A Performance and Energy Study of the Hyperbolic PDE Solver Engine ExaHyPE

Master's Thesis in Computational Science and Engineering

Fabian Güra
Department of Informatics
Technische Universität München

September 2016

Supervisor: Univ.-Prof. Dr. Michael Bader
Dr. Tobias Weinzierl
Advisor: Dr. Vasco Varduhn

A Performance and Energy Study of the Hyperbolic PDE Solver Engine ExaHyPE

Master's Thesis in Computational Science and Engineering

Fabian Güra
Department of Informatics
Technische Universität München

September 2016

Supervisor: Univ.-Prof. Dr. Michael Bader
Dr. Tobias Weinzierl
Advisor: Dr. Vasco Varduhn

Abstract

... ...

Contents

Chapter 1

Introduction

- Challenges of exascale
- The ExaHyPE project (numerics, resilience, profiling) as an answer
- On the importance of profiling and performance measuring

Chapter 2

Theory

2.1 A D-dimensional ADER-DG scheme with MUSCL-Hancock a-posteriori subcell limiting for non-linear hyperbolic conservation laws

Arbitrary High Order Derivatives Discontinuous Galerkin (ADER-DG)

2.1.1 Notation

We use vector notation whenever possible. Advantage: Complete derivation, direct conversion to code.

2.1.2 PDE

Task: Solve the PDE

$$\frac{\partial}{\partial t}\left[\boldsymbol{u}\right]_v + \frac{\partial}{\partial x_d}\left[\boldsymbol{F}(\boldsymbol{u})\right]_{vd} = \left[\boldsymbol{s}(\boldsymbol{u})\right]_v \quad \text{on } \boldsymbol{\Omega} \times (0,T) \tag{2.1}$$

with initial conditions

$$\left[\boldsymbol{u}(\boldsymbol{x},0)\right]_v = \left[\boldsymbol{u}_0(\boldsymbol{x})\right]_v \ \forall \boldsymbol{x} \in \boldsymbol{\Omega}, \tag{2.2}$$

and boundary conditions

$$\left[\boldsymbol{u}(\boldsymbol{x},t)\right]_v = \left[\boldsymbol{u}_B(\boldsymbol{x},t)\right]_v \ \forall \boldsymbol{x} \in \partial\boldsymbol{\Omega}, t \in (0,T), \tag{2.3}$$

for all $v \in \{1,2,\ldots,V\}$, where V is the number of quantities involved in the system, $\boldsymbol{\Omega} \subset \mathbb{R}^D$ is the spatial domain, D the number of space dimensions, and $(0,T)$ a time interval. The function $\boldsymbol{F} : \mathbb{R}^V \to \mathbb{R}^{V\times D}, \boldsymbol{u} \mapsto \boldsymbol{F}(\boldsymbol{u}) = \left[\boldsymbol{f}_1(\boldsymbol{u}), \boldsymbol{f}_2(\boldsymbol{u}), \ldots, \boldsymbol{f}_D(\boldsymbol{u})\right]$ is called the flux function.

For the problem to be hyperbolic we require that all Jacobian matrices $\boldsymbol{A}_d(\boldsymbol{x}, t)$, $d \in \{1, 2, \ldots, D\}$, defined as

$$[\boldsymbol{A}_d]_{ij} = \frac{\partial \left[\boldsymbol{f}_d\right]_i}{\partial x_j}, \tag{2.4}$$

have D real eigenvalues in each admissible state $(\boldsymbol{x}, t) \in \boldsymbol{\Omega} \times (0, T)$.

2.1.3 Mesh

Let $\mathcal{T}_h$ be a quadrilateral partition of $\boldsymbol{\Omega}$, i.e.

$$K \cap J = \emptyset \, \forall K, J \in \mathcal{T}_h, K \neq J \tag{2.5}$$

$$\bigcup_{K \in \mathcal{T}_h} K = \boldsymbol{\Omega}. \tag{2.6}$$

Let $\{t_i\}_{i=0,1,\ldots I}$ be a partition of the time interval $(0, T)$ such that

$$0 = t_0 < t_1 < \ldots < t_I = T, \tag{2.7}$$

where I is the number of sub intervals. We furthermore define

$$\Delta t_i = t_{i+1} - t_i, \; i \; in \; \{0, 1, \ldots, I-1\}, \tag{2.8}$$

so that the interval (t_i, t_{i+1}) can be written as $(t_i, t_i + \Delta t_i)$.

