



ERKENNUNG VON ANOMALIEN IN SENSORDATEN EINES DEEP LEARNING DEVELOPMENT SERVER CLUSTERS

PROJEKTZIEL

Der Kern des Projektes ist es Anomalien im Verhalten eines Deep Learning Development Server Clusters anhand von Sensordaten zu identifizieren. Der Cluster besteht aus High-End GPU Servern, die dazu benutzt werden große Datenmengen im Rahmen von Deep Learning Projekten zu prozessieren. Durch lange Berechnungsdauern und parallelen Zugriff vieler Entwickler, ist der Cluster eine Engpassressource. Um die Ausnutzung des Clusters besser planen zu können und Benachrichtigungen beim Start von großen Batch-Prozessen zu erkennen, war ein geeignetes Monitoring-System, mit integrierter Erkennung von Anomalien notwendig. So können Abweichung vom normalen Verhalten des Clusters erkannt und automatisierte E-Mail Benachrichtigungen versandt werden.

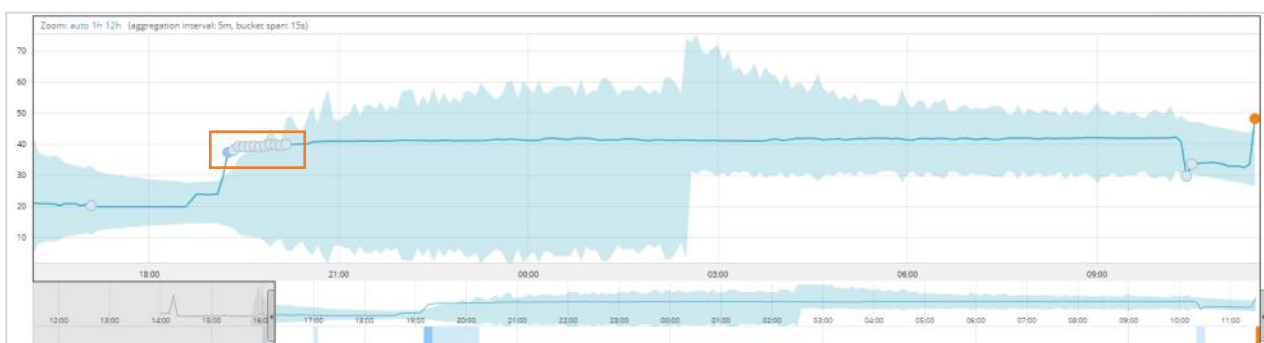
DATENSATZ

Für jeden Server werden sekundlich 58 Datenpunkte verschiedenster Sensoren aufgezeichnet und jeweils in eine CSV-Datei gespeichert. Diese CSV-Dateien werden täglich gesplittet, um die Dateigrößen klein zu halten, da jeden Tag über 65.000 Datensätze pro Server generiert werden.

Die Sensoren erfassen verschiedene Leistungskennzahlen der Hardwarekomponenten. Mit der Speicherauslastung, Temperaturen, Taktfrequenzen und dem Energieverbrauch werden alle wichtigen Aspekte zum Clusterzustand und Auslastungen abgedeckt.

HERAUSFORDERUNGEN

Die Daten werden auf den Entwicklungsservern gesammelt und gespeichert, während die Analyse auf einem dedizierten Server durchgeführt wird. Der Datenstrom muss gewährleisten, dass Transferverzögerungen minimal sind. Da die Deep Learning Anwendungen die meiste Zeit entweder hohe oder keine Auslastung verursachen, reichen einfach Grenzwerte nicht aus, um Anomalien zu detektieren. Das Machine Learning Modell muss Trends und Saisonalitäten berücksichtigen und unerwartete Trendveränderungen oder Ausreißer erkennen. Der Algorithmus muss also in der Lage sein den Trend eigenständig zu erkennen und sich dahingehend anzupassen.



