Future Proof Parallelism for Electron-atom Scattering Codes on Hector

Andrew Sunderland, Cliff Noble, Martin Plummer

Advanced Research Computing and Atomic and Molecular Physics Groups

STFC Daresbury Laboratory

Science & Technology Facilities Council

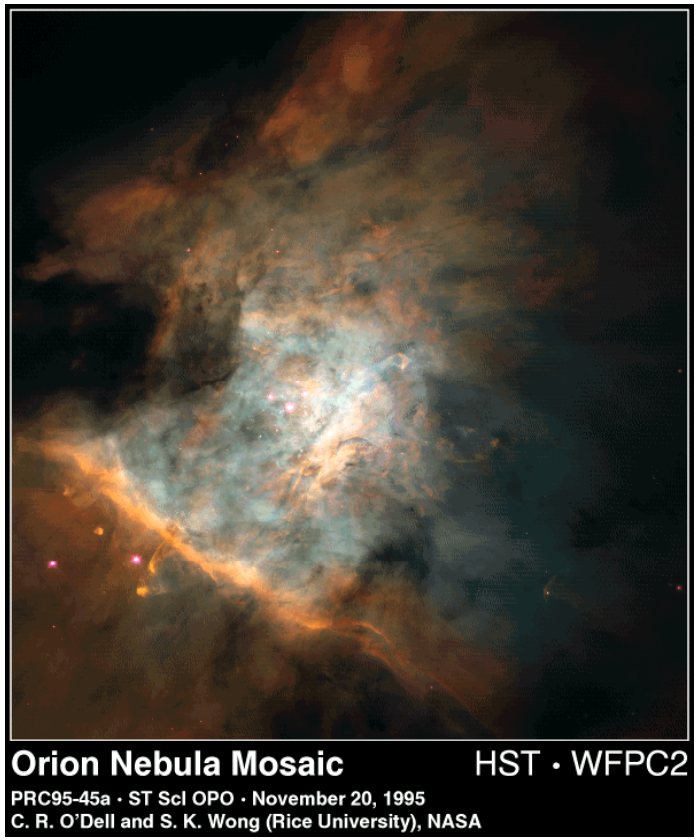
Summary

- Very Brief Background to R-matrix approach
  - Baluja-Burke Morgan method
  - Results of interest
- Background to PFARM code
  - Design Features
  - Performance characterization
- EPSRC DCSE project for code optimization on Hector
  - Performance analysis on the XT4
  - Optimizations performed to date
  - Ongoing / future plans to expand the code

Update on ALD propagator program

Science & Technology Facilities Council

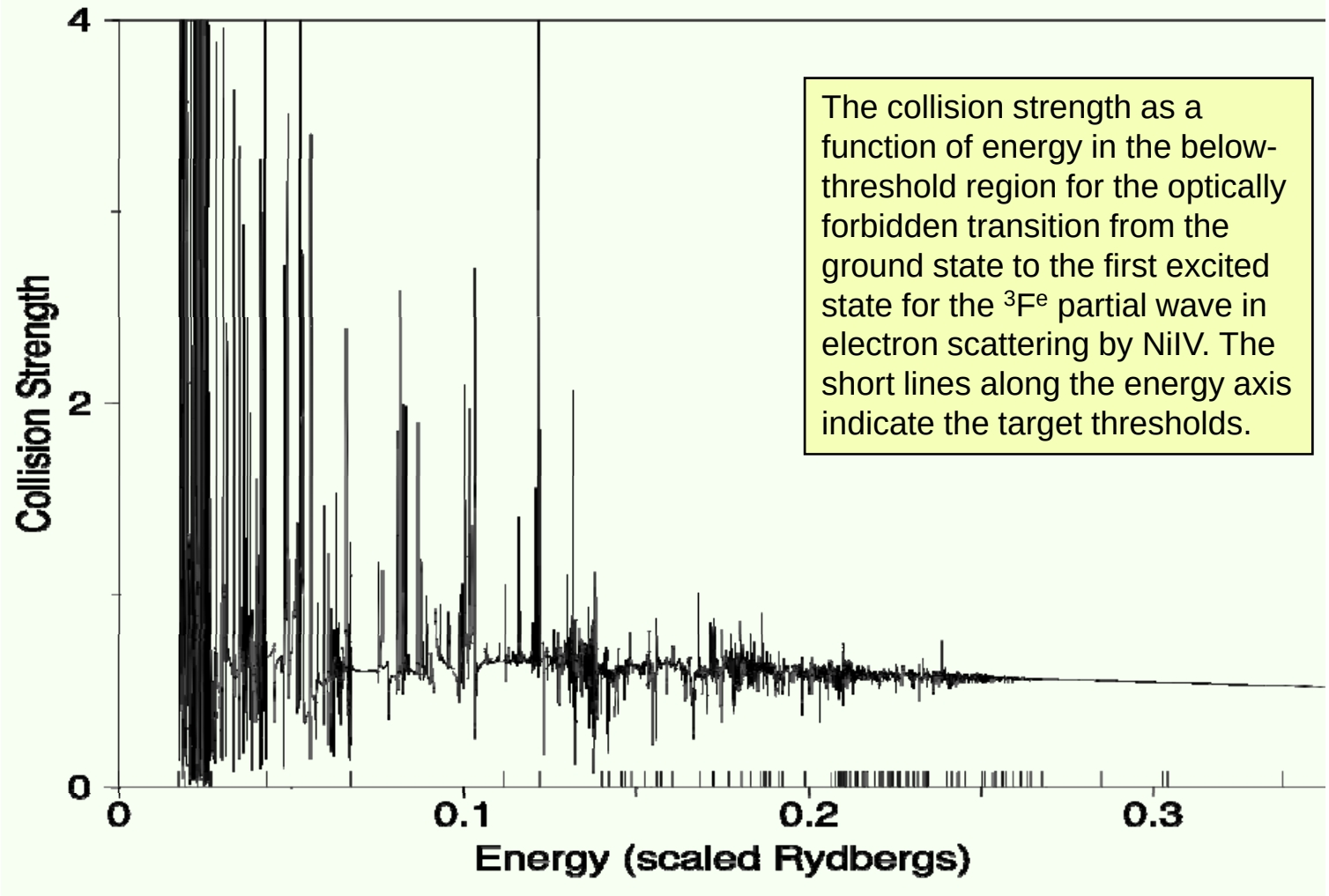
Electron-Atom Collisions



- Detailed electron-atom collision data is essential for understanding the behaviour of plasmas such as
  - Identifying forbidden lines such as those corresponding to the excitation of Ni^+ seen in observations of the Orion nebula (NGC 1976).
  - Plasma diagnostics of impurities in plasma fusion tokamaks.
  - Tin ions in next-generation nanolithography tools.
- R-matrix theory provides efficient computational methods for investigating electron-atom and electron-molecule collisions.
- Calculation involves integration of very large sets of coupled second-order linear differential equations. This presents huge computational challenges.

Science & Technology Facilities Council

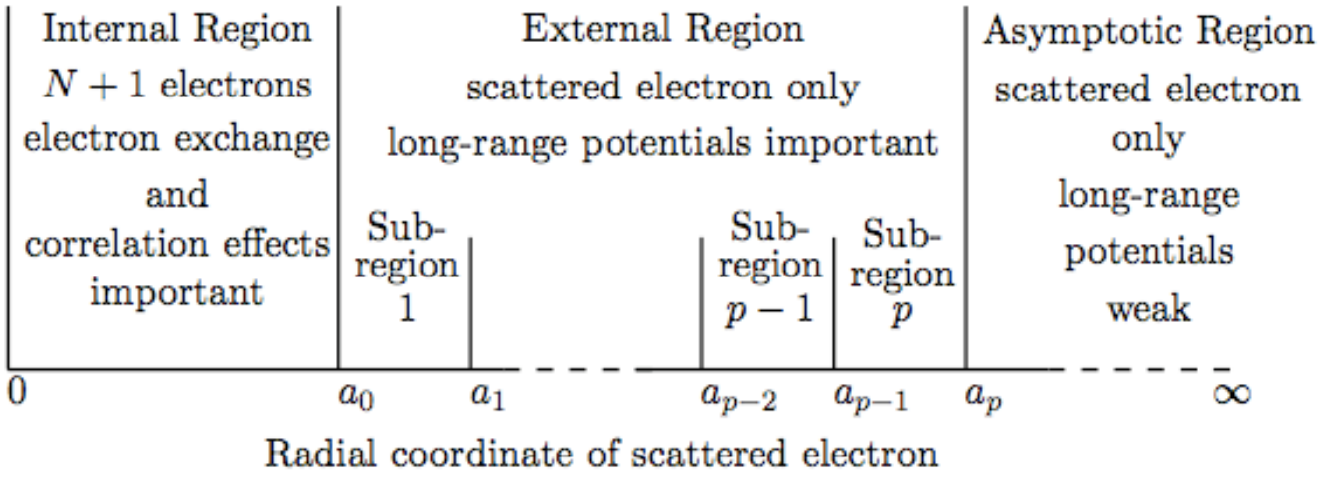
Results: Collision Strengths



The collision strength as a function of energy in the below-threshold region for the optically forbidden transition from the ground state to the first excited state for the $^3F^e$ partial wave in electron scattering by NiIV. The short lines along the energy axis indicate the target thresholds.

