

ClusterCockpit

-

A web application for job-specific performance monitoring

Eitzinger et al.



Universität Hamburg

DER FORSCHUNG | DER LEHRE | DER BILDUNG

Inhalte

- Was ist Monitoring?
 - Monitoring im HPC
 - Motivation
 - Mögliche Probleme beim Monitoring
- Der Einsatz von Cluster Cockpit
 - Cluster Cockpit
 - Anforderungen an ClusterCockpit
 - Motivation hinter der Entwicklung
 - Welche Probleme können durch ClusterCockpit gelöst werden?
 - Aufbau und Anwendung von ClusterCockpit
 - Was wird überwacht?
 - Alternativen zu ClusterCockpit
 - XDMoD als Konkurrenzprodukt zu Cluster Cockpit
- Fazit

Was ist Monitoring?

//

[Monitoring] ist im Allgemeinen [ein] Überbegriff für alle Arten der systematischen Beobachtung, Erfassung, Messung und Protokollierung eines Prozesses oder Vorgangs, wobei bestimmte technische Hilfsmittel zum Einsatz kommen.

Was ist Monitoring?

Monitoring im HPC

- Überwachung der HP Rechenzentren
- Analyse der Systemgesundheit & Ressourcennutzung
- Analyse & Visualisierung von Job scheduler Informationen
- Monitoring von HPC Cluster Umgebungen komplizierter
- Fehlersuche bei steigender RZ Größe komplexer
- Auswahl der richtigen Software essentiell
- Ggf. Erweiterungsmöglichkeit der Software

Motivation

- Feststellen möglicher Abweichungen zwischen Ist- & Sollzuständen
- Zuwachs von Know-How über Systeme
- Kapazitätsplanung für effizientere Systeme
- Ermöglicht maximale Performance und Verfügbarkeit

Mögliche Probleme beim Monitoring

- Kein Einbinden von Experten beim Monitoring
- Nicht genug Zeit zum Einrichten des Monitoring nehmen
- Planung über Struktur und Einsatzgebiet des Monitoring
- Setzen fester Schwellenwerte in der Monitoring Software
- Monitoring Software nicht kontrolliert und überwacht
- Sofortwarnungen in Monitoring Software eingestellt
- Keine Labor- oder Test- Umgebungssysteme erstellt

Der Einsatz von Cluster Cockpit

//

Cluster Cockpit soll die Darstellung der Monitoring Daten, sowie das User-Interface der Monitoring-Oberfläche vereinfachen und somit die Bedienung für den Endbenutzer erleichtern.

Eitzinger et al. //

Anforderungen an ClusterCockpit I

- Zugriff auf Monitoringsysteme durch Authentifizierung gesichert
- Nutzerrollen mit unterschiedlichen Zugriffsrechten
- Jobs sollen getagged werden können
- Statistische Daten über Cluster, Projekt, Nutzer, Job & Node-Daten
- Ansicht aktuell laufender Jobs
- Suchfunktion für bereits beendete Jobs

Anforderungen an ClusterCockpit II

- Interaktive und willkürliche Filterung
- Sortierung nach Jobs
- Jobsortierung basierend auf Dringlichkeit
- Detaillierte Einzelansicht auf Jobs
- Automatisches Tagging
- Verschiedene Level an Expertise auswählbar (z.B. basic, advanced, expert)
- Vollständige Konfigurierbarkeit in Web-Ansicht

Motivation hinter der Entwicklung

- Bisherige Lösungen erfüllen nicht Sicherheits- und Performancestandards
- Oberfläche und Funktionalität nicht an eigene Bedürfnisse anpassbar
- Meist komplizierte Installation
- Unübersichtliche Präsentation der Daten
- Erschwerte Skalierbarkeit

