Cyberinfrastructure Breakout Session

Looking 5 years ahead...

***Broad themes that emerged*:**

A) Calibration tools and skill assessment (MInt, as)
B) More examples and info for developers (Hello world)
C) Component builder idea (linked to above; View Source)
D) Need for more education on using HPC resources
E) How to get observational/experimental data into CSDMS models?
F) Can we set things up so models can use own GUIs?

1. Models and modules must be tested against data.
How should we prioritize connecting to data centers?

-calibration would be another benefit of connecting to data centers

-we one component is using data from another, and one is calibrated, turns into a feedback system

-calibrated a 'linked' system is a complex task. leave calibration data with components might make this task easier

-need to track down what work has already been done on topics like this, and invite them to CSDMS meetings.

-Sharing CSDMS output with others: could make cluster openDAP server, has connection to netCDF, use DEMs for some CSDMS stuff
-NOAA has a 'live access server' (google this)

-may be more fruitful to look at validation using this data

-some models have tuning parameters (with no physical analogue) that need to be calibrated

2. Should we invest in a 'component builder', that allows users to develop and compile code on beach?

-model repository has tons and tons of models, but few components

-how many people are planning on building a component?

-is there anything you feel you need that would help you build that component?

-a 'hello world!' type component would be useful to get started

-library of skeleton functions that could build over time, well documented looking for solutions can find something to get them started

-more interactive training material would be useful

-powerpoints with voiceover

-screencasts, net meetings (WebEx)

-should contact SMF for insights regarding getting scientists to write components

-'toy sandbox' are would allow developers to test how changes to their models would affect their behaviour in the connected environment -> similar to component builder ideas

-people could develop their own prototype components, turnaround time could decrease

-good if GUI had a shell component that would allow you to build things on beach (usable by everyone), would allow to see what's going on behind the scenes, give more control. Give fine control over GUI code

-is there a cloning feature for developing a model similar to one that already exists?

-can a busy person look at the website and within an hour get their head around the idea, and then have something out within a day or so? The initial hello world tutorial needs a very small time commitment, so busy people can understand what's going on without substantial time investment

-how to motivate busy community members to write public components?

General Discussion

Identified a need for more education on using HPC resources:

Q. CSDMS provides resources we're not used to having. now, offering machine with exceptional capabilities. not sure how to best utilize these resources. how can CSDMS offer support on HOW to use these resources (eg. code parallelization, etc.)

A. can parallelize in different ways. some talks at the meeting on using MPI and parallelization strategies. commercial applications available that parallelize matlab array operations.

-could CSDMS provide template libraries that show examples of parallel communication?
-use code that already has parallelism built in (like PetSc)

-how easy is it for people who are not a part of CSDMS to connect to databases for comparison purposes?

    -climate community has conventions for netCDF, validators exist online
    -use of standards facilitates community use

interface, IRF. Agent based models could be explored. What issues would arise? Configuration might be an issue. Agent based models are often stochastic models.

-What is being done to ensure sustainability of the CSDMS modeling approach?

  -IRF seems to be standard outside of this community
  -XML also designed to persist into the future
  -VISIT has persisted
  -CCA also has community traction and projected longevity

-more influence CDSMS, more critics. What could infrastructure do to convince the world of the validity of it's model?

-What is the flexibility of connecting components? eg. what if my coastal model needs two rivers? What if my water quality model in Delft3D has 600 inputs/outputs?

-Perhaps groups developing models could communicate with other groups within CSDMS to share knowledge about parallelizing codes, share postdocs, events and workshops to parallelize code

-Limiting step in morphodynamic modeling is transporting sediment. If going in the direction of including data for testing, should include small scale data as well (LTER program, NSED), like sediment transport

    -sediment data from 6" wide flume to field scale measurements

-making data distributed across institutions at various scales can be very challenging, need to be aware of this. Also, in terms of sediment transport even values that are measured are subject of argument

-Some models already have a user interface. by putting the code into CSDMS, this interface can be lost.

    -guiML could be used to provide a language-agnostic method of describing an interface

    -haven't been many components with gui's to date, so this issue hasn't really come up

    -but for delft3d, their gui has been engineered and provides more functionality than the standard CSDMS model would