Sparse Days and ANR SOLSTICE Final Workshop

June 15-17th, 2010

CIC Météo-France, Toulouse, France

Tuesday, June 15th 2010

9.30 - 10.00 Registration and coffee break

Session I

10.00 - 10.40 *Computing the diagonal of the inverse of a matrix*
Y. Saad

10.40 - 11.10 *Generalized Golub-Kahan bidiagonalization and stopping criteria*
M. Arioli

11.10 - 11.25 Break

11.25 - 11.55 *Risolv - Robust Iterative Solver*
H. Tal-Ezer

11.55 - 12.25 *Efficient Gram-Schmidt orthogonalization with CUDA for iterative eigensolvers*
A. Tomas

12.25 - 12.55 *On the residual norm in FOM and GMRES*
G. Meurant

Lunch

Session II

14.00 - 14.40 *Multilevel block preconditioning for shifted Maxwell equations*
M. Bollhöfer

14.40 - 15.10 *Approximate Factoring of the Inverse*
M. Byckling

15.10 - 15.40 *The multicore challenge: the sparse indefinite case*
J. Scott

15.40 - 16.10 Coffee break

Session III

16.10 - 16.40 *Combinative preconditioning based on Relaxed Nested Factorization and Tangential Filtering preconditioner*
P. Kumar

16.40 - 17.10 *A Parallel Neutron Transport Solver based on Domain Decomposition*
S. Van Criekingen

17.10 - 17.40 *3D CFD Simulation Using a Parallel Hybrid Approach of the Block Cimmino Iterative Method*
M. Zenadi

Wednesday, June 16th 2010

Session IV

9.00 - 9.40 *Sparse Matrix Computation in Large-Scale Scientific Applications*
E. Ng

9.40 - 10.10 *The MUMPS library: work done during the Solstice project*
MUMPS team

10.10 - 10.40 *TBA*
PaStiX team

10.40 - 11.00 Coffee break

11.00 - 11.30 *High performance Methods to solve Large 3D Electromagnetic Problems*
D. Goudin

11.30 - 12.00 *EDF participation to the ANR Project SOLSTICE*
O. Boiteau

12.00 - 12.30 *Amélioration des schémas de transport, nouveau solveur chimique et parallélisation en Z du solveur de pression du modèle "MésoNH"*
J.-P. Pinty

Lunch

Session V

14.00 - 14.30 *Preconditioning the Schur complement in a hybrid direct/iterative solver*
J. Gaidamour, P. Hénon

14.30 - 15.00 *Parallel scalability and complexity analysis of the sparse hybrid linear solver MaPHYS*
E. Agullo, L. Giraud, A. Guermouche, A. Haidar, J. Roman,Y. Lee-Tin-Yen

15.00 - 15.30 Coffee break

Session VI

15.30 - 16.00 *Parallel preprocessing in MUMPS*
A. Buttari, B. Ucar

16.00 - 16.30 GRID-TLSE Project
P. Amestoy, F. Camillo, M. Daydé, R. Guivarch, A. Hurault, J.-Y. L'Excellent, C. Puglisi

20.00 Banquet at the restaurant "Chez Fazoul" in Toulouse.

Thursday, June 17th 2010

Session VII

9.00 - 9.40 *Preconditioners for stabilized finite element discretizations of the steady incompressible Navier-Stokes equations*
M. Benzi

9.40 - 10.10 *Sparse matrix partitioning, ordering, and visualisation by Mondriaan 3.0*
R. Bisseling

10.10 - 10.40 Coffee break

Session VIII

10.40 - 11.10 *PT-Scotch in Solstice and beyond: where to go now*
F. Pellegrini

11.10 - 11.50 *A Supernodal Approach to Incomplete LU Factorization with Partial Pivoting*
S. Li

11.50 - 12.00 *Conclusions*
Participants of the workshop

Closure

1 June 15th - Session I

Computing the diagonal of the inverse of a matrix

Y. Saad (University of Minnesota, USA). Several emerging application require to compute the diagonal of the inverse of a (sparse) matrix. These include Density Functional Theory in elecronic structure, Dynamic Mean Field Theory (DMFT), and uncertainty quantification. To solve this problem, we consider a method based on probing which leads to the solution of many linear systems with different right-hand sides. This technique is only applicable when certain conditions are met, so we also consider several other methods including techniques based on domain decomposition.

Generalized Golub-Kahan bidiagonalization and stopping criteria

Mario Arioli (Rutherford Appleton Laboratory (RAL), UK). The Golub-Kahan bidiagonalization algorithm has been widely used in solving least-squares problems and in the computation of the SVD of rectangular matrices. Here we propose an algorithm based on the Golub-Kahan process for the solution of augmented systems that minimizes the norm of the error and, in particular, we propose a novel estimator of the error similar to the one proposed by Hestenes-Stiefel for the conjugate gradient. This estimator gives a lower bound for the error, and can be used as a reliable stopping criterion for the whole process. We also propose an upper bound of the error based on Gauss-Radau quadrature. Finally, we show how we can transform and optimally precondition augmented systems rising from the mixed finite-element approximation of differential problems.

