In-Memory Database Optimizations for Modern Hardware

Motivation

Seit ca. 1,5 Jahren wird am Lehrstuhl von Prof. Plattner / Dr. Zeier - Enterprise Platform and Integration Concepts - in Kooperation mit dem MIT Institut CSAIL ein Hauptspiecherdatenbankprototyp namens HYRISE entwickelt. Die Motivation für HYRISE ist die für moderne Unternehmensanwendungen immer wichtiger werdende Anforderung transaktionale und analytische Konzepte miteinander zu verbinden. Unter transaktionalen Konzepten wird die typische Dateneingabe und Prozessverarbeitung verstanden, während analytischen Konzepte die Betrachtung von Zusammenhängen dieser transaktionalen Daten versteht. Da jedes dieser Konzepte ein spezifisches Zugriffsmuster erzeugt, werden heute die Anfragen auf verschiedene Systeme verteilt, die optimierte physikalische Datenablagen je Anwendungsfall verwenden. Neben der damiteinhergehenden Redundanz führt diese Trennung in der traditionellen Architektur zu einer Verzögerung in der Auswertung der Daten. Des Weiteren ist es damit nicht möglich analytische Anfragen direkt in transaktionalen Anwendungen zu verwenden um z.B. Simulationen während der Dateneingabe zu ermöglichen. Moderne Server mit mehr als 512GB Hauptspeicher erlauben es komplette Datenbanken darin vorzuhalten. Daneben bieten Multicore-CPU Architekturen genügend Rechenkapazität um auf Basis der transaktionalen Events analytische Anfragen direkt zu berechnen ohne das spezifische Materialisieren notwendig sind. Die sich aus diesen Hardwaretrends ergebenden potentiellen Optimierungen sind, unter Berücksichtigung der Anforderungen von Unternehmensanwendungen, im Fokus der Forschung und protypischen Implementierung von HYRISE.

Projektbeschreibung

Ziel des Projektes ist es, neue Konzepte und Algorithmen aus der Forschung theoretisch und praktisch im Prototyp HYRISE anzuwenden. Der Fokus liegt dabei auf den Möglichkeiten und Trends moderner Hardwarearchitekturen, wie der wesentlich größere Hauptspeicher und die erhöhte Rechenkapazität der Multicore-CPU's. Es werden dabei Konzepte evaluiert, die besondere Eigenschaften von Intel-Prozessoren ausnutzen und damit einen zusätzlichen Geschwindigkeitsvorteil erzielen – z.B. durch spezielle Single Instruction, Multiple Data (SIMD) oder Streaming SIMD Extensions (SSE) Operationen. Evaluiert werden die gewonnen Erkenntnisse mit modernster Hardware im Enterprise Application Architecture Lab des Lehrstuhls und Kundendaten von Unternehmensanwendungen um möglichst realistische Szenarien abzubilden.
Die Umsetzung erfolgt dabei in enger Kooperation und technologischer Unterstützung durch unseren Projektpartner SAP.

Projektvorbereitung

Zur Vorbereitung des Projektes werden die relevanten Themen aus dem Bereich der Hauptspeicherdatenbanken vorgestellt. Im Verlauf der Vorbereitung werden diese von den Projektteilnehmern theoretisch und praktisch bearbeitet.

Projektpartner

Dieses Projekt findet in Kooperation mit unserem Projektpartner SAP AG statt. Die SAP AG ist der weltweit größte Hersteller von Unternehmenssoftware. Das Kerngeschäft liegt in der Entwicklung und Wartung von Unternehmenssoftware für alle Unternehmens-größen und -arten, z.B. im Finanzwesen und der Unternehmensplanung.

Kontakt

Für weiterführende Informationen stehen wir gerne zur Verfügung. Entweder in der Villa im Raum V2.05 oder per E-Mail an Jens Krüger oder Martin Grund. Das Projekt wird geleitet von Dr. Alexander Zeier, Jens Krüger und Martin Grund.

Dr. Alexander Zeier (alexander.zeier@hpi.uni-potsdam.de)
Jens Krüger (jens.krueger@hpi.uni-potsdam.de)
Martin Grund (martin.grund@hpi.uni-potsdam.de)

Teilnehmer und Projektbeginn

Zwischen 6 und 8 Teilnehmer können in diesem Projekt mitarbeiten. Projektbeginn ist der 19. Oktober 2010.