On some robust solvers and their scalable implementation

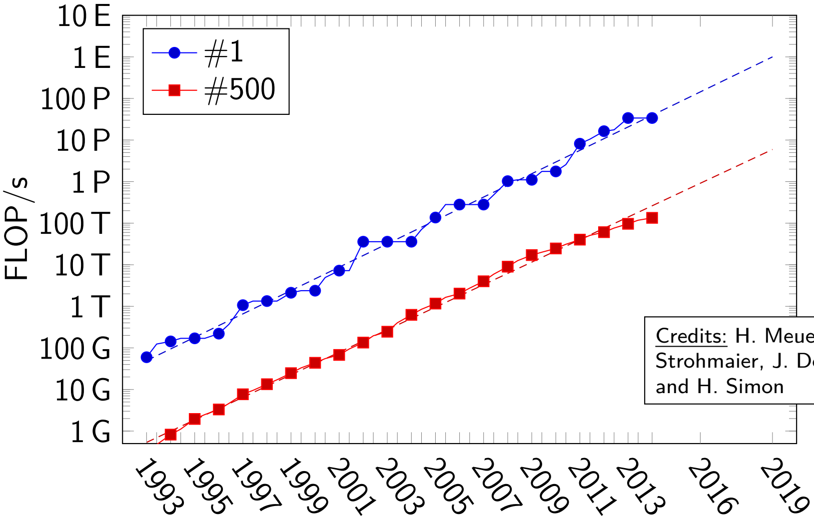
Pierre Jolivet
ETH Zürich

Journée DDM - parallélisation

June 11, 2015

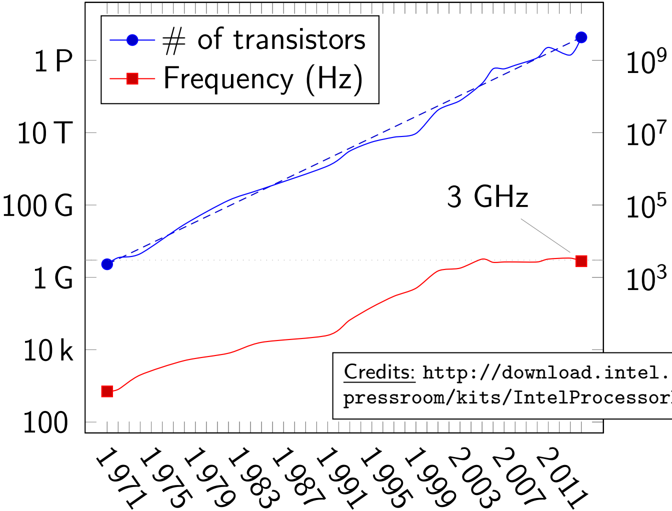
High-performance computing context

Today's large-scale supercomputers are tomorrow's desktops:



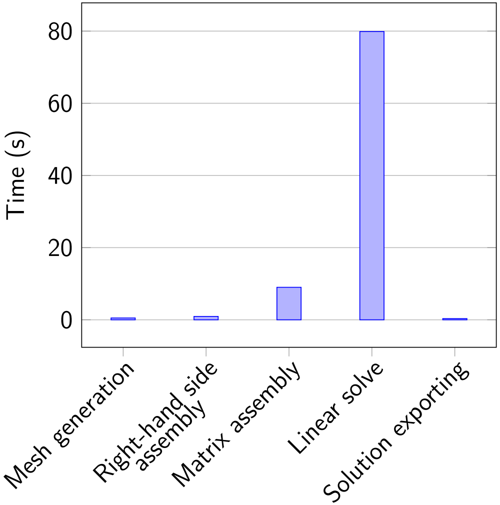
Credits: H. Meuer, E. Strohmaier, J. Dongarra and H. Simon

2005 and the end of f-scaling



Credits: `http://download.intel.com/pressroom/kits/IntelProcessorHistory.pdf`

The big picture of implicit methods



State of the art

Many theoretical work on DDM. [Mathew 2008; Pechstein 2012; Quarteroni and Valli 1999; Smith, Bjørstad, and Gropp 2004; Toselli and Widlund 2005]

State of the art

Many theoretical work on DDM. [Mathew 2008; Pechstein 2012; Quarteroni and Valli 1999; Smith, Bjørstad, and Gropp 2004; Toselli and Widlund 2005]

Not many available implementations:

- Klawonn & Rheinbach,
- Badia & Principe,
- Šístek, Mandel, and Sousedík,
- Trilinos SALINAS.

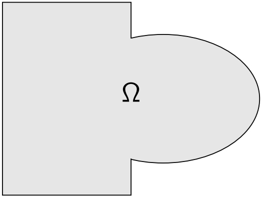
Mostly application dependant solvers.

High-Performance DDM

- https://github.com/hpddm/hpddm [Jolivet and Nataf 2014],
- C++ open-source,
- supported by many work [Dolean, Jolivet, and Nataf 2015; Jolivet, Hecht, Nataf, and Prud'homme 2013; Spillane, Dolean, et al. 2013; Spillane and Rixen 2013]...
- interfaced with FE libraries (FreeFem++ and Feel++),
- scales on 16k threads to solve billion d.o.f. problems,
- Stokes equation, Maxwell equations...

Overlappings methods I

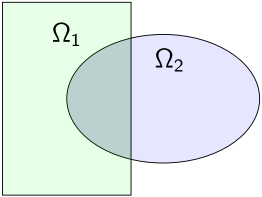
Consider the linear system: $Au = f \in \mathbb{R}^n$.



Overlappings methods I

Consider the linear system: $Au = f \in \mathbb{R}^n$.
Given a decomposition of $[\![1; n]\!]$, $(\mathcal{N}_1, \mathcal{N}_2)$, define:

- the restriction operator R_i from $[\![1; n]\!]$ into $\mathcal{N}_i$,
- R_i^T as the extension by 0 from $\mathcal{N}_i$ into $[\![1; n]\!]$.



