Arbeitsbereich Wissenschaftliches Rechnen
Fachbereich Informatik
Fakultät für Mathematik, Informatik und Naturwissenschaften
Universität Hamburg

Seminar Supercomputer: Forschung und Innovation

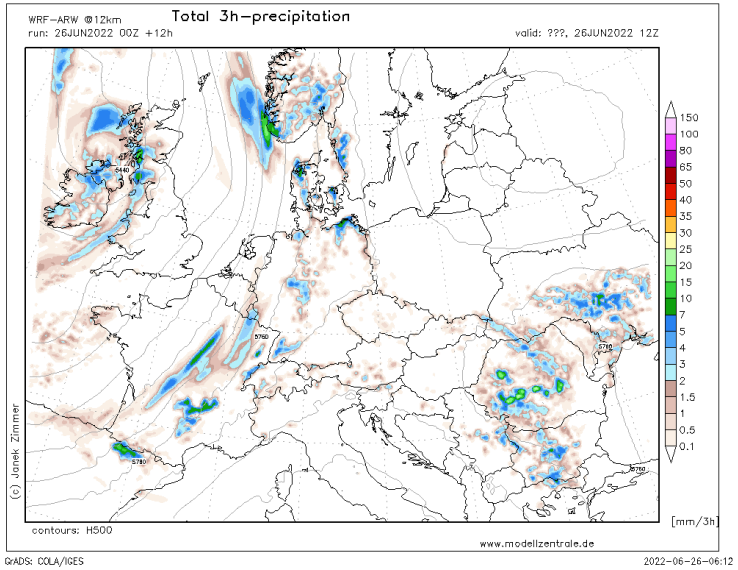
High Performance Parallel I/O and In-Situ Analysis in the WRF Model with ADIOS2

Inhalt

- Einleitung
- WRF: Weather Research and Forecasting Model
- ADIOS2: The Adaptable Input/Output System
  - Schreibleistung
  - Node-local Burst-Buffer
  - Aggregatoren
  - In-line Kompression
  - In-situ Methoden
- Optimierung der Gesamtabläufe
- Zusammenfassung

Einleitung

- Wetter- / Klimavorhersagen benötigen riesige Rechenkapazitäten und Datenspeicher
  - Faktor: Auflösung
  - Komplexe Zusammenspiele vieler Variablen
- Qualität der Vorhersage hängt stark von der Leistungsfähigkeit der Computer ab
- Beispiel Datenaufkommen (Mistral, 2015)
  - Deutschland mit 25 Milliarden Gitterzellen (Wolkenauflösend)
  - Ergebnisdaten: 1 bis 20 Petabyte



Einleitung

„Ein Supercomputer ist ein Gerät, mit dem Berechnungsprobleme in I/O-Probleme umgewandelt werden können.“

Prof. Ken Batcher

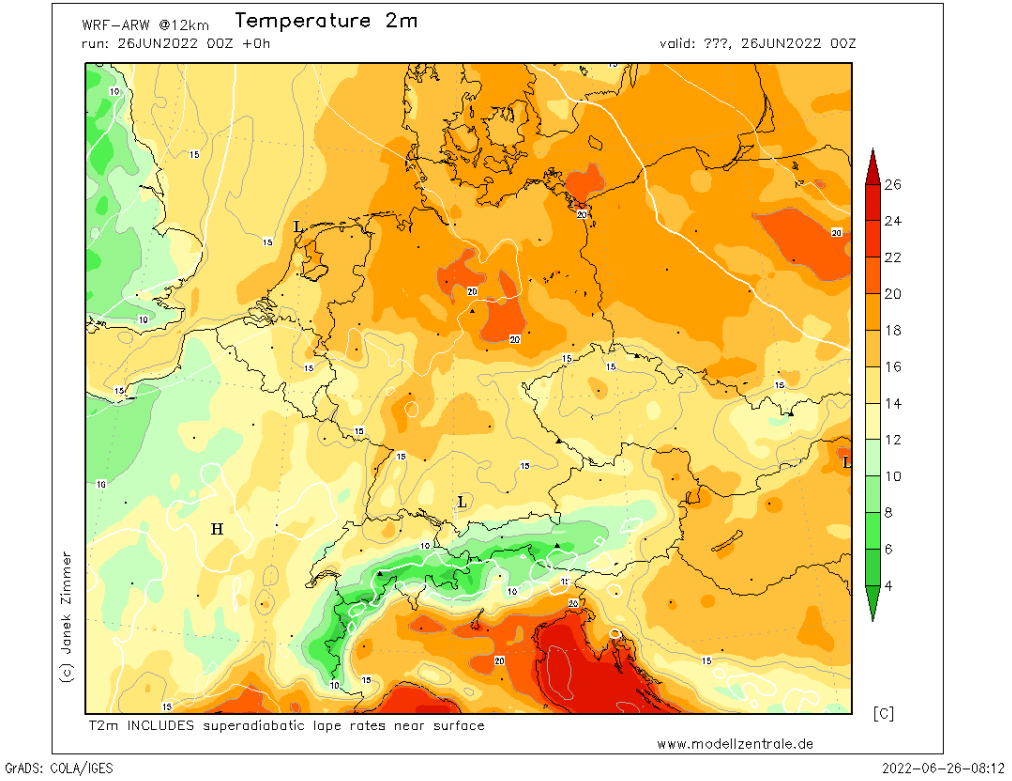
Einleitung

- Berechnung vs. Speicherung
- Ziel der Arbeit
  - Datenoperationen im WRF beschleunigen
  - „time-to-solution“ optimieren
- WRF: Weather Research and Forecast Model
- ADIOS2: Adaptable Input/Output System

