IMA SUMMER SCHOOL
IN
KENTUCKY

Multifrontal methods

Iain S Duff
Rutherford Appleton Laboratory
CERFACS
Strathclyde University

`http://www.cse.clrc.ac.uk/Person/I.S.Duff`

Major Limitation of Frontal Methods

Mflop/s are high but flops are also!

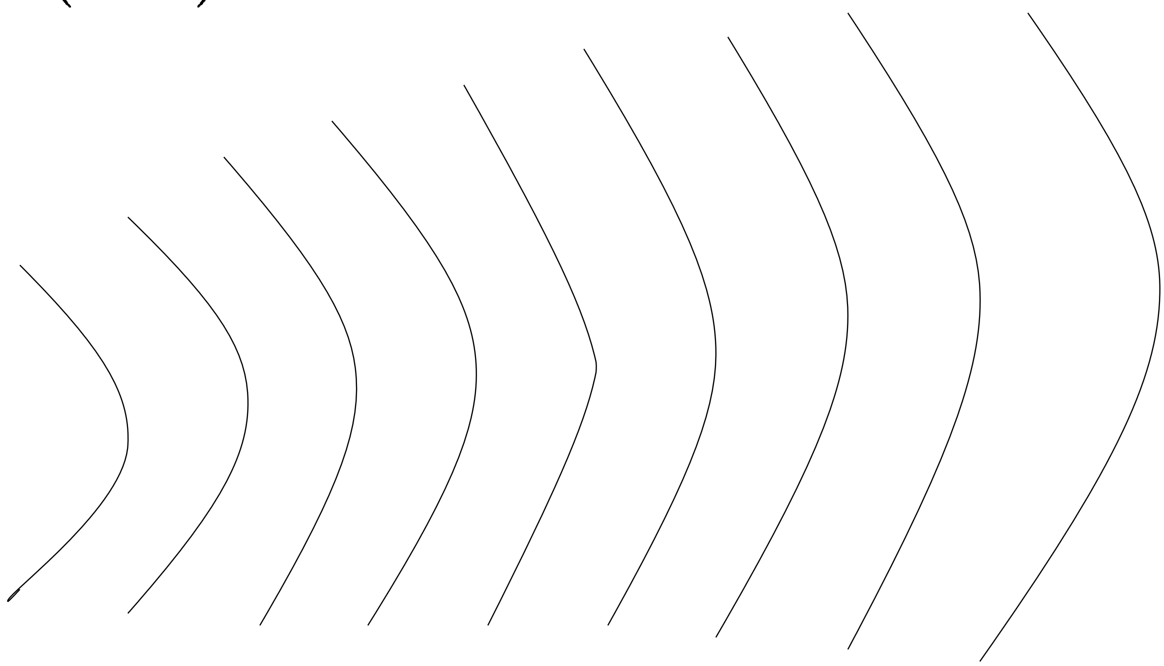
Can extend applicability by reordering but still not entirely suitable
for general systems

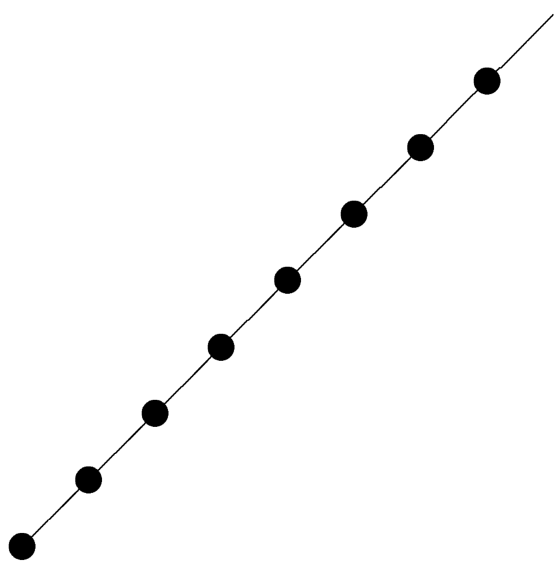
Second problem with frontal schemes relates to