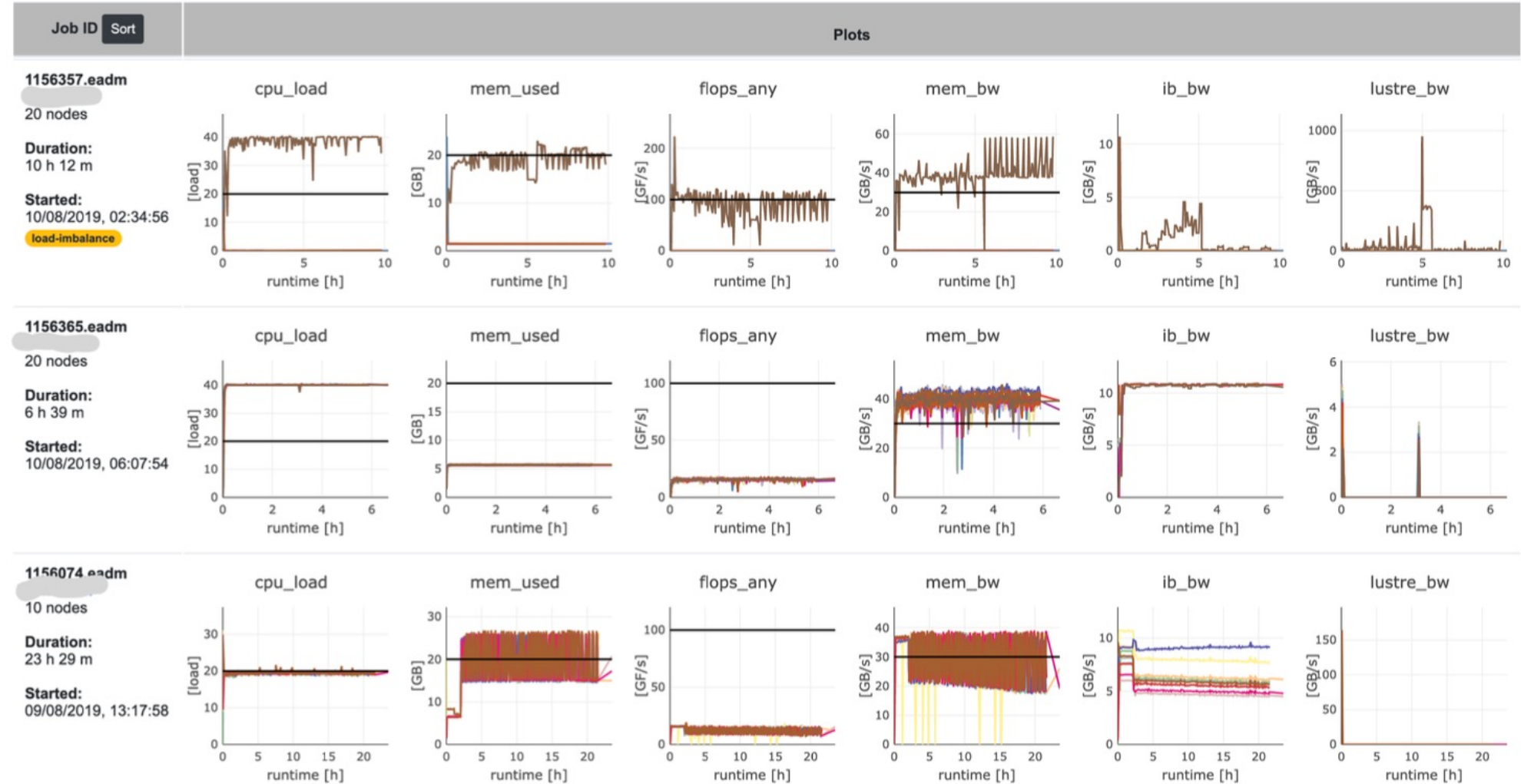
Welche Probleme können durch ClusterCockpit gelöst werden?

- Kleinere Rechenzentren ohne passende Monitoring-Software
- Übersichtliches Dashboard für Insights
- Web-Applikation mit Nutzerspezifischen Berechtigungen
- Tags für übersichtlichere Job-Auflistung und Job-Suche
- Entdecken und reparieren fehlerhafter Teile
- Unterstützung bei diversen Use Cases kleinerer Rechenzentren

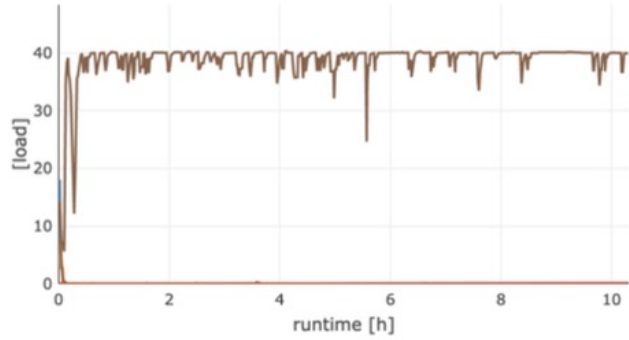
Aufbau und Anwendung von ClusterCockpit

- Live Job-, sowie Detailansicht zu jedem Job
- Job Liste als zentrale UI-Komponente
- Jobs können getagged werden
 - 'Pathological'
 - 'Low utilization'
 - 'Ticket Id'
- Zwei Varianten der Primärjobansicht:
 - Tabellen
 - Einzelansichten für Jobs
- Authentifizierung über lokale Accounts oder LDAP-Verzeichnis