WRF

Weather Research and Forecasting Model

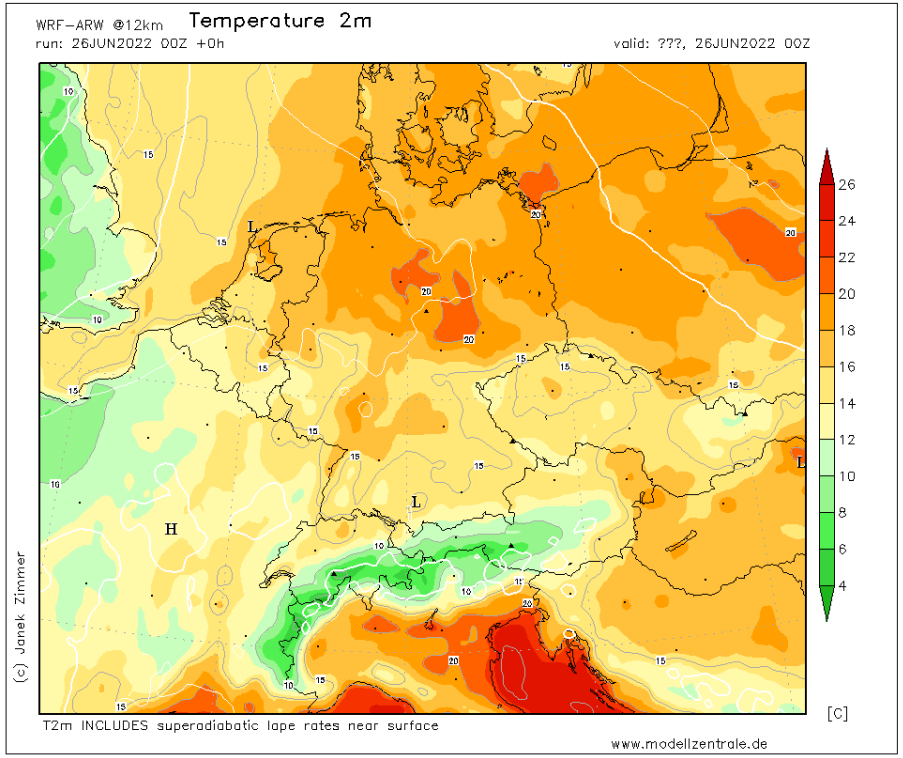
- Mesoskaliges numerisches Wettervorhersagemodell
- Mittlerweile auch zur Klimavorhersage genutzt
- Open Source
- Weltweite Verbreitung
  - Verwendung auch am DKRZ



WRF

Weather Research and Forecasting Model

- Unterstützt massiv-parallele Berechnung
- Unterstützt verschiedene I/O-Systeme
- Hohe Anforderungen für Berechnung und Speicherung
  - Viele Variablen im 3-dimensionalen Raum

