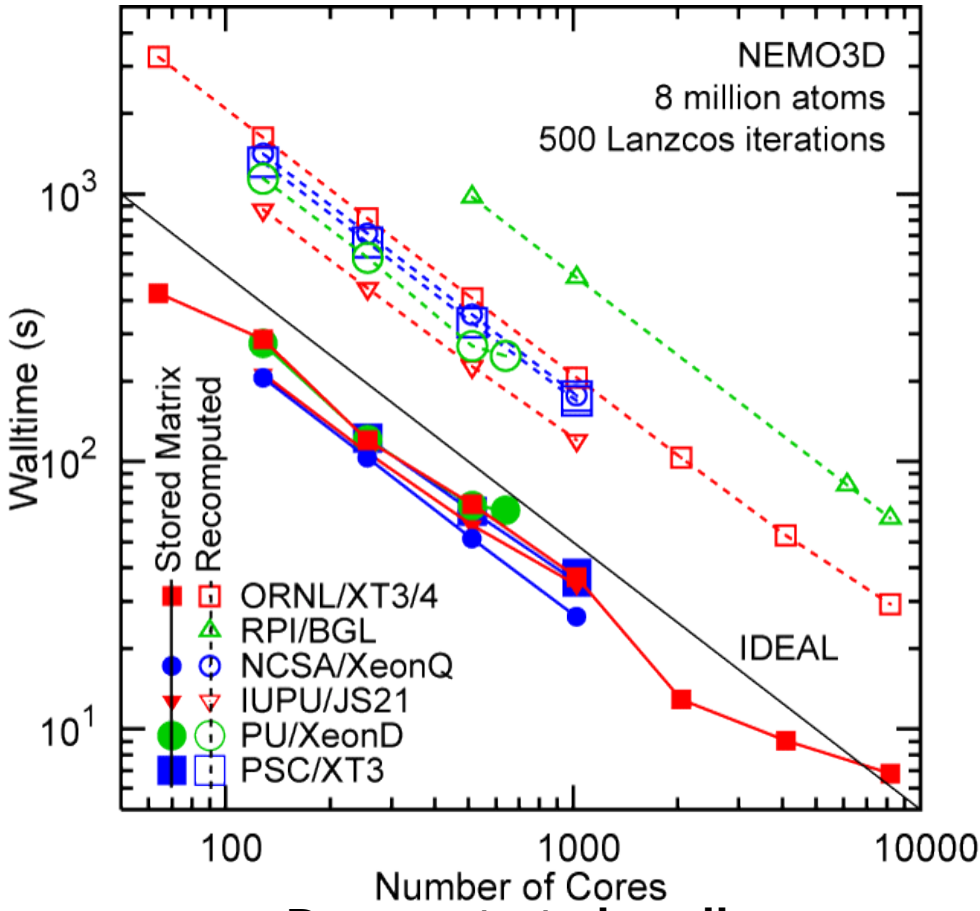
Is a bit of a cheat!
NEMO3D does a 1D decomposition!!

The domain is
a very thin and very long nanowire
NOT a Quantum Dot!

Need a 2D and 3D Decompositions



2007
IBM BGL, Cray XT3,
Intel Clusters, IBM JS21



Demonstrated scaling
to 8,192 cores

NEMO Development and Scalability Study

Sunhee Lee, Hoon Ryu, Gerhard Klimeck

Objective

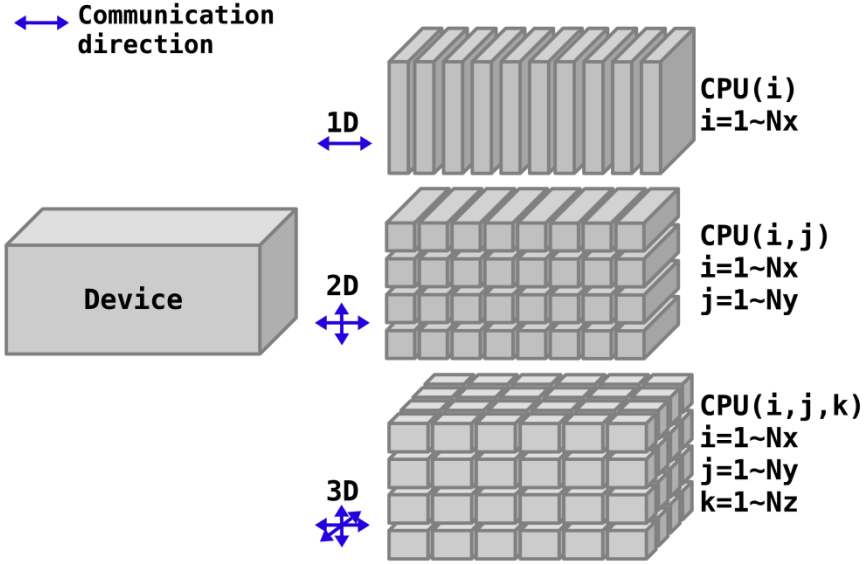
- Run large atomistic electronic structure simulations in minutes
- Represent realistically large structures atomistically
- Utilize available peta-scale computers