2.1.4 Weak formulation

Let $L^2(\boldsymbol{\Omega})^V$ be the space of vector-valued, square-integrable functions on $\boldsymbol{\Omega}$, i.e.

$$L^2(\boldsymbol{\Omega})^V = \left\{ \boldsymbol{w} : \Omega \to \mathbb{R}^V \mid \int_{\boldsymbol{\Omega}} \|\boldsymbol{w}\| \, d\boldsymbol{x} < \infty \right\}. \tag{2.9}$$

Let $\boldsymbol{w} \in L^2(\boldsymbol{\Omega})^V$ be a spatial test function. Multiplication of the original PDE (2.1) and integration over a space-time cell $K \times (t_i, t_i + \Delta t_i)$ yields a weak, element local formulation of the problem

$$\begin{gathered} \int_{t_i}^{t_i+\Delta t_i} \int_K \frac{\partial}{\partial t} \left[\boldsymbol{u}\right]_v \left[\boldsymbol{w}\right]_v \, d\boldsymbol{x} \, dt + \int_{t_i}^{t_i+\Delta t_i} \int_K \frac{\partial}{\partial x_d} \left[\boldsymbol{F}(\boldsymbol{u})\right]_{vd} \left[\boldsymbol{w}\right]_v \, d\boldsymbol{x} \, dt = \\ \int_{t_i}^{t_i+\Delta t_i} \int_K \left[\boldsymbol{s}(\boldsymbol{u})\right]_v \left[\boldsymbol{w}\right]_v \, d\boldsymbol{x} \, dt, \end{gathered} \tag{2.10}$$

which we require to hold for $v \in \{1, 2, \ldots, V\}$, $\boldsymbol{w} \in L^2(\Omega)^V$, $K \in \mathcal{T}_h$ and $i \in \{0, 1, \ldots, I-1\}$.

Integration by parts of the spatial integral in the second term yields

$$\begin{aligned}&\int_K \frac{\partial}{\partial x_d} \left[\boldsymbol{F}(\boldsymbol{u})\right]_{vd} [\boldsymbol{w}]_v \, d\boldsymbol{x} = \\ &\int_K \frac{\partial}{\partial x_d} \left(\left[\boldsymbol{F}(\boldsymbol{u})\right]_{vd} [\boldsymbol{w}]_v\right) d\boldsymbol{x} - \int_K \left[\boldsymbol{F}(\boldsymbol{u})\right]_{vd} \frac{\partial}{\partial x_d} [\boldsymbol{w}]_v \, d\boldsymbol{x}.\end{aligned} \tag{2.11}$$

Application of the divergence theorem to the first term on the right-hand side of (2.11) yields

$$\int_K \frac{\partial}{\partial x_d} \left(\left[\boldsymbol{F}(\boldsymbol{u})\right]_{vd} [\boldsymbol{w}]_v\right) d\boldsymbol{x} = \int_{\partial K} \left[\boldsymbol{F}(\boldsymbol{u})\right]_{vd} [\boldsymbol{w}]_v [\boldsymbol{n}]_d \, ds(\boldsymbol{x}), \tag{2.12}$$

where $\boldsymbol{n} \in \mathbb{R}^D$ is the unit-length, outward-pointing normal vector at a point $\boldsymbol{x}$ on the surface of K, which we denote by ∂K.

Inserting eqs. (2.11) and (2.12) into eq. (2.10) yields the following weak, element-local formulation of the original equation (2.1):

$$\begin{aligned}&\int_{t_i}^{t_i+\Delta t_i} \int_K \frac{\partial}{\partial t} [\boldsymbol{u}]_v [\boldsymbol{w}]_v \, d\boldsymbol{x}\, dt - \int_{t_i}^{t_i+\Delta t_i} \int_K \left[\boldsymbol{F}(\boldsymbol{u})\right]_{vd} \frac{\partial}{\partial x_d} [\boldsymbol{w}]_v \, d\boldsymbol{x}\, dt + \\ &\int_{t_i}^{t_i+\Delta t_i} \int_{\partial K} \left[\boldsymbol{F}(\boldsymbol{u})\right]_{vd} [\boldsymbol{w}]_v [\boldsymbol{n}]_d \, ds(\boldsymbol{x})\, dt = \int_{t_i}^{t_i+\Delta t_i} \int_K \left[\boldsymbol{s}(\boldsymbol{u})\right]_v [\boldsymbol{w}]_v \, d\boldsymbol{x}\, dt.\end{aligned} \tag{2.13}$$