Science & Technology Facilities Council

Partition of Configuration Space



The parallelization of the code maps closely to this partitioning

Science & Technology Facilities Council

Baluja-Burke-Morgan Method in the External Region Calculation

Used to solve the non-relativistic Schrodinger equation:

$$\mathbf{H}_{N+1}\Psi = \boldsymbol{E}\Psi$$

- R-matrices (inverse log-derivative matrices) at successively larger radial distances are obtained using Green's functions defined within finite radial sectors
- Green's functions are obtained using a shifted-Legendre basis.
- Diagonalize representative of the Green's function $(H + L - E\mathrm{I})^{-1}$ within a basis.

R-Matrix Propagation Program For Solving Coupled Second-order Differential Equations, *K.L. Baluja, P.G.Burke and L.A.Morgan, Computer Physics Communications 27 (1982) 299-307*

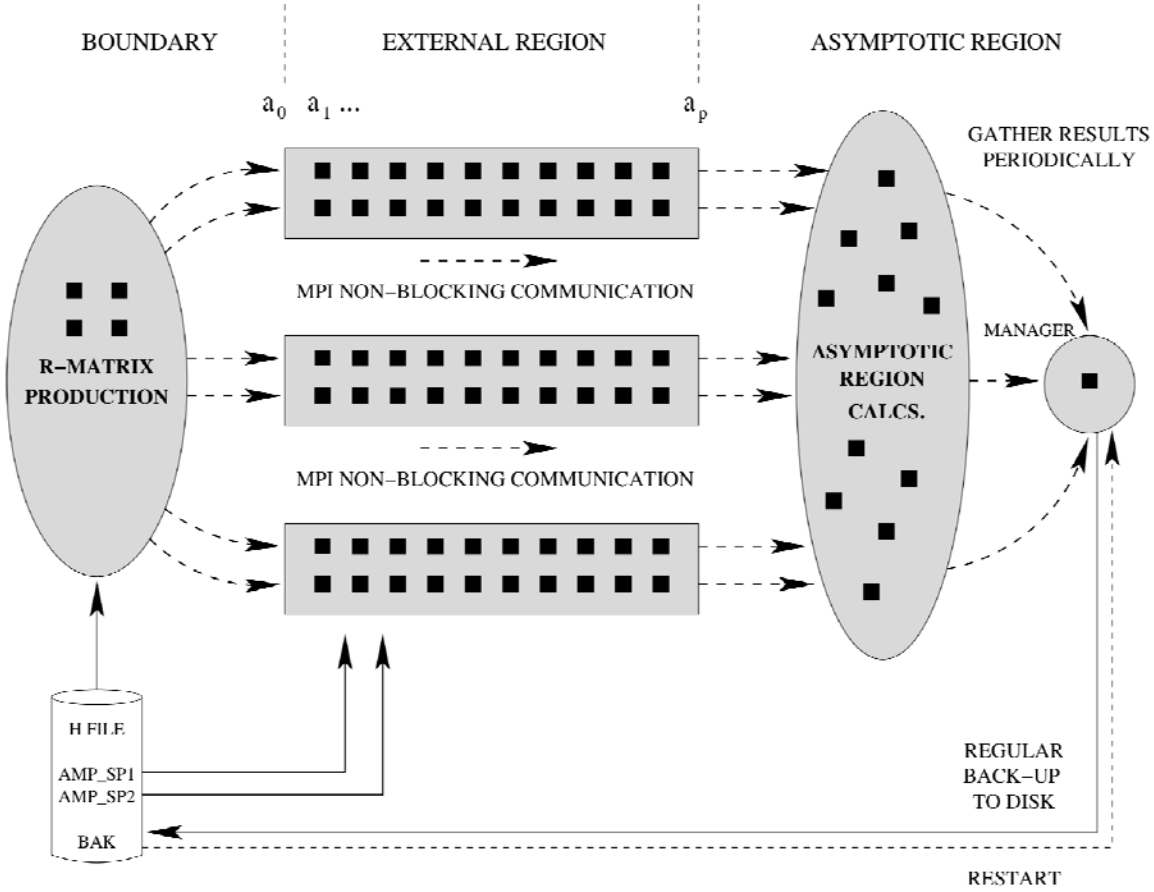
Baluja-Burke-Morgan (BBM)-based Implementation

2 Stage Parallelization of BBM approach in the external region:

- EXDIG Program:
  - Diagonalize Sector Hamiltonian matrices using ScaLAPACK (Blacs-based Data decomposition).
- EXAS Program:
  - For each scattering energy propagate using 3 functional groups:
  - Generate initial R-Matrix (Data decomposition).
  - Propagate R-Matrix across each sector in pipeline (Control decomposition).
  - Calculate thermally averaged collision strengths (Task Farmed).

Science & Technology Facilities Council

EXAS Parallelization



Features of 2-stage PFARM Code

Scalable performance for a wide range of problem sizes:

- Replication of propagation pipelines (facilitated by MPI Groups and MPI Communicators)
- Fully flexible configuration to achieve optimal load balancing.
- Asynchronous characteristics permit effective overlapping of communication and computation (MPI Asynchronous, NonBlocking Send/Receive).
- Code is developed using standard Fortran95, MPI and numerical library routines from BLAS, LAPACK, PBLAS, ScaLAPACK
  - Produces portable and optimized code

