

Ein Konzeptansatz für das zukünftige Daten- und Informations-Management in der Landbewirtschaftung mit Hilfe verteilter Datenbanken

Hans-R. Langner

Institut für Agrartechnik Potsdam (ATB)
Max-Eyth-Allee 100
14469 Potsdam
hlangner@atb-potsdam.de

Abstract: Der Beitrag erläutert einen Konzeptansatz für das „Precision Agriculture“ auf der Grundlage eines verteilten Datenbanksystems und einer Datenreplikation über verfügbare drahtlose Netzwerke WLAN, GSM oder UMTS. Im Vergleich zum Stand der Technik erhält man eine höhere Aktualität der Steuerdaten für die Landmaschinen, eine verbesserte Integrität der Rohdaten sowie eine größere Datensicherheit. Der Einsatz einer verteilten Datenbank direkt vor Ort auf den Landmaschinen ermöglicht eine effektivere Speicherung und Verarbeitung der Rohdaten. Der Entwurf eines Entity-Relationship-Modells (ERM) für ortsbezogene landwirtschaftlichen Daten wird dargelegt und verschiedene Möglichkeiten zur Replikation der landwirtschaftlichen Daten zwischen den Landmaschinen bzw. zwischen den Speicherorten der Teildatenbanken werden betrachtet (teilweise Replikation, Snapshots).

1. Einleitung

Eine zukunftsorientierte Landbewirtschaftung erfordert auch die Entwicklung moderner Hard- und Softwarelösungen für die Landwirte, mit deren Hilfe eine effektivere, umweltschonende und auch kostengünstigere Bewirtschaftung der Ressource Land möglich wird. Ein wichtiger Ansatz für eine präzisere Landwirtschaft, das Precision Agriculture, wird in der Zukunft verstärkt zum Einsatz von Sensoren führen. Durch eine Landbewirtschaftung mit Precision Agriculture werden im Vergleich zu heute erheblich größere Datenmengen generiert, woraus die Forderung nach einer ausreichend leistungsfähigen Hard- und Software für Precision Agriculture resultiert. Die großen Datenmengen entstehen durch das ortsbezogene Erfassen von Bodendaten, pflanzenbaulichen Daten sowie aus der durchgehenden Protokollierung aller landwirtschaftlichen Bearbeitungsschritte. Weiterhin ist eine mehrjährige Erfassung und Auswertung von Precision Agriculture Daten notwendig, um die Bearbeitungsschritte über saisonale Schwankungen hinaus auswerten zu können [Ehl97],[EJ97].

Für das Precision Agriculture wird nachfolgend eine verteilte Datenbank-Lösung betrachtet, die alle landwirtschaftlichen Maßnahmen ortsbezogen von der Planung über die Durchführung bis hin zur Abrechnung speichern und verwalten kann. Verfügbare Soft-

warelösungen, wie z.B. Ackerschlagdateien oder landwirtschaftliche Abrechnungs-Programme umfassen nur einen kleinen Bereich der Bewirtschaftungskette und müssen für Precision Agriculture in ihrem Funktionsumfang und speziell für das ortsbezogene Datenmanagement angepasst werden. Für die zukünftige Hard- und Softwarelösungen auf den Landmaschinen ergibt sich die Forderung, dass sie offen sein müssen für eine Online-Datenerfassung, die direkt von den Landmaschinen ausgehen wird. Die Bordelektronik der heutigen Landmaschinen, basierend auf 8Bit-/16Bit-Jobcomputern, Anzeigeterminals und ISO-Bus, ist noch nicht ausreichend auf die Anforderungen des Precision Agriculture vorbereitet und besitzt keine ausreichenden Kapazitäten für die Speicherung großer Datenmengen. Das Planen, Steuern und Auswerten ortsbezogener Bewirtschaftungs-Maßnahmen sowie das Dokumentieren und Abrechnen landwirtschaftlicher Maßnahmen wird zukünftig zu den zentralen Aufgaben des Precision Agriculture gehören. Wichtige Leistungsmerkmale der Datenbank-Lösung sind die automatische Überprüfung der Eingangsdaten, die effektive Kontrolle über die Datenzugriffe und die Art der Replikation von Ergebnisdaten zwischen verschiedenen Standorten.