Again we require the weak formulation to hold for all $v \in \{1, 2, \dots V\}$, $\boldsymbol{w} \in L^2(\boldsymbol{\Omega})^V$, $K \in \mathcal{T}_h$ and $i \in \{0, 1, \dots, I-1\}$.

2.1.5 Restriction to finite-dimensional function spaces

To discretize eq. (2.13) we need to impose the restriction that both test and ansatz functions come from a finite-dimensional space. First, let $\mathbb{Q}_N(K)^V$ and $\mathbb{Q}_N(K \times (t_i, t_i + \Delta t_i))^V$ be the space of vector-valued, multivariate polynomials of degree less or equal N in each variable on K and $K \times (t_i, t_i + \Delta t_i)$, respectively. We then make the following choices:

- For spatial functions we restrict ourselves to

$$\mathbb{W}_h = \left\{\boldsymbol{w}_h \in L^2(\boldsymbol{\Omega})^V : \boldsymbol{w}_h|_K := \boldsymbol{w}_h^K \in \mathbb{Q}_N(K)^V \, \forall K \in \mathcal{T}_h\right\}. \tag{2.14}$$

- For space-time functions we restrict ourselves to

$$\begin{aligned}\tilde{\mathbb{W}}_h = \Big\{&\tilde{\boldsymbol{w}}_h \in L^2(\boldsymbol{\Omega} \times (t_i, t_i + \Delta t_i) : \\ \tilde{\boldsymbol{w}}_h|_{K \times (t_i, t_i+\Delta t_i)} &:= \tilde{\boldsymbol{w}}_h^{Ki} \in \mathbb{Q}_N\left(K \times (t_i, t_i + \Delta t_i)\right) \\ &\forall K \in \mathcal{T}_h, i \in \{0, 1, \dots, I-1\}\Big\}.\end{aligned} \tag{2.15}$$

Replacing $\boldsymbol{w}$ by $\boldsymbol{w}_h \in \mathbb{W}_h$ and $\boldsymbol{u}$ by $\tilde{\boldsymbol{u}}_h \in \tilde{\mathbb{W}}_h$ in eq. (2.13) yields a finite-dimensional approximation of the weak formulation

$$\begin{aligned}&\int_{t_i}^{t_i+\Delta t_i}\int_K \frac{\partial}{\partial t}\left[\tilde{\boldsymbol{u}}_h^{Ki}\right]_v \left[\boldsymbol{w}_h^K\right]_v \, d\boldsymbol{x}\, dt - \int_{t_i}^{t_i+\Delta t_i}\int_{\partial K}\left[\boldsymbol{F}(\tilde{\boldsymbol{u}}_h^{Ki})\right]_{vd}\frac{\partial}{\partial x_d}\left[\boldsymbol{w}_h^K\right]_v \, d\boldsymbol{x}\, dt + \\ &\int_{t_i}^{t_i+\Delta t_i}\int_{\partial K}\left[\boldsymbol{\mathcal{G}}(\tilde{\boldsymbol{u}}_h^{Ki},\tilde{\boldsymbol{u}}_h^{K^+i},\boldsymbol{n})\right]_v\left[\boldsymbol{w}_h^K\right]_v \, ds(\boldsymbol{x})\, dt = \\ &\int_{t_i}^{t_i+\Delta t_i}\int_K\left[\boldsymbol{s}(\tilde{\boldsymbol{u}}_h^{Ki})\right]_v\left[\boldsymbol{w}_h^K\right]_v \, d\boldsymbol{x}\, dt,\end{aligned} \tag{2.16}$$

which now has to hold for all $\boldsymbol{w}_h \in \mathbb{W}_h$, $K \in \mathcal{T}_h$ and $i \in \{0, 1, \ldots, I-1\}$. Since for a cell $K \in \mathcal{T}_h$ and one of its Voronoi neighbors $K' \in \mathcal{V}(K)$ one has