Science & Technology Facilities Council

Physics-based optimizations: Scattering Energy Mesh Partition & Spin-splitting

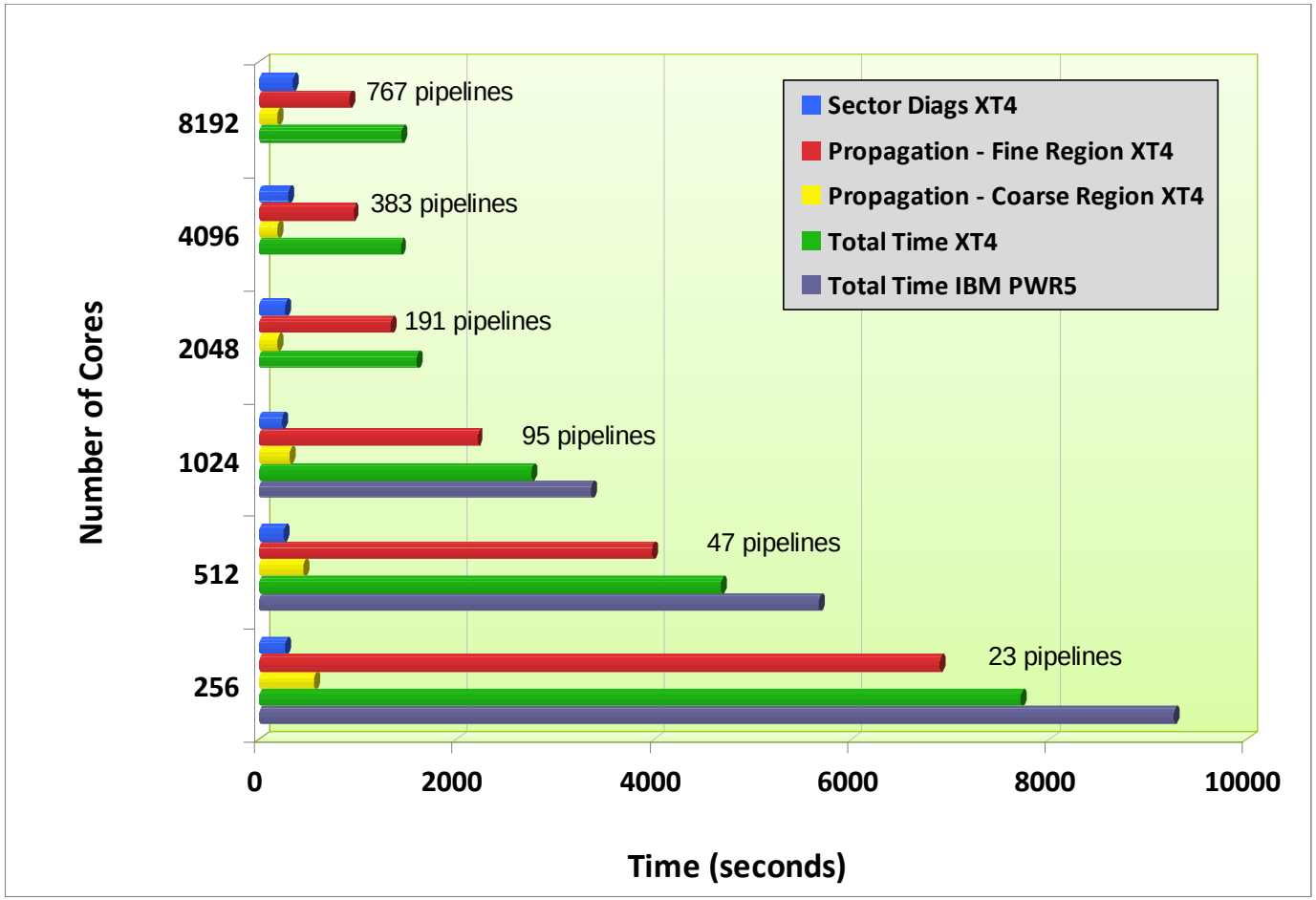
- Vast majority of scattering energy points reside in the *fine* section of the energy mesh in the resonance region below the final threshold.
  - Highest energy in the *coarse* section of the energy mesh is usually around 4 *x* highest scattering energy in the fine section.
  - **Separate the fine and coarse mesh calculations** and maximize sector length, thereby reducing pipeline processors for each run.

- **Decouple channels associated with targets with different spin** *'spin-splitting'.*
  - Results in two smaller streams of data through the external region calculation thereby reducing computational and memory requirements

Science & Technology Facilities Council

Initial Performance on Hector XT4

FeIIIJ030, 10667 scattering energies



Optimization (i) – Sector Hamiltonian diagonalization

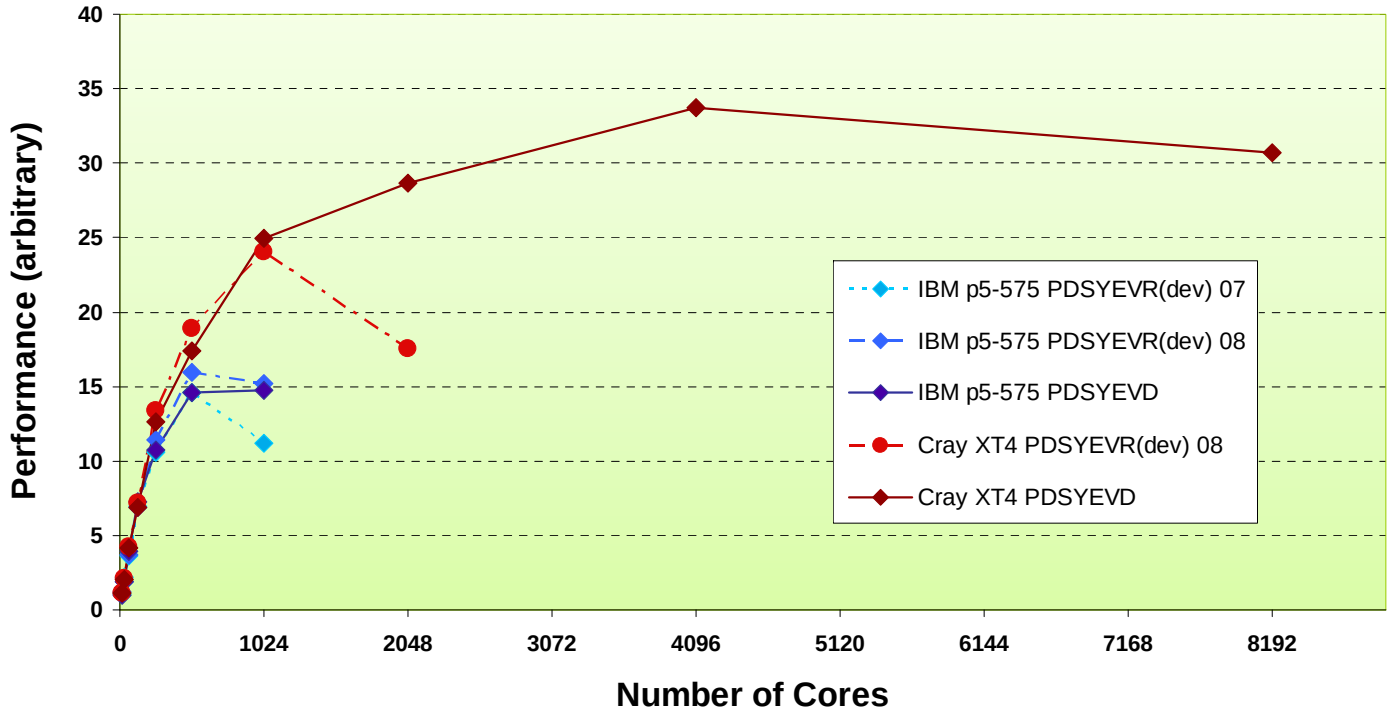
The 1st stage of the calculation EXDIG involves diagonalizing multiple sector Hamiltonian matrices:

- Dimension of sector Hamiltonian matrix can vary from 5000-100000+ depending upon the problem.
- Code currently steps through the sectors, undertaking a parallel diagonalization on each Hamiltonian using ScaLAPACK routines using **all the processors**.
- Optimizations:
  1. Performance analysis of available diagonalization routines to determine best method (e.g. Divide-and-Conquer vs Multiple Relatively Robust Representation)
  2. Divide the processors into sub-groups (split Blacs grids)
     - Sector matrices are distributed to each **sub-group where a parallel diagonalization** takes place independently of other sub-groups.