Overlappings methods I

Consider the linear system: $Au = f \in \mathbb{R}^n$.
Given a decomposition of $[\![1; n]\!]$, $(\mathcal{N}_1, \mathcal{N}_2)$, define:

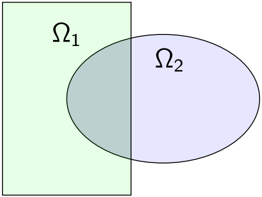
- the restriction operator R_i from $[\![1; n]\!]$ into $\mathcal{N}_i$,
- R_i^T as the extension by 0 from $\mathcal{N}_i$ into $[\![1; n]\!]$.

Then solve concurrently:

$$u_1^{m+1} = u_1^m + A_{11}^{-1} R_1 (f - Au^m) \qquad u_2^{m+1} = u_2^m + A_{22}^{-1} R_2 (f - Au^m)$$

where $u_i = R_i u$ and $A_{ij} := R_i A R_j^T$.

[Schwarz 1870]

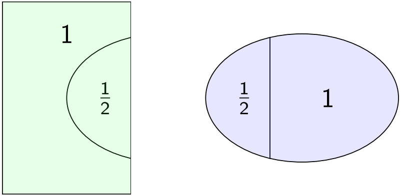


Overlappings methods II

Problem is effectively divided, but yet to be conquered.

Duplicated unknowns coupled via a *partition of unity*:

$$I = \sum_{i=1}^{N} R_i^T D_i R_i.$$

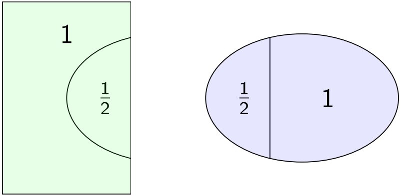


Overlappings methods II

Problem is effectively divided, but yet to be conquered.

Duplicated unknowns coupled via a *partition of unity*:

$$I = \sum_{i=1}^{N} R_i^T D_i R_i.$$



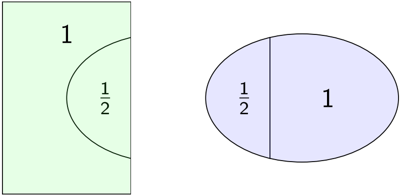
Then, $u^{m+1} = \sum_{i=1}^{N} R_i^T D_i u_i^{m+1}$.

Overlappings methods II

Problem is effectively divided, but yet to be conquered.

Duplicated unknowns coupled via a *partition of unity*:

$$I = \sum_{i=1}^{N} R_i^T D_i R_i.$$

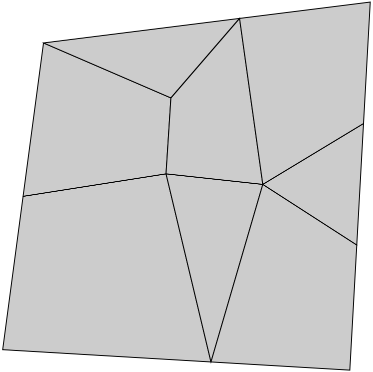


Then, $u^{m+1} = \sum_{i=1}^{N} R_i^T D_i u_i^{m+1}.$

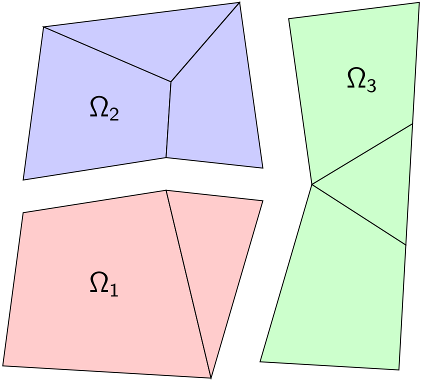
$$M_{\text{RAS}}^{-1} = \sum_{i=1}^{N} R_i^T D_i A_{ii}^{-1} R_i$$

[Cai and Sarkis 1999]

Substructuring preconditioners I

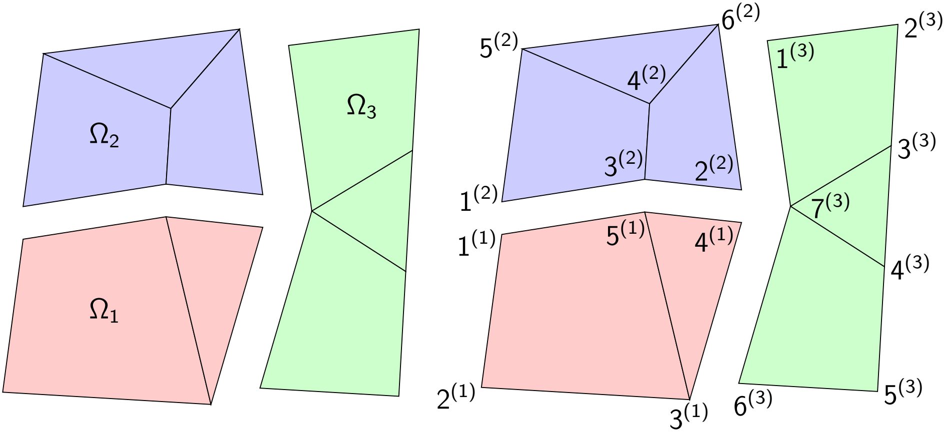


[Gosselet and Rey 2006]



Subdomain tearing

Substructuring preconditioners I

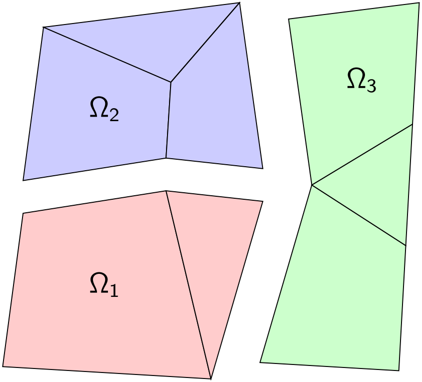


[Gosselet and Rey 2006]