Parallelism

or rather lack of it

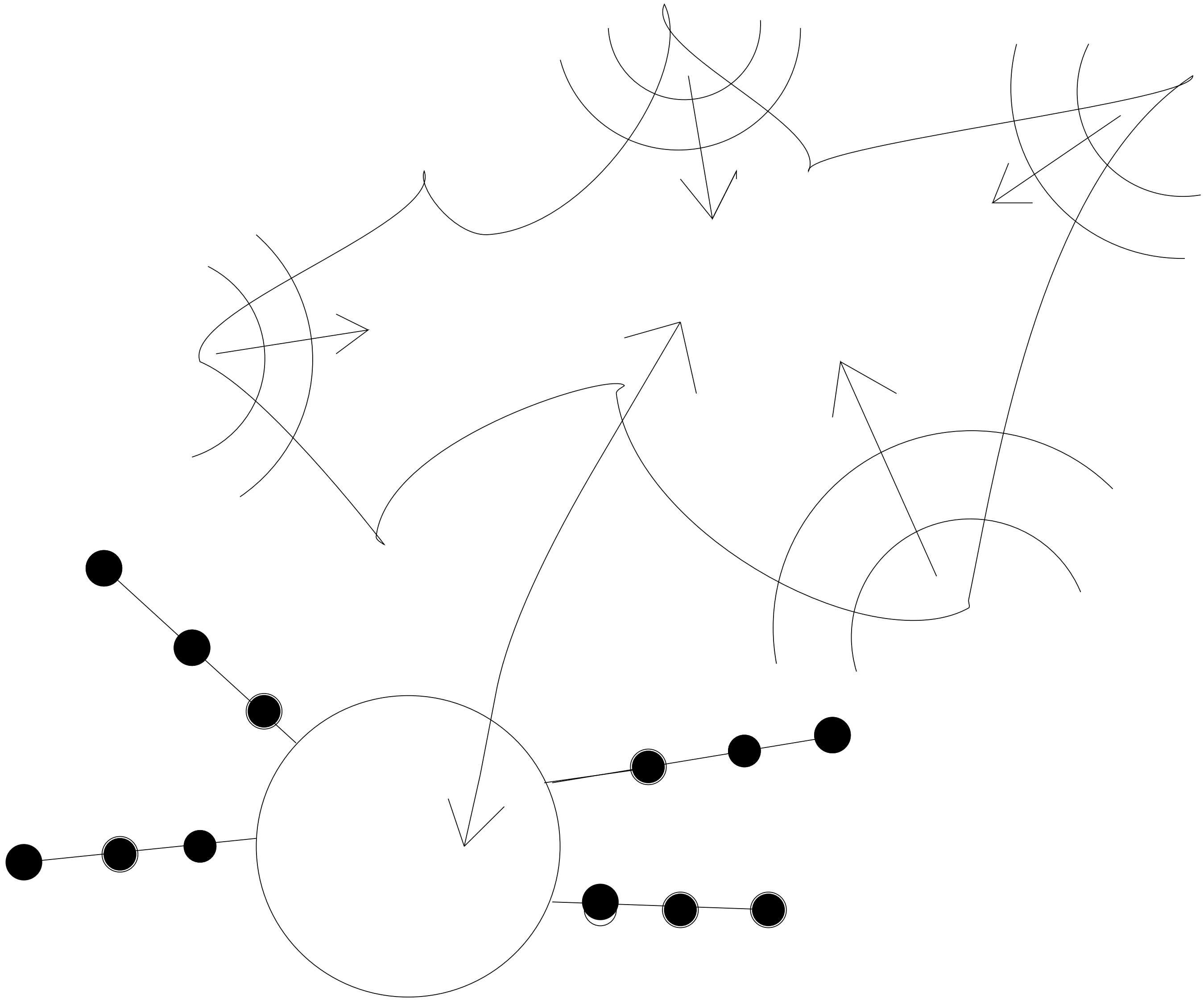
(UNI) FRONTAL





Computational tree

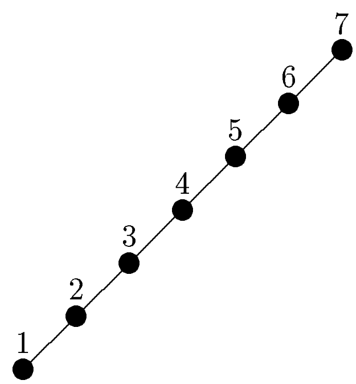
MULTIFRONTAL



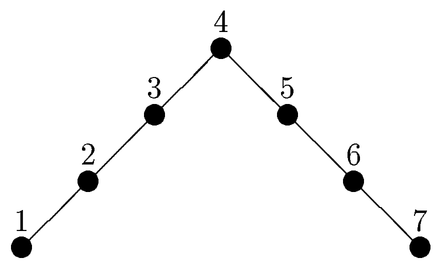
TRIDIAGONAL MATRIX

$$\begin{bmatrix} X & X & & & & & \\ X & X & X & & & & \\ & X & X & X & & & \\ & & X & X & X & & \\ & & & X & X & X & \\ & & & & X & X & X \\ & & & & & X & X \end{bmatrix}$$

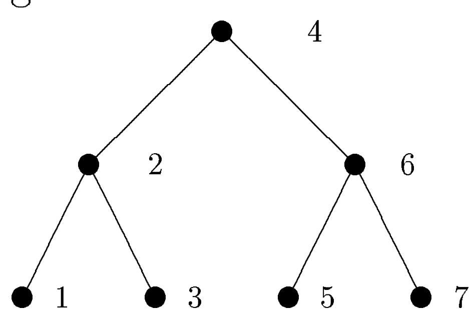
Frontal ... chain



