

IDEA Center Tsunami Use-Case Kickoff Meeting Opening Remarks

This is a quick recap of some of our discussions to date, hopefully nothing new. IDEA, NGDC, NDBC have developed a clear vision for the tsunami use-case. In the big picture, the use-case is to be realized through the implementation of a concept “Enterprise Service Bus” to be referred to as “bus” in discussion. Through a bus, we hope to avoid the pitfalls of the other models of systems/data integration namely vertical (or “silo”) and star (or “spaghetti”) types of operations, both of which have merits, but generally have limitations in a loosely coupled distributed environment where collaboration is expected between independent actors that need to serve and consume data in many-to-many relationships. The use-case, however, will focus on “getting off the dime” by facilitating a one-to-one data transfer relationship, specifically from NDBC to NGDC, the prize being the resulting data schema and extensible remote procedures architecture that may be readily adapted to serve other players (CO-OPS, PTWC, outside NOAA and international) achieving an ever increasing level of sophistication in each subsequent cycle of development.

This effort is informed by an independent body of work at the IDEA center, which has demonstrated a sea level station service, enabling multiple clients. The relevant lessons learnt seem blindingly obvious in hindsight. The first lesson was that developing a web service, however useful or sophisticated it may potentially be, is completely useless unless there are operational client(s) that are regularly using (consuming) this service. In fact, taking the approach of “build it and they will come” simply did not work for us. It is only when we took the “additional step” of engaging a functioning organization (PTWC) and enabled a widely used client, namely Google Earth/Maps through a KML transformation, did the service become truly useful. The second obvious lesson was dealing with “real data” on the server side. Without going into too much detail about evolution of the sea level station schema, the take away lesson is really just incremental development or “crawl, walk, run”. This hopefully explains why we view working with NDBC and NGDC as absolutely essential pre-requisite to realizing a bus. As mentioned above, if we set ourselves a manageable target and do it right, we should be able to extend to any level of complexity in the future as we add more and more service nodes enabling active clients.

Naturally, we have encountered doubts from ourselves and others. Among the most popular ones is “Why are we/you developing a new schema/standard when there is XYZ that already does ABC?” Our evolving position is that every schema is as unique as the problem it serves. There is no single schema that can address all the problems by itself. In fact, the best schema is actually really a “hierarchy of schemas” starting out at the very top, closet to the application, being totally faithful and unique to that application, while extending some of its data types into robust standards/frameworks. For example, we may choose to define the location element in the schema per OGC standards. Another possible question is “Why are you guys doing this when agency EFG is vested with the authority to do this?” We certainly want to be aware of such overlaps within the organization (and without) and be complementary. By design, the bus promotes independence of nodes that participate in the integration. In fact, once the source publishes through a bus, it becomes loosely coupled with all its clients since it becomes the responsibility of the bus to provide a consistent interface to clients. Also, we hope that all code being developed is going to be in the public domain for anyone to use as they see fit, including hosting the bus itself in their own agency. Either way, we strive to be complementary to all similar data flows. Somewhat like a bike lane that adds versatility to the transportation system.

With the above background, seems like the ideal outcome of this meeting would be that we are able to agree to discuss a feasible scope of work based on excellent information already provided by participants, form a core implementation team with roles and responsibilities and perhaps agree on success benchmarks, relative to existing conditions, that may be monitored as we proceed in the next weeks and months.