$$\tilde{\boldsymbol{u}}_h^{Ki}(\boldsymbol{x}) \neq \tilde{\boldsymbol{u}}_h^{K'i}(\boldsymbol{x}),\ \boldsymbol{x} \in K \cap K', \tag{2.17}$$

i.e. $\tilde{\boldsymbol{u}}_h$ is double-valued at the interface between K and K', in order to compute the surface integral we need to introduce the numerical flux function $\boldsymbol{\mathcal{G}}(\tilde{\boldsymbol{u}}_h^{Ki}, \tilde{\boldsymbol{u}}_h^{K'i}, \boldsymbol{n})$. The numerical flux at a position $\boldsymbol{x} \in K \cap K'$ on the interface is obtained by (approximately) solving a Riemann problem in normal direction.

Integration by parts in time of the first term of eq. (2.16) and noting that $\boldsymbol{w}_h$ is constant in time yields the following one-step update scheme for the cell-local time-discrete solution $\tilde{\boldsymbol{u}}_h^{Ki}$:

$$\begin{aligned}\int_K\left[\tilde{\boldsymbol{u}}_h^{Ki}\Big|_{t_i+\Delta t_i}\right]_v\left[\boldsymbol{w}_h^K\right]_v d\boldsymbol{x} = \int_K &\left[\tilde{\boldsymbol{u}}_h^{Ki}\Big|_{t_i}\right]_v\left[\boldsymbol{w}_h^K\right]_v d\boldsymbol{x} + \\ &\int_{t_i}^{t_i+\Delta t_i}\int_K\left[\boldsymbol{F}(\tilde{\boldsymbol{u}}_h^{Ki})\right]_{vd}\frac{\partial}{\partial x_d}\left[\boldsymbol{w}_h^K\right]_v d\boldsymbol{x}\, dt - \\ &\int_{t_i}^{t_i+\Delta t_i}\int_{\partial K}\left[\boldsymbol{\mathcal{G}}(\tilde{\boldsymbol{u}}_h^{Ki},\tilde{\boldsymbol{u}}_h^{K^+i},\boldsymbol{n})\right]_v\left[\boldsymbol{w}_h^K\right]_v ds(\boldsymbol{x})\, dt + \\ &\int_{t_i}^{t_i+\Delta t_i}\int_K\left[\boldsymbol{s}(\tilde{\boldsymbol{u}}_h^{Ki})\right]_v\left[\boldsymbol{w}_h^K\right]_v d\boldsymbol{x}\, dt.\end{aligned} \tag{2.18}$$

Again we require eq. (2.18) to hold for all $v \in \{1, 2, \ldots, V\}$, $\boldsymbol{w}_h \in \mathbb{W}_h$, $K \in \mathcal{T}_h$ and $i \in \{0, 1, \ldots, I-1\}$.

Problem: We only have $\tilde{\boldsymbol{u}}_h|_t$ for $t \in \{t_i, t_i + \Delta t_i\}$, not for $t \in (t_i, t_i + \Delta t_i)$.

Idea: Replace $\tilde{\boldsymbol{u}}_h$ in $K \times (t_i, t_i + \Delta t_i)$ by an approximation $\tilde{\boldsymbol{q}}_h \in \tilde{\mathbb{W}}_h$ which we call space-time predictor.

2.1.6 Space-time predictor

To derive a procedure to compute the space-time predictor $\tilde{\boldsymbol{q}}_h \in \tilde{\mathbb{W}}_h$ we again start from the original PDE (2.1)

2.2 Profiling and Energy-aware Computing

Chapter 3

A profiling infrastructure for ExaHyPE

- General architecture
- Architecture profiling
- Functionality

Chapter 4

Preliminary profiling results, case studies

- Analytic benchmark: Introduction, derivation
- Pie-chart per kernel
- Case-study: Cache-misses, compile-time ($\rightarrow$ Toolkit philosophy)
- Degree $\rightarrow$ Wallclock, Energy (AMR)
- Static mesh Δx $\rightarrow$ Error for polynomials (convergence tables)

Chapter 5

Conclusion and Outlook

- PA is important
- ExaHyPE as an answer to exascale challenges
- Applications

Chapter 6

Acknowledgment