BABE



Multifrontal ... general tree



M

Ultifrontal

Parallel

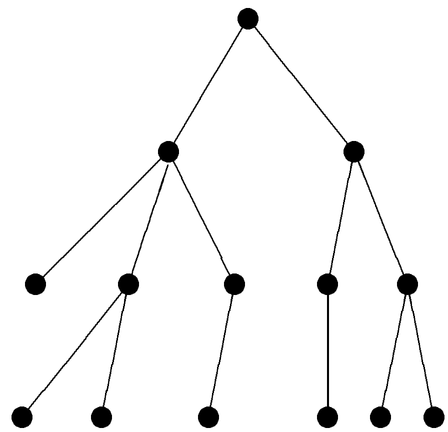
Solver

Amestoy + Duff

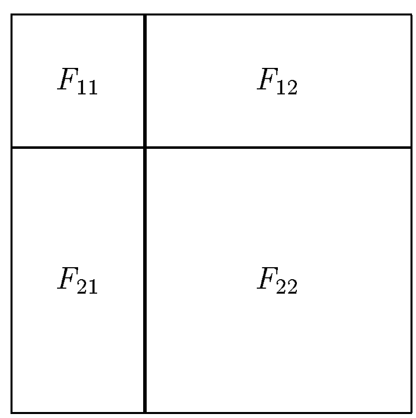
CERFACS

Designed for *nearly* symmetric systems

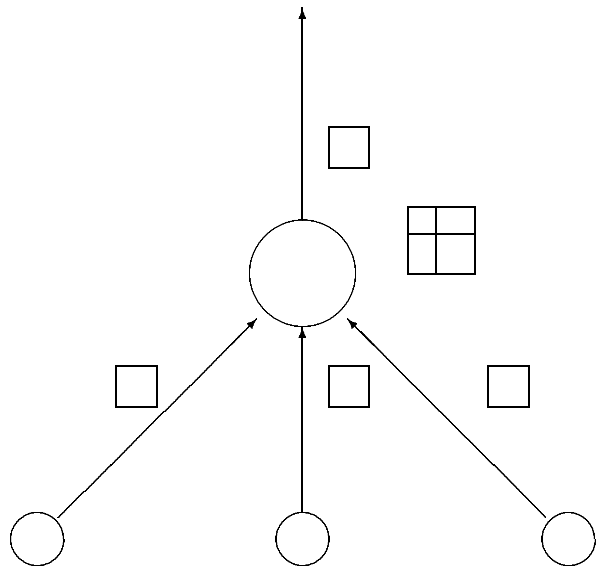
ASSEMBLY TREE



AT EACH NODE
Perform some steps of Gaussian Elimination of Full Matrices
(Frontal Matrices)



$$F_{22} \leftarrow F_{22} \; - \; F_{21} F_{11}^{-1} F_{12}$$



Assembly ...

