Ideas for alleviating GSI computer limitations

David F. Parrish[1] and Rahul Mahajan[2]

[1] NOAA/NWS/NCEP/Environmental Modeling Center
[2] IMSG

Computational limitations for GSI

In this short presentation, I will discuss two specific computational limitations in the GSI code.

1) Significant additional run time (~2-5 min) was added to global and regional GSI due to change of format for global ensemble members from spectral coefficients to NEMS format grid files.
2) GSI is limited to O(400) processors for parallel computing, primarily because of how the background error covariance is computed.

We already have a fix for item 1, which is now in operations for global and regional GSI. Item 2 is much more challenging.

Computational limitations for GSI

1) Significant additional run time (~2-5 min) was added to global and regional GSI due to change of format for global ensemble members from spectral coefficients to NEMS format grid files.

NEMS format global ensemble input more expensive than spectral for 3d/4d EnVar GSI

Essentially the same code was used to read global spectral format ensemble for both global and regional runs.

The read was done one ensemble member at a time. Reading was done in parallel over all processors, with every processor reading in a single horizontal variable/level, and all processors simultaneously reading from the same ensemble file.

Not so bad, since read contention was likely masked by significant computation to convert from spectral to grid independently on every processor.

global ensemble input very expensive for 3d/4d EnVar GSI

In operations, the use of spectral files for GFS has been replaced by NEMS format grid files. This is now operational for global and regional GSI.

A nemsio version of the ensemble read routine for global model now exists in GSI, which reads grid files instead of spectral files. The initial version increased run times substantially, which could not be improved by adding more processors.

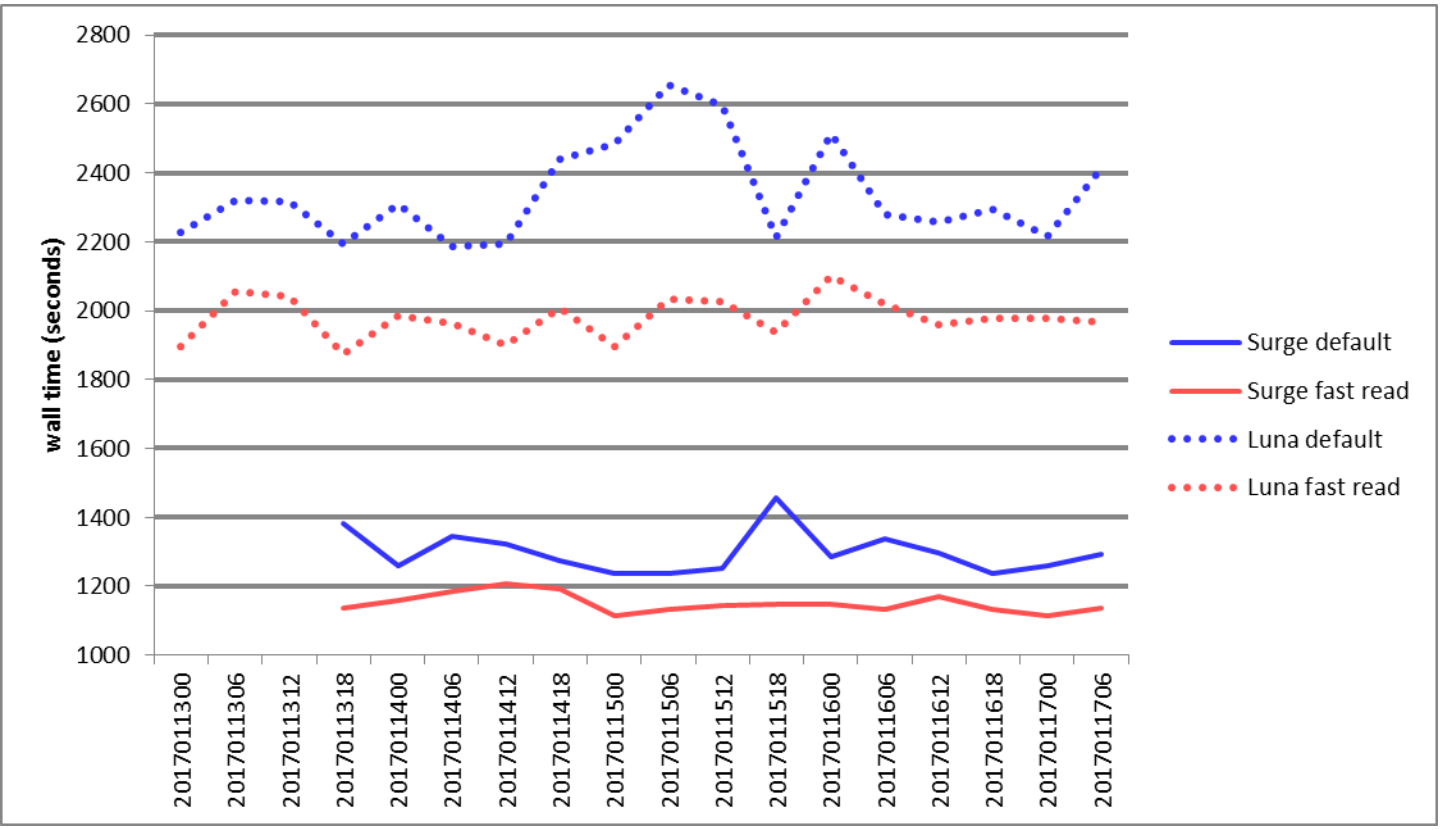
The new fast ensemble read saves 2-5 minutes of run time compared to the first version. It is now operational for both global and regional GSI.

