



# Seminar Distributed Data Management

Conny Kühne  
Christian von der Weth  
Stephan Schosser



- P2P-Systeme als Infrastruktur
  - lose gekoppelter Rechnerverbund
  - dezentrale Organisation
  - P2P != File-Sharing
- Ziel dieses Vortrags:
  - Abgrenzung zu Client/Server (Vor- und Nachteile)
  - geeignete bzw. natürliche Anwendungsgebiete
  - Herausforderungen
- Literatur:
  - diverse Bücher...
  - L. Ji: *Computation in Peer-to-Peer Networks*



- Fragen
  - Wie werden Daten verteilt bzw. gefunden?
  - Wie werden Verbundoperationen durchgeführt?
  - Wo sind die Faktoren für eine Optimierung?
- Ziel des Vortrags:
  - Vorstellung strukturierte P2P-Systeme am Beispiel des CAN
  - Vorstellung einer Anfrageverarbeitung für relationale Daten
  - Grenzen solcher Systeme
- Literatur:
  - S. Ratnasamy, P. Francis, M. Handley, R. Karp, S. Shenker: *A Scaleable Content-Addressable Network*
  - P. Roesch, C. von der Weth, K.-U. Sattler, E. Buchmann: *Best Effort Query Processing in DHT-based P2P Systems*



- Untertitel: "Verteilte Speicherung unter Berücksichtigung unzuverlässiger Infrastruktur"
  - hohe Dynamik in P2P-Systemen (Ausfall von Knoten)
  - Knotenausfall = Datenverlust?
  - sinnvolle Lösung: Redundanz
- Ziel des Vortrags:
  - Einführung in die redundante Speicherung
  - Organisation von Replikaten
  - Vor- und Nachteile
- Literatur:
  - G. On, J. Schmitt, R. Steinmetz: *The Effectiveness of Realistic Replication Strategies on Quality of Availability for Peer-to-Peer Systems*



- Sensornetzwerke als "natürliche" P2P Systeme
  - Sensoren als autonome Knoten
  - dezentrale Koordination
- Ziel des Vortrags:
  - Einführung in Sensornetzwerke
  - Gemeinsamkeiten & Unterschiede zu P2P Systemen
  - Anfrageverarbeitung
- Literatur:
  - Y. Yao, J. Gehrke: *Query Processing in Sensor Networks*
  - A. Deshpande, C. Guestrin, S. R. Madden, J. M. Hellerstein, W. Hong: *Model-Driven Data Acquisition in Sensor Networks*



- Zwei Kategorien von Angriffen auf P2P Systeme:
  - Attacken
  - Whitewashing und Free-Riding
- Ziel dieses Vortrags
  - Aufzeigen der Ursachen
  - Erklärung möglicher Gegenmaßnahmen
- Literatur:
  - E. Sit, R. Morris: *Security Considerations for Peer-to-Peer Distributed Hash Tables*
  - M. Feldman, C. Papadimitriou, J. Chuang, I. Stoica: *Free-Riding and Whitewashing in Peer-To-Peer Systems*



- Definitionen von Vertrauen
  - ... basierend auf Kontakten zu Nachbarn
  - ... basierend auf sozialen Netzwerken
  - ... basierend auf Feedbacknachrichten
- Ziel des Vortrags
  - Vorstellung,
  - Vorteile und
  - Nachteile der unterschiedlichen Ansätze
- Literatur:
  - K. Sankaralingam, S. Sethumadhavan, J. Browne: *Distributed Pagerank for P2P Systems*
  - B. Ooi, C. Liao, K. Tan: *Managing Trust in Peer-to-Peer Systems Using Reputation-Based Techniques*
  - J. Sabater, C. Sierra: *Reputation and Social Network Analysis in Multi-Agent Systems*



- Anreize zum Verhindern von Free-Riding
  - Austausch von Leistungen (bartering)
  - Micropayments
  - Reputationssysteme
- Ziel des Vortrags
  - Vorteilen,
  - Nachteile
  - ...und geeignete Anwendungsbeispiele dieser Ansätze
- Literatur:
  - V. Vishnumurthy, S. Chandrakumar, E. Sirer: *KARMA: A Secure Economic Framework for Peer-to-Peer Resource Sharing*
  - J. Ioannidis, S. Ioannidis, A. Keromytis, V. Prevelakis: *Fileteller: Paying and getting Paid for File Storage*
  - T. Papaioannou, G. Stamoulis: *Effective Use of Reputation in Peer-To-Peer Environments*