Gather/scatter operations

Indirect addressing

Elimination ...

Full Gaussian elimination

Level 3 BLAS (TRSM, GEMM)

Perform the analysis on the pattern of $A + A^T$

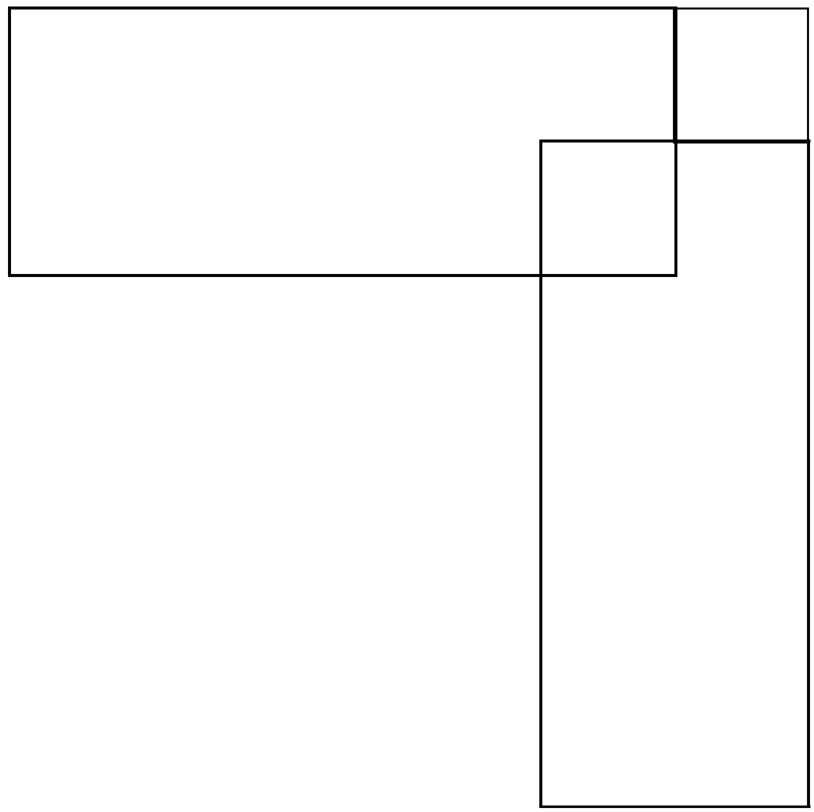
Can generate assembly tree symbolically using any sparsity preserving ordering

Then perform numerical pivoting *a posteriori* on the tree.

Performance in Mflop/s of multifrontal code MA41 on matrix BCSSTK15 on range of RISC processors

Computer	Peak	DGEMM	MA41
DEC 3000/400 AXP	133	49	34
HP 715/64	128	55	30
IBM RS6000/750	125	101	64
IBM SP2 (Thin node)	266	213	122
MEIKO CS2-HA	100	43	31

Unsymmetric Multifrontal



Unsymmetric Multifrontal

- Assembly tree no longer suitable model for computation.

  Use directed acyclic graph *dag*.

- Frontal matrices no longer square.

  Use rectangular frontal matrices.

- Fronts persist in dag rather than being absorbed by parent node.

  More complicated data structures.

- Very sensitive to numerical pivoting.

  Pay attentions to numerical values when doing analysis.

Code

UMFPACK/MA38 ..... Davis and Duff

See Report:

A combined unifrontal/multifrontal method for unsymmetric sparse matrices.
Timothy A. Davis and Iain S. Duff.
Report RAL-TR-97-046, Rutherford Appleton Laboratory.

