

High-Order Finite Element Methods for Parallel Atmospheric Modeling

Amik St.-Cyr and Stephen J. Thomas

National Center for Atmospheric Research,
1850 Table Mesa Drive, Boulder, 80305 CO, USA
{amik, thomas}@ucar.edu

Abstract. High-order finite element methods for the atmospheric shallow water equations are reviewed. The accuracy and efficiency of nodal continuous and discontinuous Galerkin spectral elements are evaluated using the standard test problems proposed by Williamson et al (1992). The relative merits of strong-stability preserving (SSP) explicit Runge-Kutta and multistep time discretizations are discussed. Distributed memory MPI implementations are compared on the basis of the total computation time required, sustained performance and parallel scalability. Because a discontinuous Galerkin method permits the overlap of computation and communication, higher sustained execution rates are possible at large processor counts.

1 Introduction

High-order finite element methods are well-suited to atmospheric modeling due to their desirable numerical properties and inherent parallelism. A spectral element atmospheric model received an honorable mention in the 2001 Gordon Bell award competition, Loft et al (2001). Discontinuous Galerkin approximations are an extension of low order finite-volume techniques for compressible flows with shocks (Cockburn et al 2000). Either nodal or modal basis functions can be employed in high-order finite elements and the methods are spectrally accurate for smooth solutions. To avoid excessive memory requirements, global assembly of finite element matrices is avoided in the continuous Galerkin method by applying a direct-stiffness summation, Deville et al (2002). Computations within an element are based on tensor-product summations, taking the form of dense matrix-matrix multiplications. These are naturally cache-blocked and can be unrolled to expose instruction level parallelism to processors containing multiple floating point units.

The shallow water equations are a prototype for atmospheric general circulation models. The parallel performance of a 3D model can be estimated by solving identical 2D shallow water problems on multiple layers. To evaluate the efficiency of various time integrators, schemes of equivalent order will be compared on the basis of the total wall-clock time required to solve a given initial value problem (i.e. time to solution). The amount of computation required by

a method depends on the Courant number or equivalently the time step size. Moreover, the parallel performance depends on the number of right-hand side evaluations per time step and the associated parallel communication. In the case of the discontinuous Galerkin method, communication of conserved variables can be overlapped with the weak divergence and source term computations. The continuous Galerkin spectral element model developed by Loft et al (2002) will serve as the parallel performance baseline for our simulations.

2 Shallow Water Equations

The shallow water equations contain the essential wave propagation mechanisms found in atmospheric general circulation models. These are the fast-moving gravity waves and nonlinear Rossby waves. The latter are important for correctly capturing nonlinear atmospheric dynamics. The flux form shallow-water equations in curvilinear coordinates are described in Sadourny (1972).

$$\begin{aligned} \frac{\partial u_1}{\partial t} + \frac{\partial}{\partial x^1} E &= \sqrt{G} u^2 (f + \zeta), \\ \frac{\partial u_2}{\partial t} + \frac{\partial}{\partial x^2} E &= -\sqrt{G} u^1 (f + \zeta), \\ \frac{\partial}{\partial t} (\sqrt{G} \Phi) + \frac{\partial}{\partial x^1} (\sqrt{G} u^1 \Phi) + \frac{\partial}{\partial x^2} (\sqrt{G} u^2 \Phi) &= 0, \end{aligned}$$

where

$$E = \Phi + \frac{1}{2} (u_1 u^1 + u_2 u^2), \quad \zeta = \frac{1}{\sqrt{G}} \left[\frac{\partial u_2}{\partial x^1} - \frac{\partial u_1}{\partial x^2} \right]$$

h is the height above sea level. u^i and u_j are the contravariant and covariant velocities. $\Phi = gh$ the geopotential height. f is the Coriolis parameter. The metric tensor is G_{ij} and $G = \det(G_{ij})$.

3 Space Discretization

The computational domain Ω is partitioned into finite elements Ω_k . An approximate solution u_h belongs to the finite dimensional space $\mathcal{V}_h(\Omega)$. u_h is expanded in terms of a tensor-product of the Lagrange basis functions defined at the Gauss-Lobatto-Legendre points

$$u_h^k = \sum_{i=0}^N \sum_{j=0}^N u_{ij} h_i(x) h_j(y)$$

A weak Galerkin variational problem is obtained by integrating the equations with respect to a test function $\varphi_h \in \mathcal{V}_h$. In the continuous Galerkin spectral element method, integrals are directly evaluated using Gauss-Lobatto quadrature and continuity is enforced at the element boundaries.

