Distributed Data Management Architecture for Embedded Computing

Hans-Werner Braun[*], *Todd Hansen*[*], *Kent Lindquist*, *Bertram Ludäscher*[*], *John Orcutt*[†], *Arcot Rajasekar*[*], *Frank Vernon*[†]

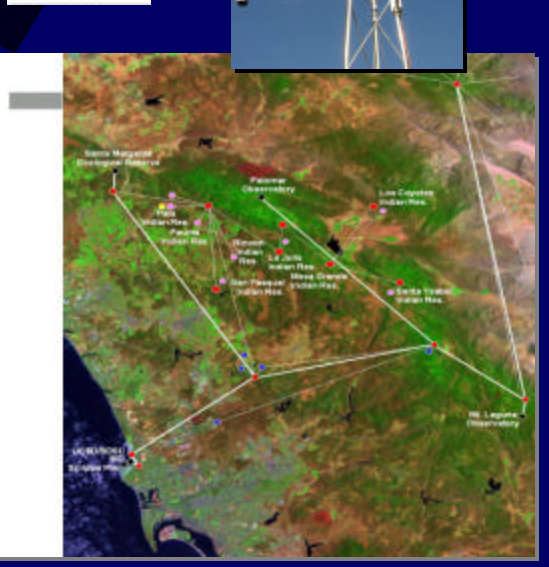
[*]San Diego Supercomputer Center [†]Scripps Institution of Oceanography

University of California, San Diego

- The Problem:
  - ***Integrated real-time management of large, distributed, heterogeneous data streams from sensor networks and embedded systems***
- … e.g.
  - integration: handling of heterogeneous data from multiple (signal) domains, interfacing with sensors and data loggers to acquire streams
  - real-time processing: monitor streams of seismic data, detect events, trigger computations and other actions
  - distribution: federate multi-domain sensor and embedded system networks for transparent access
  - persistent large-scale archival and querying of "consolidated" data packets

Real-time Observatories, Applications, and Data management Network

- Autonomous field sensors
  - Seismic, oceanic, climate, ecological, …, video, audio, …
- RT Data Acquisition:
  - ANZA Seismic Network (1981-present):13 Broadband Stations, 3 Borehole Strong Motion Arrays, 5 Infrasound Stations, 1 Bridge Monitoring System; Kyrgyz Seismic Network (1991-present): 10 Broadband Stations; IRIS PASSCAL Transportable Array (1997-Present):15 - 60 Broadband and Short Period Stations; IDA Global Seismic Network (~1990 -Present): 38 Broadband Stations
- High Performance Wireless Research Network (HPWREN)
  - High performance backbone network: 45Mbps duplex point-to-point links, backbone nodes at quality locations, network performance monitors at backbone sites; High speed access links: hard to reach areas, typically 45Mbps or 802.11radios, point-to-point or point-to-multipoint



Virtual Object Ring Buffer (VORB) Architecture

- VORB = SRB + ORB + AR
  - apply Storage Resource Broker (SRB) Data Grid Technology and Adaptive Rule (AR) based programming to federate real-time ORB systems
- Data Grid Technology (SRB)
  - collaborative access to distributed heterogeneous data, single sign-on authentication and seamless authorization,data scaling to Petabytes and 100s of millions of files, data replication, etc.