Science & Technology Facilities Council

Parallel Diagonalization Performance Analysis

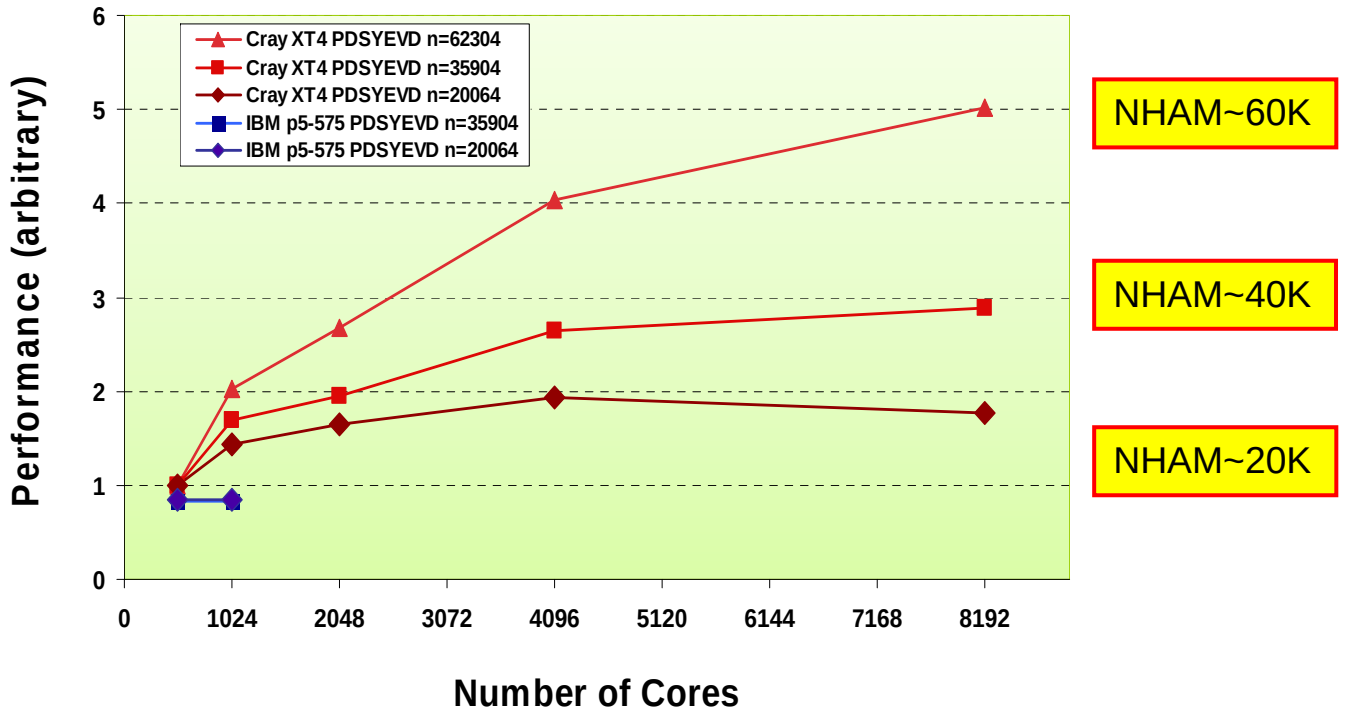
FeIII case NHAM=20064



Science & Technology Facilities Council

Parallel Diagonalization Performance Analysis for a range of problem sizes

FeIII cases

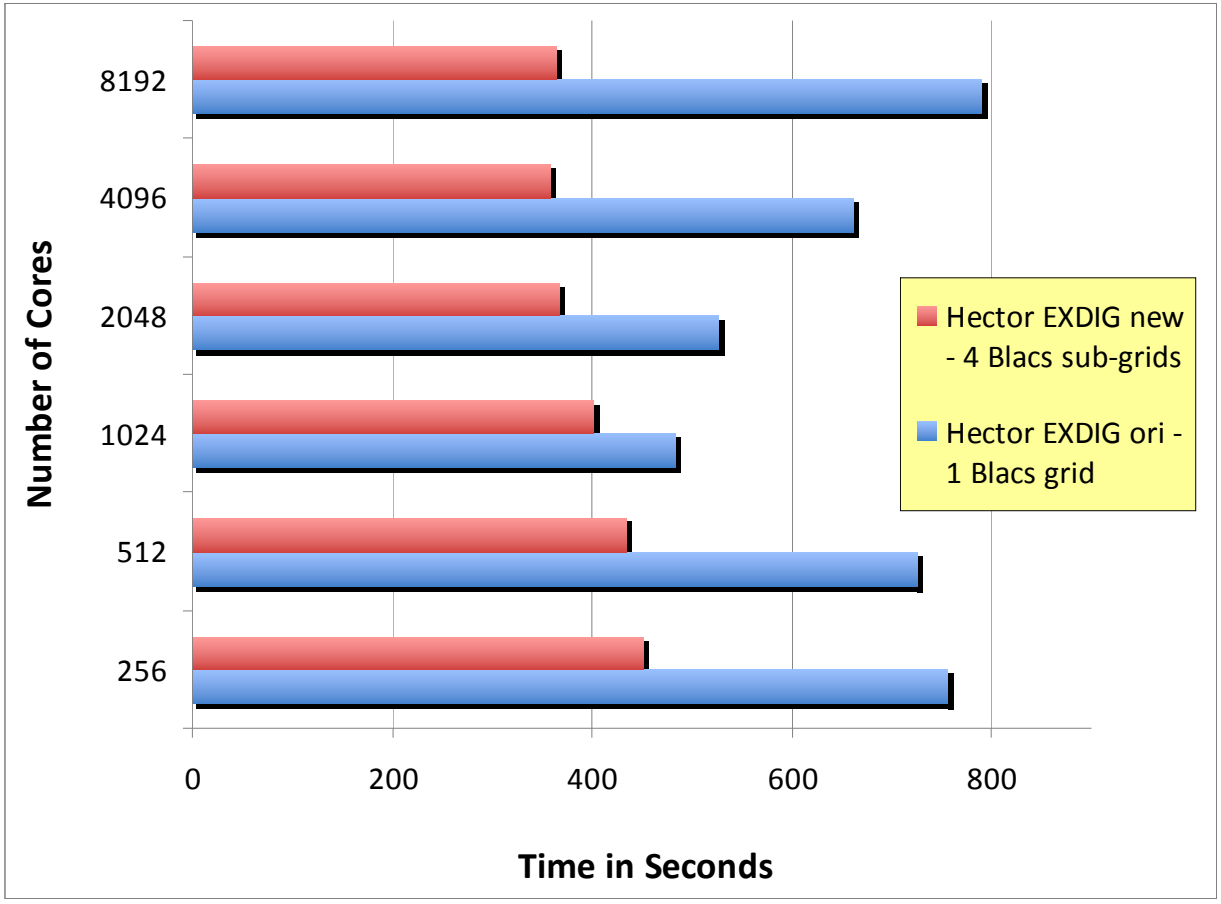


Parallel Scaling is relatively poor on high core counts unless the problem size is large. Splitting the global Blacs-grid into sub-groups for the sector diags helps mitigate this performance tail-off.