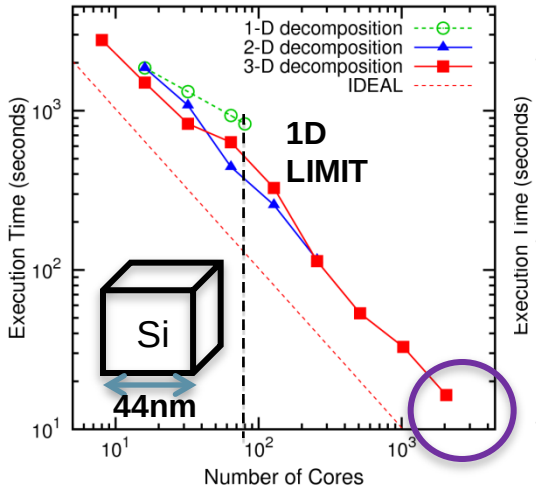
Approach

- Improve decomposition parallelism
- Include self-consistent calculations

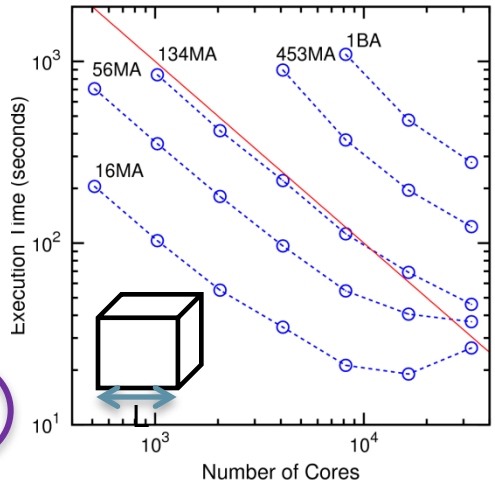
Performance / Impact

- Strong scaling of 3D decomposition => have overcome 1D limitation
- Utilized over 10,000 cores for realistic simulation domain size
- Expecting dramatic reduction of simulation time => 10hours to 20 minutes
- Published in SCIDAC proceedings, 2009, Haley, Lee et al.