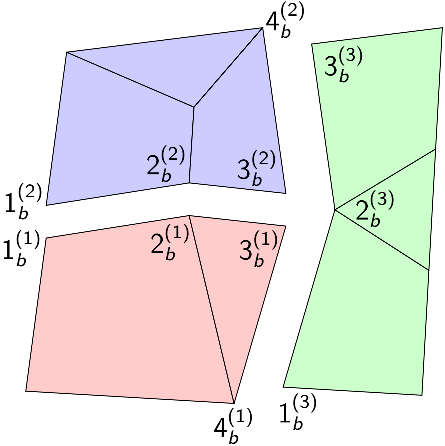
Local numbering

$$A^{(k)} = \begin{bmatrix} A_{ii} & A_{ib} \\ A_{bi} & A_{bb} \end{bmatrix}$$

Substructuring preconditioners I



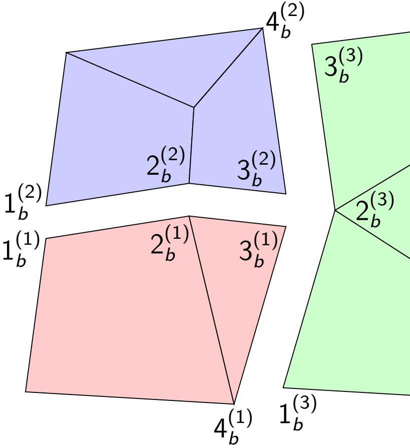
[Gosselet and Rey 2006]



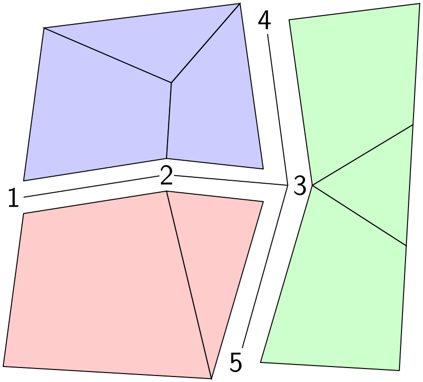
Elimination of interior d.o.f.

$$S^{(k)} = A_{bb} - A_{bi}A_{ii}^{-1}A_{ib}$$

Substructuring preconditioners I

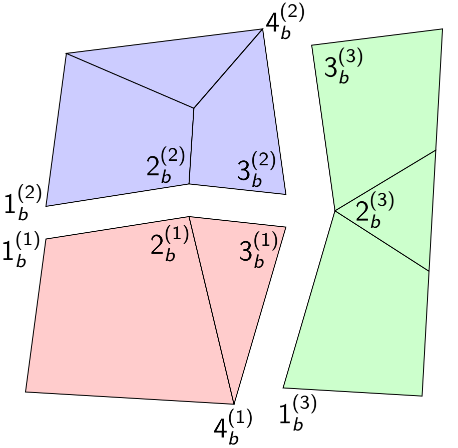


Jump operators: $\{B^{(i)}\}_{i=1}^{3}$

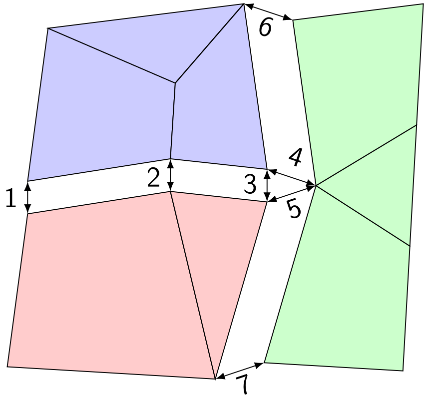


Primal constraints
[Mandel 1993]

Substructuring preconditioners I



Jump operators: $\{\underline{B}^{(i)}\}_{i=1}^{3}$



Dual constraints
[Farhat and Roux 1991]

Substructuring preconditioners II

The new system reads:

$$\forall i \in [\![1; N]\!],\ S^{(i)} u_b^{(i)} = g_i + \lambda_b^{(i)}$$

Substructuring preconditioners II

The new system reads:

$$\forall i \in [\![1; N]\!],\ S^{(i)} u_b^{(i)} = g_i + \lambda_b^{(i)}$$

$$\sum_{i=1}^{N} \underline{B}^{(i)} u_b^{(i)} = 0$$

Substructuring preconditioners II

The new system reads:

$$\begin{aligned}\forall i \in [\![1; N]\!],\ S^{(i)} u_b^{(i)} &= g_i + \lambda_b^{(i)}\\ \sum_{i=1}^{N} \underline{B}^{(i)} u_b^{(i)} &= 0\\ \sum_{i=1}^{N} B^{(i)} \lambda_b^{(i)} &= 0\,.\end{aligned}$$

Efficient preconditioners (based on scaled sum): [Dohrmann 2003; Farhat, Mandel, and Roux 1994; Rixen and Farhat 1997]

Limitations of one-level methods

One-level methods don't require exchange of global information.

This hampers numerical scalability of such preconditioners:

$$\kappa(M^{-1}A) \leqslant C\frac{1}{H^2}\left(1+\frac{H}{\delta}\right)$$

- level of overlap δ,
- characteristic size of a subdomain H.

[Le Tallec 1994; Toselli and Widlund 2005]

Two-level preconditioners

A common technique in the field of DDM, MG, deflation:

introduce an auxiliary “coarse” problem.

Let Z be a rectangular matrix. Define

$$E := Z^T A Z\,.$$

Z has $\mathcal{O}(N)$ columns, hence E is much smaller than A.

Two-level preconditioners

A common technique in the field of DDM, MG, deflation:

introduce an auxiliary “coarse” problem.

Let Z be a rectangular matrix. Define

$$E := Z^T A Z\,.$$

Z has $\mathcal{O}(N)$ columns, hence E is much smaller than A.
Enrich the original preconditioner, e.g. additively

$$P^{-1} = M^{-1} + Z E^{-1} Z^T,$$

cf. [Tang, Nabben, Vuik, and Erlangga 2009].

Formal definition of Z

The convergence is hindered by extreme eigenvalues.
$\hookrightarrow$ use the associated eigenvectors for the deflation matrix.

Formal definition of Z

The convergence is hindered by extreme eigenvalues.
$\hookrightarrow$ use the associated eigenvectors for the deflation matrix.

A local threshold criterion selects γ_i local vectors $\{\Lambda_{i_k}\}_{k=1}^{\gamma_i}$.

Formal definition of Z

The convergence is hindered by extreme eigenvalues.
$\hookrightarrow$ use the associated eigenvectors for the deflation matrix.

