(Section contribution for klf river paper)

1. Parallel computational performance

The Muskingum-Cunge algorithm is characterized by a computational pattern wherein at each timestep, a single flow computation is performed at each cell on the routing grid. However, the computations are required to be ordered so as to resolve the dependency of downstream points on cells upstream. Thus, the computation flows from headwater cells toward ocean discharge cells. The longest single sequence of computations following this dependency pattern is quite simply the longest river in the network. For the 30-minute routing grid of Vörösmarty et al. (2000) used in this study, this and other characteristics of the grid are listed in Table 1.

For a routing grid discretized into a 2D array, parallelism introduces specific problems not generally faced in other more conventional finite-difference methods. In particular, simple 2D domain decomposition is unlikely to be ideal, as at any step in the computation there is potentially some flow trajectory crossing a domain boundary and requiring parallel communication.

We have instead chosen a parallel distribution of a 2D grid where entire basins are guaranteed to stay on-processor. Basins on the routing grid are sorted by size descending, and distributed round-robin among available processors, with each subsequent basin assigned to the processor currently containing the smallest number of cells. This method provides maximal load balance given a particular distribution of basin sizes. It is also limited in scaling by the size of the largest basin relative to the global grid. As seen in Table 1, this is $\sim$30 for our chosen grid. Parallel communication occurs only at times when I/O is performed, when we gather data from the global grid onto a single root processor for output. The scaling limit on the river grid does not limit any other component of the GFDL Earth System Model, where each component can be discretized and parallelized independently and run concurrently, exchanging data using an *exchange grid* (Balaji et al. 2005).

We have chosen to discretize the routing grid as a conventional two-dimensional grid. While reordering the grid within each basin following headwater-to-discharge order is likely to improve performance further, we have so far chosen to retain conventional ordering in the interests of code readability and maintenance. Current performance for the 30-minute grid is about 45**(? Kirsten/Lori, please substitute correct number)** seconds per month of model integration on 30 SGI/Altix processors. This is currently considered to be sufficient for use in the context of the GFDL CM2 model (Delworth et al. 2005; Gnanadesikan et al. 2005), where the entire coupled model, of which this will be one component, executes at roughly 20 minutes per month of integration. This may be revisited in the future (such as when multiple nutrients are transported) if the river model assumes a significant computational load in the context of its parent Earth system model.

Field	Value
Total cell count	59132
Longest path	72
Largest basin	1916

Table 1: Characteristics of the 30-minute river routing grid of Vörösmarty et al. (2000). The *total cell count* is the number of cells on the global routing grid. The *longest path* is the longest cell path that must be computed in upstream-to-downstream order, and is associated with the Nile(?). The *largest basin* (Amazon?) is the maximum cell count associated with a single basin: the total cell count divided by the largest basin size defines the maximum parallelism.

References

Balaji, V., J. Anderson, I. Held, M. Winton, J. Durachta, S. Malyshev, and R. J. Stouffer, 2005: The Exchange Grid: a mechanism for data exchange between Earth System components on independent grids. *Parallel CFD*.

Delworth, T., A. Broccoli, A. Rosati, R. Stouffer, V. Balaji, J. Beesley, W. Cooke, K. Dixon, J. Dunne, K. Dunne, J. Durachta, K. Findell, P. Ginoux, A. Gnanadesikan, C. Gordon, S. Griffies, R. Gudgel, M. Harrison, I. Held, R. Hemler, L. Horowitz, S. Klein, T. Knutson, P. Kushner, A. Langenhorst, H.-C. Lee, S.-J. Lin, J. Lu, S. Malyshev, P. D. Milly, V. Ramaswamy, J. Russell, M. Schwarzkopf, E. Shevliakova, J. Sirutis, M. Spelman, W. Stern, M. Winton, A. Wittenberg, B. Wyman, F. Zeng, and R. Zhang, 2005: GFDL's CM2 Global Coupled Climate Models. Part 1: Formulation and Simulation Characteristics. *Journal of Climate*, **in press**.

Gnanadesikan, A., K. Dixon, S. Griffies, V. Balaji, J. Beesley, W. Cooke, T. Delworth, R. Gerdes, M. Harrison, I. Held, W. Hurlin, H.-C. Lee, Z. Liang, G. Nong, R. Pacanowski, A. Rosati, J. Russell, B. Samuels, S. Song, M. Spelman, R. Stouffer, C. Sweeney, G. Vecchi, M. Winton, A. Wittenberg, F. Zeng, and R. Zhang, 2005: GFDL's CM2 global coupled climate models, Part 2. The baseline ocean simulation. *Journal of Climate*, **in press**.

Vörösmarty, C. J., B. M. Fekete, M. Meybeck, and R. B. Lammers, 2000: Geomorphometric attributes of the global system of rivers at 30-minute spatial resolution. *Journal of Hydrology*, **237**, 17–39.