Science & Technology Facilities Council

Optimized EXDIG performance

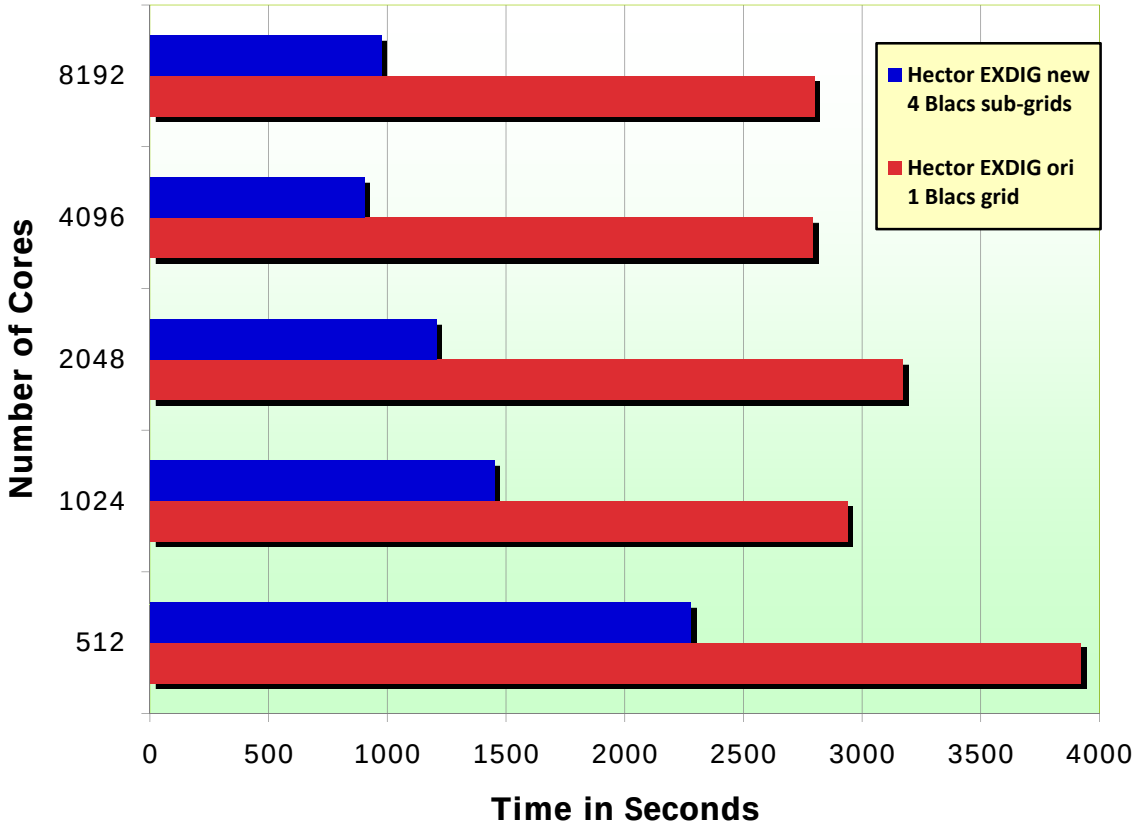
FeIII case, NHAM = 11810



Science & Technology Facilities Council

Optimized EXDIG performance

FeIII case NHAM = 44878

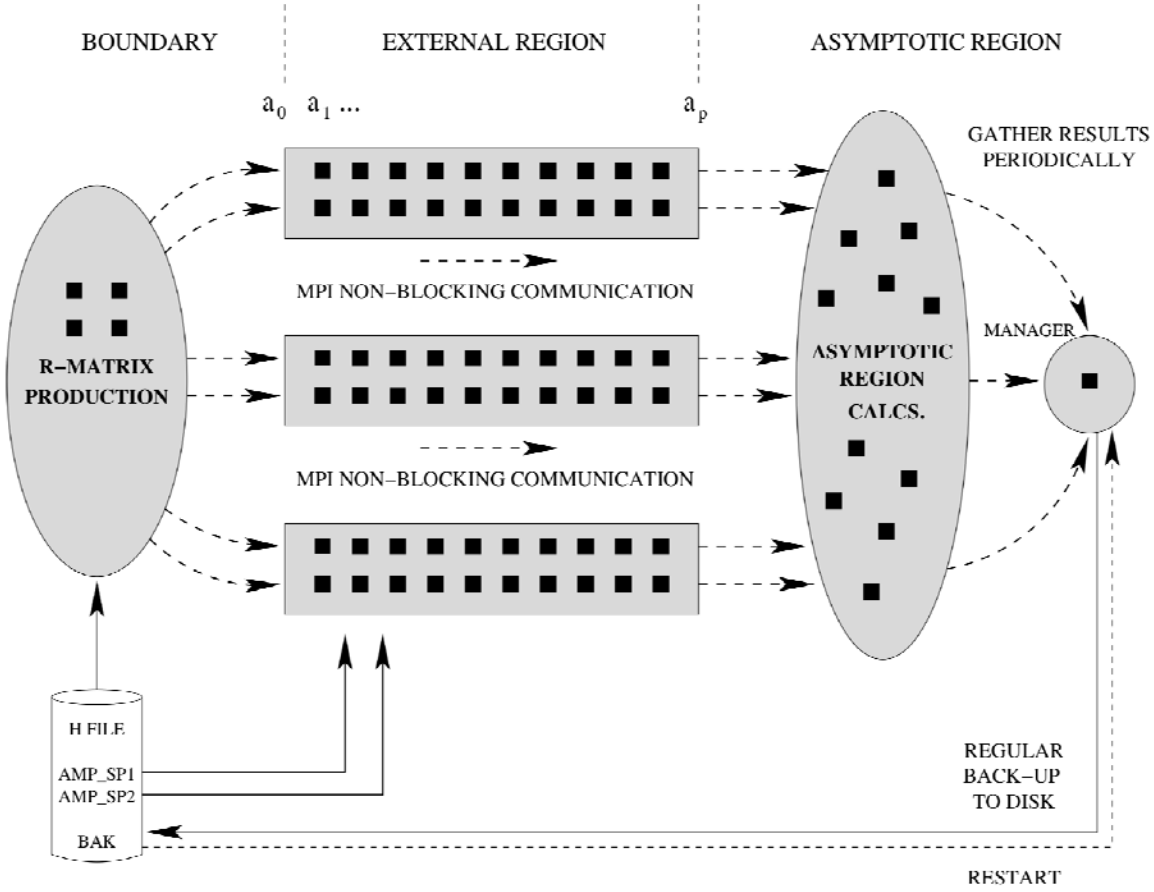


Optimization (ii): Load-balancing PFARM for the XT4

- Load-balancing the functional groups is key to good parallel performance in the propagation stage (EXAS).
  - **Initial R-matrices** need to be **produced** at a sufficient rate to satisfy **demand** from the processor pipelines.
  - Asymptotic calculations must be processed at a sufficient rate to deal with the **supply** of **Final R-matrices** from the processor pipelines.
  - Pipelines must not be held up!
- The size of each group is currently determined by the user (difficult to judge).
- Basic wrapper Perl scripts have been developed to help automate this process, based on performance analyses of the functional groups with different problem sizes.
  - Performance analyses were based on the T3E and need updating for the XT4
  - More sophisticated automation required

