

Bericht vom Herbsttreffen der GI-Fachgruppe Datenbanksysteme

Alfons Kemper · Tobias Mühlbauer · Thomas Neumann ·
Angelika Reiser · Wolf Rödiger

the date of receipt and acceptance should be inserted later

Das diesjährige Herbsttreffen der GI-Fachgruppe Datenbanksysteme fand am 22. und 23. November an der Technischen Universität München statt. Das Treffen hatte sich “Scalable Analytics” zum Thema gesetzt und stand unter dem Motto “Industry meets Academia”. Mit über 80 Teilnehmern aus der deutschsprachigen Datenbank-Community war das Treffen gut besucht. Insgesamt 10 Firmen sowie 36 Universitäten und Forschungseinrichtungen waren dabei vertreten. Vor diesem Publikum haben Datenbankler und Datenbanklerinnen aus Industrie und Forschung in insgesamt 14 Vorträgen neueste Entwicklungen und Technologien für skalierbare analytische Anfragebearbeitung präsentiert. Im Folgenden wollen wir kurz von diesen berichten.

Zum Auftakt gab es am ersten Tag einen Vortrag von **Jens Teubner (ETH Zürich)** zur “Ausnutzung moderner Hardware für optimierte Datenverarbeitung”. Dabei stellte er heraus, dass Multi-CPU-Systeme heute de facto Netzwerke darstellen, für welche Algorithmen benötigt werden, die die Topologie des Systems berücksichtigen. Es schloss sich eine Präsentation von **Alexander Hall (Google)** an. Er stellte die Arbeiten zu seiner diesjährigen VLDB-Publikation “Processing Trillion Cells per Mouse Click” vor. Weiter ging es mit Franz Färber (SAP), welcher über die Unterstützung skalierbarer Analysen mit HANA DB berichtete. **Kostas Zumas (TU Berlin)** gab einen Überblick über Stratosphere, einem System zur Analyse von großen Datenmengen. Mit “Complex Event Processing” sprach **Bernhard Seeger (Universität Marburg)** über Er-



Abb. 1 Im Hofbräuhaus

(Foto: Silke Prestel)

Alfons Kemper · Tobias Mühlbauer · Thomas Neumann ·
Angelika Reiser · Wolf Rödiger

Institut für Informatik – Lehrstuhl III, TU München,
Boltzmannstraße 3, 85748 Garching bei München,
Deutschland

weiterungen von Systemen zur Ereignisverarbeitung um dynamische Event Processing Agents. Nach einer Pause und einem Rutschwettbewerb auf der berühmten Parabelrutsche im Garching Informatik-Gebäude berichtete der aus Amsterdam angereiste **Peter Boncz (Vectorwise)** über “Micro Adaptivity in a Vectorized Database System” (siehe Abb. 2). Diese setzt sich die Beschleunigung der Ausführung von Anfrageplänen durch eine adaptive Wahl von Implementierungen zum Ziel. Auch **Wolfgang Lehner (TU Dresden)** griff das Thema Adaptivität auf und stellte in seinem Vortrag “Adaptive Query Processing on Prefix Trees” unter anderem den KISS-Tree, eine auf Präfixen basierende Indexstruktur, vor. Zum Abschluss der Vorträge am ersten Tag stellten **Martina Albutiu, Tobias Mühlbauer und Wolf Rödiger (TU München)** Entwicklungen zum scale-out der hybriden Hauptspeicher-Datenbank HyPer vor.



Abb. 2 Peter Boncz bei seinem Vortrag

(Foto: Silke Prestel)

Der Tag wurde durch ein Social Event im Münchener Hofbräuhaus zünftig abgerundet (siehe Abb. 1). Der Dank für dieses Event geht an unseren Sponsor Google, der diese Veranstaltung großzügig unterstützte. Der zweite Vortragstag startete mit einem Vortrag von **Rainer Gemulla (MPI Informatik)** über “Matrixfaktorisierung auf sehr großen Datenmengen als Grundlage für Empfehlungssysteme”. **Jens Dittrich (Universität des Saarlandes)** sprach in seiner Präsentation “Optimierung von Hadoop MapReduce” über Erweiterungen des auf Skalierbarkeit ausgelegten Hadoop für analytische Anfragebearbeitung. Über die Skalierbarkeit von Analysen in DB2 auf Mainframes wie dem System Z von IBM sprach **Namik Hrle (IBM)**. Insbesondere ging er dabei auf den IBM DB2 Analytics Accelerator ein, der die OLAP-Bearbeitung von DB2 beschleunigt. Mit **Johannes Meier und Oleksandr Kozachuk** waren zwei Vertreter von **EXASOL** angereist, um über “Skalierbare In-Memory-Analysen mit EXASolution” zu sprechen. Neben Vectorwise, welches den Non-Clustered TPC-H Benchmark anführt, war mit EXASOL somit auch das Unternehmen vertreten, das derzeit den Clustered TPC-H Benchmark

dominiert. Im Anschluss stellte **Ansgar Scherp (Universität Mannheim)** mit seinem Vortrag “Scalable Extraction of Implicit and Explicit Schema Information in Linked Open Data” einen Index für RDF-Datenquellen vor. Am Abschluss des Herbsttreffens ging **Peer Kröger (LMU München)** auf den Fluch der Dimensionalität im Zusammenhang mit Data Mining ein.

Wir bedanken uns an dieser Stelle bei allen Teilnehmern und insbesondere bei den Vortragenden für ihr Kommen und die zahlreichen Diskussionen.

Die Folien der Vorträge und weitere Informationen zum Herbsttreffen finden Sie auf der Webseite <http://www-db.in.tum.de/scalableanalytics/>.