A local threshold criterion selects γ_i local vectors $\{\Lambda_{i_k}\}_{k=1}^{\gamma_i}$.

$$W_i = \begin{bmatrix} D_i\Lambda_{i_1} & D_i\Lambda_{i_2} & \cdots & D_i\Lambda_{i_{\gamma_i}} \end{bmatrix}.$$

Formal definition of Z

The convergence is hindered by extreme eigenvalues.
$\hookrightarrow$ use the associated eigenvectors for the deflation matrix.

A local threshold criterion selects γ_i local vectors $\{\Lambda_{i_k}\}_{k=1}^{\gamma_i}$.

$$W_i = \begin{bmatrix} D_i\Lambda_{i_1} & D_i\Lambda_{i_2} & \cdots & D_i\Lambda_{i_{\gamma_i}} \end{bmatrix}.$$

Then, define the global deflation matrix as:

$$Z = \begin{cases} \begin{bmatrix} R_1^T W_1 & R_2^T W_2 & \cdots & R_N^T W_N \end{bmatrix} & \text{for Schwarz,} \\ \begin{bmatrix} B_1 W_1 & B_2 W_2 & \cdots & B_N W_N \end{bmatrix} & \text{for BDD,} \\ \begin{bmatrix} \underline{B}_1 W_1 & \underline{B}_2 W_2 & \cdots & \underline{B}_N W_N \end{bmatrix} & \text{for FETI.} \end{cases}$$

Not in assembled in reality since they require a global ordering.

Construction of Z

GenEO for Schwarz methods

Solved by ARPACK concurrently:

$$A_i^N \Lambda_{i_k} = \lambda_{i_k} D_i R_{i,0}^T R_{i,0} A_i^N R_{i,0} R_{i,0}^T D_i \Lambda_{i_k} \,,$$

where

- A_i^N is the local unassembled matrix (bilinear form with Neumann boundary conditions on the interfaces),
- $R_{i,0}$ is the restriction from Ω_i to $\Omega_i \bigcap \left(\bigcup_{j \in \mathcal{O}_i} \Omega_j \right)$.

Construction of Z

GenEO for Schwarz methods

Solved by ARPACK concurrently:

$$A_i^N \Lambda_{i_k} = \lambda_{i_k} D_i R_{i,0}^T R_{i,0} A_i^N R_{i,0} R_{i,0}^T D_i \Lambda_{i_k} \,,$$

where

- A_i^N is the local unassembled matrix (bilinear form with Neumann boundary conditions on the interfaces),
- $R_{i,0}$ is the restriction from Ω_i to $\Omega_i \bigcap \left(\bigcup_{j \in \mathcal{O}_i} \Omega_j \right)$.

The "algebraicity" of Schwarz methods is lost here.

Construction of Z

GenEO for Optimized Schwarz methods

Same framework as in [St-Cyr, Gander, and Thomas 2007].

Construction of Z

GenEO for Optimized Schwarz methods

Same framework as in [St-Cyr, Gander, and Thomas 2007].
Solved by ARPACK concurrently:

$$A_i^N \Lambda_{i_k} = \lambda_{i_k} B_i \Lambda_{i_k} \,,$$

where

- A_i^N is the local unassembled matrix (bilinear form with Neumann boundary conditions on the interfaces),
- B_i is the local matrix with optimized interface conditions.

Cf. [Conen, Dolean, Krause, and Nataf 2014; Haferssas, Jolivet, and Nataf 2015].

Construction of Z

GenEO for substructuring methods

For BDD, solved by LAPACK concurrently:

$$S^{(i)}\Lambda_{i_k} = \lambda_{i_k} D_i {B^{(i)}}^T S B^{(i)} D_i \Lambda_{i_k}\,.$$

Construction of Z

GenEO for substructuring methods

For BDD, solved by LAPACK concurrently:

$$S^{(i)}\Lambda_{i_k} = \lambda_{i_k} D_i {B^{(i)}}^T S B^{(i)} D_i \Lambda_{i_k} \,.$$

The Schur complement S is never assembled:

$${B^{(i)}}^T S B^{(i)} = {B^{(i)}}^T \sum_{j \in \overline{\mathcal{O}}_i} B^{(j)} S^{(j)} {B^{(j)}}^T B^{(i)}$$

Construction of Z
GenEO for substructuring methods

For BDD, solved by LAPACK concurrently:

$$S^{(i)}\Lambda_{i_k} = \lambda_{i_k} D_i {B^{(i)}}^T S B^{(i)} D_i \Lambda_{i_k}\,.$$

The Schur complement S is never assembled:

$${B^{(i)}}^T S B^{(i)} = {B^{(i)}}^T \sum_{j \in \overline{\mathcal{O}}_i} B^{(j)} S^{(j)} {B^{(j)}}^T B^{(i)}$$

With both GEVP, null eigenvalues yield rigid body modes.
No efficient solution for FETI.

Distribution of E

How can one apply E^{-1} to a vector in $\mathbb{R}^m$?

Some constraints:

1. E can't be centralized on one MPI process, cf. DD 20 proceedings,
2. E can't be distributed on all MPI processes,
3. the solution must be computed fast and reliably.

Distribution of E

How can one apply E^{-1} to a vector in $\mathbb{R}^m$?

Some constraints:

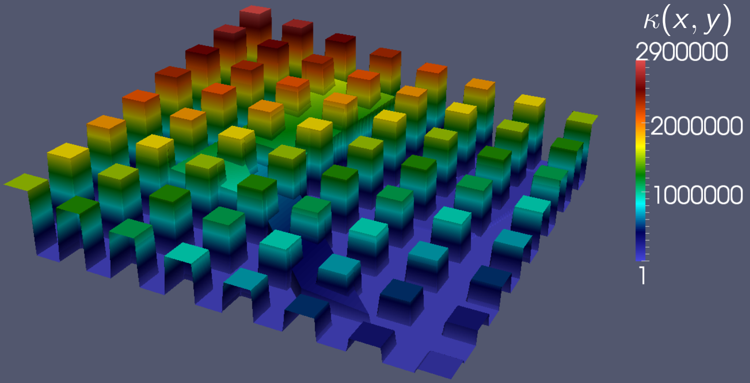
1. E can't be centralized on one MPI process, cf. DD 20 proceedings,
2. E can't be distributed on all MPI processes,
3. the solution must be computed fast and reliably.