2. Konzeptansatz für eine verteilte Datenbank

Üblicherweise wird das konzeptionelle Modell für eine verteilte Datenbank in drei Schritten generiert [BG92] [PM96]. Zuerst legt man in einem Gesamtentwurf das Entity-Relationship-Modell (ERM) und die notwendigen Entitäten und ihre Beziehungen fest, listet ihre Merkmale auf und überprüft die Relationen des Modells. Im zweiten Schritt wird aus dem ERM ein physikalisches Datenmodell für eine konkrete Datenbank generiert. In einem dritten Schritt unterteilt man die Tabellen des physikalischen Entwurfs in mehrere Fragmente für die Datenspeicherung und realisiert diese Fragmente in jeweils separaten Teildatenbanken.

3. Grundkonzepte zur Fragmentierung des ERM-Ansatzes

Es gibt erste Voruntersuchungen darüber, wie der Datenfluss für Precision Agriculture innerhalb der Landwirtschaftsbetriebe sowie zwischen den Landmaschinen zukünftig strukturiert und unter Einbeziehung verschiedener Netzstrukturen realisiert werden kann. In einem Forschungsansatz am Institut für Agrartechnik Bornim (ATB) geht es um die Einbeziehung der verfügbaren Kommunikationsnetze GSM und UMTS. Als wesentlicher Schritt zur Praxiseinführung von Precision Agriculture ist eine Automatisierung der Erfassung, Speicherung und Verarbeitung von landwirtschaftlichen Daten notwendig. Eine weitestgehend automatisch ablaufende Datenverarbeitung kann schrittweise durch die Umsetzung folgender Forderungen erreicht werden:

- Automatisierung des Datenflusses zwischen den Bordrechnern der Landmaschinen und den Client-Server-Rechnern des Landwirtschaftsbetriebes, d.h. kein manuelles Weitergeben von Daten in Form von EXCEL-, ACCESS- oder anderen Files.

- Erhöhung der Datensicherheit und der Datenintegrität durch die Einführung von Datenbanken. Abkehr vom Speichern betriebsrelevanter Daten in diversen Fileformaten.
- Kontrolle der Zugriffe auf die Betriebsdaten durch ein Datenbank Management System (DBMS) und automatische Übertragung (Replikation) von Daten zwischen mehreren Teildatenbanken.
- Anwendungsorientierte Aufteilung der Betriebsdaten (Fragmentierung) sowie Speicherung an den Orten, wo sie am häufigsten angewendet und weiterverarbeitet werden.

Die Datenreplikation und der Bedarf an Kommunikationstechnik sind stark abhängig von der Art der Fragmentierung der Daten sowie von ihrer anwendungsspezifischen Nutzung durch Auswerteprogramme. Allgemein kommen drei grundsätzliche Konzepte zur verteilten Datenhaltung in Betracht[CP84].

3.1 Fragmentierte Datenhaltung ohne Replikation

In diesem Fall wird der Datenbestand in disjunkte Fragmente aufgeteilt und die Fragmente auf mehrere Datenbanken verteilt, ohne dass ein automatischer Austausch zwischen den Teildatenbanken stattfindet. Das Zusammenführen der Datenfragmente muss manuell durch den Anwender erfolgen. Für eine mögliche Fragmentierung des Datenbestandes seien zwei unterschiedliche Lösungsansätze kurz skizziert:

- Ortsbezogene Fragmentierung des gesamten Datenbestandes auf der Grundlage von GPS-Koordinaten bzw. Ackerschlaggrenzen. In diesem Fall entstehen in Analogie zu üblichen Ackerschlagdateien Datenfragmente, die jeweils komplette Datensätze zu den Messpunkten der Ackerschläge enthalten. Diese ortsbezogene Aufteilung der Daten ist speziell bei mehreren landwirtschaftlichen Teilbetrieben mit getrennter Flächenbewirtschaftung sinnvoll.
- Attributbezogene Fragmentierung des gesamten Datenbestandes auf der Grundlage tätigkeitsbezogener Merkmale. Hierbei entstehen Fragmente (Teildatensätze), die jeweils eine bestimmte Bearbeitung der Flächen erfassen und speichern, z.B. das Fragment Ernteerträge. Die tätigkeitsbezogene Aufteilung hat den Vorteil, dass die Speicherung der Datenfragmente den Landmaschinen genauer zugeordnet werden kann (das Fragment Ernteerträge z.B. dem Mähdrescher).

