

# Profiling Toolkit für HPC Anwendungen in Tier-2 und Tier-3 Rechenzentren

Vanessa End

Gesellschaft für wissenschaftliche Datenverarbeitung mbH Göttingen

7. HPC-Status-Konferenz der Gauß Allianz  
04.12.2017

Gefördert durch die Deutsche Forschungsgemeinschaft (DFG)  
Funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation)  
KO 3394/14-1, OL 241/3-1, RE 1389/9-1, VO 1262/1-1, YA 191/10-1

- Projektlaufzeit 01.02.2017 - 31.01.2020
- Partner



## 1 Motivation

## 2 ProfiT-HPC

- AP 1: Umfrage an Rechenzentren
- AP 2: Tools und Metriken
- Weg zum Toolkit

## 3 Ausblick

## Hintergrund

- Immer mehr Forschungsgebiete entdecken HPC für ihre Forschung:
    - Sozialwissenschaften
    - Geisteswissenschaften
    - ...
  - Unterschiedliche HPC-Erfahrungen bei den Nutzern:
    - keine Programmiererfahrung
    - sequentielle Programmiererfahrung
    - parallele Programmiererfahrung
    - Experten
- Funktionsfähiges Programm im Fokus des Nutzers - Performance und Ressourcennutzung nur nebensächlich.



## Allgemeine Projektziele

- Bewusstsein für Wichtigkeit von Performance Analysen schaffen und erweitern.
- Zugang zu Performance Analysen und Ergebnissen vereinfachen.
- Bereitstellung von Analysen in allgemeinverständlichem Format und Hinweise zu weiterer Vorgehensweise.
- Portabilität des Toolkits zur einfachen Verbreitung in möglichst vielen Rechenzentren.

# HPC Infrastruktur in Deutschland



- Umfrage an Rechenzentren
  - Welche Hardware und Software ist vorhanden?
  - Welches Batchsystem wird genutzt?
  - Welche Performanceanalyse Werkzeuge sind vorhanden?



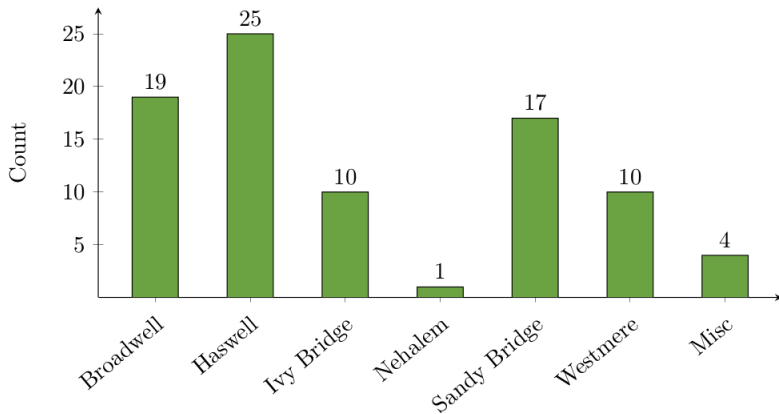
# HPC Infrastruktur in Deutschland

- Umfrage an Rechenzentren
  - Welche Hardware und Software ist vorhanden?
  - Welches Batchsystem wird genutzt?
  - Welche Performanceanalyse Werkzeuge sind vorhanden?

## Ergebnisse

- Antwort von 33 Rechenzentren.
  - Sehr heterogene Architektur- und Softwarelandschaft.
  - D1.2: Results of a Survey concerning the Tier-2 and Tier-3 HPC-Infrastructure in Germany
- Modularisierter Ansatz für das Toolkit

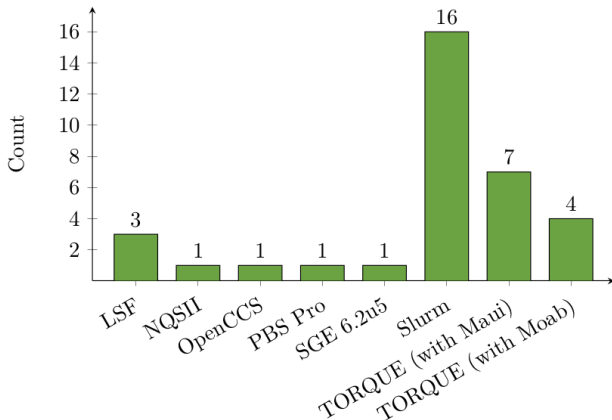
## Ergebnisse zu den Prozessoren



Und weitere 9 verschiedene AMD Architekturen...

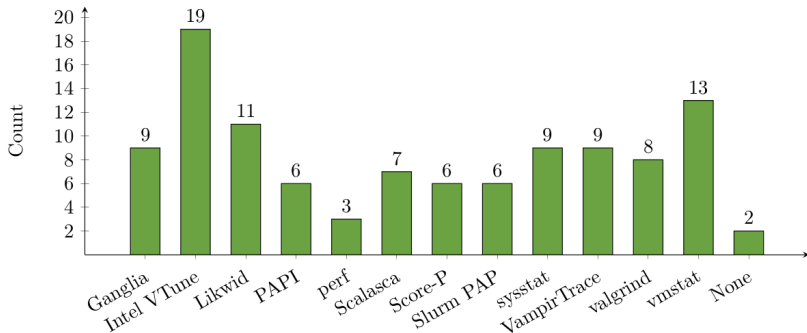


## Ergebnisse zu den Batchsystemen



... mit shared Knoten an 27/33 Zentren, non-shared nur in 14/33 ...

