

Quick Cluster Overview

Matthias Bach

Physikalische Fakultät der Universität Heidelberg

Projektpraktikum Informatik, SS 06

Aufgabenstellung

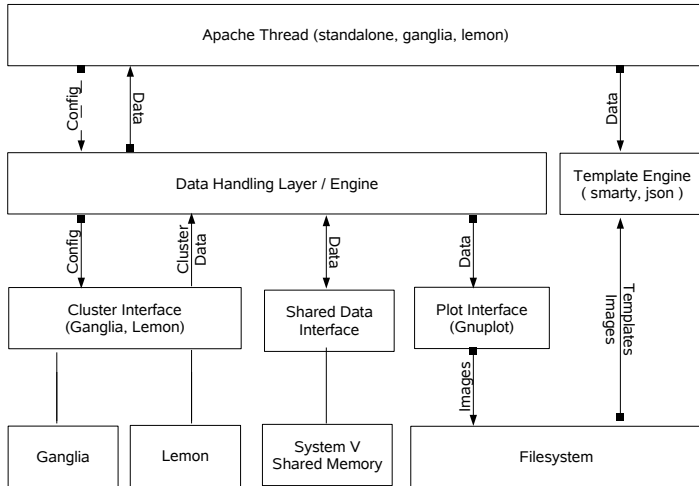
Problem:

- Visualisierung von Clusterdaten
- Vermeidung der Schwächen von Ganglia und Lemon
- Nutzung von Ganglia bzw. Lemon zur Datenerfassung

Fokus:

- Übersicht
 - ▶ Nur aktuelle Daten
 - ▶ Alle Knoten in einem Graphen
 - ▶ Möglichst wenige Metriken pro Graph
- Geschwindigkeit
 - ▶ Ausführungsgeschwindigkeit
 - ▶ Begrenzte Systemlast

Architektur



Allgemeines

- PHP 4 und aufwärts
- Apache Webserver
- Unix/Linux Betriebssystem
- Throtteling durch konfigurierbares Aktualisierungsintervall
- Modular
- Standardfunktionalität wo möglich

Plugin-System

- Einbindung verschiedener Montoringssysteme
- Ausgabe mithilfe verschiedener Plotprogramme
- Anpassung an API-Änderungen ohne Änderungen anderen Teilen des Programms
- Realisierung über PHP-Dateien mit konventionsbasierten Funktionsnamen
- Verzicht auf Objekte wegen Änderungen zwischen PHP4/5

Gemeinsamer Speicher

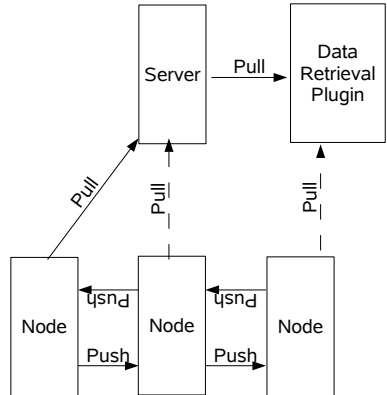
- Gewährleistet gemeinsamen Datenkontext
- Realisiert über System-V-Shared-Memory
- Datenkontext wird serialisiert gespeichert
- Semaphore zur Konsistenzgewährleistung
- Abstrahiert angebunden

Data-Retrieval

- Prüfung des Shared Memory auf aktuelle Daten
- Laden von Configuration falls notwendig
- Präparation einer Liste benötigter Metriken
- Laden des Plugins
- Abfrage des Plugins

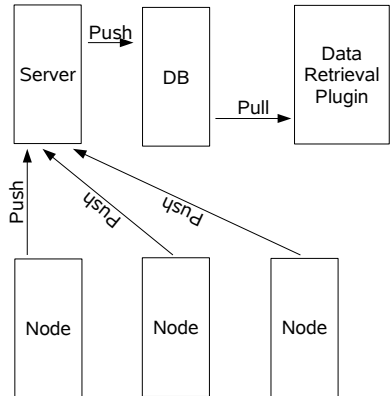
Ganglia

- Abfrage via TCP
- Ganglia stellt XML aller Werte aller Knoten
- Verarbeitung des XML mit Expat-basiertem Parser
- Benötigte Werte eines Knoten werden an Metrik gehängt
- Nicht benötigte Werte werden verworfen



Lemon

- Lemon APIs nicht stabil
- Nutzung interner Funktionalität des Lemon-UI
- Direktes parsen der Flatfile- bzw. Oracle-Datenbank
- Abfrage der Werte aller Knoten für jede benötigte Metrik



Visualisierung - Gnuplot

- Basiert auf PHP-GNUPlot
- Started Gnuplot in neuem Prozess
- Öffnet Pipe zu Gnuplots STDIN
- Übergibt Daten über Pipe
- Schreibt ausschließlich die vom Browser genutzten Dateien
- Faktor 2-4 durch Modifikationen an PHP-GNUPlot

Zusammenfassung

Realisiert:

- Schnell erfassbare Übersicht
- Sekündliche Aktualisierung möglich
- Skaliert auch für mehrere Nutzer

Was man noch machen könnte:

- Anpassung an neue Lemon APIs
- Plotten mehrerer Zeitintervalle
- Genauere Performancemessungen