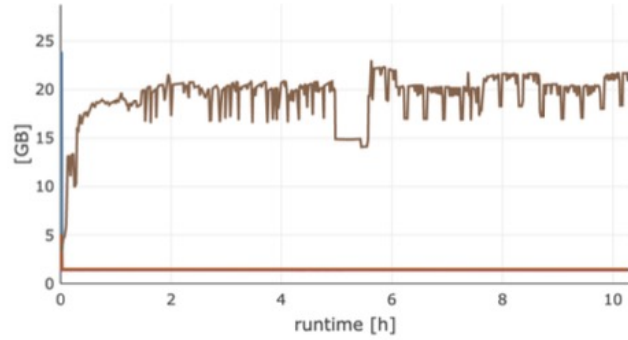
Showing 267 jobs.

Show entriesSearch: 

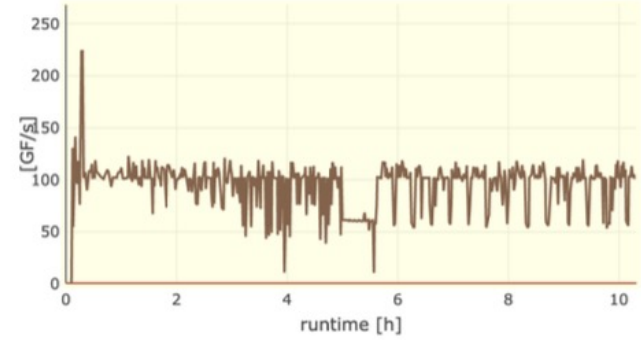
cpu_load



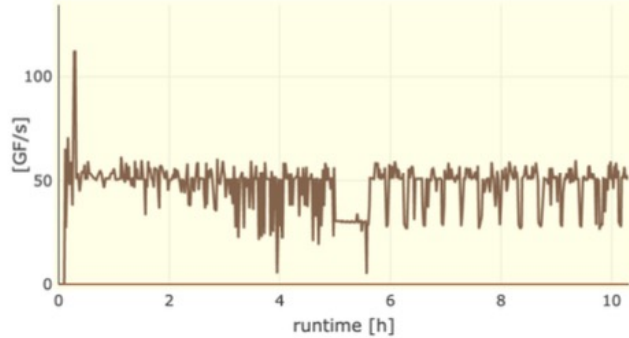
mem_used



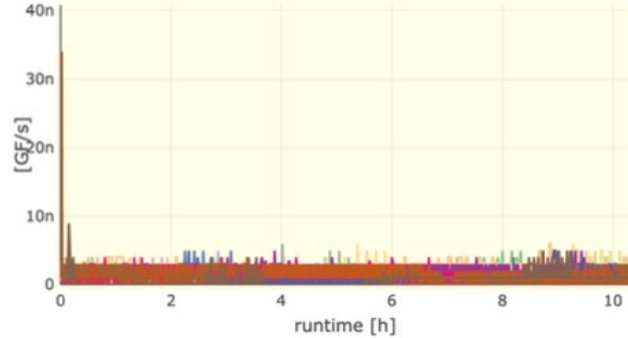
flops_any



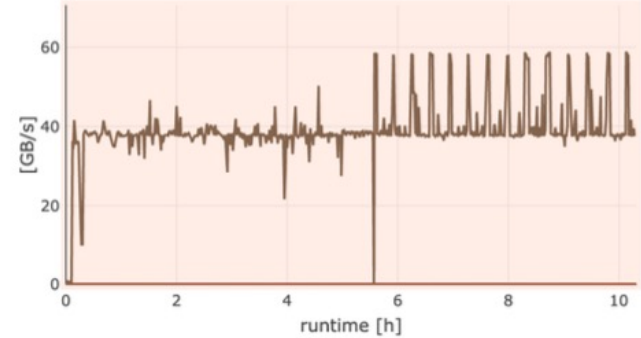
flops_dp



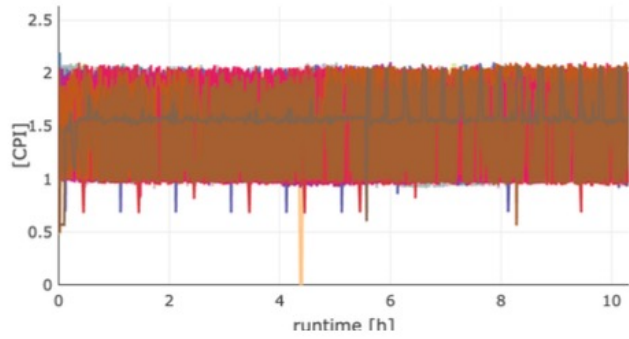
flops_sp



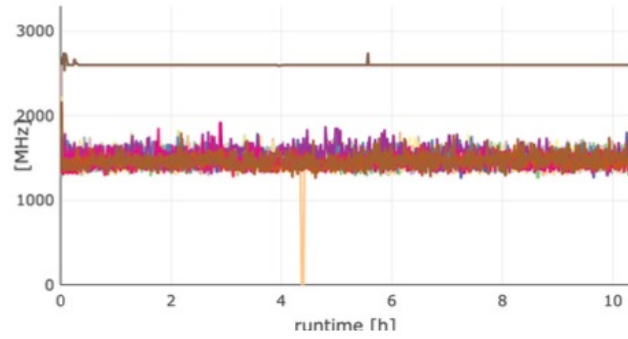
mem_bw



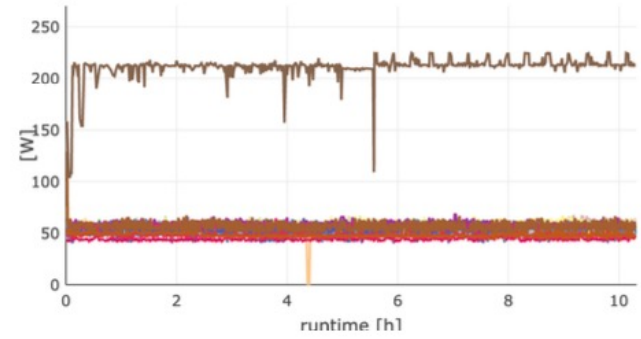
cpi



clock



rapl_power



Was wird überwacht?

CPU-Auslastung

Speichernutzung

Flops

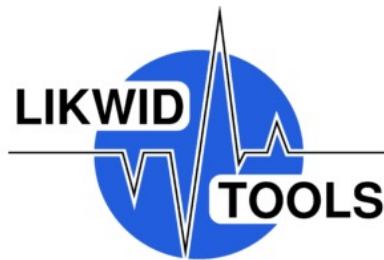
Netzwerknutzung

Energieverbrauch

Aufbau und Anwendung von ClusterCockpit



Symfony



Alternativen zu Cluster Cockpit I

- **LDMS** entwickelt von Sandia Labs
 - Modularer software stack für system monitoring
 - Vielzahl von Analysetools
 - Kein visualisierungs- bzw. Zusammenfassungstool
- **Tacc_stats** entwickelt von Texas Advanced Computer Center
 - Sammelt Standardmetriken von den HPCs
 - Automatisiertes flagging von Jobs
 - Hardware performance Data von diversen Quellen

Alternativen zu Cluster Cockpit II

- **PerSyst** hat baumähnliche Struktur
 - collection, aggregation und synchronization agents;
 - synchronization agents kontrollieren Datenansammlung
 - aggregation & synchronization agents berechnen Prozentdokumente
 - Daten durch Website präsentiert
- **DCDB** vom Leibniz-Rechenzentrum
 - Baumstruktur
 - Nodes synchronisieren sich eigenständig
 - Daten durch publish/subscribe layer hochgeladen
 - Grafana zur Datenpräsentation

XDMoD als Konkurrenzprodukt zu ClusterCockpit

- Einsatz in akademischen, Industrie- oder Regierungsrechenzentren
- Identifizieren schlechter Hard- & Software Performance
- Reporting von Job-Level Performancedaten
- Unterstützung für optimale Nutzung von HPC-Ressourcen
- Monitoring, Diagnose und Anpassung der Systemperformance
- Detaillierte, anschauliche Analyse der Anwendungsperformance
- Einfache Ablaufplanung durch Diagnostic-Tool
- Berechnung des „Return of Investment“