Risolv - Robust Iterative Solver

H. Tal-Ezer (Academic College of Tel-Aviv Yaffo, Israel). One of the most effective iterative algorithms for solving large, sparse, non-symmetric linear system is GMRES. Nevertheless, it suffers from lack of robustness. The algorithm can exhibit very slow rate of convergence or complete stagnation. In this talk we would like to present a new iterative algorithm, named Risolv, which overcomes this drawback. The algorithm is based on polynomial interpolation of the function $\frac{1}{z}$ at Leja points in the domain D which includes the eigenvalues of the matrix. Whenever a theoretical polynomial algorithm can solve the linear system, so does Risolv. The algorithm is specially efficient in cases where there is a need to solve many systems which share the same matrix and differ by the right hand side vectors. In this case, implementing Risolv for most of the systems is almost free of inner-products. This feature can result in significant saving of cpu, specially in parallel computing.

Efficient Gram-Schmidt orthogonalization with CUDA for iterative eigensolvers

Andres Tomas (Universidad Politécnica de Valencia, Spain) joint work with Vicente Hernandez (Universidad Politécnica de Valencia, Spain). The Gram-Schmidt orthogonalization procedure is widely used by iterative eigensolvers, either for building a Krylov subspace or deflating against already converged eigenvectors. The classical Gram-Schmidt variant with refinement is usually preferred because it can be easily implemented with matrix vector products (BLAS level 2). The throughput of these operations is mainly lim- ited by memory access speed in current computer architectures. Therefore, this procedure can become the most time-consuming step in the eigensolver, even when computing a small fraction of all the solutions from a large eigenproblem. Graphical processors are very interesting for this task because they offer higher memory bandwidth than CPU's. However, current available implementations of the matrix vector product (like the CUBLAS library) show poor performance with the rectangular matrices used in the Gram-Schmidt procedure. In this work, we propose a GPGPU parallelization scheme for the matrix vector prod- uct well suited for rectangular matrices with a larger number of rows than columns. Performance results for the CUDA implementation of this scheme in the thick restart

Lanczos method will be presented and discussed.

On the residual norm in FOM and GMRES

G. Meurant (formerly CEA). In this paper we derive expressions for the l_2 norm of the residual which involve the Hessenberg matrix H_k constructed by the Arnoldi process during the FOM or GMRES iterations. From these expressions we can obtain bounds which show that the FOM or GMRES convergence depend on the smallest singular value of an upper triangular submatrix of H_k. The norm of the residual goes to zero when this smallest singular value goes to zero.

Numerical examples show that even though these bounds are not sharp they describe quite well the rate of decrease of the residual norms.

2 June 15th - Session II

Multilevel block preconditioning for shifted Maxwell equations

M. Bollhöfer (TU Braunschweig, Germany) joint work with S. Lanteri (INRIA Sophia Antipolis). The time-harmonic Maxwell equations lead to large scale complex-symmetric systems which are highly indefinite. Moreover, discretization based on high order discontinous Galerkin discretizations lead to an underlying block structure with dense blocks. In this talk we discuss algebraic multilevel block preconditioning for block-structured complex symmetric systems such as Maxwell equations using Level-3-BLAS and a suitably shifted system. The preconditioner itself turns out to be complex symmetric which allows for the use of the simplified QMR method.

Approximate Factoring of the Inverse

M. Byckling (Aalto University, Finland) joint work with M. Huhtanen (Aalto University, Finland). We consider an algebraic approach to preconditioning linear systems by computing approximate factors. We associate the approximate factoring problem with a minimization problem involving sparse matrix subspaces.

Let $\mathcal{W}$ and $\mathcal{V}_1$ be sparse matrix subspaces of $\mathbb{C}^{n\times n}$ over $\mathbb{C}$ containing invertible elements and assume that nonsingular elements of $\mathcal{V}_1$ are readily invertible. To approximately factor the inverse of a large and sparse nonsingular matrix $A \in \mathbb{C}^{n\times n}$ into the product WV_1^{-1}, we consider the problem $AW \approx V_1$ with non-zero matrices $W \in \mathcal{W}$ and $V_1 \in \mathcal{V}_1$ regarded as variables both. This equation can be inspected by examining the nullspace of the linear map

$$W \longmapsto (I - P_1)AW, \text{ with } W \in \mathcal{W}, \tag{1}$$

where P_1 is the orthogonal projection onto $\mathcal{V}_1$. We have $AW \approx V_1 = P_1AW$ if and only if $(I - P_1)AW \approx 0$. This formulation leads to the optimality criterion

$$\min_{W\in\mathcal{W},||W||_F=1} ||(I - P_1)AW||_F \tag{2}$$

for generating factors W and $V_1 = P_1AW$ in terms of the singular values of the linear map (1).