$\Longrightarrow$ use a direct solver with a distributed matrix on few *master* processes (number chosen at runtime).

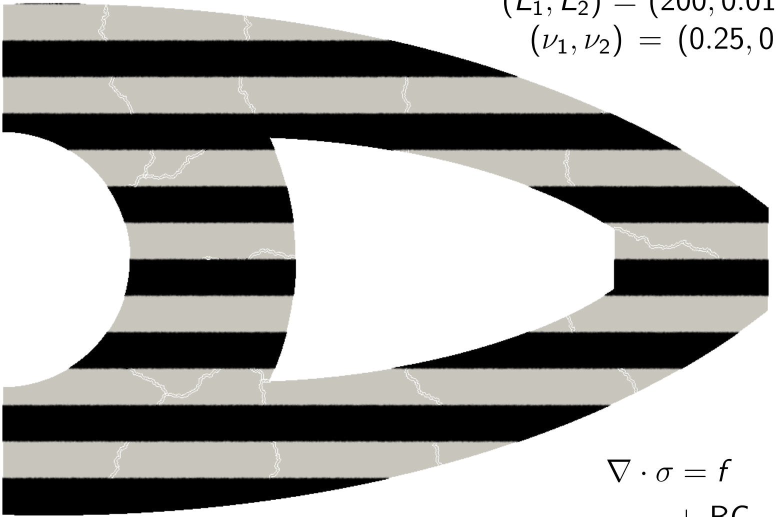
SC13 best paper award nominee [Jolivet et al. 2013].

Example of heterogeneous coefficients

$$\nabla \cdot (\kappa \nabla u) = f$$
$$+ \text{BC}$$



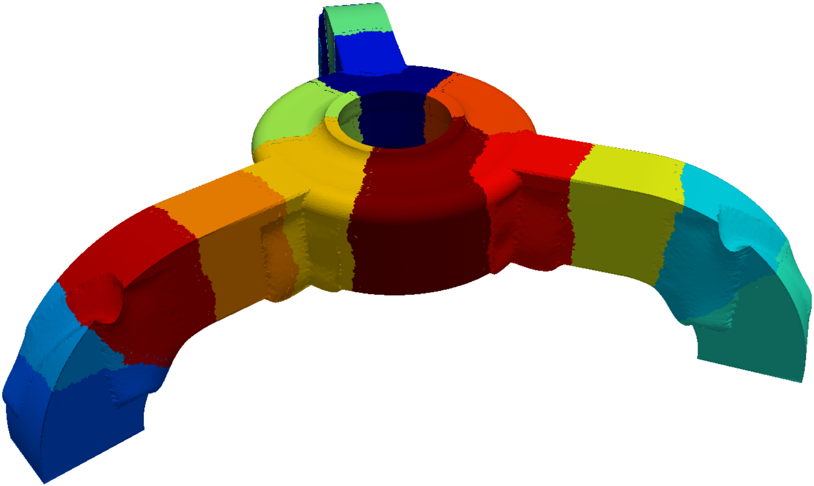
2D geometry



$$(E_1, E_2) = (200, 0.01) \text{ GPa}$$
$$(\nu_1, \nu_2) = (0.25, 0.45)$$

$$\nabla \cdot \sigma = f$$
$$+ \text{ BC}$$

3D geometry



Machines used for scaling runs

Curie Thin Nodes

- 5 040 compute nodes (2 eight-core Intel Sandy Bridge).
- IB QDR full fat-tree.
- BullxMPI based on an old OpenMPI (1.6.x).

Turing

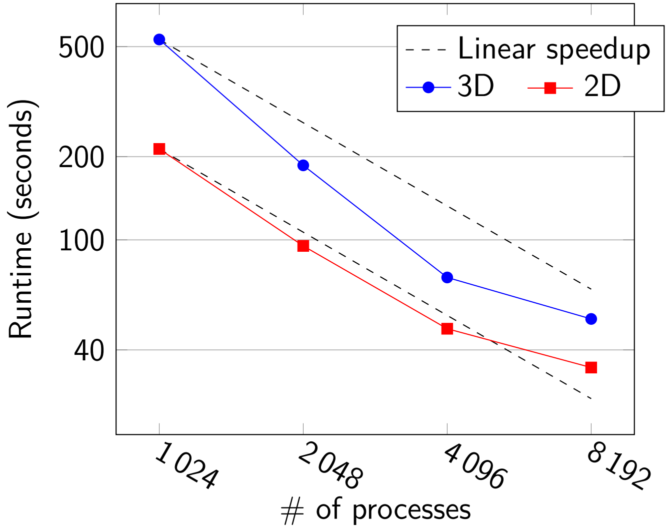
- 6 BlueGene/Q racks.

PRACE

GENCI
GRAND EQUIPEMENT NATIONAL DE CALCUL INTENSIF

Strong scaling (linear elasticity)

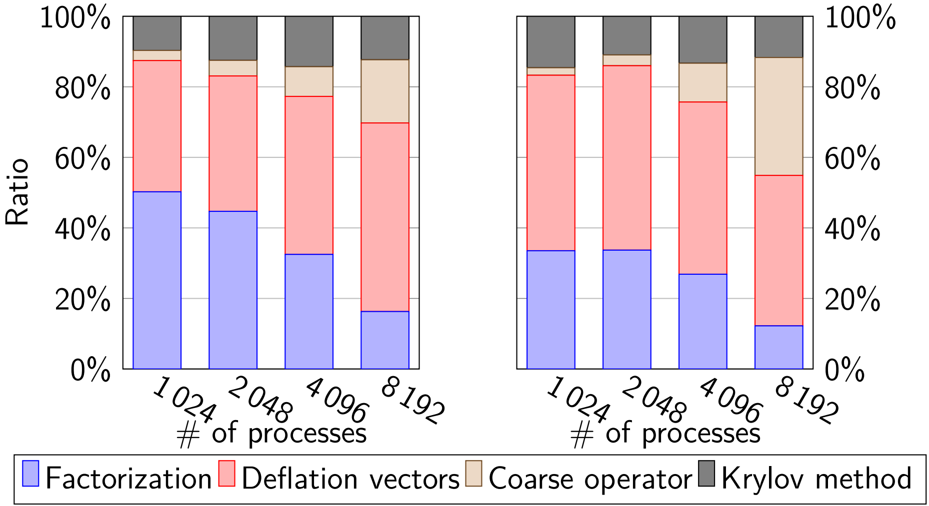
1 subdomain/MPI process, 2 OpenMP threads/MPI process.