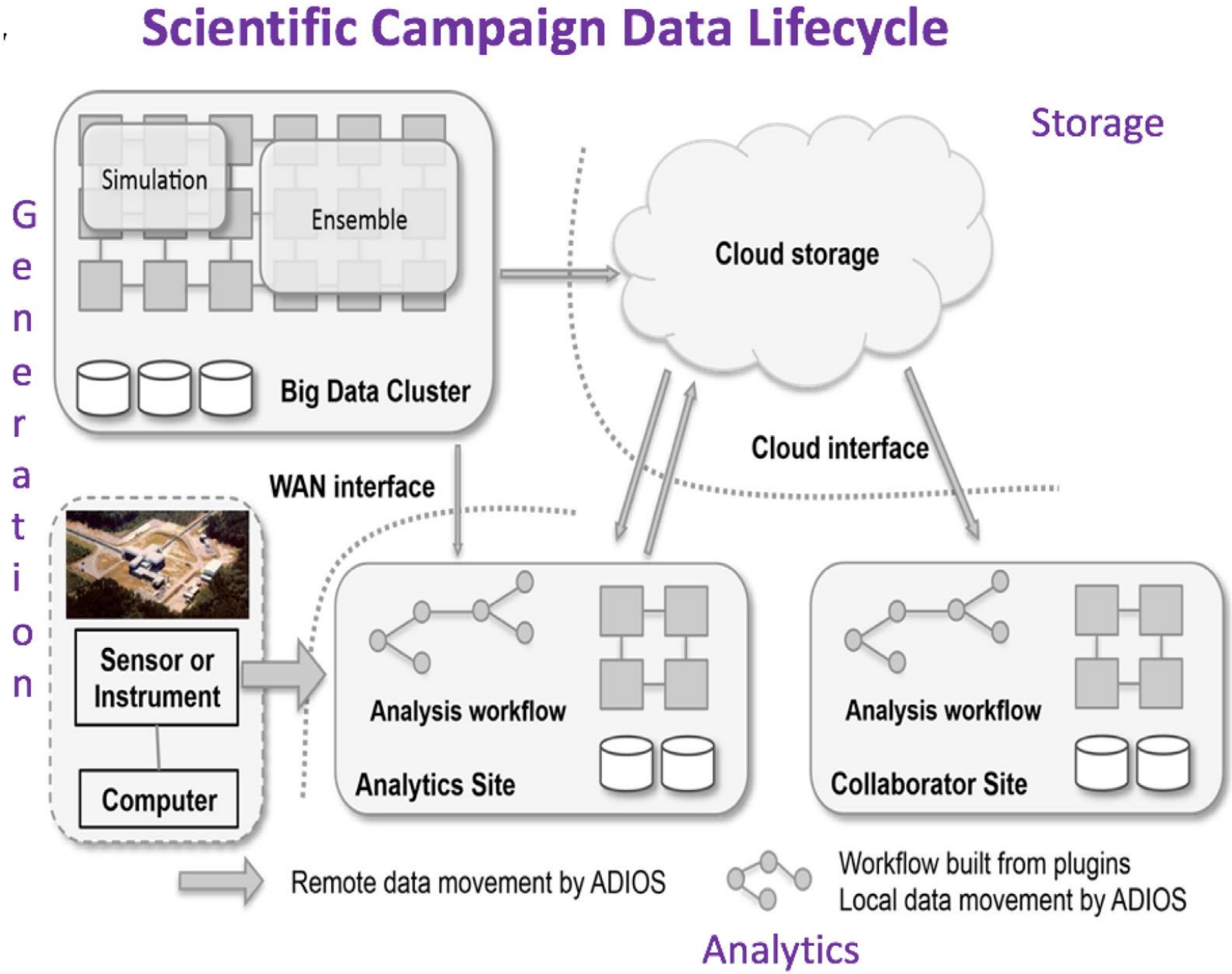


ADIOS
The Adaptable Input/Output System

- Datenverwaltung von Supercomputer I/O bis zu PC- oder Cloud-Systemen
- Einheitliche Schnittstelle für Datenübertragung über Dateien, WAN-Netzwerke oder direkten Speicherzugriff
- Unterstützte Sprachen: C++, C, Python, Matlab, …
- Ziel
  - Rahmenwerk für Datenverwaltung im Hinblick auf die Exascale-Ära

ADIOS

Einsatzgebiete



Bildquelle: ADIOS 2: The Adaptable Input Output System. A framework for high-performance data management

ADIOS

The Adaptable Input/Output System

- Eigenes Dateiformat
- „Self-describing“ Daten
- ADIOS2 ist „step-based“
  - ADIOS: Daten pro Zeitschritt
- Schreibt mehrere Sub-Dateien
  - Metadaten-Algorithmus zum Verwalten der Dateien
  - Sub-Dateien werden später rekombiniert

ADIOS
Vergleich der Schreibleistung

- Testaufbau
  - Bis zu 8 Nodes mit jeweils zwei 18 Core Prozessoren und 1 TB SSDs
  - Bis zu 288 Ranks
  - Ein paralleles Zieldateisystem (BeeGFS, striped auf 8 Platten)
- WRF Benchmark
  - „CONUS“ 2.5km Wetterberechnung (Continental US, 2.5km Auflösung)

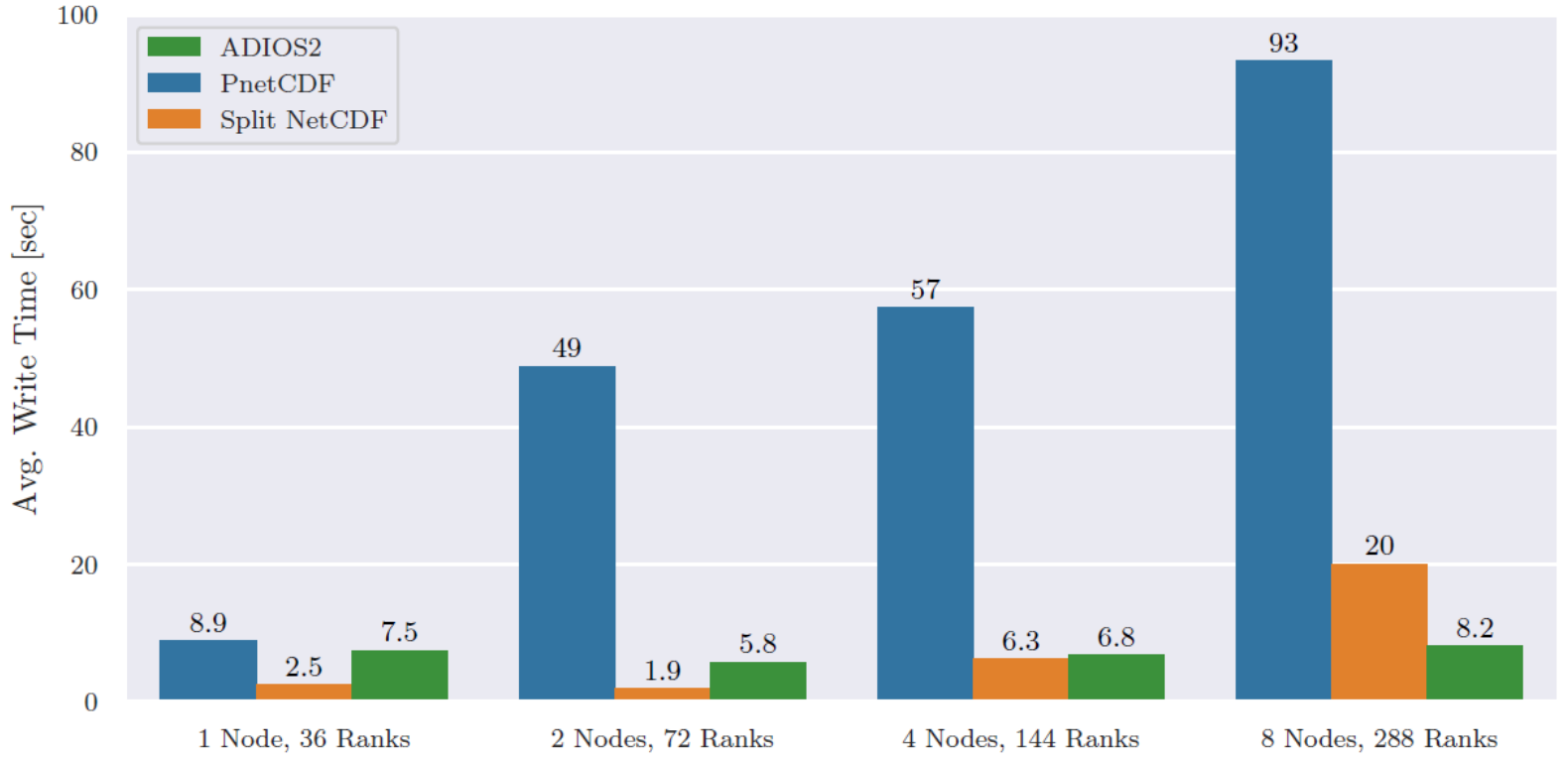
ADIOS
Vergleich der Schreibleistung

Vergleich mehrerer I/O Lösungen

- PnetCDF
  - Defaultkonfiguration für paralleles Schreiben im WRF
- Split NetCDF
  - Variante des seriellen NetCDF
  - Schnell bei wenig Prozessen
- ADIOS2
  - Mischung beider Varianten