To approximately solve the minimization problem (2) we approximately compute the smallest singular values of the linear operator (1) by using the power method with sparse-sparse operations. We then consider the eigenvalue problem involving the Hermitian positive semidefinite operator

$$W \longmapsto LW = P_{\mathcal{W}}A^*(I - P_1)AW \tag{3}$$

on the matrix subspace $\mathcal{W}$, where $P_{\mathcal{W}}$ denotes orthogonal projection onto $\mathcal{W}$. Since we are interested in the extreme eigenvalues located in the left end of the spectrum, we apply the power method to

$$\alpha I - L \text{ on } \mathcal{W}$$

with $\alpha = r\,||A||^2$ having $1/2 < r \leq 3/4$.

In our presentation, we address the choice of subspaces $\mathcal{W}$ and $\mathcal{V}_1$. We also consider parallel implementation and computational complexity of the approximate factors algorithm. Numerical comparisons with ILU and SPAI preconditioners are given when the computed approximate factors are used to precondition restarted GMRES.

The multicore challenge: the sparse indefinite case

J. Scott (Rutherford Appleton Laboratory (RAL), UK) joint work with J. Hogg (RAL, UK). The rapid emergence of multicore machines has led to the need to design new algorithms that are efficient on these architectures. In this talk, we consider the design and development of the direct solver HSL MA87 for the efficient solution of large sparse symmetric linear systems on multicore architectures. We were motivated by the successful division of the computation in the dense positive-definite case into tasks on blocks and use of a task manager to exploit all the parallelism that is available between these tasks, whose dependencies may be represented by a directed acyclic graph (DAG). Our algorithm is built on the assembly tree and subdivides the work at each node into tasks on blocks, whose dependencies may again be represented by a DAG. To limit memory requirements, updates of blocks are performed directly. The first release of the code in 2009 was for positive-definite systems. In this talk, we concentrate on indefinite systems and look at the modifications to the approach that were needed to solve these tougher systems. The main changes stem from the inclusion of threshold partial pivoting for numerical stability. In particular, we use a new combined factorize column task that replaces the separate tasks of the Cholesky factorization. We also highlight the problem of the memory-bound solve phase and explore a number of different approaches that try to improve its performance on multicore machines. Problems arising from a range of practical applications are use to illustrate the performance of HSL_MA87.

References

[1] J. Hogg and J. Scott, *A note on the solve phase of a multicore solver*, Technical Report RAL-TR-2010-007, Rutherford Appleton Laboratory, 2010.

[2] J. Hogg and J. Scott, *An indefinite sparse direct solver for large problems on multicore machines*, Technical Report RAL-TR-2010-011, Rutherford Appleton Laboratory, 2010.

3 June 15th - Session III

Combinative preconditioning based on Relaxed Nested Factorization and Tangential Filtering preconditioner

P. Kumar (INRIA Saclay). The problem of solving block tridiagonal linear systems arising from the discretization of PDE is considered. The nested factorization preconditioner introduced by [J. R. Appleyard and I. M. Cheshire, *Nested Factorization*, SPE 12264, presented at the Seventh SPE Symposium on Reservoir Simulation, San Francisco, 1983] is an effective preconditioner for certain class of problems and a similar method is implemented in Schlumerger's Eclipse oil reservoir simulator. In this paper, a relaxed version of Nested Factorization preconditioner is proposed as a replacement to ILU(0). Indeed, the proposed preconditioner is SPD and leads to a stable splitting if the input matrix is S.P.D.. For ILU(0), equivalent properties hold if the input matrix is a M-matrix. Moreover it has no storage cost. Effective multiplicative/additive preconditioning is achieved in combination with Tangential filtering preconditioner with the filter vector chosen as vector of *ones*. Numerical tests are carried out with both additive and multiplicative combinations. With this setup the new preconditioner is as robust as the combination of ILU(0) with tangential filtering preconditioner.