To illustrate the discontinuous Galerkin approach, consider a scalar hyperbolic equation in flux form,

$$u_t + \nabla \cdot \mathcal{F} = S.$$

By applying the Gauss divergence theorem, the weak form becomes

$$\frac{d}{dt} \int_{\Omega_k} \varphi_h u_h d\Omega = \int_{\Omega_k} \varphi_h S d\Omega + \int_{\Omega_k} \mathcal{F} \cdot \nabla \varphi_h d\Omega - \int_{\partial\Omega_k} \varphi_h \mathcal{F} \cdot \hat{n} ds$$

The jump discontinuity at an element boundary requires the solution of a Riemann problem where the flux function $\mathcal{F} \cdot \hat{n}$ is approximated by a Lax-Friedrichs numerical flux. The resulting semi-discrete equation is given by

$$\frac{du_h}{dt} = L(u_h).$$

4 Time Discretization

Strong-stability preserving (SSP) time discretization methods were developed for semi-discrete method of lines approximation of hyperbolic PDE's in conservative form, Gottlieb et al (2001). Strong stability is a monotonicity property for the internal stages and the numerical solution. A general m -stage SSP Runge-Kutta method is given by

$$\begin{aligned} u^{(0)} &= u^n \\ u^{(i)} &= \sum_{k=0}^{i-1} \alpha_{ik} u^{(k)} + \Delta t \beta_{ik} L(u^{(k)}), \quad i = 1, \dots, m, \\ u^{n+1} &= u^{(m)}. \end{aligned} \quad (1)$$

To compute $u^{(i)}$ for each stage requires up to m evaluations of the right-hand side $L(u^{(k)})$. Thus, higher-order SSP Runge-Kutta methods can be expensive, in terms of the number of floating point operations, memory to store intermediate stages and parallel communication overhead per time step. Linear multistep methods (LMM) substitute time levels for stages.

Higueras (2004) discovered second order SSP Runge-Kutta methods with three stages and having Courant number $C = 2$. Therefore, the most efficient second order explicit integrator for the discontinuous Galerkin approximation would appear to be the three stage SSP Runge-Kutta scheme with an efficiency factor of $C/3 = 2/3$. Indeed, our numerical experiments confirmed this to be the case. In fact it was found in practice that this scheme integrates twice as fast as the two stage method. Moreover, the maximum time step for the three stage SSP RK2-3 matches that of the second order leap frog integrator employed in the continuous Galerkin spectral element model of Loft et al (2001).

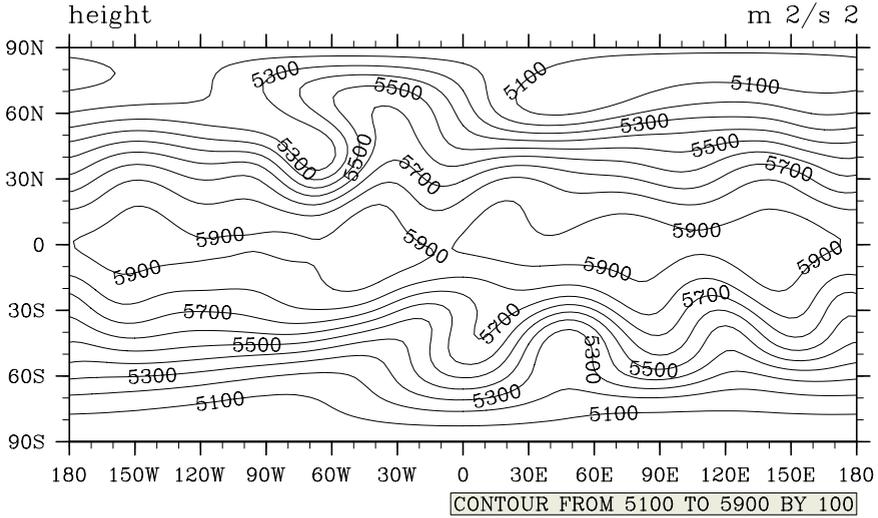


Fig. 1. Shallow water test case 5: Flow impinging on a mountain. 150 spectral elements, 8×8 Gauss-Legendre Lobatto points per element Geopotential height field h at fifteen days produced by discontinuous Galerkin method