global ensemble input very expensive for 3d/4d EnVar GSI

Rahul learned about this new method for faster read of the global gridded ensembles from a meeting late in 2016. Mark Buehner presented results where 256 ensemble members were read in and scattered to subdomains in O(5) seconds. This was accomplished by simultaneous reads of 256 entire members by 256 processors, followed by a scatter to subdomains on all processors.

global ensemble input very expensive for 3d/4d EnVar GSI

Here are some run times for parallel NEMS global gfs (now operational with fast read—thanks to Russ Treadon for gathering this data).

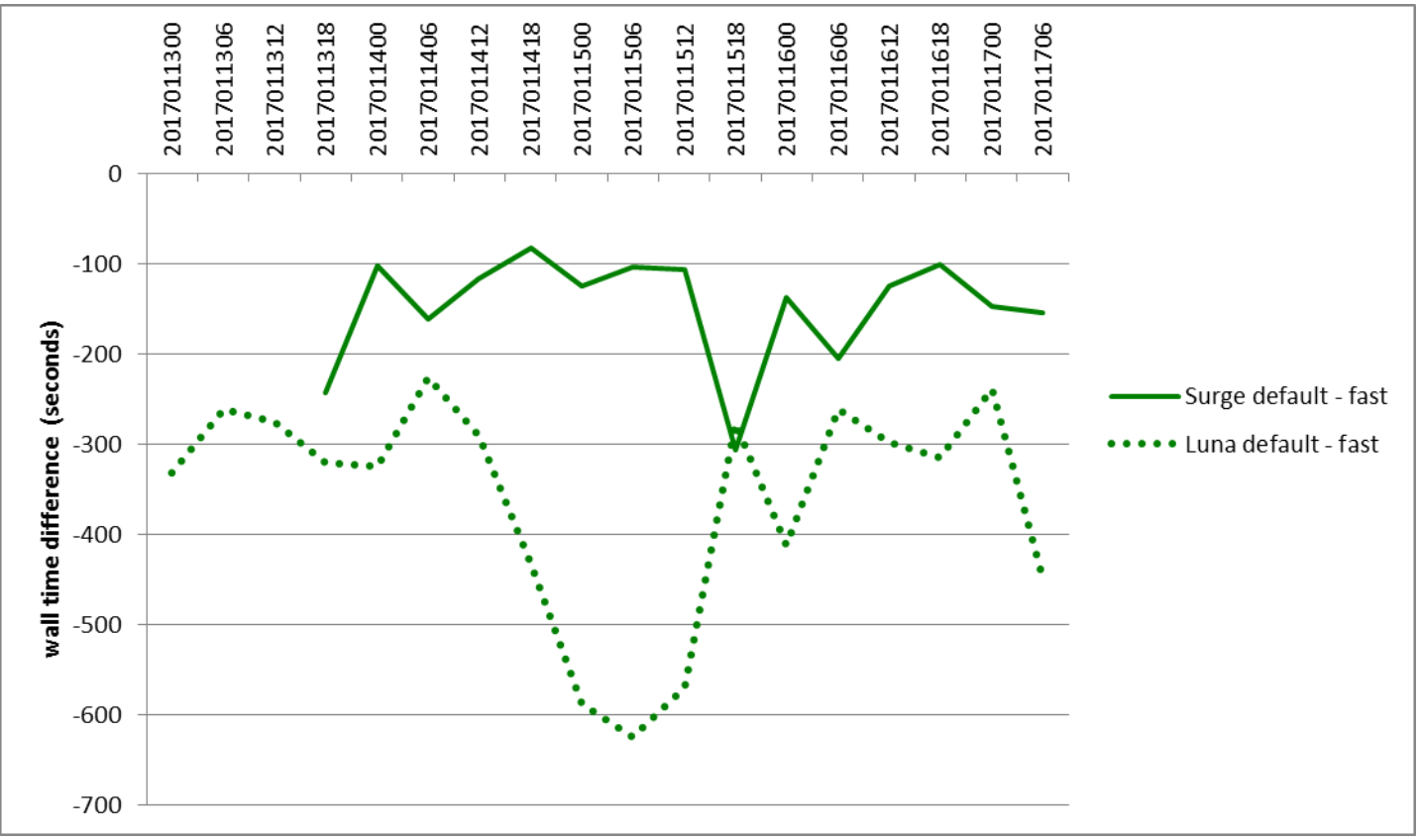


Luna (development):
60 nodes:
360 tasks,
ptile=6,
threads=4.

Surge (operational):
240 nodes
480 tasks,
ptile=2,
threads=12.

global ensemble input very expensive for 3d/4d EnVar GSI

Reduction in times for parallel NEMS global gfs (original - fast read)



Computational limitations for GSI

2) GSI is limited to O(400) processors for parallel computing, primarily because of how the background error covariance is computed.

Background error

The current operational global GSI is 4denvar, but the background error **B** is still the same, so is illustrated with this basic 3dvar cost function:

$$J(\mathbf{x}) = \frac{1}{2}\left[\, \mathbf{x}^T\mathbf{B}^{-1}\mathbf{x} + (\mathbf{H}\mathbf{x} - \mathbf{y})^T\mathbf{R}^{-1}(\mathbf{H}\mathbf{x} - \mathbf{y}) \,\right]$$

where

$\mathbf{x} = \mathbf{x}_a - \mathbf{x}_b$, $\mathbf{x}_a$ = analysis, $\mathbf{x}_b$ = background,

$\mathbf{y} = \mathbf{y}_o - \boldsymbol{H}(\mathbf{x}_b)$, $\mathbf{y}_o$ = observations,

$\boldsymbol{H}$ = observation forward operator,

$\mathbf{H}$ = tangent linear of $\boldsymbol{H}$,

$\mathbf{B}$ = background error covariance,

$\mathbf{R}$ = observation error covariance.

Background error