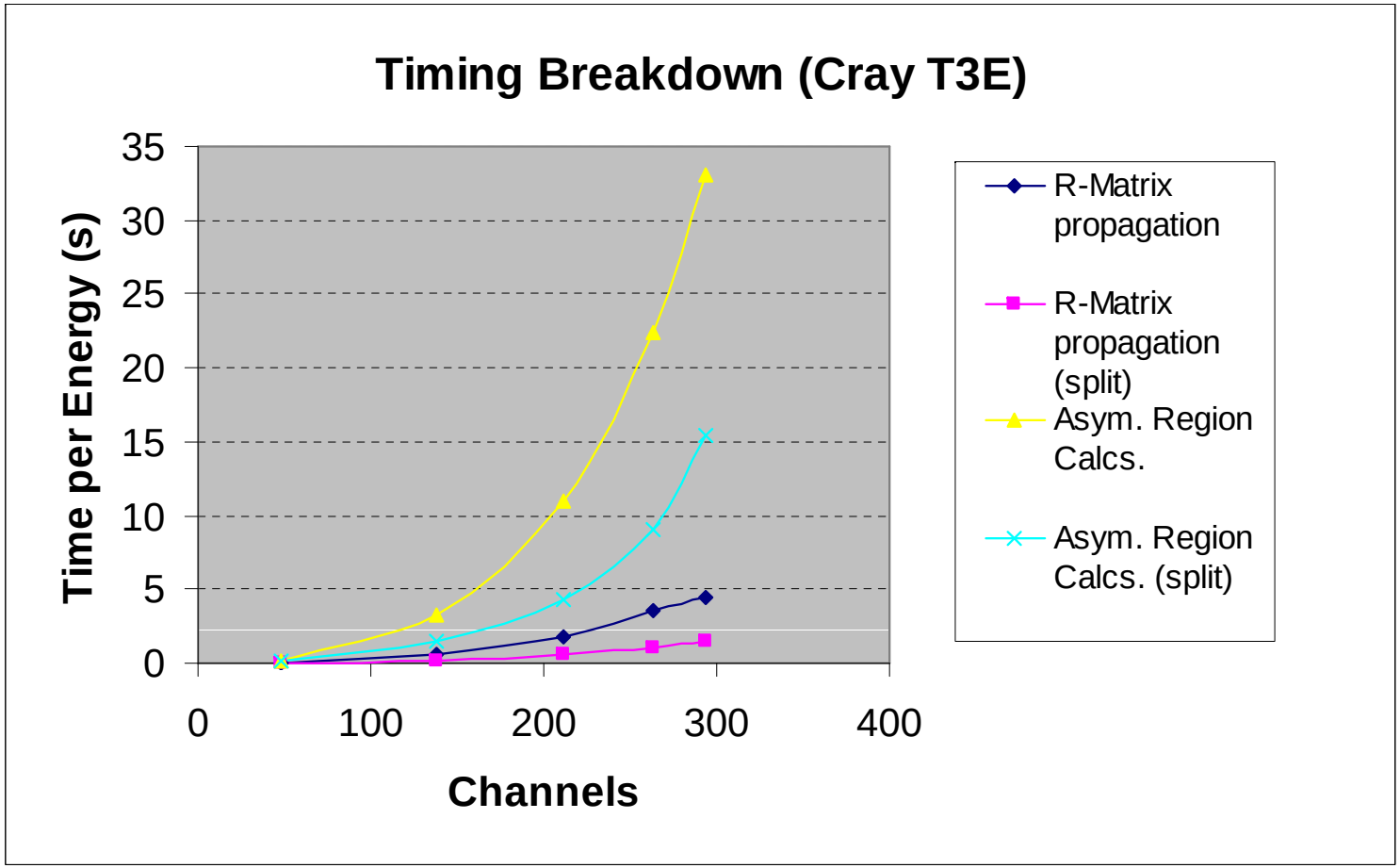
Science & Technology Facilities Council

EXAS Parallelization

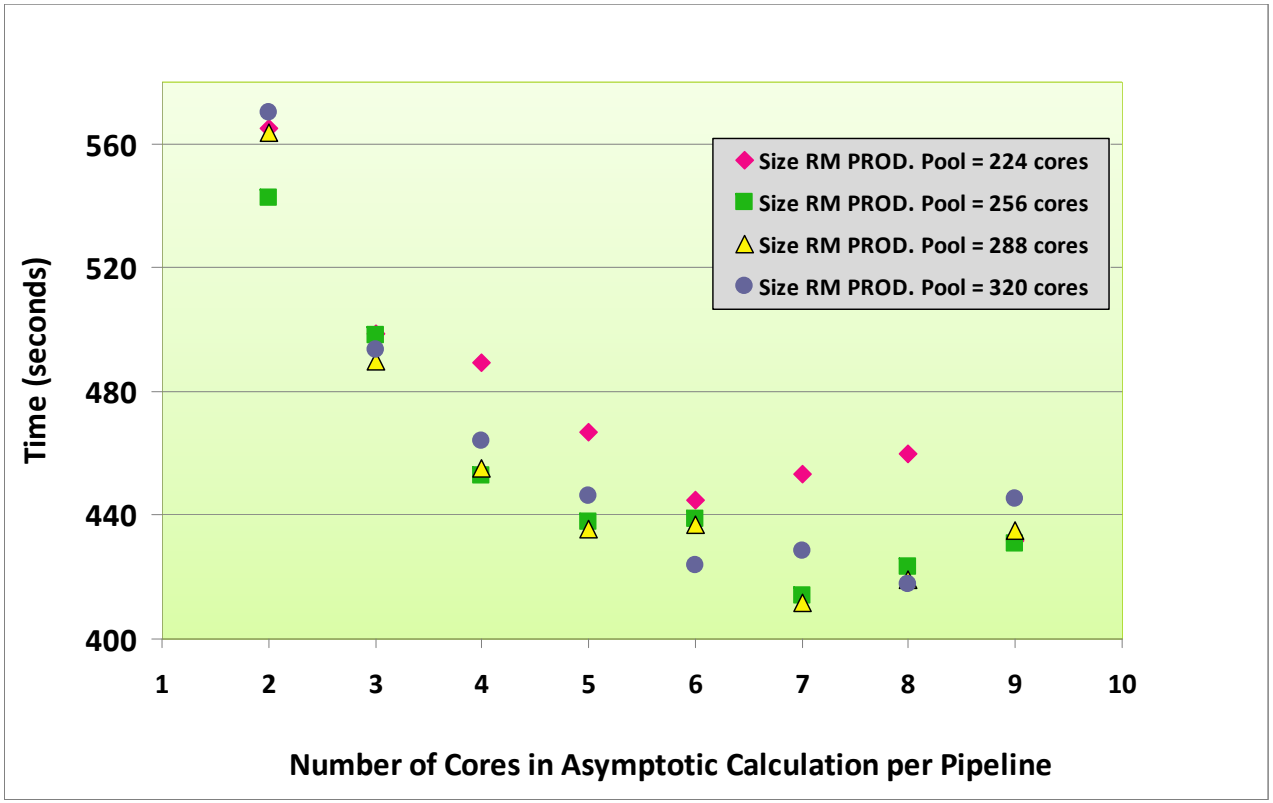


Science & Technology Facilities Council

Example: Load-balancing the pipelines with the asymptotic pool



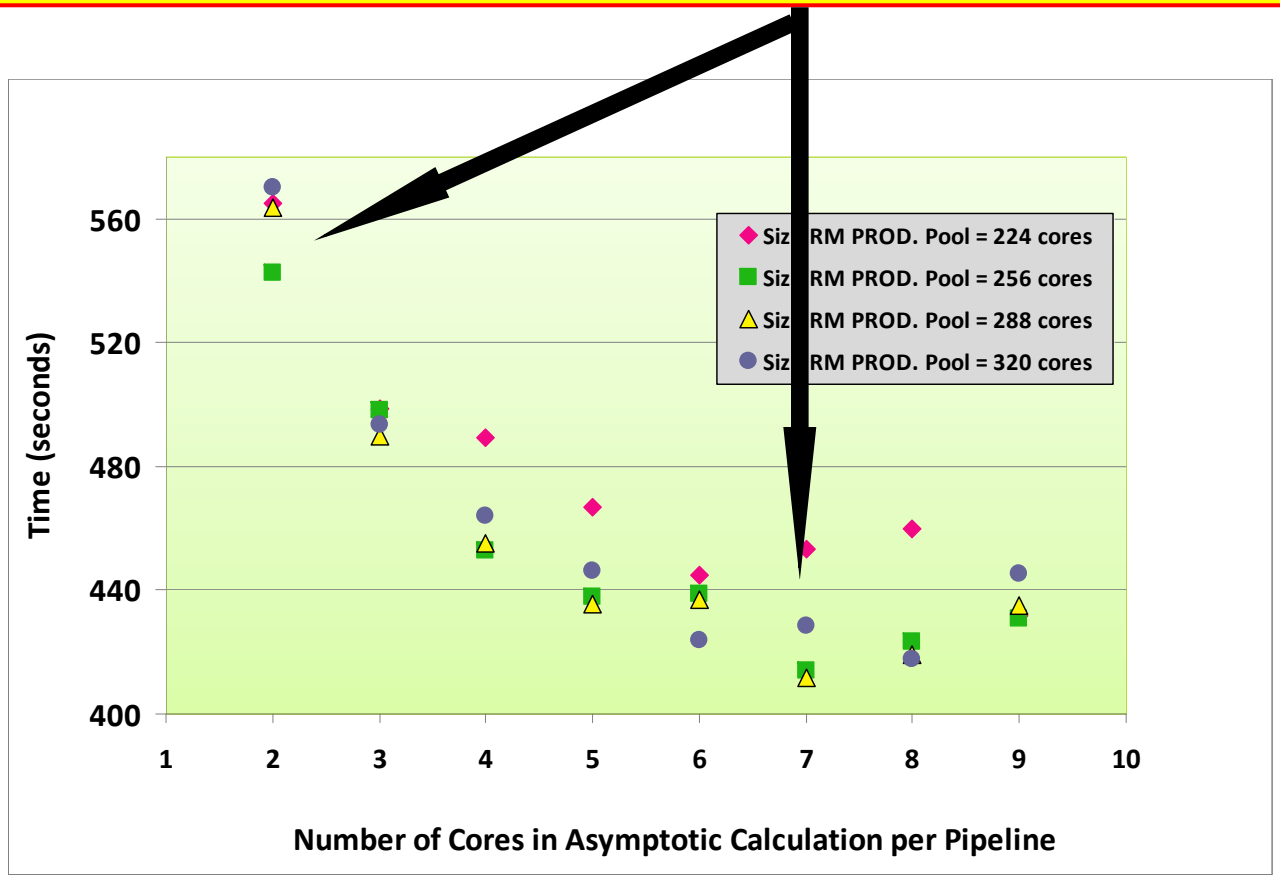
Load-balancing a 1024 core job on the XT4



