

Leerkilometerreduktion in der Logistikbranche mit Data Science



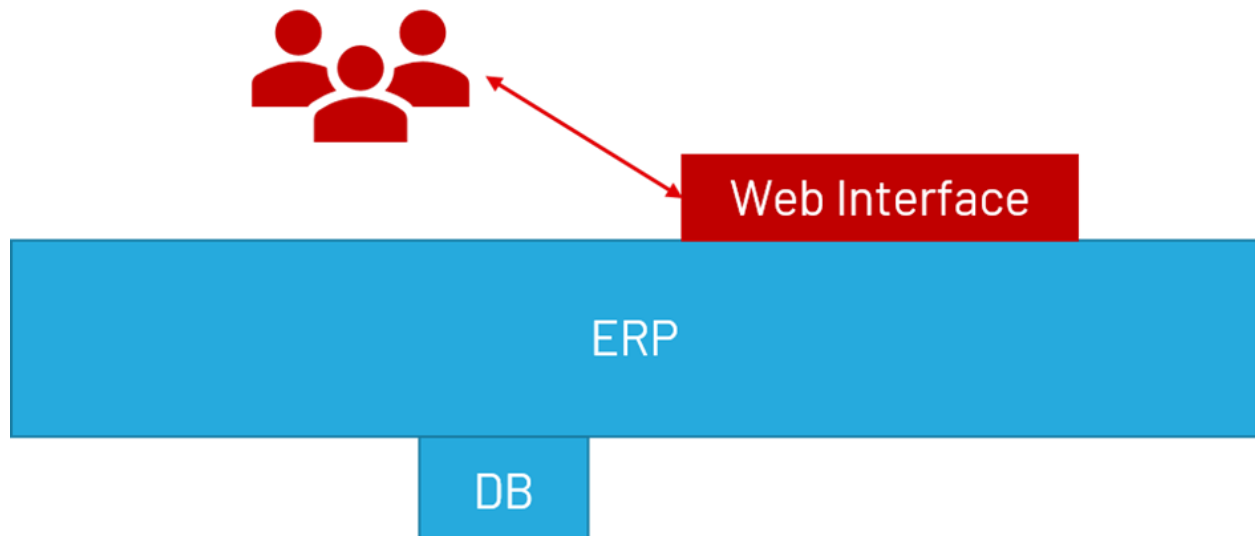
Ziel jedes Transportunternehmens ist es, effiziente Routen zu planen und leere Kilometer zu vermeiden. Das spart nicht nur Kosten, sondern ist auch besser für die Umwelt. Wir zeigen, wie man mit Data Science die Fragestellung erfolgreich lösen kann.

Das Data Science Team der IT-PS durfte im vergangenen Jahr ein spannendes Projekt gemeinsam mit der Gartner KG vom Konzept bis zur produktiven Integration der Lösung ins ERP durchführen.

Die verschiedenen Teams in der Disposition versuchen Aufträge und verfügbare LKWs möglichst so zuzuteilen, dass möglichst wenige Leerkilometer zwischen den Aufträgen entstehen. Dabei ist es jedoch sehr schwer den Überblick über die gesamte Flotte und alle Abteilungen zu wahren. Deswegen war es das Ziel, die Disponierenden im Prozess zu unterstützen, damit bessere Entscheidungen schneller getroffen werden können. Mit Echtzeitdaten sollten automatisiert, auf Knopfdruck, Vorschläge generiert werden, um die Suche nach dem besten LKW für einen Auftrag (oder den besten Auftrag für einen LKW) zu erleichtern.

Datengrundlage: DB2 auf IBM Power

Die Problemstellung, und damit die grundlegende Use Case Idee, war bekannt. Die Details und speziellen Anforderungen mussten noch geklärt werden. Zusätzlich mussten die Datenqualität und Datenverwendbarkeit für den Use Case bewertet werden. Die Datengrundlage für den Use Case lieferten zwei unterschiedliche IT-Systeme: das ERP (enterprise resource planning), das unter IBM i läuft und das TMS (transport management system), das die Daten über Schnittstellen in das ERP schickt. Die Daten werden dabei zentral in einer DB2 Datenbank gespeichert.