A Parallel Neutron Transport Solver based on Domain Decomposition

S. Van Criekingen (Karlsruhe Institute for Technology, Germany) joint work with F. Nataf (Paris VI), P. Havé (IFP). The Parafish (parallel finite-element spherical harmonic) code is a new 3-D C++ parallel (MPI) neutron transport code being developed at the Karlsruhe Institute for Technology. It aims at deterministically solving the time-independent linear Boltzmann Transport Equation (BTE) which models the behavior of neutrons within a nuclear reactor core. The stability of a reactor is determined through an eigenvalue calculation yielding a multiplication factor whose value determines whether the nuclear chain reaction makes the number of neutrons increase exponentially (super-criticality), is selfsustained (criticality), or dies out over time (sub-criticality). Parafish employs the so-called even parity formulation of the BTE, where only the even part of the angular dependence is computed. The discretization process uses finite elements for the spatial dependence, and spherical harmonic expansions for the angular dependence. For the solution process, an algebraic non-overlapping domaindecomposition method is introduced, along the lines of modified Schwarz methods [1, 2]. An important feature - novel in the neutron transport field - of Parafish is that it enables one processor to handle more than one domain. Numerical results show that the domain-decomposition method has an inherent acceleration potential: acceleration is already obtained in sequential calculations, and more generally obtained without increasing the number of processors. Moreover, enabling one processor to handle more than one domain allows proper speed-up evaluations. Encouraging efficiency results were obtained, and will be presented at the meeting.

References

[1] P.-L. Lions, *On the Schwarz alternating method III: a variant for nonoverlapping subdomains.* In Third International Symposium on Domain Decomposition Methods for Partial Differential Equations, Philadelphia, 1990.

[2] F. Nataf, *Domain decomposition methods for non-symmetric problems. In Computational Mechanics using High Performance Computing* - Proceedings of the Third Euro-Conference on Parallel and Distributed Computing for Computational Mechanics, Weimar, 1999.

3D CFD Simulation Using a Parallel Hybrid Approach of the Block Cimmino Iterative Method

M. Zenadi (University of Toulouse, INPT-IRIT) joint work with C. Balsa (CEsA-FEUP, Porto, Portugal), R. Guivarch (University of Toulouse, INPT-IRIT), D. Ruiz (University of Toulouse, INPT-IRIT), A. Silva Lopes (CEsA-FEUP, Porto, Portugal). The Cimmino method is a row projection method in which the original linear system is divided into subsystems. At every iteration, it computes one projection per subsystem and uses these projections to construct an approximation to the solution of the linear system. In our approach we do not perform explicitly the block distribution to processors within the code, but let the multi-frontal sparse solver MUMPS handle the data distribution and parallelism. The data coming from the subsystems defined by the block partition in the Block Cimmino method are gathered in an unique block diagonal sparse matrix which is analyzed, distributed and factorized in parallel by MUMPS. We present the first results of our work, in particular the memory con- sumption and the cputime performance results. Moreover, we discuss the possible orientations of the hybrid approach.

4 June 16th - Session IV

Sparse Matrix Computation in Large-Scale Scientific Applications

E. G. Ng (Lawrence Berkeley National Laboratory, USA). The U.S. Department of Energy's Scientific Discovery through Advanced Computing (SciDAC) program is a multidisciplinary and diverse R&D program that brings together application scientists, applied mathematicians, and computer scientists to tackle challenging scientific problems. In this talk, we will describe some of sparse matrix work that we are involved in SciDAC at Lawrence Berkeley National Laboratory. We will also discuss some of the challenges we are facing in some of the large-scale scientific applications.

The MUMPS library: work done during the Solstice project

MUMPS team () joint work with. We present the work done during the SOLSTICE project within or around the MUMPS library, a numerical library to solve sparse systems of linear equations of the form A x = b by direct methods. MUMPS is based on a multifrontal approach and uses message passing for parallelism. It is available free of charge and more information can be obtained from the URLs http://mumps.enseeiht.fr or http://graal.ens-lyon.fr/MUMPS. One of the originalities of MUMPS is its numerical behaviour and accuracy relying on runtime pivoting involving dynamic data structures and dynamic distributed schedulers with an asynchronous approach to parallelism. MUMPS has a wide range of features resulting from collaborations and feedback by its community of users. In this talk, we summarize some of the work done for and with the partners of the Solstice project during the last 3.5 years, focusing on out-of-core issues during both factorization and solution phases, null pivots and null space basis detection, performance behaviour and memory usage.

Working with industrial partners providing challenging problems has contributed to identifying new research directions that need to be worked on. To conclude we will thus briefly comment on-going work and projects related to sparse direct solvers within or around MUMPS library.

TBA

PastiX team (University of Bordeaux, INRIA Bordeaux Sud-Ouest).

5 June 16th - Session V

High performance Methods to solve Large 3D Electromagnetic Problems

David Goudin (CEA-CESTA). The numerical treatment of high frequency electromagnetic scattering in inhomogeneous media is very computationally intensive. For scattering, the electromagnetic field must be computed around and inside 3D complex bodies composed of inhomogeneous media. Because of this, accurate numerical methods must be used to solve Maxwell's equations in the frequency domain. In this talk, we consider the hybrid integral equation and the finite element techniques. For some high frequency applications, these numerical approaches lead to linear systems that are too large for current computer architecture. In order to solve these very large systems, typically tens of millions of Degrees Of Freedom (DOF), we have combined modern numerical methods with very efficient parallel algorithms.