The analysis increment **x** is determined by a conjugate gradient minimization iteration (up to 250 total iterations for global GSI). Each iteration requires the matrix-vector multiplication **Bx**. **B** is implemented in the horizontal using a recursive filter (Wu et al, 2002). This algorithm requires the entire horizontal field. This is why there is an upper limit of O(400) processors for parallel computing. The operational **x** contains 6 3d variables with 64 layers each and 2 2d variables, for a total of 386 2d fields.

Wu,W.-S., D. F. Parrish, and R. J. Purser, 2002: Three-dimensional variational analysis with spatially inhomogeneous covariances. *Mon. Wea. Rev.* , **130**, 2905–2916.

Background error

A proposal for **Bx** on FV3 cube grid

Covariance operators on the equiangular gnomonic cubic grid,
R. James Purser, Miodrag Ranˇci´c, Manual de Pondeca, and David F. Parrish, NOAA/NCEP/EMC, College Park, MD 20740-3818, U.S.A.
(Email: jim.purser@noaa.gov)

Above submitted to 2017 WGNE Blue Book:

This is a new application of the recursive filter specifically to the FV3 gnomonic cube sphere. I asked Jim if this could be applied in a subdomain context. He didn't think so, but the 6 faces could be extended with overlap to adjacent faces.

Background error

A proposal for **Bx** on FV3 cube grid

Covariance operators on the equiangular gnomonic cubic grid,
R. James Purser, Miodrag Ranˇci´c, Manual de Pondeca, and David F. Parrish, NOAA/NCEP/EMC, College Park, MD 20740-3818, U.S.A.
(Email: jim.purser@noaa.gov)

Above submitted to 2017 WGNE Blue Book:

I suppose that this could allow 6 streams of computation across 6 processors per variable, so the number kept busy would be O(2400) processors. This would be a significant advance, with a method similar to what is already used in GSI, so has good probability of success. Also, it can still be made bit reproducible for different number of processors.

Background error

Subdomain version of **Bx**?