Strong scaling (linear elasticity)

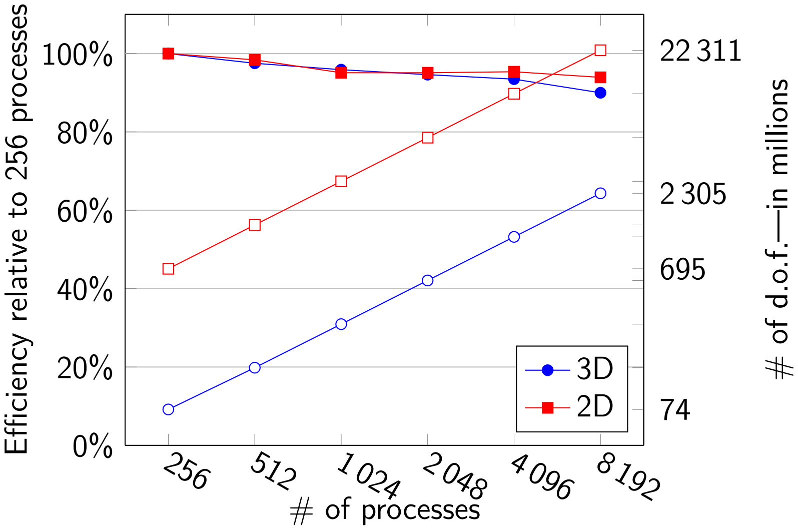
1 subdomain/MPI process, 2 OpenMP threads/MPI process.

2.1B d.o.f. in 2D ($\mathbb{P}_3$ FE) 300M d.o.f. in 3D ($\mathbb{P}_2$ FE)



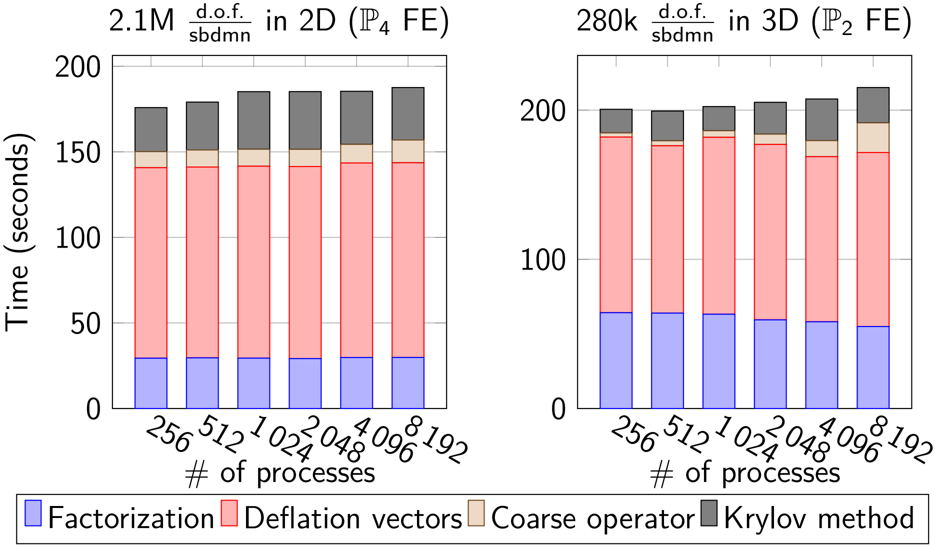
Weak scaling (scalar diffusion equation)

1 subdomain/MPI process, 2 OpenMP threads/MPI process.



Weak scaling (scalar diffusion equation)

1 subdomain/MPI process, 2 OpenMP threads/MPI process.



Schwarz preconditioners vs. MG solvers

Performance of setup and solution phases of algebraic solvers:

- Hypre BoomerAMG, selective algebraic multigrid (LLNL)

- GAMG, aggregative algebraic multrigrid (ANL/LBL)

and:

- Schwarz GenEO

Schwarz preconditioners vs. MG solvers

Performance of setup and solution phases of algebraic solvers:

- Hypre BoomerAMG, selective algebraic multigrid (LLNL)

- GAMG, aggregative algebraic multrigrid (ANL/LBL)

and:

- Schwarz GenEO

$\hookrightarrow$ number of deflation vectors (increasing it only "improves" the preconditioner).

Schwarz preconditioners vs. MG solvers

Performance of setup and solution phases of algebraic solvers:

- Hypre BoomerAMG, selective algebraic multigrid (LLNL)

↪ `-pc_hypre_boomeramg_agg_nl 2 -pc_hypre_boomeramg_P_max 4 -pc_hypre_boomeramg_coarsen_type HMIS -pc_hypre_boomeramg_interp_type ext+i -pc_hypre_boomeramg_no_CF`

- GAMG, aggregative algebraic multrigrid (ANL/LBL)

↪ `-pc_gamg_agg_nsmooths 1 -mg_levels_pc_type jacobi -mg_levels_ksp_type chebyshev -mg_levels_ksp_max_it 1 -pc_gamg_threshold 0.1 -mg_coarse_pc_type svd -mg_coarse_pc_svd_monitor -gamg_est_ksp_type gmres -gamg_est_pc_type jacobi`

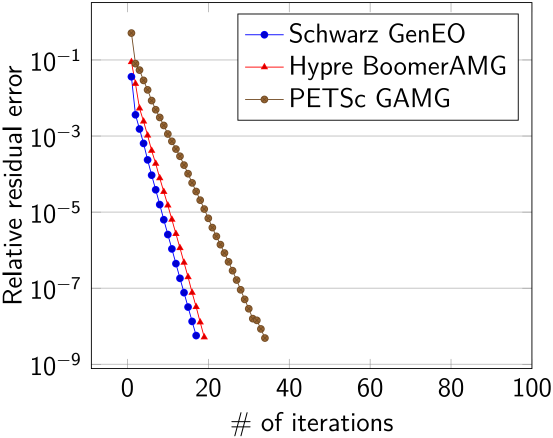
and:

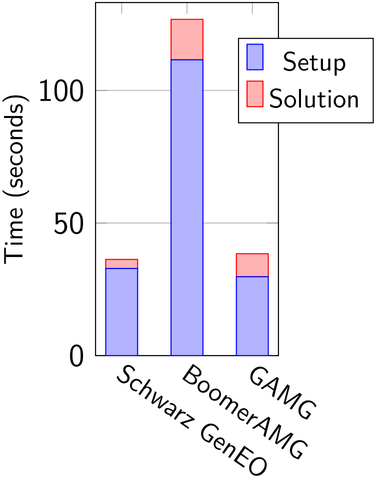
- Schwarz GenEO

↪ number of deflation vectors (increasing it only "improves" the preconditioner).

