

# Arbeitsgruppe Betriebssysteme



HEINRICH HEINE  
UNIVERSITÄT DÜSSELDORF

Prof. Dr. Michael Schöttner

## 2. Betriebssysteme

- **Projekt Plurix: Natives Betriebssystem für PC-Cluster (1996 – 2002)**
  - Ziel: Aggregation von günstigen PCs, statt teurer Super-Computer
  - Verteilter Transaktionaler Speicher (DSM)
  - An der Universität Ulm bei Prof. Peter Schulthess



### 3. Grid-System: XtreamOS

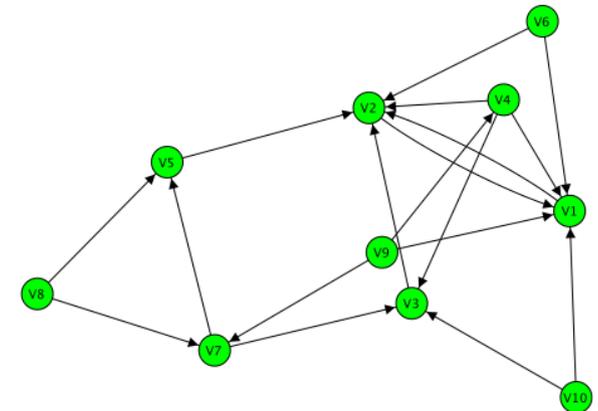
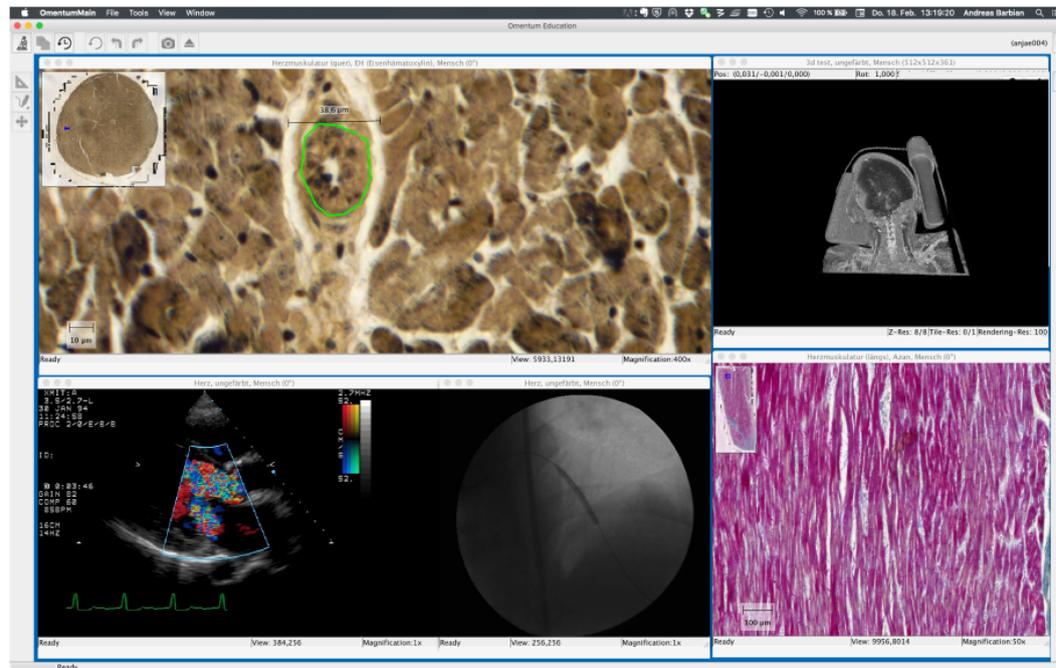
- Linux-basiertes System zur Bündelung und Verwaltung großer heterogener Grid-Infrastrukturen (2006 – 2010)



- **Data Sharing Service:**
  - Page-basierte transparente Speicherzugriffserkennung
  - Verschiedene Konsistenzmodelle (auch transaktionaler Speicher)
- **Map Reduce Framework** aufbauend auf In-Memory Service
- **Verteiltes Checkpointing und Restart für lang laufende Grid-Jobs**
  - heterogene Hard- und Software
  - Berücksichtigt auch Kommunikationskanäle

## 4. Verteilte Mikroskopie: Omentum

- Kooperation mit der Anatomie an der HHU (Prof. Filler, Dr. Barbian)
  - P2P-Anwendung für die Lehre (Studierende & Ärztefortbildung)
  - große Datenmengen (50 GB pro Datensatz)
  - Viele interaktive Benutzer



## 5. Verteilte In-Memory Systeme

### 5.1 Anwendungskontext

- Interaktive Anfragen an sehr große Graphen
  - **Sehr viele kleine Datenobjekte** (20 – 64 Byte), anspruchsvoll für die Kommunikation und die Speicherverwaltung
  - **Irreguläre Zugriffsmuster**
  - **Lesezugriffe dominieren, aber auch ständig hinzukommende Daten**



**> 1 billion  
active users**

## Lösung mithilfe von Caches

- Problem 1: **Synchronisierung Cache & Sekundärspeicher**  
→ Aufgabe des Programmierers (evt. Datenverluste)
- Problem 2: Bei **irregulären Zugriffsmustern** sind sehr große Caches notwendig → Facebook hält fast 75% aller Daten in Caches
- Problem 3: **Befüllen der sehr großen Caches nach einem Stromausfall kann Stunden dauern**

## 5.3 DXRAM-Projekt

- Seit 2012 an der HHU (Dr. F. Klein, K. Beineke, S. Nothaas)
- **Verteiltes In-Memory System (Key-Value Store)**
  - **Ziel: Alle Daten immer komplett im Arbeitsspeicher**
  - Aggregation von vielen Computern in einem Rechenzentrum mit sehr schneller Vernetzung (Infiniband,  $\geq 40$  Gbit/s,  $5\mu\text{s}$ )
  - Sehr kurze Zugriffszeiten für alle Daten
- Transparente persistente Sicherung der Daten im Hintergrund
  - Keine Synchronisierung von Cache & Sekundärspeicher
  - **Schnelle parallele Wiederherstellung** von ausgefallenen Knoten

## 6. Zusammenfassung: Kernkompetenzen

- Verteilte und parallele In-Memory Systeme
- Replikation & Konsistenz (u.a. Transaktionen)
- Schnelle Datenwiederherstellung
- (P2P-Anwendungen)