ADIOS

Vergleich der Schreibleistung

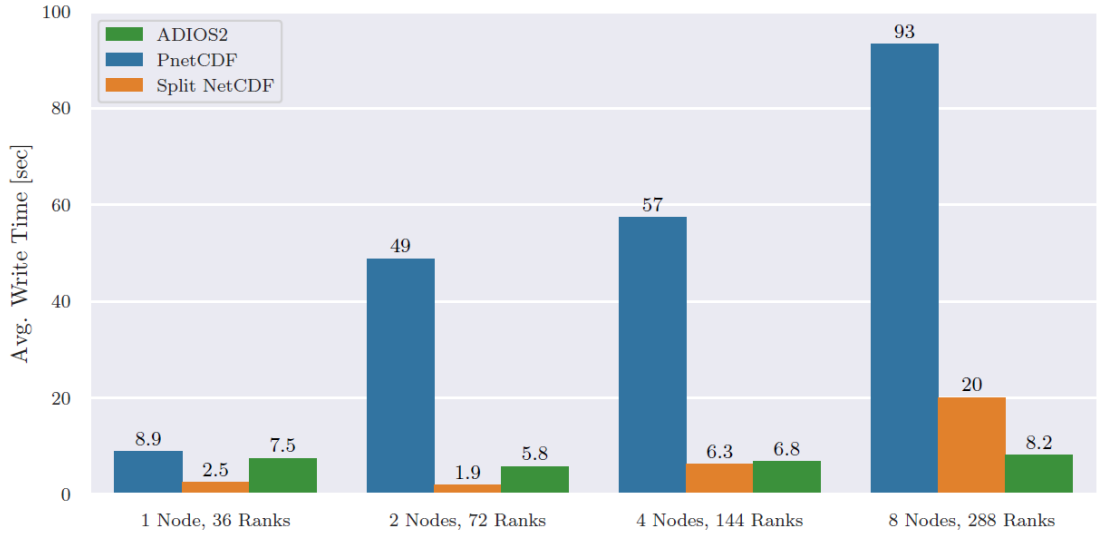


Bildquelle: High Performance Parallel I/O and In-Situ Analysis in the WRF Model with ADIOS2

ADIOS

Vergleich der Schreibleistung

- Split NetCDF
  - Gute Leistung bei wenig Nodes und Ranks
- PnetCDF
  - Deutlicher Anstieg der Zeit
- ADIOS2
  - Konstante Leistung



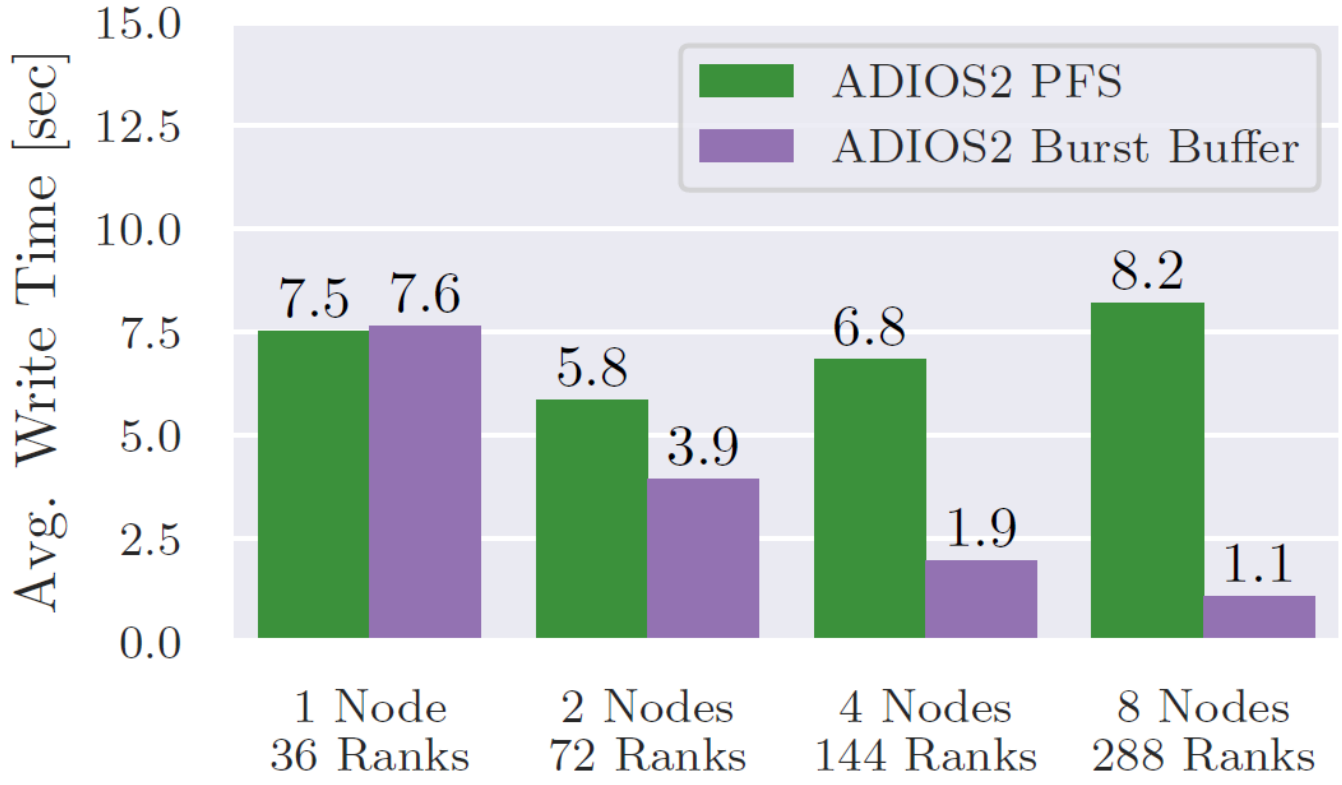
Bildquelle: High Performance Parallel I/O and In-Situ Analysis in the WRF Model with ADIOS2