Solution of a linear system I

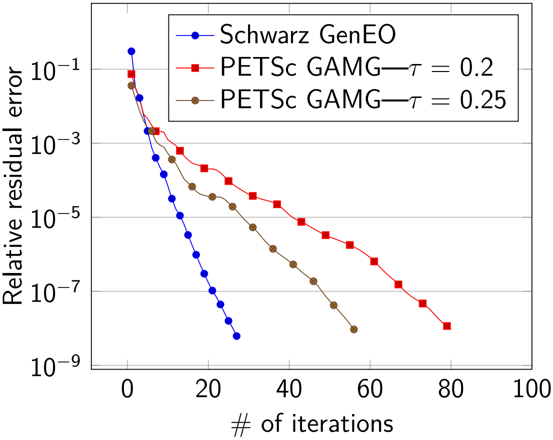
Homogeneous 3D Poisson equation discretized by $\mathbb{P}_1$ FE solved on 4 096 MPI processes, 217M d.o.f.

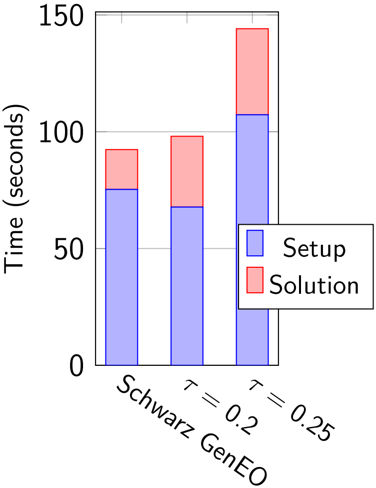




Solution of a linear system II

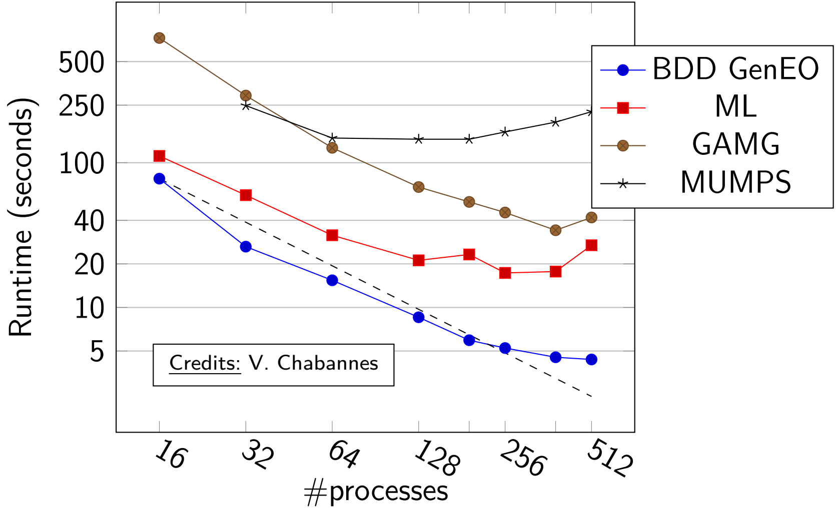
Heterogeneous 3D linear elasticity equation discretized by $\mathbb{P}_2$ FE solved on 4 096 MPI processes, 262M d.o.f.





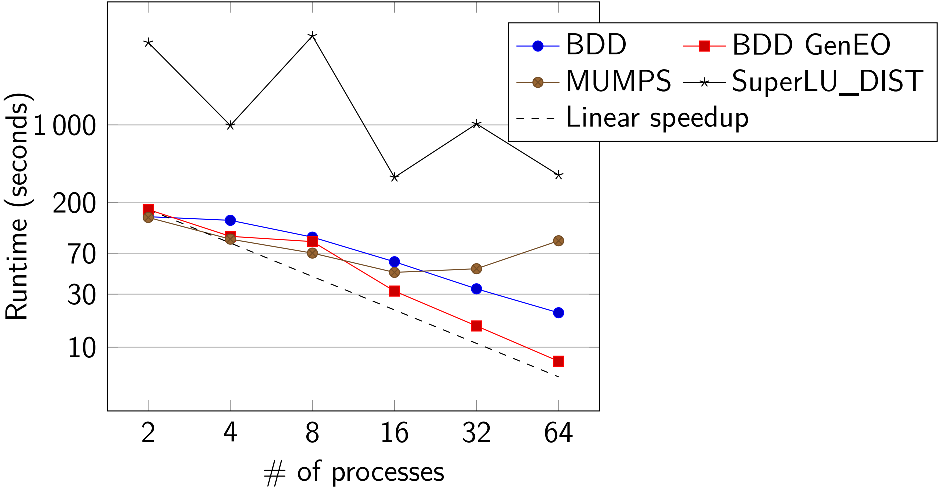
Substructuring preconditioners vs. MG solvers

Homogeneous 2D elasticity discretized by $\mathbb{P}_3$ FE, 23M d.o.f.