EDF participation to the ANR Project SOLSTICE

O. Boiteau, (EDF R&D/SINETICS). As a major european energy utility, EDF must guarantee in time the technical and economic control of its numerous and various means of production and transport. The maintenance and optimization of these facilities and materials need, among

several things, reliable and powerful simulation softwares like Code_Aster (www.code-aster.org) and TELEMAC (www.telemacsystem.com). These are EDF physical simulation codes. The former is an all-purpose FEM code for structural analysis while the latter is a group of numerical modelling softwares for free surface water, sedimentology, waves and underground flows. To achieve good performances, the bottleneck of these codes is the solve of linear systems. So the HPC issue leads naturally to optimize and parallelize this crucial step through in-house algorithms as well as external ones.

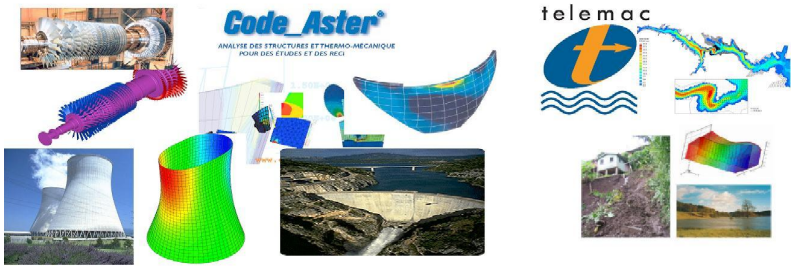

Figure 1: Code_Aster and TELEMAC are two examples of EDF simulation codes that test and integrate some tools promoted by this ANR.

Therefore, in the frame of the ANR SOLSTICE, weve strengthened our partnership with the academic teams that developed the tools MUMPS, Scotch and PaStiX. We tested, benchmarked and, sometimes, integrated these products in the pre-cited EDF codes. Obviously, this work generated lots of feedbacks concerning the installation of these products, the use of their new functionnalities, their coherence with the previous options, some numerical tricks and bugs. But, beyond the improvement of the performance, this questioning/debugging about exterior librairies induces often a productive thinking on the framework of the EDF caller code. Moreover, we provided almost 60 industrial tests to feed the database of sparse matrices of the TLSE Project. Thereby, some of these configurations could be used by the academic teams to test their tools: for example, the hybrid solvers HIPS/MaPHyS and the null space detection algorithms.

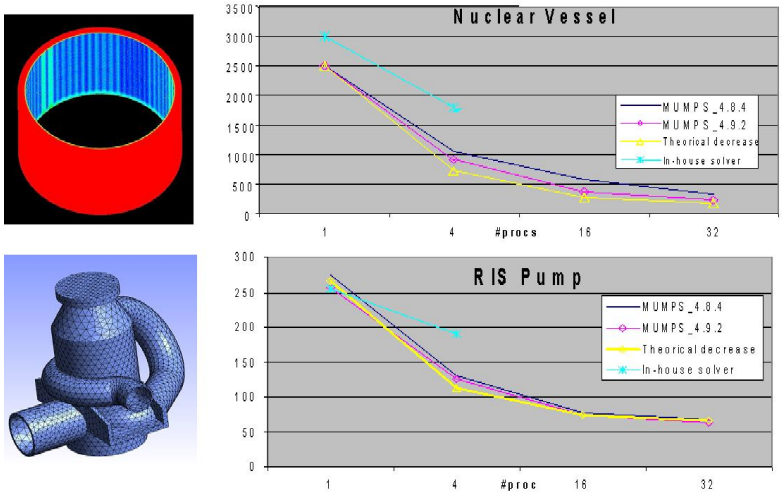

Figure 2: Decreases of the elapsed time as a function of the number of cores. Comparison of two linear thermomechanical simulations on industrial geometries.

During this ANR, a particularly productive collaboration was born between the direct solver MUMPS and the EDF software Code_Aster. Numerous feedbacks, advices and matrices were shared between the two teams. Thanks to this solver, Code_Aster can benefit from a modest parallelism (¡ 128 cores), but this one is relatively robust, all-purpose and user-friendly. So, for two years, this solvers been used daily by EDF R&D/Engineering, for industrial simulations,

through the HPC functionalities of Code_Aster. Moreover, the EDF code TELEMACs been industrializing this best-in-class direct solver in some of its software components. In short, at EDF, this ANRs really accelerated the transfer academic-industry in the domain of high-performance parallel linear solver. But, much more important than performances, we particulary appreciate the software quality and the reactivity/friendliness of the academic teams.