`http://www.numerical.rl.ac.uk/reports/reports.html`

Comparison of "symmetric" (MA41) and "unsymmetric" (MA38) code

Matrix	FIDAPM11	EX11	WANG4	RIM	TWOTONE	ONETONE1
Order	22294	16614	26068	22560	120750	36057
Nonzeros	623554	1096948	177196	1014951	1224224	341088
Index of asymmetry	0.00	0.00	0.00	0.35	0.72	0.90
Floating-point ops (10^9)						
MA41	9.7	6.7	10.5	0.6	29.6	2.3
MA41 (zero exploiting)	9.7	6.7	10.5	0.5	18.5	1.7
MA38/UMFPACK V2	156.7	56.2	90.3	7.3	9.2	2.4
UMFPACK V3	19.5	6.6	29.5	7.8	10.8	2.1
Factorization time (*seconds on Origin 2000*)						
MA41	41.3	29.5	43.9	4.6	162.9	15.3
MA41 (zero exploiting)	41.4	29.6	44.0	4.9	87.7	9.6
MA38/UMFPACK V2	1107.5	368.3	602.2	59.0	91.5	20.4
UMFPACK V3	151.4	54.3	188.5	56.8	126.9	19.8

$$\text{Index of asymmetry} = \frac{\text{Number of pairs such that } a_{ij}=0, a_{ji}\neq 0}{\text{Total number of off-diagonal entries}}$$

Results from

Amestoy, P.R. and Puglisi, C. (2000).
An unsymmetrized multifrontal LU factorization.
Technical report ENSEEIHT-IRIT RT/APO/00/3.

Comparison of multifrontal against supernodal approach

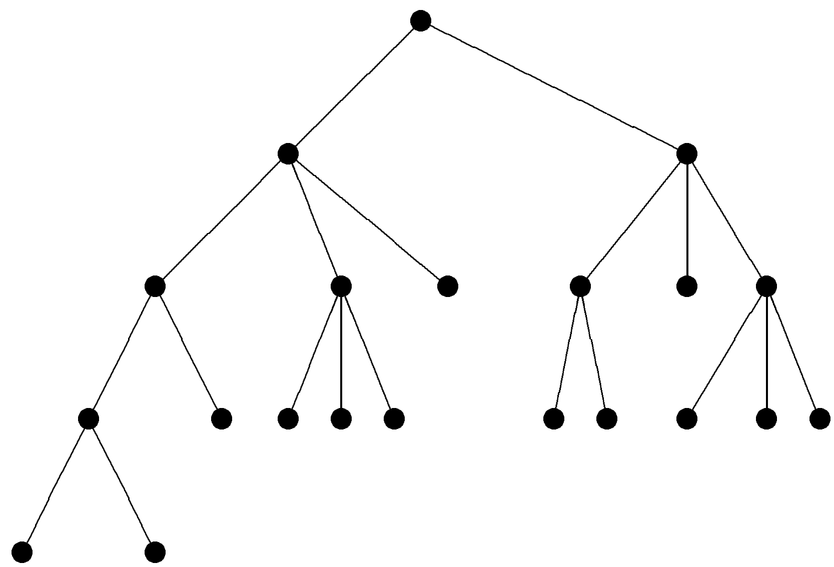
Times in seconds on an Origin 2000

Matrix	Order	Entries	Analyse and Factorize Time (secs)		Entries in LU factors (10^6)	
			SuperLU	MA41	SuperLU	MA41
ONETONE2	36057	227628	2.8	2.9	1.2	1.9
TWOTONE	120750	1224224	198.4	87.7	21.2	17.0
WANG4	26068	177196	323.2	44.0	26.2	11.6
RIM	22560	1014951	93.1	4.9	19.1	4.0
LHR71C	70304	1528092	17.5	25.0	7.3	11.7

Results from

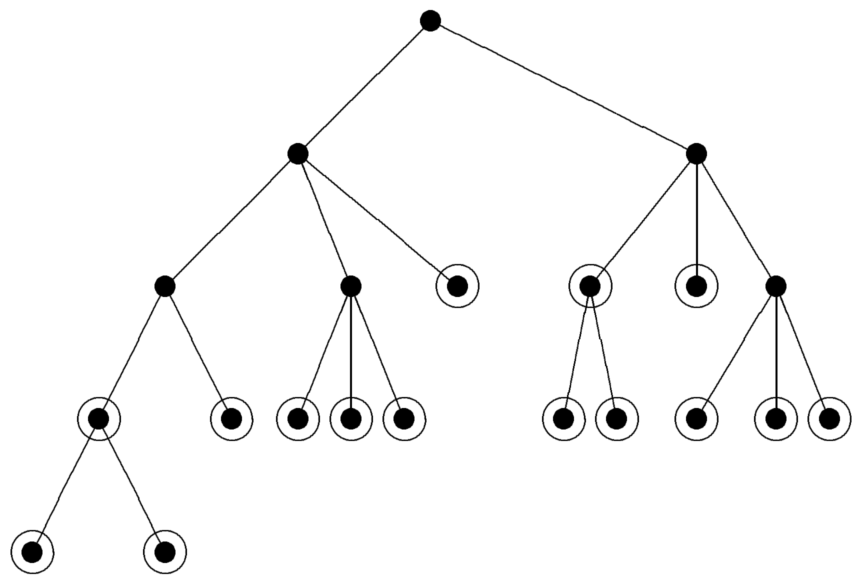
Amestoy, P.R. and Puglisi, C. (2000).
An unsymmetrized multifrontal LU factorization.
Technical report ENSEEIHT-IRIT RT/APO/00/3.