Das bestehende ERP System

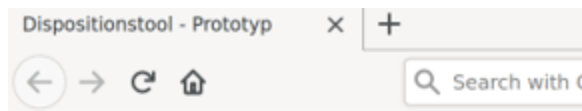
Entwicklung eines Prototyps mit dem IT-PS Data Lab auf IBM Power

Das IT-PS Data Lab integriert unterschiedliche Entwicklungsumgebungen, wie Rstudio Server und JupyterLab, und schafft reproduzierbare Umgebungen für Data Scientists mit allen wichtigen Open Source Frameworks für Python und R. Durch konsequenten Einsatz von Docker-Containertechnologie kann das IT-PS Data Lab sehr rasch in Kundenumgebungen im Rechenzentrum des Kunden oder von IT-PS aufgebaut werden. Auch ein Einsatz in der Cloud ist möglich. Ein eigener Docker-Build-Framework ermöglicht es Docker Images für verschiedene CPU-Architekturen, wie Intel x86 und IBM Power ppc64le, zu erstellen.

In diesem Projekt wurde das IT-PS Datalab in einer Linux on Power LPAR neben der IBM i LPAR installiert. Das hat den Vorteil, dass der Datentransfer schnell und sicher am

bestehenden Server stattfindet und auch keine zusätzlichen Startinvestitionen in Infrastruktur anfallen.

Der Prototyp wurde nach dem bewährten IT-PS Wocheniterationen bearbeitet. Zu Beginn wurden erste Visualisierungen gemacht, um das Verständnis für den Prozess zu schärfen und um die Daten gegen den Use Case zu evaluieren. Auch erste Lösungsansätze wurden intensiv diskutiert. Ein wichtiger Erfolgsfaktor war die Entwicklung eines zunächst nicht-funktionellen Prototyps, um die Anforderungen an die Endlösung zu präzisieren. Dabei wurden einfache HTML Seiten für die Suche und die Ergebnisse entworfen.



Suche

LKW Kennzeichen:

Die prototypische Suchmaske



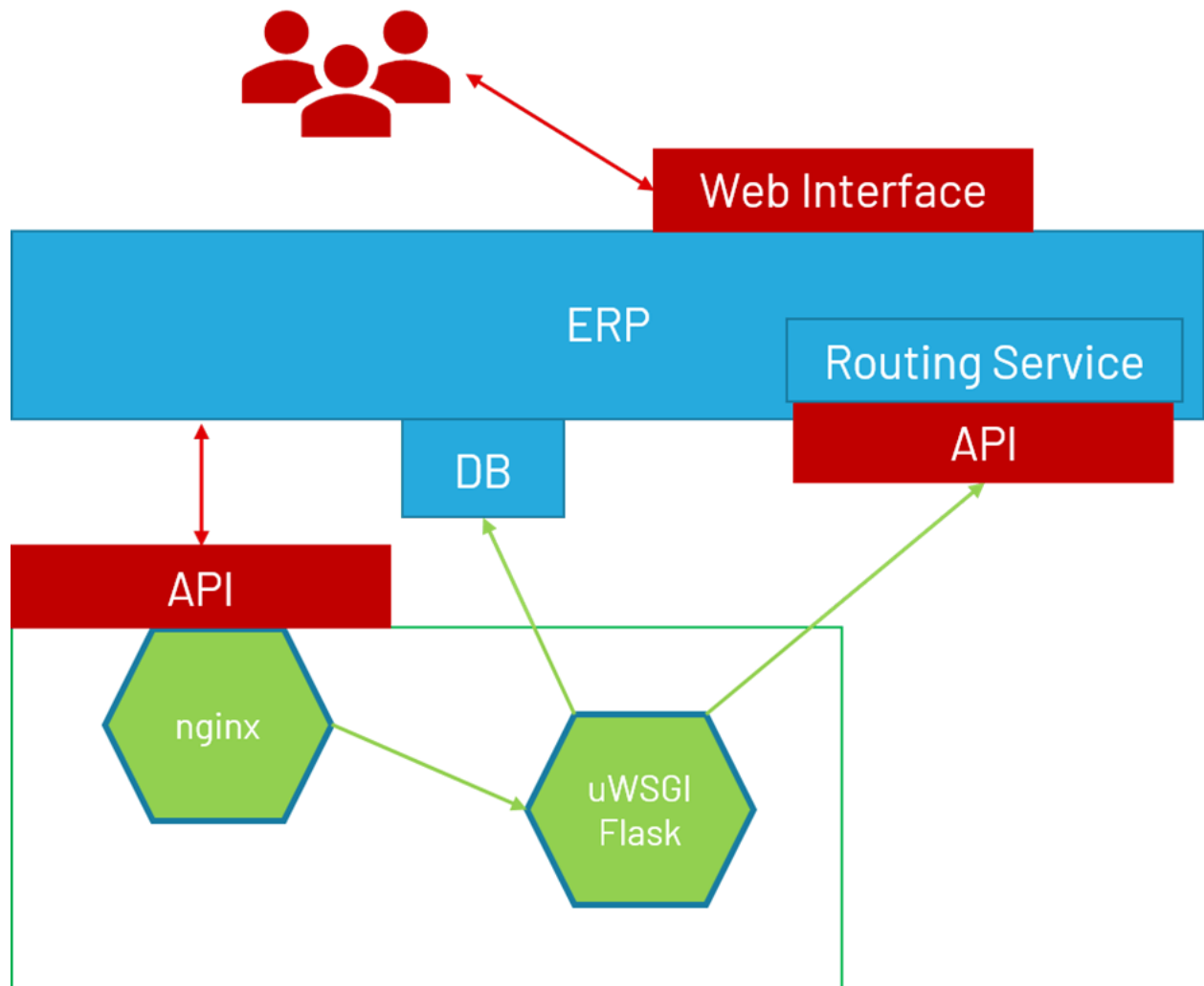
Die prototypischen Ergebnisse

Das hilft mögliche Stolpersteine früh zu erkennen und zu beheben und obendrein ermöglicht es, den Quellcode später in der Produktionsalisierung wiederzuverwenden.

In den späteren Iterationen wurden Modellergebnisse diskutiert und der Prototyp mit Flask funktional gemacht, was einen ersten Test des entscheidungsbaumbasierten Empfehlungssystems ermöglichte.

Integration der Lösung in das ERP System

Nach einer Testphase wurde beschlossen, den entwickelten Prototypen in das bestehende ERP System zu integrieren, um ihn allen Disponierenden zur Verfügung stellen zu können. Dabei konnte die prototypische Umsetzung als Basis wiederverwendet werden. Wie bei jeder Produktionsialisierung wurde der Quellcode gehärtet und sichergestellt, dass das System fehlertolerant skaliert und wartbar ist.



Architekturskizze des integrierten Datenprodukts

Der prototypische Entwicklungsserver wurde durch nginx und uWSGI ersetzt. Das Deployment verwendet einen single-node Docker Swarm Cluster zur Orchestrierung, damit auch die Produktivumgebung nach dem "Infrastructure as Code"-Gedanken versioniert werden kann und reproduzierbar ist.

Bestehende Ressourcen als Startvorteil

So wie bei diesem konkreten Anwendungsfall bestehen auch in anderen Situationen oft freie Kapazitäten auf vorhandenen IBM Power und IBM i Plattformen, die genutzt werden können, um die ersten Data Science und Machine Learning Use Cases umzusetzen, ohne dass eine Investition in neue Hardware notwendig wird.

Damit ist ein schneller und kostengünstiger Einstieg in die Welt von Data und AI garantiert!

Platinum
Business
Partner