Amélioration des schémas de transport, nouveau solveur chimique et parallélisation en Z du solveur de pression...

J.-P. Pinty (Météo-France) joint work with C. Lac (Météo-France), T. Maric (Météo-France), F. Visentin (Météo-France) and J. Escobar (Météo-France). Le projet initial portait sur la rénovation des schémas numériques de transport dans le modèle de simulation de l'atmosphère à haute résolution, le modèle MésoNH. Cet objectif a été complété par l'introduction d'un solveur chimique capable de résoudre des gros systèmes creux et raides, typiques de la chimie en phase aqueuse dans MésoNH. Enfin l'accès de MésoNH aux machines massivement parallèles ($Nproc > 10^3$) a été renforcé par une intervention pointue sur l'algorithme de parallélisation des FFT utilisé pour la résolution de l'équation elliptique de pression (MésoNH est basé sur un système d'équations anélastique).

L'amélioration du transport dans MésoNH s'est traduite par la mise au point suivie de nombreux tests de schémas 3D exclusivement conservatifs, monotones, définis positifs et parallélisables pour diverses conditions aux limites. Notre choix s'est porté sur le schéma PPM (Piecewise Parabolic Method) pour le transport des champs scalaires et sur la famille de schémas WENO (Weighted Essentially Non Oscillating) pour le transport des composantes de la quantité de mouvement. Le recours à ces 2 types de schéma est nécessaire car MésoNH est discrétisé sur un système de 4 grilles imbriquées. La configuration stable actuellement obtenue combine PPM et WENO d'ordre 3, intégrés avec un schéma temporel simple (Forward in Time). L'effort se poursuit pour substituer un schéma Runge-Kutta d'ordre 3 et passer à un schéma WENO d'ordre 5.

Le nouveau solveur chimique introduit dans MésoNH est basé sur une résolution semi-implicite à pas de temps adaptatif d'un système d'équations différentielles non-linéaires. L'algorithme s'appuie sur la méthode de Rosenbrock. S'agissant de systèmes creux, le codage astucieux des indices non-nuls d'une matrice jacobienne permet d'inverser très rapidement le système. Des exemples de simulation montrent qu'un système chimique $100 * 100$ peut être résolu efficacement et directement sur un sous-ensemble de points consécutifs d'une grille de calcul habituelle $200 * 200 * 50$ de MésoNH.

Enfin la parallélisation du solveur de pression a été améliorée (parallélisation en Z) de manière à pouvoir intégrer MésoNH sur des grilles de calcul de l'ordre de $Nx * Ny * Nz$ sur un ensemble de $Nz * Min(Nx, Ny)$ processeurs. La limitation précédente était $Nproc < Min(Nx, Ny)$. La parallélisation de MésoNH, muni de sa propre bibliothèque de fonctions de haut niveau utilisant MPI, est basée sur la technique de décomposition de domaine sur le plan X, Y.

6 June 16th - Session VI

Preconditioning the Schur complement in a hybrid direct/iterative solver

P. Hénon (INRIA Bordeaux Sud-Ouest) joint work with J. Gaidamour (Sandia National Laboratories, USA). Nowadays, three dimensional numerical simulations often require a tremendous amount of resource. On one hand, direct methods can be mandatory to solve very ill-conditioned systems. But for large 3D simulations, they are constrained by prohibitive memory requirements and they also need a high amount of floating point operations. Iterative methods on the other hand require much less memory and are more scalable in general. Methods such as algebraic multi-grids or multilevel solvers are optimal for some problems : a linear complexity in term of

memory and number of operations can be obtained on elliptic problems for example but they lack adaptability in the general case. In the ANR project "Solstice", we have developed an hybrid solver that tries to combine the assets of direct methods and iterative methods. More precisely, we will give some algorithms to build a scalable hybrid direct/iterative solver based on a Schur complement approach. Our method is based on a domain decomposition of the matrix adjacency graph and we use a special ordering and partitioning of the interface between the domain in order to compute a robust ILU preconditioner of the Schur complement. An important aspect in our method is to avoid as much as possible the high memory consumption required to compute the Schur complement preconditioner. To this end, we will present different algorithms based on a fine coupling between supernodal algorithms (coming from direct factorization) and block ILUT algorithms.