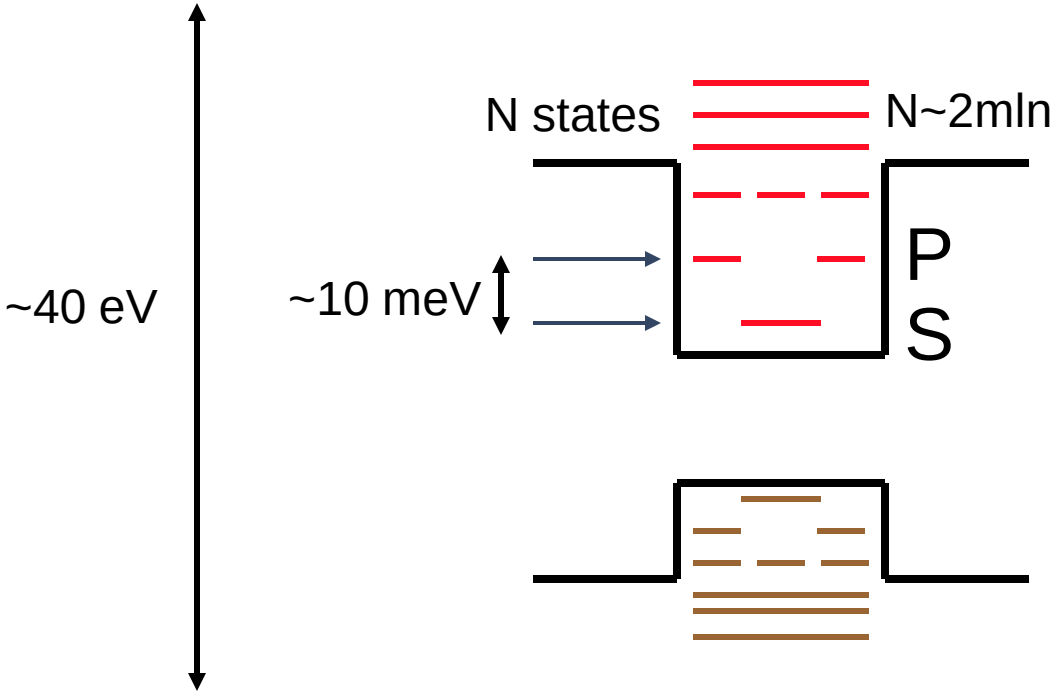
Tested on Ranger



Tested on Kraken

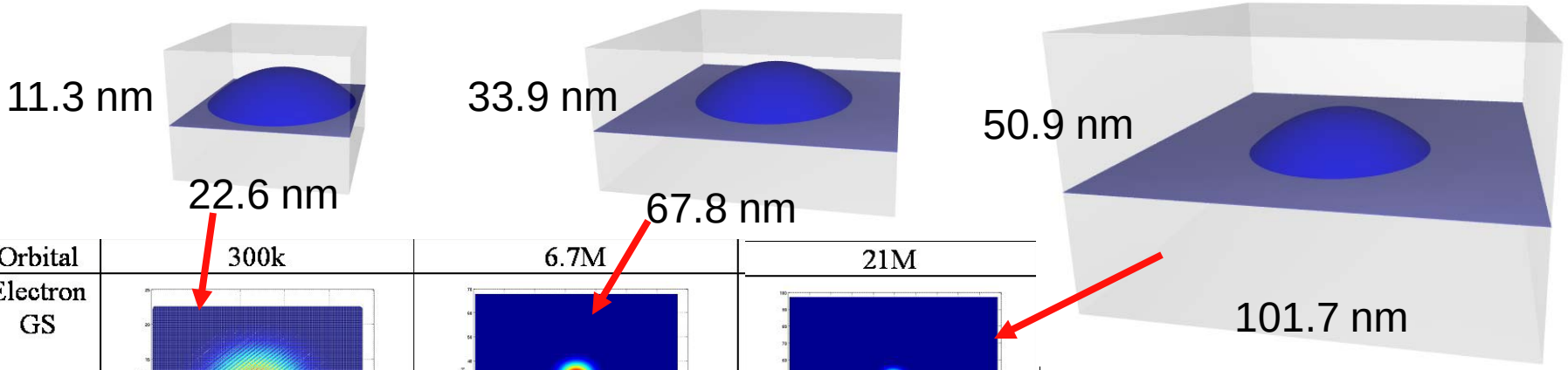
Proof of Concept

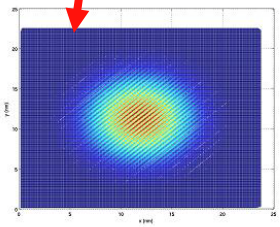
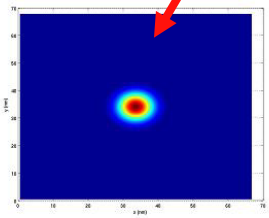
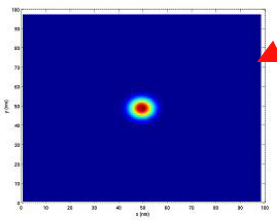
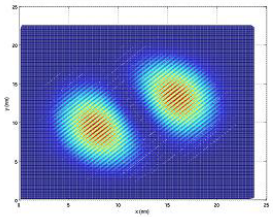
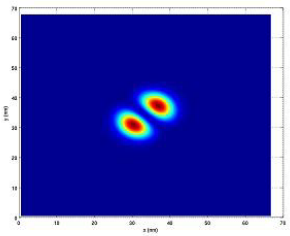
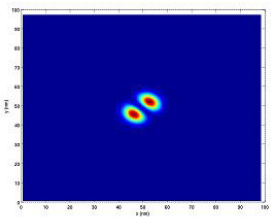
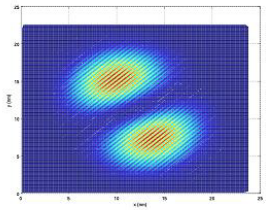
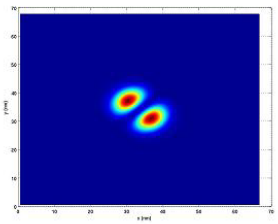
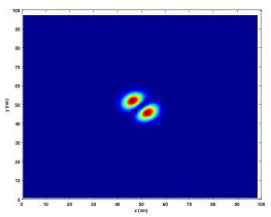
Extraction of Targeted Interior Eigenstates