PARALLEL IMPLEMENTATION



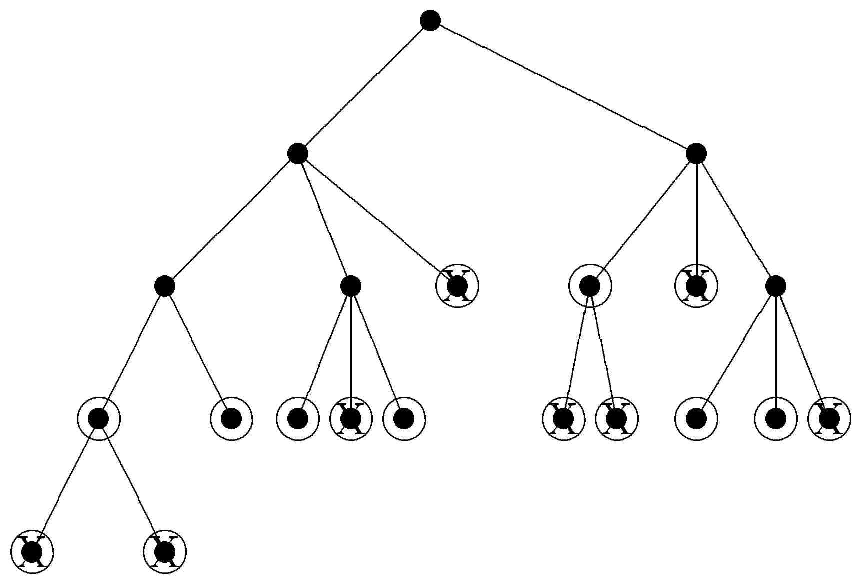
Assemby tree for multifrontal method

PARALLEL IMPLEMENTATION



Work corresponding to leaf nodes can proceed immediately and independently

PARALLEL IMPLEMENTATION



Work corresponding to leaf nodes can proceed immediately and independently

When all children complete then work can commence at parent node.

Statistics on front sizes in tree

			Leaf nodes		Top 3 levels	
Matrix	Order	Tree nodes	Number	Av. size	Number	Av. size
BCSSTK15	3948	576	317	13	10	376
BCSSTK33	8738	545	198	5	10	711
BBMAT	38744	5716	3621	23	10	1463
GRE1107	1107	344	250	7	12	129
SAYLR4	3564	1341	1010	5	12	123
GEMAT11	4929	1300	973	10	112	148

Typically 75% work in top three levels

Two forms of parallelism

1. Tree parallelism [from assembly tree]
2. Node parallelism

$$\begin{bmatrix} \mathbf{F}_{11} & \mathbf{F}_{12} \\ \mathbf{F}_{21} & \mathbf{F}_{22} \end{bmatrix}$$

as in full linear algebra (eg LAPACK)

Shared memory parallel performance

MA41 on 7 processors of a CRAY C98

In columns (1) we exploit only parallelism from the tree

In columns (2) we combine the two levels of parallelism

			(1)		(2)	
Matrix	order	entries	Mflop/s	(speedup)	Mflop/s	(speedup)
WANG3	26064	177168	1062	(1.42)	3718	(4.98)
WANG4	26068	177196	1262	(1.70)	3994	(5.39)
BBMAT	38744	1771722	2182	(3.15)	3777	(5.46)

DISTRIBUTED MEMORY