Parallel scalability and complexity analysis of the sparse hybrid linear solver MaPHYS

E. Agullo (INRIA) joint work with L. Giraud (INRIA), A. Guermouche (Université de Bordeaux, LaBRI UMR 5800), A. Haidar (ICL, University of Tennessee, USA), J. Roman (INRIA, Université de Bordeaux, LaBRI UMR 5800), Y. Lee-Tin-Yen (INRIA). In many large scale numerical simulations the inner most and most time consuming kernel is the solution of a sparse linear systems. Sparse direct solvers have been for years the methods of choice because of their reliable numerical behaviour. However, it is nowadays admitted that such approaches are not scalable in terms of computational complexity or memory for large problems such as those arising from the 3D modelling. Iterative methods, on the other hand, generate sequences of approximations to the solution. These methods have the advantage that the memory requirements are small. Also, they tend to be easier to be parallelized than direct methods. However, the main problem with this class of methods is the rate of convergence, which depends on the properties of the matrix. One way to improve the convergence rate is through preconditioning, which is another difficult problem. Our approach to high-performance, scalable solution of large sparse linear systems in parallel scientific computing is to combine direct and iterative methods [1].

Sparse hybrid solvers are a trade-off between direct methods and iterative methods. Part of the computation is first performed with a direct method in order to ensure numerical robustness; the algorithm then switches to an iterative method to alleviate the computational complexity and memory usage. The convergence and number of iterations in the second step depend on the amount of computation performed in the direct part. In this talk, we first describe the parallel numerical behaviour of the additive Schwarz preconditioner, and its variants, for a Schur complement approach. We illustrate its numerical performance on both academic test problems as well as on matrices provided by the Solstice end-user partners. We also discuss a computatianal model that enables us to study the asymptotic memory and floating point arithmetic cost of our approch [2]. Finally, we expose various possible API's for the forthcoming MaPHYS package in order to get feed-back of future users.

References

[1] L. Giraud and A. Haidar and L. T. Watson, *Parallel scalability study of hybrid preconditioners in three dimensions*, Parallel Computing, 34 (2008), pp. 363-379.

[2] E. Agullo and L. Giraud and A. Guermouche and J. Roman, *On the complexity analysis of sparse hybrid linear solvers*, INRIA Technical Report, in preparation, 2010.

Parallel preprocessing in MUMPS

Alfredo Buttari (University of Toulouse, INPT-IRIT) joint work with Bora Ucar (ENS-Lyon, LIP). The preprocessing phase of direct solvers, also commonly known as analysis phase, has often been considered not worth of parallelization due to its low complexity and storage requirements compared to the factorization phase. As the scalability of the factorization phase has

been considerably improved and as the dimension of problems from real life applications has significantly grown over the last years, the relative cost of sequential preprocessing came to be higher and higher and its storage requirements unsustainable even for high-end supercomputers. These limitations become even more penalizing in consideration of the fact that data has to be centralized in order to perform the sequential analysis and that the storage requirements of the factorization phase can be reduced by employing out-of-core techniques.

This talk presents the parallelization of two operations in the preprocessing phase of the MUMPS software package:

- symbolic factorization: generally, the first step in the preprocessing of the input problem is the computation of the fill-minimizing pivotal order. The parallel execution of this step is delegated to an external partitioning tool such as PT-SCOTCH or ParMetis. Once this ordering has been computed, the structure of the factors must be computed or estimated in order to define the computational pattern and storage requirements of the subsequent factorization phase. The proposed approach to the parallelization of the symbolic factorization relies on the use of quotient graphs and restarting techniques.

- scaling: For a given the input matrix, the scaling step computes two positive diagonal matrices D and E such that the rows and columns of the scaled matrix DAE have the same magnitude in some norm. We have parallelized the scaling routines originally proposed by Ruiz and plugged these routines into the MUMPS solver. These routines iteratively compute the scaling matrices D and E. We have carefully analyzed the communication and computation requirements in a distributed memory environment and showed that they are closely related to the sparse matrix-vector multiplication operation. We have fine-tuned certain parameters of the scaling routines such that we do only a few iterations and compute approximate scaling matrices which improve the solver's performance in practice.

GRID-TLSE Project

P. Amestoy (University of Toulouse, INPT-IRIT), F. Camillo (University of Toulouse, INPT-IRIT), M. Daydé (University of Toulouse, INPT-IRIT), R. Guivarch (University of Toulouse, INPT-IRIT), A. Hurault (University of Toulouse, INPT-IRIT), J.Y. L'Excellent (ENS-Lyon, LIP, INRIA), C. Puglisi (University of Toulouse, INPT-IRIT). The GRID-TLSE web site (http://gridtlse.org/) provides facilities to compare sparse direct solvers and to share sparse matrices.

After an introduction on the goals and the purposes of this site, we present the different functionalities already available. We detail one of the key notion of this project, the concept of scenario. The scenarios allows us to compare sparse direct solvers and in particular those used in the project SOLSTICE.

We end the presentation with a complex scenario that shows the power of the tools developed in the project. The example is a scenario that couples solvers. Precisely, we show how the pre-treatments computed by one solver are used by another solver.