- Warum entwickeln Programmierer unentgeltlich Open Source Software?
- Wieso beteiligen sich Millionen von Internet Benutzern an Communities wie Wikipedia oder SlashDot?
- Ziel des Vortrages:
  - Gemeinsamkeiten/Unterschiede zw. Open Source Entwicklung und Peer-to-Peer Wissensmanagement
  - Einfluss extrinsische/intrinsische Motivation
  - geeignete Anreizmechanismen für Mitarbeit an P2P-Communities
- Literatur:
  - M. Osterloh, B. Kuster, S. Rota: *Open Source Software Produktion: Ein neues Innovationsmodell?*
  - J. Lerner, J. Tirole: *The Simple Economics of Open Source*
  - Y. Benkler: *Coase's Penguin, or Linux and the Nature of the Firm*
  - G. Beenen et.al.: *Using Social Psychology to Motivate Contributions to Online Communities*
  - Cheng R., Vassileva, J.: *User Motivation and Persuasion Strategy for Peer-to-peer Communities*



- Motivation:
  - Worin besteht der Anreiz einem Ebayer ehrliches Feedback zu geben oder einen Slashdot Beitrag fair zu bewerten?
- Ziel des Vortrages:
  - Vorstellung und Vergleich von Mechanismen zur Unterstützung von ehrlichem Feedback
- Literatur:
  - N. Miller, P. Resnick, R. Zeckhauser: *Eliciting honest feedback in electronic markets*
  - N. Miller, P. Resnick, R. Zeckhauser: *Eliciting Informative Feedback: The Peer-Prediction Method*
  - R. Jurca, B. Faltings: "CONFESS". *Eliciting Honest Feedback without Independent Verification Authorities*



- P2P und Grid Systeme sind heute sehr populär
- Ziel des Vortrags
  - Herausarbeiten von
    - Gemeinsamkeiten
    - Unterschieden
    - Synergieeffekten
- Literatur:
  - K. Armin, G. Laszewski, A. Mikler: *Grid Computing for the Masses: An Overview*
  - I. Foster, A. Iamnitchi: *On Death, Taxes and the Convergence of Peer-To-Peer and Grid Computing*



- Relationale Datenbanken sind weit verbreitet
  - Alle unterstützen SQL als Anfragesprache
  - XQuery als innovative Anfragesprache
- Ziel des Vortrags
  - Beschreibung des Vorgehens bei der Abbildung von XQuery auf SQL
- Literatur:
  - T. Grust, J. Teubner: *Relational Algebra: Mother Tongue – XQuery: Fluent*



- XQuery: Anfragesprache für XML
  - Formale Definition nötig
- Ziel des Vortrags
  - Formale Definition von XQuery
    - Vorgehen
    - Beispiele
- Literatur:
  - W3C Technical Report: *XQuery 1.0 and XPath 2.0 Formal Semantics*



- XML Anfragesprachen erzeugen XML aus XML
  - Problem: Ist das erzeugte XML-Dokument valide?
- Ziel des Vortrags
  - Vorstellung von Algorithmen zur Überprüfung
    - ... generierter XML-Dateien auf Validität der Datentypen
    - ... von Schlüsseln in XML-Dateien
- Literatur:
  - T. Milo, D. Suci, V. Vianu: *Typechecking for XML Transformers*
  - A. Tozawa: *Towards Static Type Checking for XSLT*
  - B. Bouchou, M. Ferrari, M. Musicante: *Tree Automata to Verify XML Key Constraints*



- Grobgliederung
  - Abgabetermin: 2. Mai
- Ausarbeitung
  - 15 – 20 Textseiten
  - Latex (Vorlage über Webseite)
  - Abgabetermin: 17. Juni 2005
- Vortrag
  - Dauer: 30 Minuten + 15 Minuten Diskussion
  - Blockveranstaltung Mitte bis Ende Juni
  - Folien 14 vor Vortrag abzugeben
- Teilnahme an den anderen Vorträgen



- Kernthemen:
  - Markus Besthorn
  - Ioannis Papadakis
  - Haozhou Liu
  - Christina Hentschel
  - Thomas Karcher
  - Sabine Enge
  - Huiming Yu
  - Benno Schweiger
  - Maria Kopaigorenko
- XML-Themen:
  - Yingzhe Liu
  - Ningning Sun