Science & Technology Facilities Council

Load-balancing a 1024 core job on the XT4

Upto 30% improvement in speed from correct load-balancing

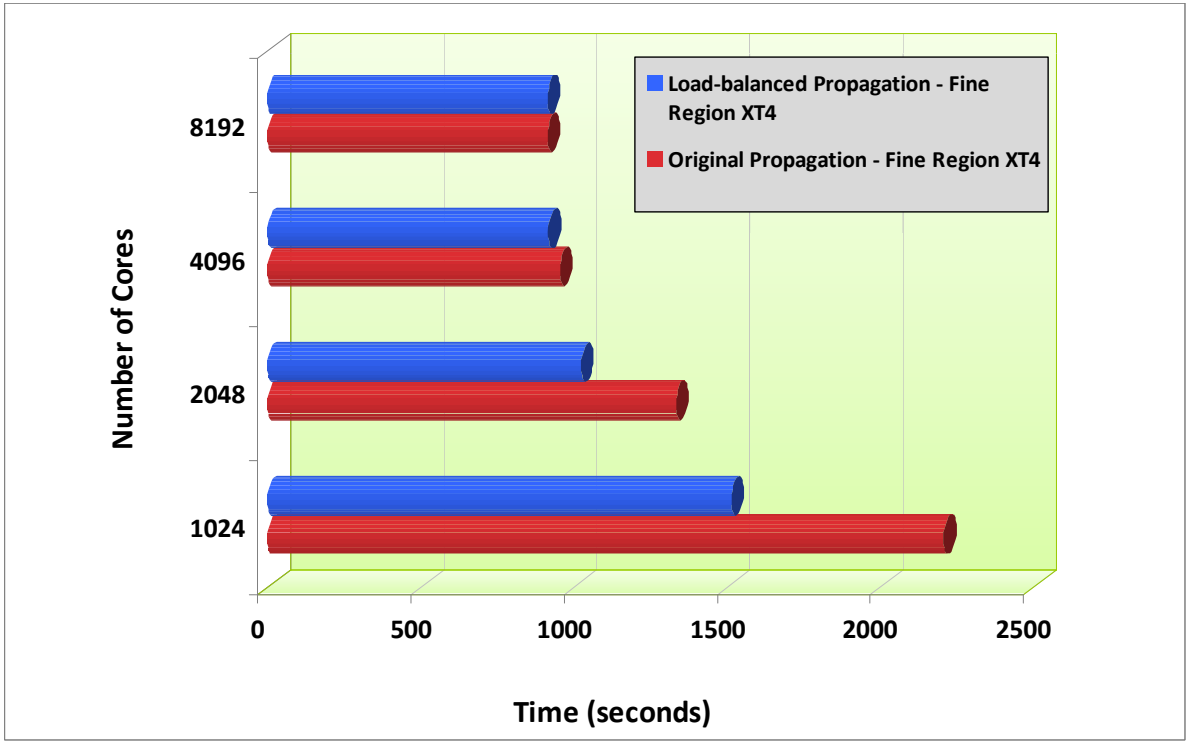


Science & Technology Facilities Council

Load-balanced EXAS performance on the XT4

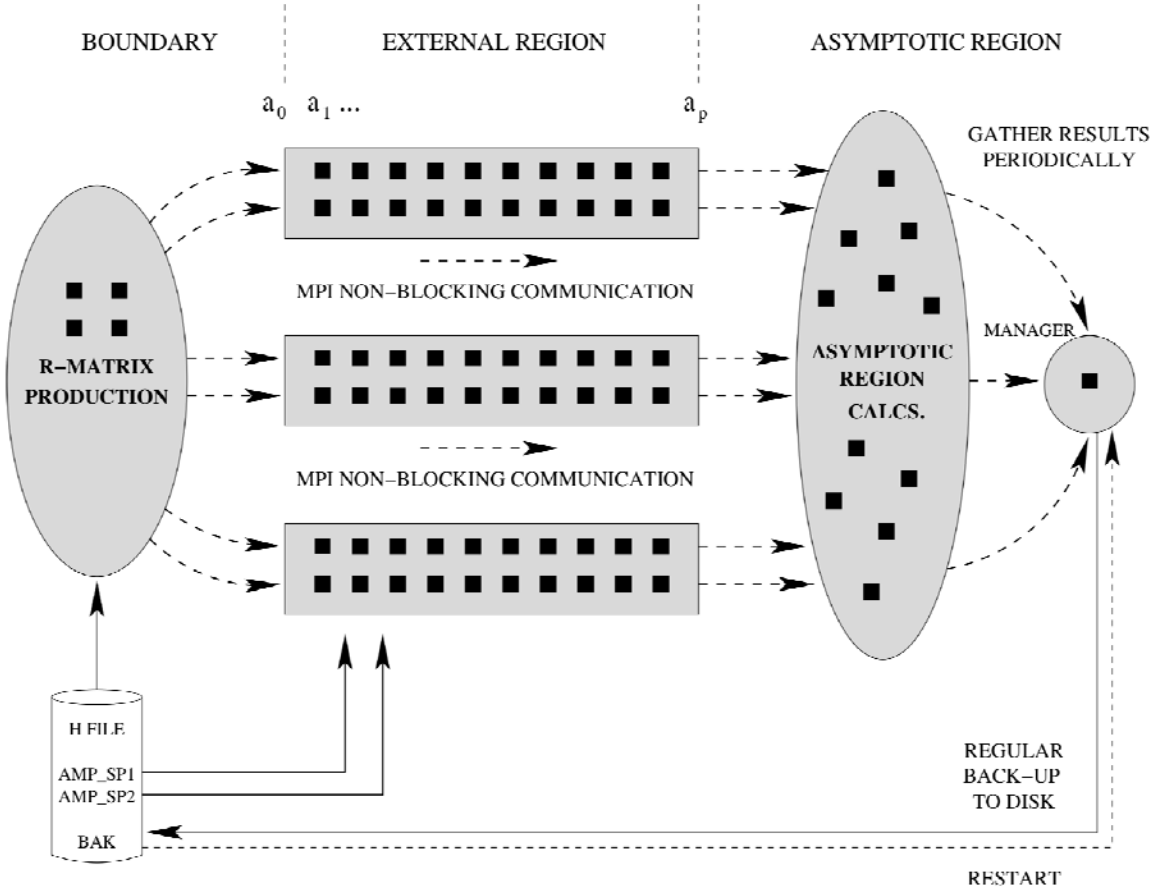
Fine region propagation for FeIIIJ030, 10667 scattering energies

Limited impact at high core counts. Here bottleneck lies elsewhere – probably with manager process



