

Verteiltes Job-System

NetDist ist ein verteiltes Job-System für .NET, das von comparis.ch AG verwendet wird. Es plant die Arbeit basierend auf dem Cron-Format auf Nodes in einem Cluster. Seit seiner Einführung im Jahr 2015 hat es keine wesentlichen Neuerungen erfahren. So wird die Integration in moderne containerisierte Orchestrierungslösungen nicht unterstützt.

Basierend darauf und auf den Anforderungen von Einzelpersonen bei comparis.ch AG schlägt diese Arbeit ein generisches (unternehmensagnostisches) Konzept und einen Proof of Concept vor, um eine moderne, skalierbare und verteilte Jobverarbeitung mit dynamischem Lastausgleich zu ermöglichen. Dabei liegt der Fokus auf der Cloud-Native-Foundation und Kubernetes. Zudem werden verschiedene Standardisierungen bezüglich Containern und die Beiträge von Firmen wie Docker und Red Hat analysiert.

Mehrere Systemarchitekturen werden vorgeschlagen und miteinander verglichen, indem ihre Vor- und Nachteile gegenübergestellt werden. Danach wird die Proof of Concept Softwarebibliothek Jobbernetes vorgestellt, welche die am besten geeignete Architektur implementiert. Schliesslich wird Jobbernetes für die Entwicklung eines Anwendungsfalles verwendet, der bei comparis.ch AG zu finden ist.

Um Jobs zu planen und horizontal zu skalieren, wird Kubernetes und dessen CronJob Workload verwendet. RabbitMQ wird als WorkQueue verwendet, um die Daten dynamisch auf die Jobs zu verteilen. Indem die Outputs eines Jobs die Inputs eines anderen sind, wird ein Digraphen zur Modellierung des Datenflusses gebildet. Darüber hinaus wird ein Grafana-Dashboard entwickelt, das die Job-Metriken anzeigt. Weiter werden diese Daten mit dem RabbitMQ-Admin-Tool korreliert, um ihre Korrektheit zu validieren. Schliesslich werden zukünftige Arbeiten identifiziert, wobei der wichtigste nächste Schritt die Entwicklung einer GUI-Anwendung zur Erstellung und Überprüfung der Pipeline ist, mit dem Ziel, Jobbernetes transparent und zugänglich zu machen.



Diplomand
Anastassios Martakos

Dozent
Matthias Bachmann



Zeigt das Grafana-Dashboard von Jobbernetes, welches die Anzahl der verarbeiteten Nachrichten und den Durchsatz visualisiert. Die linke Seite wird von oben nach unten gelesen, wobei oben die Anzahl der gestarteten Elemente angezeigt wird, weiter unten die fehlgeschlagenen und schliesslich die erfolgreich verarbeiteten Elemente. Auf der rechten Seite visualisieren wir den Verarbeitungsdurchsatz pro Sekunde. Oben wird eine Linie für jeden Job angezeigt, während unten eine aggregierte Ansicht zu sehen ist.