## Ergebnisse zu den Profiling Tools





## Tools und Metriken

- Evaluierung verschiedener Tools und Metriken
  - Ist das Tool/Programm automatisierbar?
  - Welche Metriken werden geliefert?
  - Wie sehr ist der Nutzer eingebunden?



## Tools und Metriken

- Evaluierung verschiedener Tools und Metriken
  - Ist das Tool/Programm automatisierbar?
  - Welche Metriken werden geliefert?
  - Wie sehr ist der Nutzer eingebunden?

### Ergebnisse

- Wenige Tools komplett automatisierbar nutzbar
  - D2.1: Concise Overview of Metrics and Tools
  - D2.2.1: Functional Specification of the Backend – Part 1:  
Pre-Selection of the Metric Collection Tools
- Vorerst Fokus auf Monitoringdaten



# Stufenweise zum kompletten Profil

## Datensammlung

- 1 Monitoringdaten
  - Rohdaten
  - verarbeitete Daten
- 2 Laufzeitprofile
- 3 Zeitreihe

## Report für die Nutzer

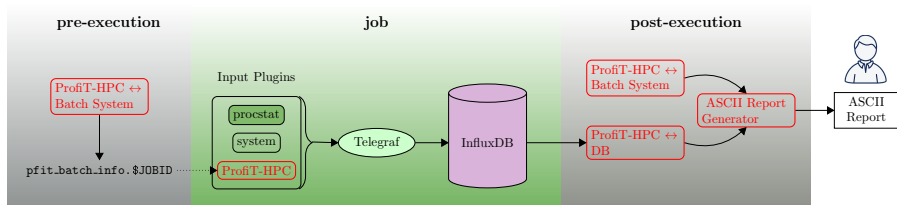
- 1 ASCII Report
  - 2 pdf Report
  - 3 Zeitreihenvisualisierung
  - 4 *Experteninterface*
- + Best Practices



## Schritte zum Toolkit

- Schnittstelle zu Batchsystemen
  - allgemeines Ausgabeformat definiert
- Wahl eines Monitoringframeworks
  - TICK Stack von InfluxData
  - *D2.2.2 bald verfügbar*
- Schnittstelle zu Telegraf
- Telegraf Plugin um JobID den Monitoring Daten zuzuordnen
- ASCII Bericht für die Nutzer
  - Eingabeformat der Monitoring- und Batchsystemdaten definiert

# Aktuelle Architektur





## ASCII Bericht Überblick

- Bericht aus den Monitoringdaten und Daten des Batchsystems
- Zusammenfassung wichtiger Jobmerkmale
- Definition von Performance Indikatoren
- Definition von Schwellwerten ab wann Warnungen angezeigt werden
  - Ressourcenanforderung
  - Ressourcenauslastung





## ASCII Bericht - Allgemeines

```
-----  
ProfiT-HPC Report  
-----
```

```
Job identified as multiple node multiple core (CPU) case
```

```
User:      <Uname>  
JobID:     <JID>  
Queue:     <QID>
```

```
Job submitted: <Time of Job Submission>
```



# ASCII Bericht - Ressourcen Check

-----  
ProfiT-HPC check of requested resources  
-----

Requested Time: <RWT>  
Used Time: <Tend-Tstart>  
Ratio: <RWT/(Tend-Tstart)>

! Warning: You requested much more time than you used.  
! This may lead to unnecessary waiting in the queue.  
! Consider reducing the requested time.

Number of processes: <nprocs>  
Number of nodes: <nodes>  
Number of CPUs: <n\_cpus>  
Number of threads per process: <n\_threads> (Average)

! Warning: You have started <nprocs>/<n\_cpus> processes  
! per CPU on node <node>

CPU Migrations: <value>

! Warning: Your application has been migrating among CPUs.  
! It may affect performance. Consider pinning.

Memory/Swapping: <yes/no>

! Warning: Swapping has been activated on node <node>.  
! Request larger memory resources.

- Vor Ende des Jahres:
  - Vollautomatische Erstellung des ASCII Reports.
  - Test des JobID-Plugins.
- 1. HJ 2018:
  - Erste Tests mit Nutzern und weiteren Rechenzentren.
  - Stufe 2 des Toolkits beginnen:
    - Festlegung der zu nutzenden Tools.
    - Modul zur vereinfachten und annähernd automatisierten Nutzung der Tools.
    - Graphische Aufbereitung der Metriken und Indikatoren.
  - Automatisierte Grafana Dashboards.

Vielen Dank für die Aufmerksamkeit!  
Fragen und Anmerkungen?

<https://profit-hpc.de>

[info@profit-hpc.de](mailto:info@profit-hpc.de)