Beispiel für eine erkannte Anomalie und anschließende Anpassung des Trends

ANGEWANDTE METHODIK

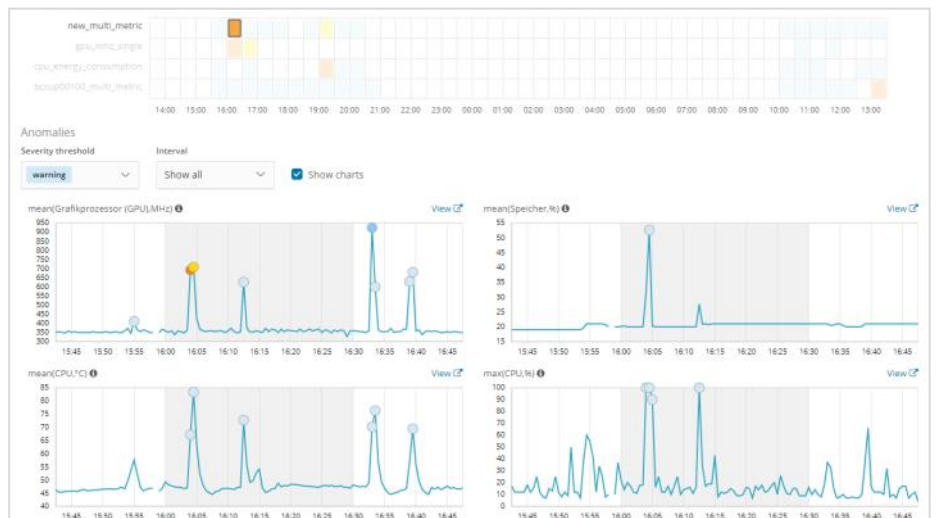
Wegen seiner Skalierbarkeit wurde der Elastic-Stack als Analyseumgebung gewählt. Mit dem Logstash Modul können Datenströme in Echtzeit direkt in den Elasticsearch-Cluster geladen werden. Die eigentliche Analyse findet im Machine Learning Modul, in der Erweiterung X-Pack von Kibana statt.

Um die Log-Daten direkt verfügbar zu machen, wurde eine direkte Netzwerkfreigabe zu den entsprechenden Dateipfaden konfiguriert. Dadurch kann Logstash direkt auf die Sensordaten zugreifen, wo es permanent nach neuen Information sucht und diese in den Elasticsearch-Cluster lädt. Die aktuellsten Sensordaten sind somit in weniger als einer Sekunde Verzögerung bereit zur Analyse.

Das Machine Learning Modul wurde so konfiguriert, dass es mehrere kritische Sensoren gleichzeitig analysiert und die Änderungen der Trends automatisch erkennt und übernimmt. Analysiert werden Temperaturen, Taktfrequenzen sowie Auslastung von CPU, GPU, RAM und Speicherkomponenten. E-Mail Benachrichtigungen für unerwartete Events wurden ebenfalls eingerichtet.

PROJEKTERGEBNIS

Ein Echtzeit-Monitoring Tool mit einem selbstlernenden Modell zur Erkennung von Anomalien wurde implementiert. Es wurde so konfiguriert, dass es bei unerwarteten Ereignissen automatisch E-Mail Alerts versendet. So kann die Nutzung des Deep Learning Server Clusters optimiert werden. Durch das Monitoring können ebenfalls potentielle Kapazitätsengpässe schnell erkannt werden.



Der Anomalie-Explorer zeigt Events der überwachten Sensoren

time	severity	detector	actual	typical	description	job ID	actions
> November 20th 2018, 16:14:00	72	Max GPU MHz	1,000	373.737	↑ 3x higher	bcsup00100_multi_metric	⚙
> November 20th 2018, 15:27:45	72	SUM Energy Consumption	0	18,252.57	↓ Unexpected zero value	bcsup00100_multi_metric	⚙
> November 19th 2018, 19:26:00	50	Mean CPU %	63.2	6.645	↑ 10x higher	bcsup00100_multi_metric	⚙
> November 19th 2018, 19:24:45	52	Mean CPU %	68.889	6.593	↑ 10x higher	bcsup00100_multi_metric	⚙
> November 19th 2018, 19:19:00	72	Max CPU Temperature	93	53.222	↑ 2x higher	bcsup00100_multi_metric	⚙

Auflistung der erkannten Anomalien eines Servers

WEITERE ANWENDUNGSFÄLLE

In dem Machine Learning Modul können in einem Kalender Einträge erstellt werden, welche zur Erklärung des Clusterverhaltens beitragen können. Diese Events treten dann nicht als Anomalie auf. In Zukunft sollen Entwickler Zeiten für die Benutzung des Servers in diesem Kalender buchen können. Der Ursprung der Anomalien ist in diesen Zeiträumen dann bekannt und es werden keine Alerts verschickt. Die Anzahl von unnötigen Alarmen wird somit reduziert.

Die Benachrichtigung können weiterhin mit exakteren Regeln versehen werden, um gezieltere Alarme zu generieren.