5 Numerical Experiments

Our numerical experiments are based on the shallow water test suite of Williamson et al (1992). Test case 5 is a zonal flow impinging on an isolated mountain. The center of the mountain is located at $(3\pi/2, \pi/6)$ with height $h_s = 2000(1 - r/R)$ meters, where $R = \pi/9$ and $r^2 = \min[R^2, (\lambda - 3\pi/2)^2 + (\theta - \pi/6)^2]$. The initial wind and height fields take the same form as test case 2 with $\alpha_0 = 0$, $gh_0 = 5960 \text{ m}^2/\text{s}^2$ and $u_0 = 20 \text{ m/s}$. A total of 150 spectral elements containing 8×8 Gauss-Lobatto-Legendre points are employed. The explicit time step was $\Delta t = 90 \text{ sec}$. A spatial filter was not applied during this integration. Figure 1 contains a plot of the geopotential height field after 15 days of integration using the discontinuous Galerkin approximation. These results compare favorably with the continuous spectral element model.

6 Parallel Performance Results

Both the continuous and discontinuous Galerkin spectral element models are implemented within a unified software framework. A hybrid MPI/OpenMP programming model is supported where the entire time step is threaded according to an SPMD shared-memory approach. MPI message passing calls are serialized in hybrid mode. The cubed-sphere computational domain is partitioned across the compute nodes of a distributed-memory machine using either Metis

or space-filling curves (Dennis 2003). The latter algorithm is applied in the case where the number of elements along a cube face edge is divisible by $2^n 3^m$. Unlike continuous Galerkin spectral elements, almost all the computations within a discontinuous Galerkin finite element do not require any information from neighboring elements. Both the weak divergence and source terms can be computed independently on each element and then the local contribution of the boundary integrals can be added later. By employing non-blocking MPI communication, these computations can be performed while the exchange of conserved variables between elements proceeds. Single processor optimizations are based on data structures designed for extensive re-use and stride-1 memory access to minimize cache misses. Finite elements are represented as Fortran 90 derived types which are allocated statically on each processor at run-time. Spherical and cartesian coordinates along with the metric tensor are defined in the element type. Extensive loop unrolling is applied to expose instruction level parallelism to the processor. An effective technique for fast computation of the maximum eigenvalue α of the flux Jacobian is to invoke vector intrinsics and the use of square roots is minimized.

An experiment was designed to test if overlapping communication and computation in a pure MPI code has any measurable effect on an IBM p690 cluster with a Colony switch. The machine consists of 32-way SMP nodes containing 1.3 Ghz Power4 processors capable of four flops per clock cycle. A node can be configured as either a single 32-way or four 8-way logical partitions. Both the

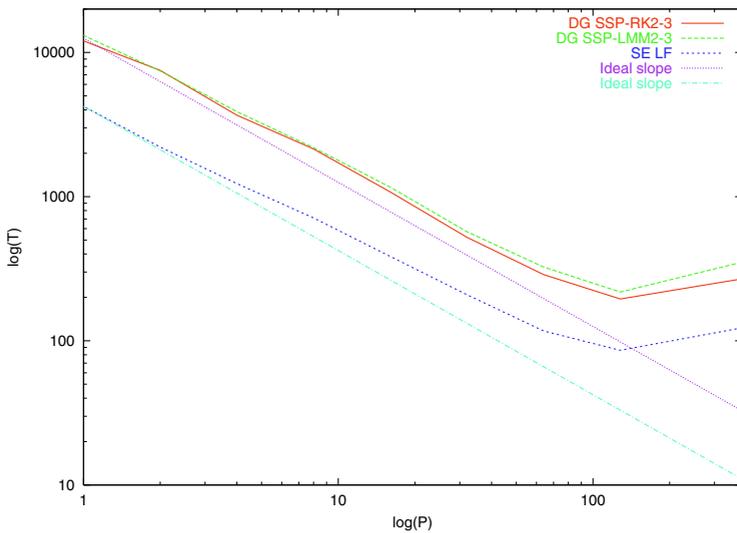


Fig. 2. Shallow water test case 5: Flow impinging on a mountain. 384 spectral elements, 10×10 Gauss-Legendre Lobatto points per element. 40 vertical levels. Integration time in seconds for IBM p690 8-way partitions

