**Big Data Challenges**

Enter Moore's Law & Storage capacity increase & But

- Volume – size matters
- Velocity – data at speed
- Variety – heterogeneity
- Veracity – handle uncertainty
- Variability – rapid change of characteristics
- Voracity – strong appetite for data

**The Vs:**

- Volume
- Velocity
- Variety
- Veracity
- Variability
- Voracity

**Data Stream Management Systems**

- **Scheduling Continuous Queries**
  - Average vs Max Response Time
  - Average vs Max Slowdown
  - Priority Classes
  - Single-, Dual-, Multi-core, Cloud
- **Aggregate Continuous Queries**
  - Optimized Processing to eliminate redundant computation

**Load Shedding**

- Worst-case delay target
- Multi-tenant environment
- Differentiated classes of service
- Synergy between scheduler and load manager
- Handle complex queries and operator sharing

**Continuous Workflows**

- **Continuous Workflow Model**
  - Extend traditional workflow model with streaming semantics to enable monitoring applications
  - Design window operators
  - Facilitate waves of events
- **CONFLuEnCE**
  - CONtinuous workFLow ExeCution Engine
  - Based on Kepler workflow system

**User-Centric Data Management**

- **Quality Contracts**
  - Capture user preferences on different dimensions of quality (QoS,QoD) and use to guide system resource allocation
- **Unified Model for User Preferences**
  - Combine quantitative and qualitative user preferences into a single model to guide query result personalization
- **Annotation Management**
  - View-based (declarative) annotations

**Big Data Projects**

- **AQSIOS** – Next Generation Data Stream Management Systems
  - NSF IIS-0534531
  - PI: Chrysanthis; Co-PIs: Labrinidis, Pruhs
  - Students: Al-Moakar, Guirguis, Pham, Sharaf

- **AstroShelf** – Understanding the Universe through scalable navigation of a galaxy of annotations
  - NSF OIA-1028162
  - PI: Labrinidis; Co-PIs: Chrysanthis, Kosowsky, Newman, Marai, Wood-Vasey
  - Students: Bao, Cherinka, Gheorghiu, Luciani, Myers, Neophytou, Rexit, Sun

- **Data Management and Visualization in Petascale Turbulent Combustion Simulation**
  - NSF CBET-1250171
  - PI: Givi; Co-PIs: Chrysanthis, Labrinidis, Marai, Yilmaz
  - Students: Ansari, Maries, Nik, Piscinler

- **NSF CAREER IIS-0746696 (PI: Labrinidis)**
  - CONFluEnCE
    - CONtinuous workFLow ExeCution Engine
    - Based on Kepler workflow system