7 June 17th - Session VII

Preconditioners for stabilized finite element discretizations of the steady incompressible Navier-Stokes equations

Michele Benzi (Emory University, USA) joint work with Maxim Olshanskii (Moscow State University, Russia) and Zhen Wang (Emory University, USA). In this talk we consider block preconditioners for stabilized saddle point systems arising from incompressible flow problems. Our focus is on block triangular preconditioners based on the augmented Lagrangian formulation. These techniques have been shown to be remarkably robust and effective when applied to linear systems arising from inf-sup stable mixed finite element discretizations of the Stokes and Oseen problems. Indeed, with an appropriate choice of the augmentation parameter these preconditioners result in rates of convergence that are independent of the mesh size and largely insensitive to the value of the viscosity; see [1, 2].

In this talk we describe a (non-trivial) extension of the augmented Lagrangian-based block preconditioners to the case of stabilized finite elements, e.g., Q1-Q1 and Q1-P0 elements. We also describe an inexpensive technique for estimating the optimal value of the augmentation parameter. This technique is applicable to both stable and stabilized finite element schemes.

References

[1] M. Benzi and M. A. Olshanskii, *An augmented Lagrangian-based approach to the Oseen problem*, SIAM J. Sci. Comput., 28 (2006), pp. 2095-2113.

[2] M. Benzi, M. A. Olshanskii and Z. Wang, *Modifed augmented Lagrangian preconditioners for the incompressible Navier-Stokes equations*, Int. J. Numerical Methods in Fluids, in press, 2010.

Sparse matrix partitioning, ordering, and visualisation by Mondriaan 3.0

Rob H. Bisseling (Mathematical Institute, Utrecht University, The Netherlands) joint work with Bas O. Fagginger Auer, and Albert-Jan N. Yzelman. This talk presents two combinatorial problems encountered in scientific computations on today's high-performance architectures, such as parallel computers with many processors and several cores on each processor, and with sophisticated memory hierarchies and multiple levels of cache.

For parallelism, the most important problem is to partition sparse matrices, graphs, or hypergraphs into nearly equal-sized parts while trying to reduce inter-processor communication. For better cache performance, the problem is to reorder sparse matrices by suitable row and column permutations. Common approaches to such problems involve multilevel methods based on coarsening and uncoarsening (hyper)graphs, matching of similar vertices, searching for good separator sets and good splittings, and two-dimensional matrix splitting methods such as incorporated in the software package Mondriaan.

We will discuss new algorithms and features included in version 3.0 of the Mondriaan package, to be released soon. By using this package, and its subprograms MondriaanPlot and MondriaanMovie, we can visualise the partitioning process of a sparse matrix by various algorithms. We can also do this in Matlab. Mondriaan has now been made into a library that can be called from other programs, such as Matlab, Mathematica, or as a standalone program. New reordering methods have been included such as Separated Block Diagonal (SBD), along with well-known methods such as Bordered Block Diagonal. Doubly separated and doubly bordered versions are also included.

8 June 17th - Session VIII

PT-Scotch in Solstice and beyond: where to go now

F. Pellegrini (University of Bordeaux, INRIA Bordeaux Sud-Ouest). The Solstice project aimed at designing and developing high-performance parallel linear solvers that would be efficient to solve complex multi-physics and multi-scale problems of very large size. Graph partitioning is a critical issue in order to achieve this goal, both to distribute problem data evenly across processes and to compute fill-reducing orderings of the linear systems to solve.

Because problem sizes keep increasing, large problems graphs cannot fit in the memory of sequential computers, and cost too much to partition. This is why the Solstice project comprised a task devoted to the conception of efficient algorithms for parallel graph partitioning, which have been implemented into the PT-Scotch software package.

This talk will present an overview of the algorithmic advances on parallel graph partitioning brought by the Solstice project. It will start by focusing on the parallelization of the ubiquitous multi-level scheme, and present the parallel coarsening and refinement algorithms used in this context for recursive bipartitioning. It will then discuss the challenges brought by massively parallel computing, from an algorithmic point of view, with respect to direct k-way partitioning and static mapping.

A Supernodal Approach to Incomplete LU Factorization with Partial Pivoting

S. Li (Lawrence Berkeley National Laboratory, USA). We present a new supernode-based incomplete LU factorization method to construct a preconditioner for solving sparse linear systems with iterative methods. The new algorithm is primarily based on the ILUTP approach by Saad, and we incorporate a number of techniques to improve the robustness and performance of the traditional ILUTP method. These include new dropping strategies that accommodate the use of supernodal structures in the factored matrix and an area-based fill control heuristic for the secondary dropping strategy. We present numerical experiments to demonstrate that our new method is competitive with the other ILU approaches and is well suited for today's high performance architectures.