Magnetic Fusion Collaboratory:

**Goals:** understanding and innovation in magnetic fusion research by more efficient use of experimental facilities, integration of experiment + theory + modeling

Specific goals: software tools for use by researchers

Networked realtime data analysis, instantaneous communication among teams , data acquisition, simulation and visualization

Fusion: is power source of the stars

Fusion energy as source of alternative energy source

**Context/experiment**: magnetic fusion experiments operate in a pulsed mode producing plasmas of upto 10 seconds duration every 10-20 minutes, 2535 pulses a day.

For each pulse upto 10000 measurements resulting in 250-500MB of data.

Decision to change the parameters of the experiments is done in the 15 minutes intervening the initiation of the pulses.

Rapid data analysis is needed; rapid rendering and visulalization:

**Benefits:** efficient utilization, transparent access to data, standard tool set for remote data access, experimentation, facilitate multi-institution collaboration, theory + experiment+ simulation

**Relevance to CS**:

Security: automatic propagation of security credentials; authorization, authentication, encryption+ Akenti (digitally signed certificates are used for all types of security issues)

Distributed computing: Globus toolkit will be used: open grid services infrastructure (OGSI), GRAM (gloobus resource allocation manager);

Scientific visualization: Access-grid extensions, tiled walls, IDL scientific visualization language

Data acquisition: MDSPlus for data acquisition and storage

Scientific software: EFIT domain-specific analysis

Can we do this on the cloud? What are the disadvantages? What are the advantages? Any newer data-intensive computing methods usable?

Data-intensive computing is called the fourth paradigm: it is never used in this set up!

(Other three paradigms are: theory, simulation and experimentation)