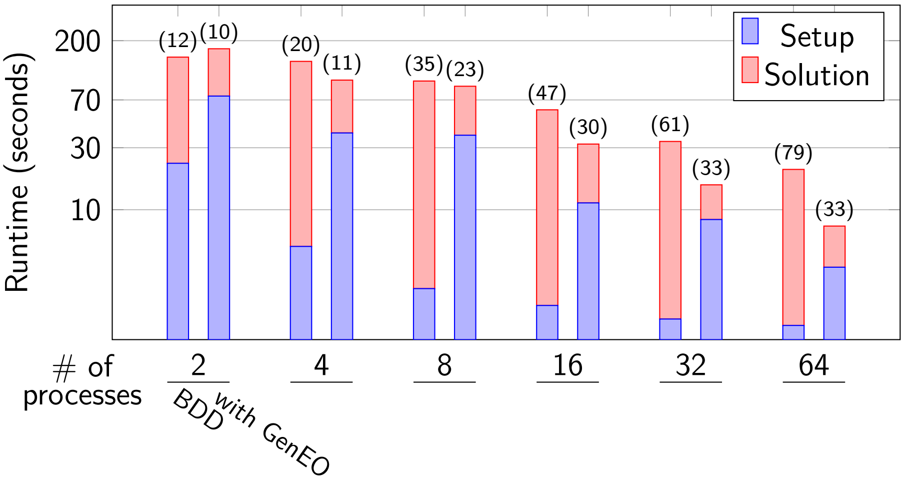


Substructuring preconditioners vs. direct solvers

Homogeneous 2D Poisson equation discretized by $\mathbb{P}_4$ FE, 5M d.o.f.



Effect of the GenEO deflation vectors



Final words

Summary:

- scalable framework for building two-level preconditioners for both Schwarz or substructuring methods (FETI, BDD),
- wide range of applications,
- still many ongoing projects:
  - assembly of $Z^\star A^\star AZ$,
  - systems with multiple RHS,
  - multilevel extension.

Perspectives

Outlooks:

- adaptive construction of the coarse operator,
- recycling for nonlinear problems.

Perspectives

Outlooks:

- adaptive construction of the coarse operator,
- recycling for nonlinear problems.

Thank you !

Cai, X-C. and M. Sarkis (1999). "Restricted additive Schwarz preconditioner for general sparse linear systems". In: *SIAM Journal on Scientific Computing* 21.2, pp. 792–797.

Conen, L., V. Dolean, R. Krause, and F. Nataf (2014). "A coarse space for heterogeneous Helmholtz problems based on the Dirichlet-to-Neumann operator". In: *Journal of Computational and Applied Mathematics* 271, pp. 83–99.

St-Cyr, A., M. Gander, and S. Thomas (2007). "Optimized multiplicative, additive, and restricted additive Schwarz preconditioning". In: *SIAM Journal on Scientific Computing* 29.6, pp. 2402–2425.

Dohrmann, C. (2003). "A preconditioner for substructuring based on constrained energy minimization". In: *SIAM Journal on Scientific Computing* 25.1, pp. 246–258.

Dolean, V., P. Jolivet, and F. Nataf (2015). *An introduction to domain decomposition methods: algorithms, theory and parallel implementation*. SIAM, in preparation.

Farhat, C., J. Mandel, and F-X. Roux (1994). "Optimal convergence properties of the FETI domain decomposition method". In: *Computer Methods in Applied Mechanics and Engineering* 115.3, pp. 365–385.

Farhat, C. and F-X. Roux (1991). "A method of finite element tearing and interconnecting and its parallel solution algorithm". In: *International Journal for Numerical Methods in Engineering* 32.6, pp. 1205–1227.

Gosselet, P. and C. Rey (2006). "Non-overlapping domain decomposition methods in structural mechanics". In: *Archives of Computational Methods in Engineering* 13.4, pp. 515–572.

Haferssas, R., P. Jolivet, and F. Nataf (2015). "A robust coarse space for Optimized Schwarz methods". In: *C. R. Math. Acad. Sci. Paris*. Submitted.

Jolivet, P., F. Hecht, F. Nataf, and C. Prud'homme (2013). "Scalable Domain Decomposition Preconditioners For Heterogeneous Elliptic Problems". In: *Proceedings of the 2013 ACM/IEEE conference on Supercomputing*. SC13. Best paper finalist. ACM, 80:1–80:11.

Jolivet, P. and F. Nataf (2014). *HPDDM*. URL: https://github.com/hpddm/hpddm.

Le Tallec, P. (1994). "Domain decomposition methods in computational mechanics". In: *Computational Mechanics Advances* 1.2, pp. 121–220.

Mandel, J. (1993). "Balancing domain decomposition". In: *Communications in Numerical Methods in Engineering* 9.3, pp. 233–241.

Mathew, T. (2008). *Domain Decomposition Methods for the Numerical Solution of Partial Differential Equations*. Vol. 61. Lecture Notes in Computational Science and Engineering. Springer.

Pechstein, C. (2012). *Finite and Boundary Element Tearing and Interconnecting Solvers for Multiscale Problems*. Vol. 90. Lecture Notes in Computational Science and Engineering. Springer.

Quarteroni, A. and A. Valli (1999). *Domain decomposition methods for partial differential equations*. Vol. 10. Clarendon Press.

Rixen, D. and C. Farhat (1997). "Preconditioning the FETI method for problems with intra- and inter-subdomain coefficient jumps". In: *Ninth International Conference on Domain Decomposition Methods*, pp. 472–479.

Schwarz, H. (1870). “Über einen Grenzübergang durch alternierendes Verfahren”. In: *Vierteljahrsschrift der Naturforschenden Gesellschaft in Zürich* 15, pp. 272–286.

Smith, B., P. Bjørstad, and W. Gropp (2004). *Domain decomposition: parallel multilevel methods for elliptic partial differential equations*. Cambridge University Press.

Spillane, N., V. Dolean, P. Hauret, F. Nataf, C. Pechstein, and R. Scheichl (2013). “Abstract robust coarse spaces for systems of PDEs via generalized eigenproblems in the overlaps”. In: *Numerische Mathematik* 126.4, pp. 741–700.

Spillane, N. and D. Rixen (2013). “Automatic spectral coarse spaces for robust FETI and BDD algorithms”. In: *International Journal for Numerical Methods in Engineering* 95.11, pp. 953–990.

Tang, J., R. Nabben, C. Vuik, and Y. Erlangga (2009). "Comparison of two-level preconditioners derived from deflation, domain decomposition and multigrid methods". In: *Journal of Scientific Computing* 39.3, pp. 340–370.

Toselli, A. and O. Widlund (2005). *Domain decomposition methods: algorithms and theory*. Vol. 34. Series in Computational Mathematics. Springer.