ADIOS
Node-local Burst-Buffer

- Node-Local Burst-Buffer
  - Jeder Node kann seine Daten auf einen eigenen Speicher (z.B. NVMe SSD) schreiben
  - Prozesse können schnell fortgesetzt werden
  - Burst-Buffer können parallel von einem anderen Prozess im Hintergrund bearbeitet werden

ADIOS

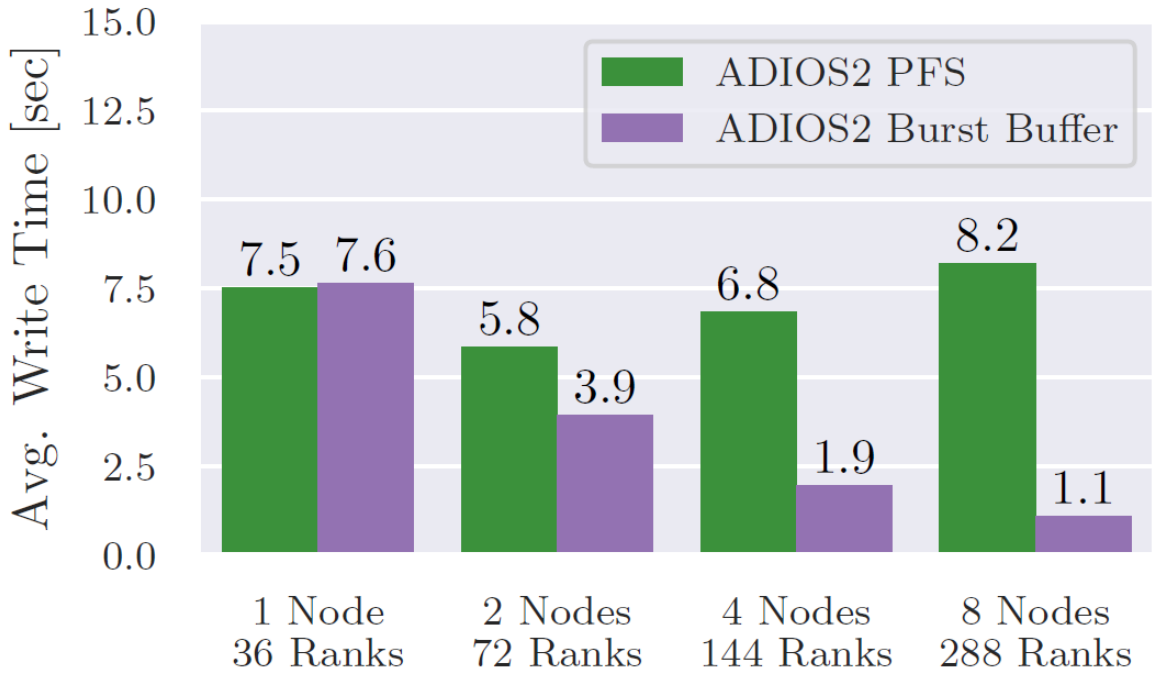
Node-local Burst-Buffer

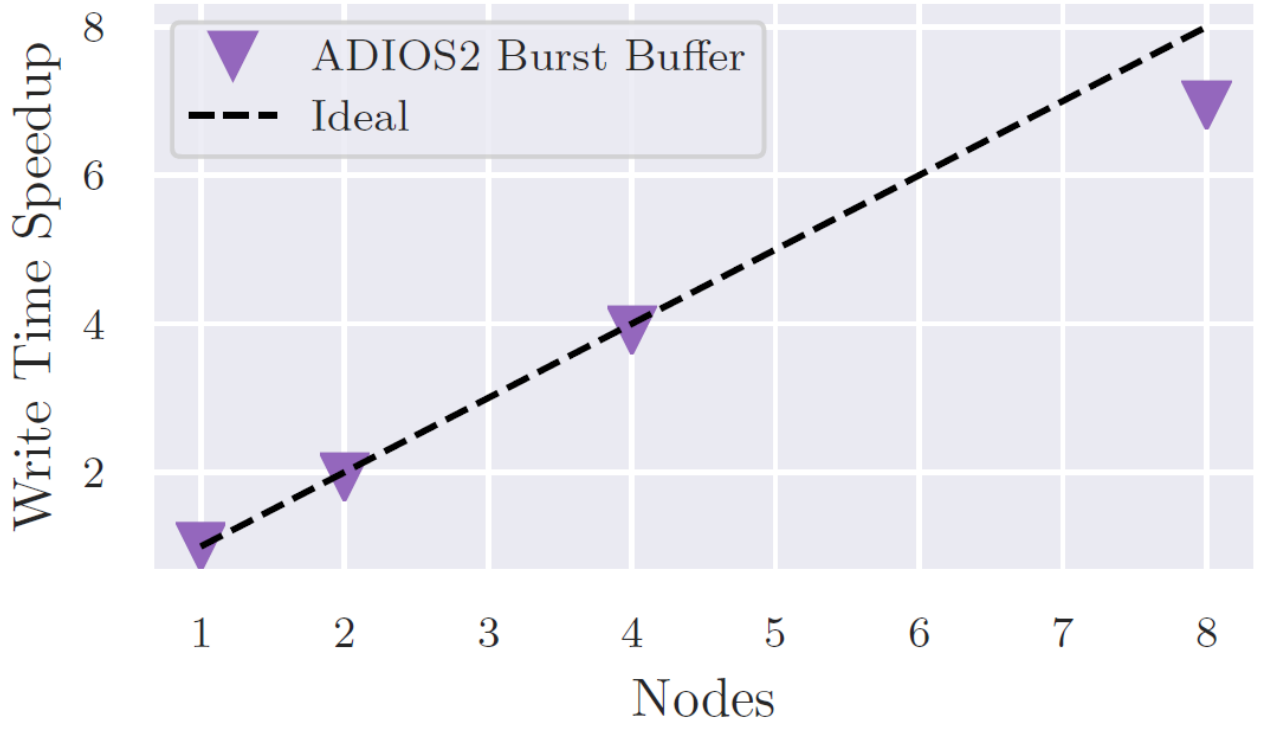


Bildquelle: High Performance Parallel I/O and In-Situ Analysis in the WRF Model with ADIOS2