Vorzugsweise sollte die Fragmentierung der Daten nach Attributen der landwirtschaftlichen Bearbeitung durchgeführt werden, d.h. nach den Bearbeitungsschritten Drillen, Düngen, Pflanzenschutz, Ernten u.a.m. Durch die Fragmentierung nach Bearbeitungsschritten können die Daten den entsprechenden Landmaschinen zugeordnet werden und sind somit an den Standorten verfügbar, wo sie mit der größten Wahrscheinlichkeit benötigt werden. Die Anforderungen an die Hardware auf den Landmaschinen und an das Speichervolumen bleiben bei dieser Datenhaltung moderat, da keine Replikation der Daten stattfindet.

3.2 Fragmentierte Datenhaltung mit Replikation ausgewählter Daten

In der Regel ist es nicht erforderlich, alle Rohdaten und alle bearbeitungs-spezifischen Daten ständig zwischen allen Teildatenbanken auszutauschen. Die Datenhaltung mit ausgewählter Replikation beschränkt den Datenaustausch auf ausgewählte Ergebnisdaten. Sie ist eine Kompromisslösung, die unter günstigen Umständen die Vorteile einer Datenhaltung ohne Replikation (geringe Kosten, hohe Verfügbarkeit der Teildaten vor Ort) mit den Vorteilen einer Datenhaltung mit Snapshot-Replikation verbinden kann (hohe Datensicherheit durch Redundanz, hohe Verfügbarkeit aller Daten, hohe partielle Leistung des Systems).

3.3 Fragmentierte Datenhaltung mit Replikation durch Snapshots

Die Systemkosten für eine Datenhaltung mit vollständiger Replikation können durch eine zeitlich eingeschränkte Replikationslösung gesenkt werden. In der landwirtschaftlichen Praxis werden nicht zu jeder Zeit alle Daten benötigt, da der Getreideanbau z.B. stark saisonal geprägt ist. Es reicht häufig aus, eine Kopie der Gesamtdaten zu bestimmten Zeiten bereitzustellen. Das Kopieren der Gesamtdaten zu bestimmten Zeitpunkten nennt man Snapshots. Die Snapshots werden in der verteilten Datenbank periodisch aktualisiert, z.B. in der Saison täglich, sonst wöchentlich. Die Verfügbarkeit der Snapshots auf den Landmaschinen ermöglicht es, die Bearbeitungsschritte der Landbewirtschaftung auf der Basis des gesamten Datenbestandes bzw. bestimmter zeitlicher Sichten auf den Datenbestand zu planen und durchzuführen.

4. Literatur

- [BG92] Bell, D.; Grimson, J.: Distributed Database Systems. Addison Wesley Longman, 1992.
- [CP84] Cheri, S.; Pelagatti, G.: Distributed Databases: Principles and Systems. McGraw-Hill New York, 1984.
- [EJ97] Ehlert, D.; Jürschik, P.: Precision Farming - a German Perspective. Proceedings of the INMA'97 "Present and Future in the Scientific Research of Machines and Installations Designed to Agriculture and Food Industry", Bucharest 18./19.9.1997, Vol.1, S. 18 – 25.
- [Ehl97] Ehlert, D.: Techniques for Determining Heterogeneity for Precision Agriculture. Papers presented at the First European Conference on Precision Agriculture, Warwick University Conference Centre, UK, 7.-10. September 1997; S. 627-634.
- [PM96] Podeameni, S.; Mittelmeir, M.: Distributed Relational Database. Cross Platform Connectivity. Prentice-Hall Englewood Cliffs NJ, 1996.