continuous and discontinuous Galerkin models were compared using test case 5, but integrated for only five days. For this test, the problem is replicated on 40 independent vertical levels. The discontinuous Galerkin code was integrated using the three step linear multistep method (LMM2-3) with $\Delta t = 15$ sec and three stage SSP Runge-Kutta (RK2-3) scheme with $\Delta t = 60$ sec. The continuous Galerkin model uses a second order leap frog integrator and $\Delta t = 60$ sec. $K = 384$ elements were employed with 10×10 Gauss-Lobatto-Legendre points. The LMM2-3 and RK2-3 single processor execution rates are 667 MFlops and 528 MFlops, respectively. The leap frog code sustains 724 MFlops. The wait time is defined as the average time for the MPI non-blocking communication to complete, summed over all time steps and all processors. We find that the wait time for 8-way is less than for 32-way partitions. Figure 2 is a plot of the integration times using 8-way partitions.

7 Conclusions

The three stage RK2-3 integrator discovered by Higuera (2004) was the most efficient SSP method examined. The discontinuous Galerkin method converges exponentially for smooth solutions and standard error metrics compare favorably with a continuous Galerkin spectral element model. Overlapping communication with computation was suggested by Baggag et al (1999). However, we only communicate u_h and compute fluxes $\mathcal{F}(u_h)$ locally. Both non-overlapping and overlapping implementations were compared. The latter was found to be clearly beneficial on SMP clusters such as the IBM p690 and leads to improved scalability in the strong sense for a fixed problem size. The per performance within a 32-way node is higher but 8-way nodes scale better. Our experience with a nodal Galerkin method indicates that a filter is required for long integrations to stabilize the scheme, thereby extending the results of Nair et al (2004).

Acknowledgements. NCAR is supported by the National Science Foundation. This work was partially supported by an NSF collaborations in mathematics and the geosciences grant (0222282) and the DOE climate change prediction program (CCPP).

References

1. Baggag, A., H. Atkins, D. Keyes, 1999: Parallel implementation of the discontinuous Galerkin method. ICASE Report 99-35, NASA/CR-1999-209546.
2. Cockburn, B., G. E. Karniadakis, and C. W. Shu, 2000: Discontinuous Galerkin Methods. Springer-Verlag, New York, 470 pp.
3. Dennis, J. M., 2003: Partitioning with space-filling curves on the cubed-sphere. Proceedings of Workshop on Massively Parallel Processing at IPDPS'03. Nice, France, April 2003.

4. Deville, M. O., P. F. Fischer, and E. H. Mund, 2002: High-Order Methods for Incompressible Fluid Flow. Cambridge University Press, 499 pp.
5. Fischer, P. F., and J. S. Mullen, 2001: Filter-Based stabilization of spectral element methods. *Comptes Rendus de l'Académie des sciences Paris*, t. 332, Série I - Analyse numérique, 265–270.
6. Giraldo, F. X., J. S. Hesthaven, and T. Warburton, 2003: Nodal high-order discontinuous Galerkin methods for spherical shallow water equations. *J. Comput. Phys.*, **181**, 499–525.
7. Gottlieb, S., C. W. Shu, and E. Tadmor, 2001: Strong stability preserving high-order time discretization methods. *SIAM Review*, **43**, 89–112.
8. Higuera, I., 2004: On strong stability preserving time discretization methods. *J. Sci. Comput.*, **21**, 193–223.
9. Loft, R. L., S. J. Thomas, and J. M. Dennis, 2001: Terascale spectral element dynamical core for atmospheric general circulation models. Proceedings of Supercomputing 01, IEEE/ACM.
10. Nair, R. D., S. J. Thomas, and R. D. Loft, 2004: A discontinuous Galerkin global shallow water model. *Mon. Wea. Rev.*, to appear.
11. Sadourny, R., 1972: Conservative finite-difference approximations of the primitive equations on quasi-uniform spherical grids. *Mon. Wea. Rev.*, **100**, 136–144.
12. Thomas, S. J., and R. D. Loft, 2002: Semi-implicit spectral element atmospheric model. *J. Sci. Comp.*, **17**, 339–350.
13. Williamson, D. L., J. B. Drake, J. J. Hack, R. Jakob, P. N. Swarztrauber, 1992: A standard test set for numerical approximations to the shallow water equations in spherical geometry *J. Comp. Phys.*, **102**, 211–224.