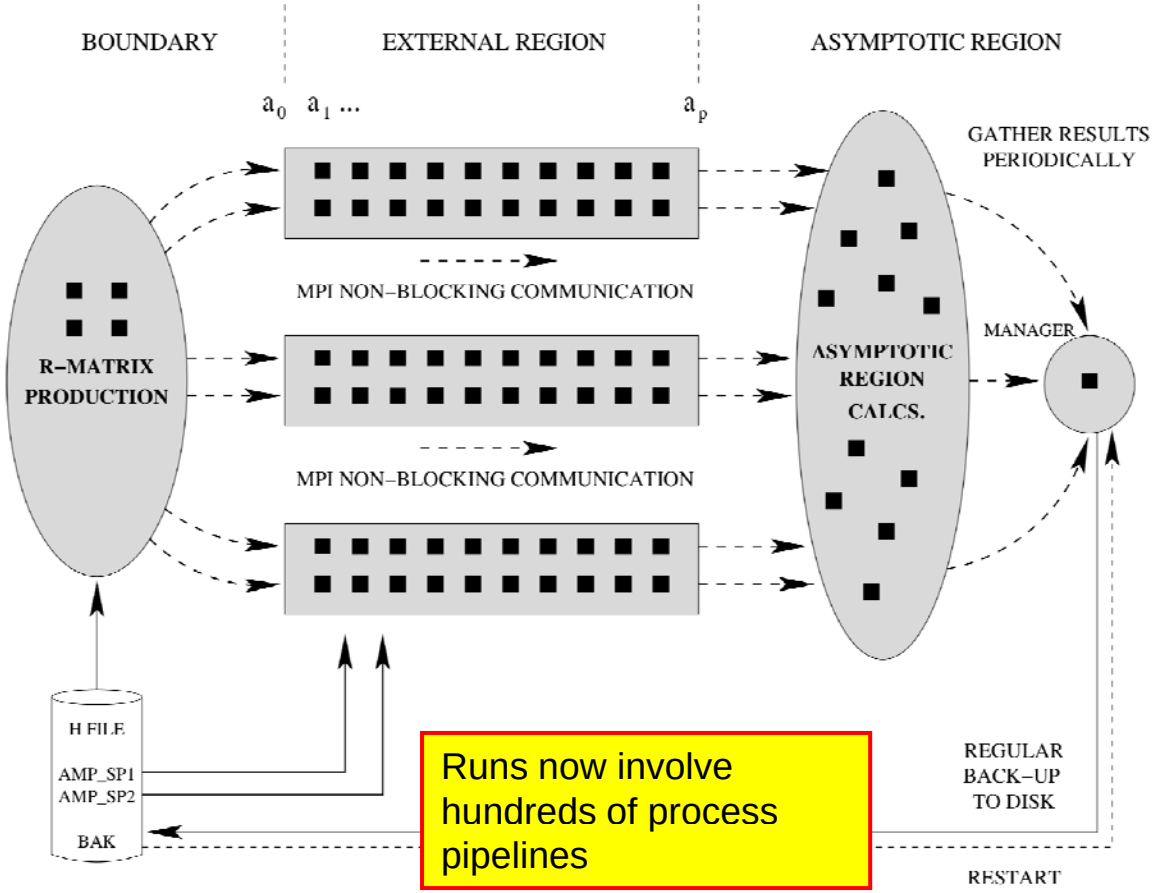
Science & Technology Facilities Council

EXAS Parallelization



Science & Technology Facilities Council

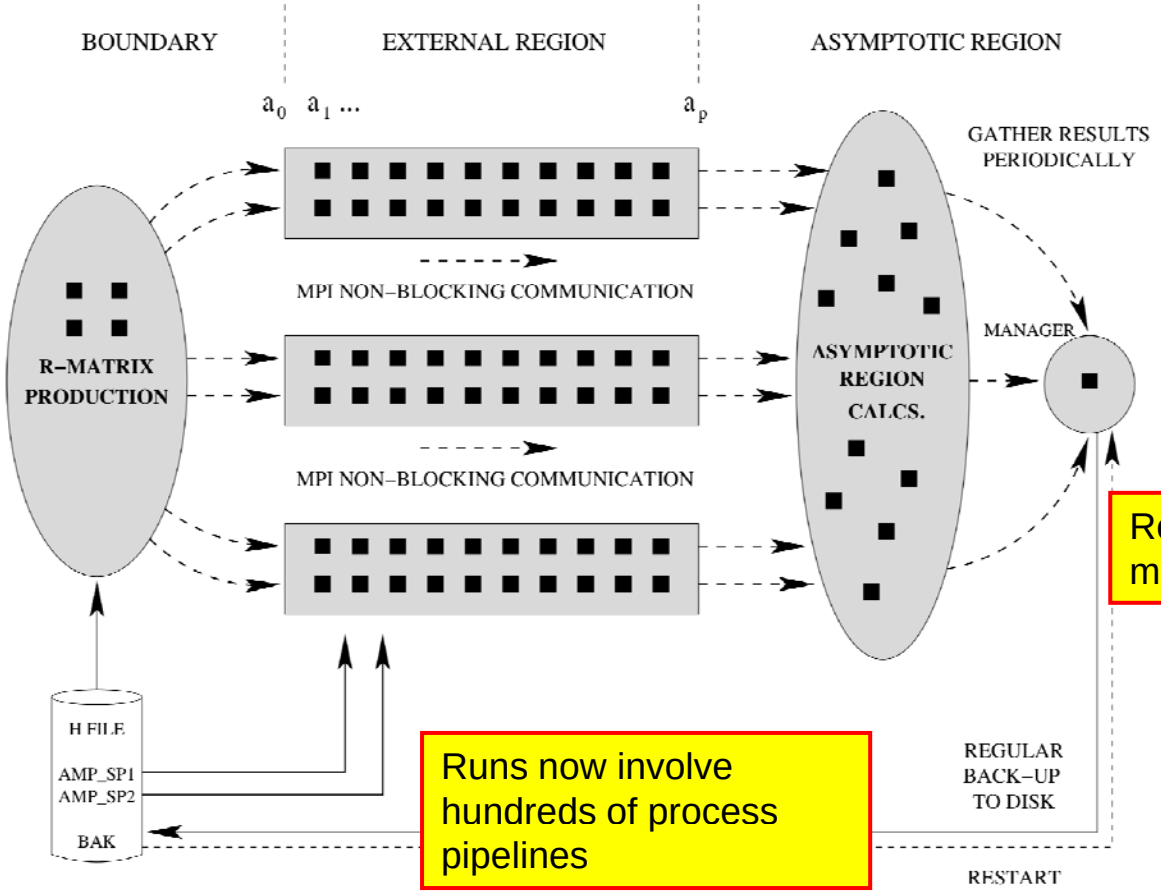
EXAS Parallelization



Runs now involve hundreds of process pipelines

Science & Technology Facilities Council

EXAS Parallelization



Runs now involve hundreds of process pipelines

Requires multiple manager processes

Future Calculations with EXDIG

- More ambitious calculations (more channels, longer range and more complex potentials (particularly for molecules and applications to multiple scattering) imply :
  - More sectors and larger sectors
  - Larger BBM basis sets (and sector Hamiltonians)
- The EXDIG optimizations are designed to ensure that these more ambitious calculations will improve scaling.
- Automate choice of number of sectors in each farm (hence the number of PEs for each ScaLAPACK diagonalization) to produce optimal combination of ScaLAPACK performance and concurrent multiple diagonalizations

Current and Future Work

- Introduce new sub-manager communicator (i.e. more than one manager process) for processing, gathering and writing final results
- Further automate load-balancing on XT4
- Update for quad-core
  - Based on runs involving a wide range of problem sizes
- Pipelines set-up currently restricted to one sector calculation on one process (legacy of limited memory availability on past machines)
  - Redesign to allow > 1 pipeline sector per core to reduce communication
  - Map efficiently to multi-core architecture to reduce communication
- I/O is now a substantial bottleneck
  - Introduce MPI/IO where appropriate
  - Cliff Noble has developed an XStream library based on MPI/IO with XDR binary format
- Interface Airy LD propagator method with existing code

Science & Technology Facilities Council

The ALD propagator (a brief run-through)

- The BBM propagator uses a basis set to approximate the full potential over a possibly large radial sector. Linear algebra/diags have: n = n channels * n basis

- The ALD propagator (MH Alexander, J Chem Phys 81 (1984) 4510; MH Alexander and DE Manolopoulos, J Chem Phys 86 (1987) 2044) is potential-following and assumes a linear reference potential for each sector (the original Light-Walker propagator assumed a constant potential in each sector).

- The Airy functions and their derivatives used as solutions to the sector equations are of the same order of difficulty to calculate as the trigonometric functions used in LW.

- Basis sizes are n_channels

- We are replacing a ‘few’ long sectors with a larger number of shorter sectors: parallelism efficiency emphasizes pipelining over Scalapack except for the most ambitious problems.

- As the method is potential-following, the only fine/coarse division is the number of energies required over the resonance region (ALD accuracy is in principle energy-independent)

Technical details from ALD work

- In the non-exchange region, channels split into two sets which are coupled separately in the outer region (they are coupled by the inner region boundary conditions).

- Derived-type for split channel operations: (spin splitting or K-spin splitting). Matrix operations including inversion are block-partitioned.

n1	n2	'dsyt-'	'dget-'	svd	ptf
500	500	0.5, 1.1	0.6, 0.6	3.9, 4.2	0.4, 0.4
2000	2000	33.6, 42.1	36.0, 33.9	257, 241	21.5, 21.3
3000	3000	123, 118	120, 113	877, 805	74, 71

- (times for inversion of n1+n2 matrix, HPCx, HECToR)
- Parallelization within the derived type as required

- XStream I/O handling:

  - R-Matrix codes consist of various stages (eg, radial, angular, hamiltonian inner region, diag and propagate outer region). The serial codes used direct access files to pass data: not always portable
  - PFARM uses XDR files to read inner region data and between stages. MPI-IO files are more efficient in parallel but not portable.
  - XStream provides a wrapper to allow either option at any given file read/write, details of the communicator are held within the derived type.
  - In use in PFARM

- Coefficient packing: data from inner region concerning the channel potentials is reordered and packed according to the channel splitting

- Intel's 'vtune' tool is proving to be useful in investigating SMP parallelism and caching and threading optimization

- Note that with ALD, in all but the largest current cases, the 'difficult' parallel work can be kept within a multicore (or serial for smaller cases) with the emphasis on the 'dull' but simpler energy parallelization and 'reasonably straightforward' pipelining. Scalapack/PFARM special expertise will come in for the most demanding new calculations.