One of us (Parrish) created a single processor multigrid algorithm for direct computation of **Bx**.

This was applied to the RTMA option of GSI and seemed to give good results and ran a little faster when RTMA runs in full horizontal slab option. However, RTMA uses the anisotropic version of recursive filter, which can also run on subdomains. It is very heavily communication bound, but the existing subdomain form runs faster than both the horizontal slab version and the multigrid **Bx**.

Multigrid correlation algorithm:

To compute **w** = **Cu** using the hierarchy of M grids **:**

Fine to coarse pass:

$\mathbf{u}_1 = \mathbf{u}$, ($\mathbf{C}_1 = \mathbf{C}$) (don't compute this—not practical)

do m = 2, M
  $\mathbf{u}_m = \mathbf{H}_m^T \mathbf{u}_{m-1}$ (adjoint of centered polynomial interpolation from fine to coarse grid—adjoint needed to preserve symmetry of **C**)
end do

$\mathbf{w}_M = \mathbf{C}_M \mathbf{u}_M$ (direct computation by full $\mathbf{C}_M$ — only O(n) elements in $\mathbf{C}_M$ — in implementation n_used << n !)

Coarse to fine pass:

do m = M-1,1,-1
  $\mathbf{w}_m = \mathbf{H}_{m+1} \mathbf{w}_{m+1} + \mathbf{E}_m \mathbf{u}_m$ (forward polynomial interpolation and correlation correction term)
end do

Multigrid correlation algorithm:

$\mathbf{H}_m$ interpolates from $\mathbf{u}_m$ to $\mathbf{u}_{m-1}$ using centered polynomial interpolation

$$\mathbf{E}_m = \mathbf{C}_m - \mathbf{H}_{m+1}\, \mathbf{C}_{m+1}\, \mathbf{H}_{m+1}^{T}$$

$\mathbf{E}_m$ is the matrix of interpolation error between the exact correlation at level m and the interpolated correlation from level m+1. In general, $\mathbf{E}_m$ is a full matrix, but by setting an error tolerance ε which is related to the discretization error on the finest grid, the actual matrix will be very sparse.

Interpolation orders and tolerances used in tests:

1^{st} , 3^{rd} , 5^{th} , 7^{th} order (2, 4, 6, 8 points respectively)

ε = 0.01, 0.001, 0.0001

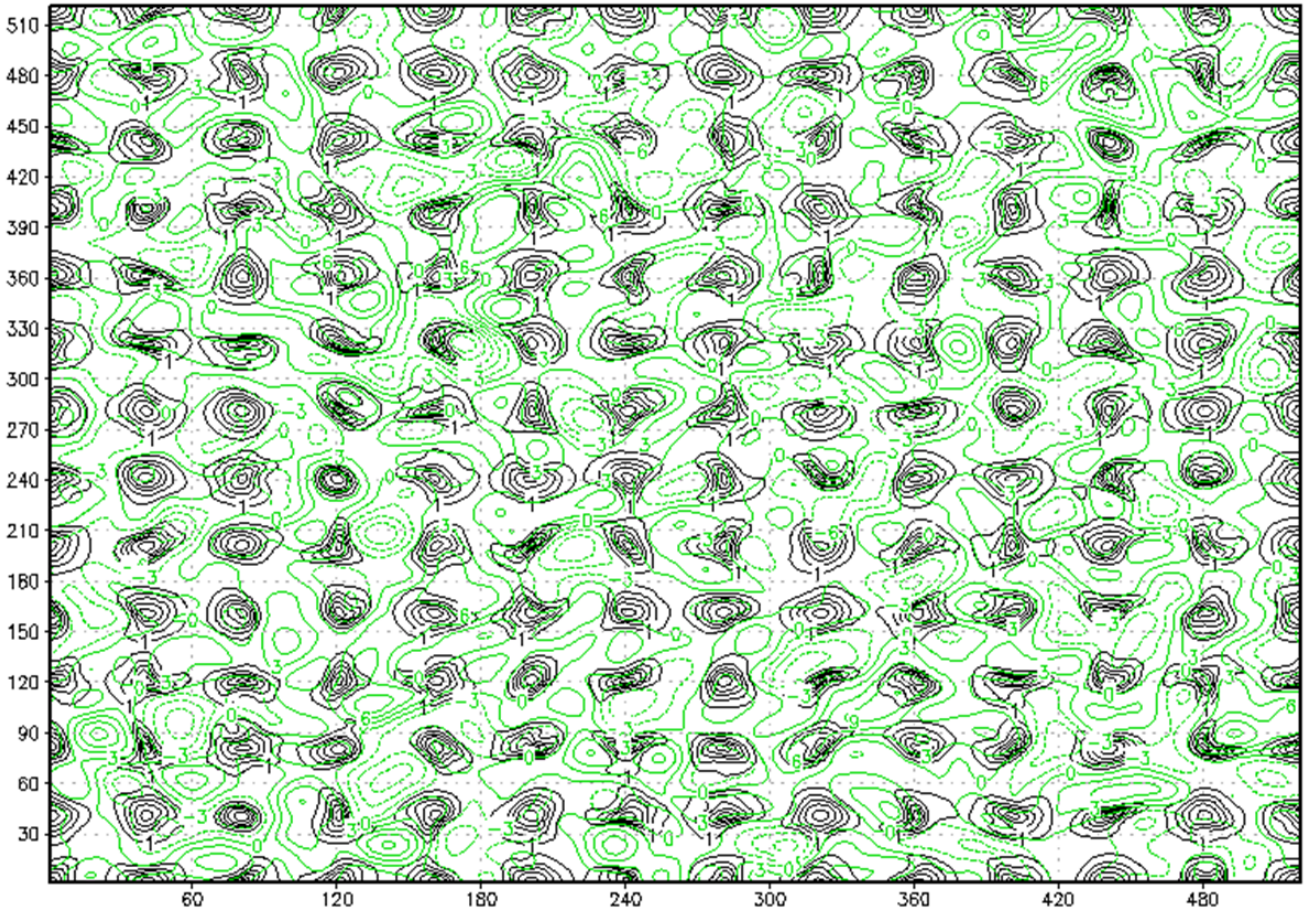
Correlation function used:

$$\mathbf{C}(\mathbf{x}_1,\mathbf{x}_2,f_1,f_2) = \exp\left[- |\mathbf{x}_2 - \mathbf{x}_1|^2 / (L_{xy}(\mathbf{x}_1) L_{xy}(\mathbf{x}_2)) - ((f_2 - f_1)/L_f)^2 \right]$$

inhomogeneous isotropic

Riishøjgaard anisotropic

Anisotropic example:



521 x 521 grid Riishøjgaard coupled correlation. Test point every 40 grid points in x and y, 194 samples all together. Green contours are for field that the correlations follow. Black contours are 0.1, 0.3, 0.5, 0.7, 0.9. Correlation computed to accuracy of 0.01 with cubic interpolation grid transfers. Cost to setup information needed for multigrid algorithm: 214 seconds. Cost for v = C*u : 0.07 seconds. Clearly need to work on the setup cost.

References (for multigrid **Bx**)

Brandt, A., and A. A. Lubrecht, 1990: Multilevel Matrix Multiplication and Fast Solution of Integral Equations. Journal of Computational Physics, vol 90, 348-370.

Parrish, D., 2015: Application of a multigrid correlation algorithm to anisotropic correlations used by the NCEP Real Time Mesoscale Analysis (RTMA). 27th Conference On Weather Analysis And Forecasting/23rd Conference On Numerical Weather Prediction, June28-July 3 2015, Chicago, IL.

Purser, R. J., W.-S. Wu, D. F. Parrish, and N. M. Roberts, 2003a: Numerical aspects of the application of recursive filters to variational statistical analysis. Part I: Spatially homogeneous and isotropic Gaussian covariances. *Mon. Wea. Rev.*, **131**, 1524-1535.

Purser, R. J., W.-S. Wu, D. F. Parrish, and N. M. Roberts, 2003b: Numerical aspects of the application of recursive filters to variational statistical analysis. Part II: Spatially inhomogeneous and anisotropic general covariances. *Mon. Wea. Rev.*, **131**, 1536-1548.

Riishøjgaard, L. P., 1998. A direct way of specifying flow-dependent background error correlations for meteorological analysis systems. *Tellus* 50, 42-57.

Background error

Subdomain version of **Bx**?

A new investigation is underway to see if a practical subdomain version of the multigrid **Bx** algorithm can be created. As a simpler alternative, a non-multigrid subdomain version is also being considered. It is too soon to know if either of these will be practical alternatives. As EMC is committed to the FV3 model, the recursive filter application on gnomonic cube sphere may end up being the practical way forward.


NOAA
NATIONAL OCEANIC AND ATMOSPHERIC ADMINISTRATION
U.S. DEPARTMENT OF COMMERCE


NATIONAL CENTERS FOR ENVIRONMENTAL PREDICTION
NCEP

Questions?