



# **SAGA Optimierung**

Rasmus Pranke

04.03.2021

Projekt Parallelrechnen



# SAGA

- **G**eo-**I**nformations-**S**ystem
- **S**ystem for **A**utomated **G**eoscientific **A**nalyses
- Vor allem ein Framework zur Implementation von GIS-Tools
- Bietet Programmier-Schnittstellen
  - Python
  - DLL



# Aufgabenstellung

- Dateien sind groß
- Tools sind oft naiv implementiert
- Daher vermutlich viel Raum für Performanzoptimierung



# Probleme

- Korrektheit sicherstellen
- Performanceprobleme finden
- Performanceprobleme beheben

# Korrektheit sicherstellen

- Regressionstests
- Geschrieben in Python
- → Nicht invasiv, einfach zu automatisieren
- Nachteil: Nicht sehr performant mit echten Daten
- → Dedizierte Testdaten sinnvoll



# Performanceprobleme finden

- Visual Studio Debugger
- ...das reicht bereits



# Performanceproblem

- Memory I/O beansprucht 60% und mehr
- Hier also:  
I/O optimieren = Performance optimieren

# Ansatz: Sliding Matrix

- Oft fließen alle angrenzenden Werte in Berechnung ein
- Für benachbarte Zellen überschneiden sich diese
- → Viele Abfragen einsparbar
- → Hat erfolgreich 20% Performancegewinn für das Tool „Morphometrie → Slope, Aspect, Curvature“ erzielt

# Sliding Matrix

1	2	3	4	5
2	3	4	5	6
3	4	5	6	7
4	5	6	7	8
5	6	7	8	9

# Verallgemeinerbarkeit

- Hilfsfunktion `Get_xTo` wird an mehr als 300 Stellen im Code verwendet
- → Jede dieser Stellen könnte von der Sliding Matrix profitieren
- → Umsetzung der Sliding Matrix als Iterator in Grid-Klasse

# Weitere Ideen

- Komprimierung der Daten im Speicher, um weniger Speicheraufrufe (dafür mehr Rechenarbeit) zu haben
- Anforderungen hierfür:
  - Aus Koordinate Speicherort ableitbar (kein Ablaufen der komprimierten Datei)
  - Komprimierung schnell und verlustfrei



# Weitere Ideen

- Vektor-Loading ist ein Tool von Compilern, um viele Daten schnell griffbereit zu halten
- Überprüfen, ob aktuelles Pattern der Abfrage pro Zelle diesem Tool im Weg steht

# Zusammenfassung

- Häufiges Laden aus dem Speicher scheint der mit Abstand größte Flaschenhals zu sein
- Zielsetzung ist, allgemeine Ansätze zu finden die entweder:
  - Leicht verständlich oder
  - Im Hintergrundsind.