ADIOS

Node-local Burst-Buffer





Bildquelle: High Performance Parallel I/O and In-Situ Analysis in the WRF Model with ADIOS2

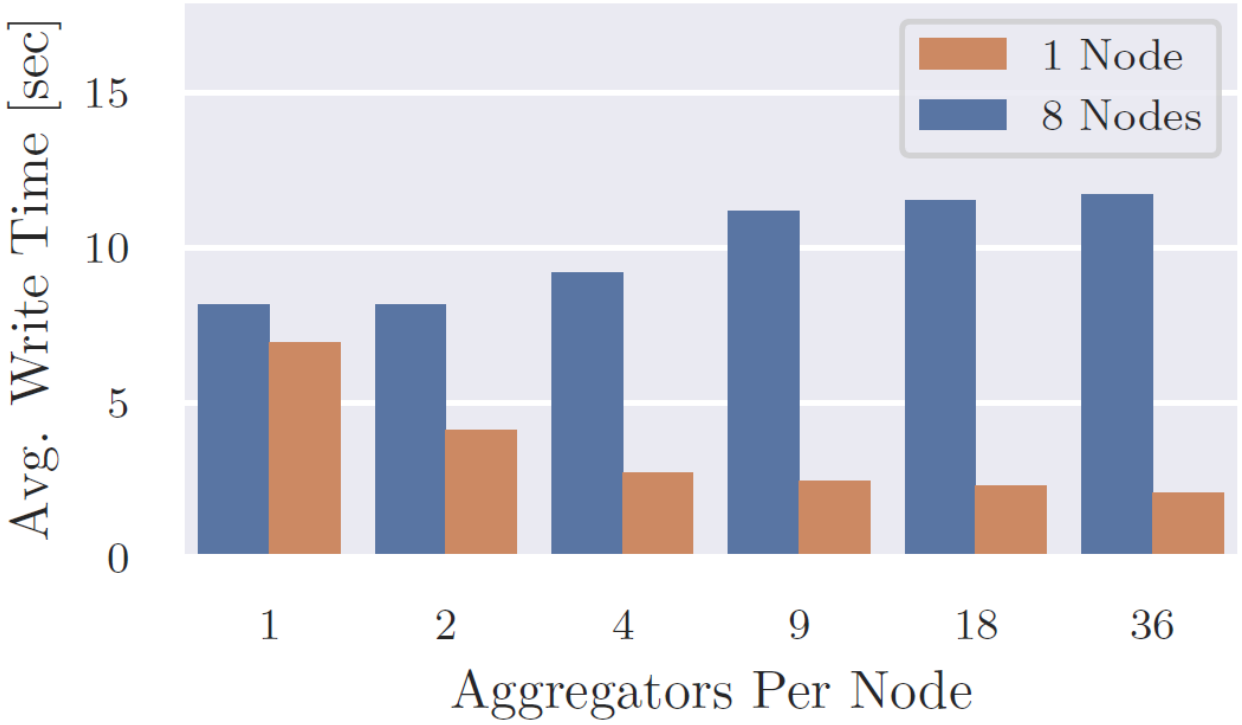
ADIOS
Aggregatoren

- Anzahl Aggregatoren = Anzahl geschriebener Dateien
- Erlaubt N-M Kopplung (N Ranks schreiben M Dateien)
- Anpassung an Systemanforderungen möglich

ADIOS

Aggregatoren

- 1 Node
  - Leistung steigt mit Aggregatoren
- 8 Nodes
  - Optimum: 1 Aggregator
- Fazit für 8 Nodes
  - 1 Aggregator / Node