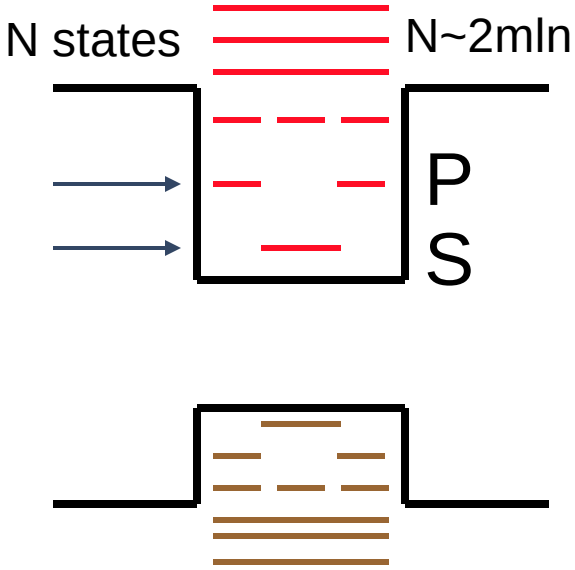


Proof of Concept

Extraction of Targeted Interior Eigenstates

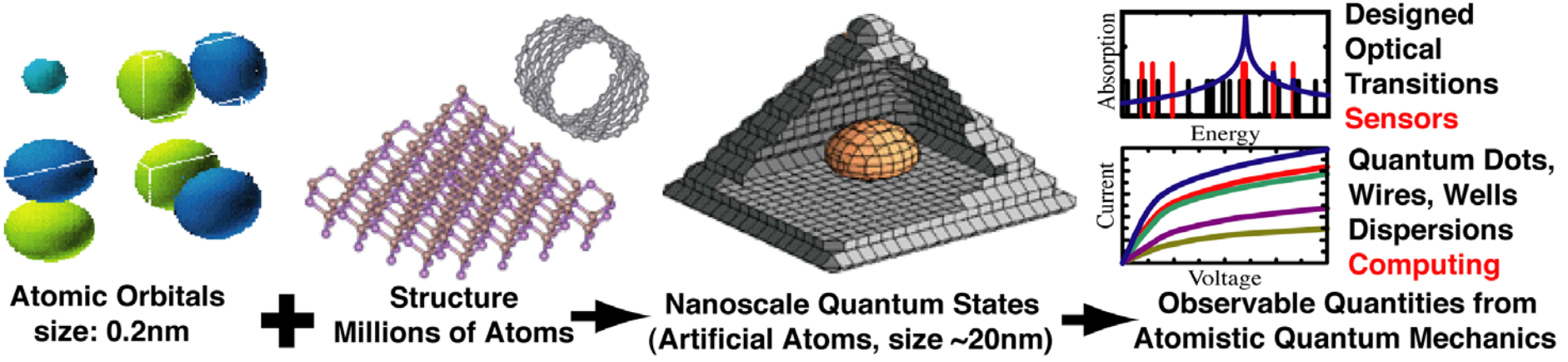


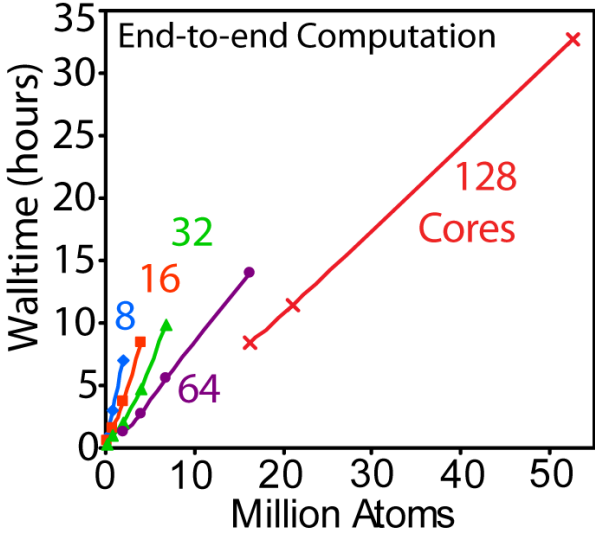
Orbital	300k	6.7M	21M
Electron GS			
Electron 1ES			
Electron 2ES			



Unique and targeted eigenstates of correct symmetry can be computed in all electronic computational domains

NEMO 3-D Technical Capabilities





Demonstration / Capability / Impact:

- 52 million atom electronic structure $(101nm)^3$.
- Quantum dots, nanowires, quantum computing…

Capabilities:

- Arbitrary 3D zincblende, wurzite crystals
- Strain in Valence Force Fields
- Piezo Electricity through potential perturbation
- Electronic Structure with empirical tight binding

Applications:

- InAlGaAs/GaAs quantum dots
- Valley splitting in Silicon
- Alloy disorder in bulk, QD, nanowires
- Single impurities in Si