Fazit

Fazit

- Monitoring Software für kleine / mittelgroße RZ
- Grafische Visualisierung der Daten
- Einfache Installation
- Authentifizierung der Nutzer
- Nutzerrollen Zuweisung
- Filtermöglichkeiten für Jobs
- Expertisenfilterung

Quellen

Quellenverzeichnis

- Bilder:

- Symfony Logo (<https://symfony.com>)
- Bootstrap Logo (https://en.wikipedia.org/wiki/Bootstrap_%28front-end_framework%29)
- JavaScript logo (<https://en.wikipedia.org/wiki/File:JavaScript-logo.png>)
- Plotly Logo (<https://en.wikipedia.org/wiki/Plotly>)
- InfluxDB Logo (<https://en.wikipedia.org/wiki/InfluxDB>)
- LIKWID performance Tools logo (<https://hpc.fau.de/research/tools/likwid/>)
- MySQL Logo (<https://kabartangsel.com/apa-itu-mysql-database-pengertian-cara-kerja-fungsi-dan-kelebihannya/>)
- Redis Logo (<https://faun.pub/the-two-ways-of-storing-nested-dictionaries-in-redis-3404bedb198b>)

- Inhalte:

- Die Präsentation stützt sich auf dem Paper von Eitzinger et al. alle weiteren Informationen stammen aus folgenden Quellen:
 - <https://www.it-administrator.de/article-241881>
 - https://metrics.access-ci.org/open_xdmod/
 - <https://akrr.xdmod.org>
 - <https://blog.uptrends.de/technologie/die-vier-hufigsten-fallen-die-zu-schlechtem-monitoring-in-einer-it-umgebung-fuehren/>
 - <https://www.ip-insider.de/so-vermeiden-sie-5-allgemeine-fehler-beim-monitoring-a-764099/>

Quellenverzeichnis

- Inhalte:

- Die Präsentation stützt sich auf dem Paper von Eitzinger et al. alle weiteren Informationen stammen aus folgenden Quellen:

- <https://www.it-administrator.de/article-241881>
- https://metrics.access-ci.org/open_xdmod/
- <https://akrr.xdmod.org>
- <https://blog.uptrends.de/technologie/die-vier-haeufigsten-fallen-die-zu-schlechtem-monitoring-in-einer-it-umgebung-fuehren/>
- <https://www.ip-insider.de/so-vermeiden-sie-5-allgemeine-fehler-beim-monitoring-a-764099/>
- <https://en.wikipedia.org/wiki/FLOPS>
- <https://www.site24x7.com/learn/linux/fix-high-cpu-usage.html#:~:text=CPU%20load%2C%20on%20the%20other,%2C%20and%2015%2Dminute%20periods.>
- <https://www.manageengine.com/network-monitoring/memory-monitoring.html>
- <https://datatables.net/#:~:text=DataTables%20is%20a%20plug%2Din,Pagination>
- [https://en.wikipedia.org/wiki/Bootstrap_\(front-end_framework\)#:~:text=Bootstrap%20is%20a%20free%20and,navigation%2C%20and%20other%20interface%20components.](https://en.wikipedia.org/wiki/Bootstrap_(front-end_framework)#:~:text=Bootstrap%20is%20a%20free%20and,navigation%2C%20and%20other%20interface%20components.)
- <https://symfony.com/what-is-symfony>
- <https://www.perf-lab.hpc.fau.de/software/likwid-performance-tools/#:~:text=LIKWID%20is%20a%20node%2Dlevel,%2C%20hardware%20configuration%2C%20and%20microbenchmarking.>
- <https://www.influxdata.com/time-series-database/#:~:text=InfluxDB%20is%20part%20of%20a,an%20open%20source%20db%20core.>
- <https://aws.amazon.com/de/rds/mysql/what-is-mysql/>
- <https://redis.io/#>
- <https://medium.datadriveninvestor.com/redis-as-object-storage-e6a084b46f8b>
- <https://www.tacc.utexas.edu/research/tacc-research/tacc-stats/>
- <https://www.admin-magazine.com/HPC/Articles/HPC-Monitoring-What-Should-You-Monitor>
- <https://par.nsf.gov/servlets/purl/10404025#:~:text=Power%20usage%20of%20an%20HPC,calculations%20done%20with%20the%20system.>
- https://en.wikipedia.org/wiki/Lightweight_Directory_Access_Protocol