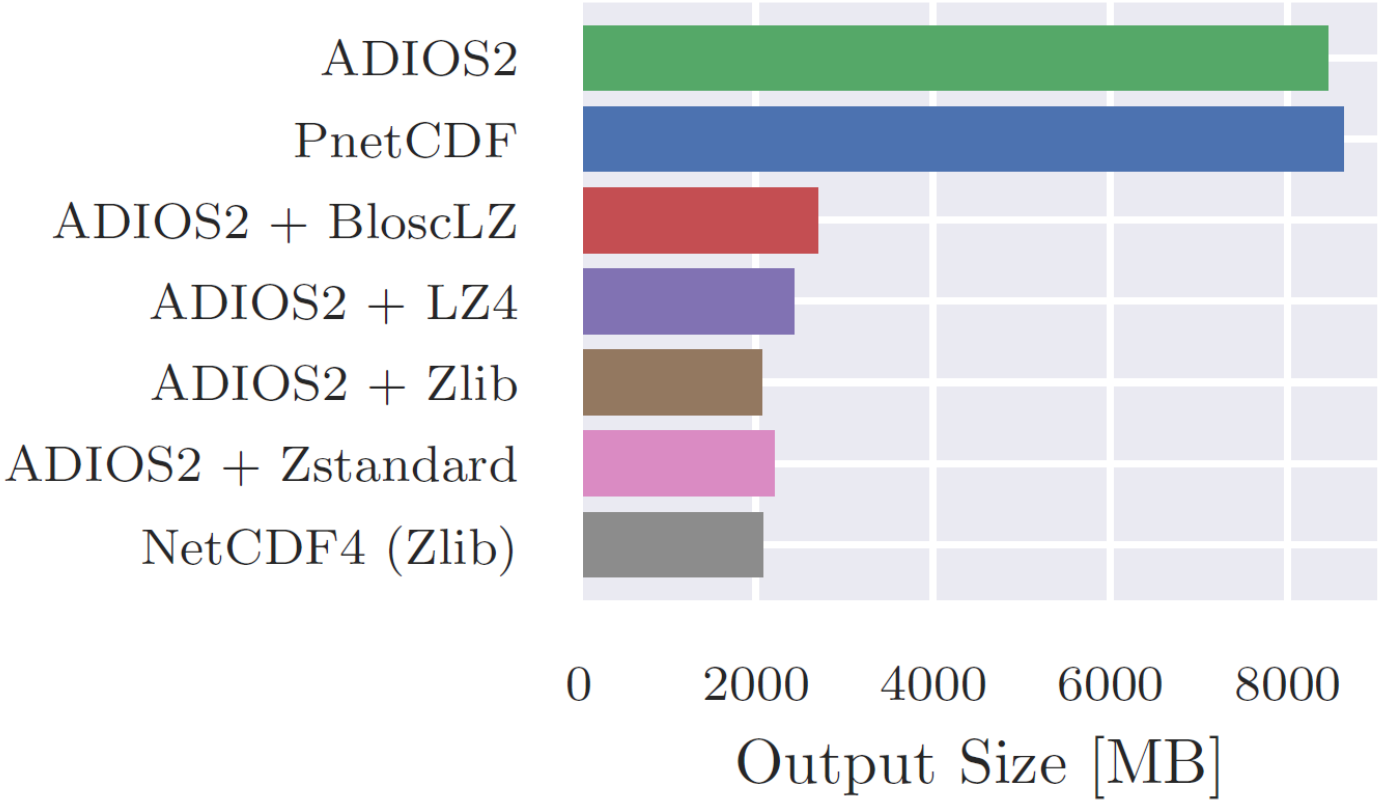
Bildquelle: High Performance Parallel I/O and In-Situ Analysis in the WRF Model with ADIOS2

ADIOS
In-line Kompression

- Kompression bevor die Daten geschrieben werden
- Kann mit unterschiedlichen Codecs kombiniert werden
  - Z.B.: BloscLZ, LZ4, Zstd
- Ziel: Verbesserung der Schreibzeit durch Komprimierung der Dateigröße vor dem Schreiben

ADIOS

In-line Kompression



Bildquelle: High Performance Parallel I/O and In-Situ Analysis in the WRF Model with ADIOS2

ADIOS

In-line Kompression

Bildquelle: High Performance Parallel I/O and In-Situ Analysis in the WRF Model with ADIOS2

ADIOS
Optimale Konfiguration

- Optimale Konfiguration:
  - 8 Nodes
  - Node-Local Burst-Buffer
  - 1 Aggregator / Node
  - In-line Kompression mit Zstd Codec

ADIOS

Optimale Konfiguration

Configuration	Write Time [s]	Speedup
PnetCDF	93	1x
ADIOS2	8.2	11x
ADIOS2 + BB	1.1	84x
ADIOS2 + BB + Zstd	0.52	179x

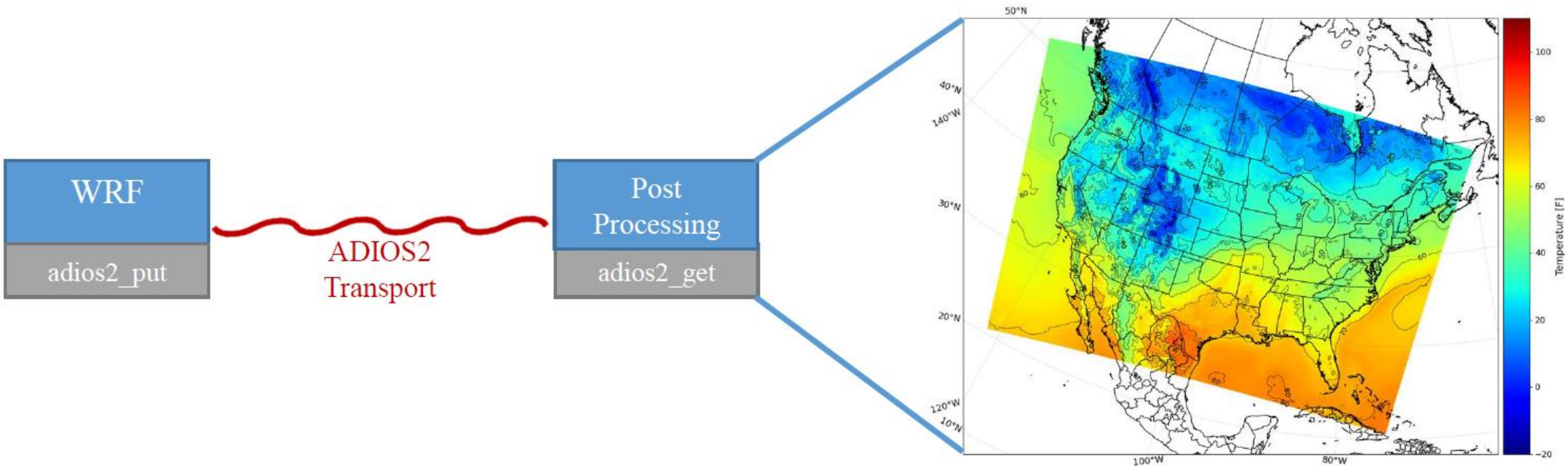
ADIOS
In-situ Methoden

- Verarbeitung der Daten während die Berechnung weiter läuft
- Mögliche In-situ Anwendungen
  - Datentransformationen
  - Analyse der Abläufe / Daten
- Hohe Anpassbarkeit an bestimmte Anforderungen
- Optimierung von Experimenten
  - „Kunst der Modellierung“

ADIOS und WRF
Optimierung der Gesamtabläufe

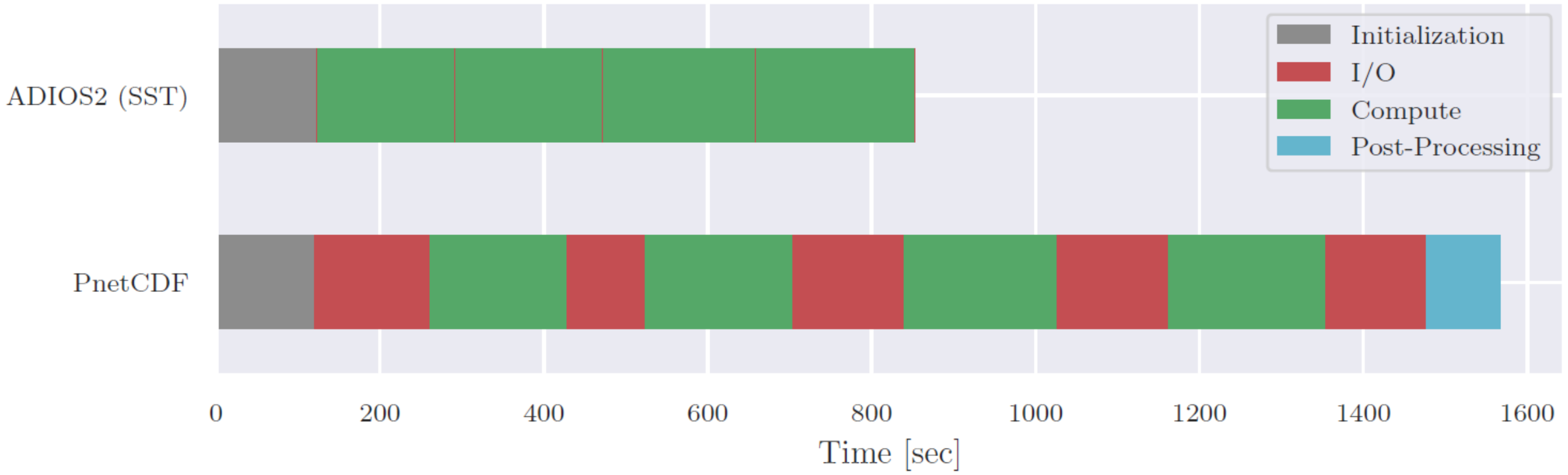
- ADIOS: Beschleunigung durch in-situ Methoden
- „Time-to-solution“
  - Benötigte Zeit für einen (vollständigen) Experimentdurchlauf
- Direkte Weiterleitung der Daten ohne Dateien kann die „time-to-solution“ deutlich verkürzen

ADIOS und WRF
Optimierung der Gesamtabläufe



Bildquelle: High Performance Parallel I/O and In-Situ Analysis in the WRF Model with ADIOS2

ADIOS und WRF
Optimierung der Gesamtabläufe



Bildquelle: High Performance Parallel I/O and In-Situ Analysis in the WRF Model with ADIOS2

ADIOS und WRF
Zusammenfassung

- Schnelle parallele I/O-Lösung
- Hoch performante Kompression möglich
- In-situ Analysen
- „Time-to-solution“ kann mithilfe von ADIOS2 deutlich optimiert werden

ADIOS und WRF

Vielen Dank für Eure Aufmerksamkeit!

Quellen

(1) Godoy, W. F., Podhorszki, N., Wang, R., Atkins, C., Eisenhauer, G., Gu, J., Davis, P., Choi, J., Germaschewski, K., Huck, K., Huebl, A., Kim, M., Kress, J., Kurc, T., Liu, Q., Logan, J., Mehta, K., Ostrouchov, G., Parashar, M., … Klasky, S. (2020). ADIOS 2: The Adaptable Input Output System. A framework for high-performance data management. SoftwareX, 12, 100561. https://doi.org/10.1016/j.softx.2020.100561

(2) Laufer, M., & Fredj, E. (2022). High Performance Parallel I/O and In-Situ Analysis in the WRF Model with ADIOS2. http://arxiv.org/abs/2201.08228

Quellen

(3) Das Deutsche Klimarechenzentrum: Partner der Klimaforschung (2015). Rechner, Daten, Wissen.

(4) Janek Zimmer (2022). Modellzentrale.de. (Letzter Zugriff: 30.06.2022)