

# 2012-04-16 DfR Technical Meeting

## Regular Attendees

- Andrew
- Bill ★
- Brad (before he sails into the sunset)
- Chris
- Dan
- Jonathan
- Danny

## General

- Call In To: [Free Conference Call HD - DuraCloud Line](#)  
Please note that we will be using the Free Conference Call HD line for this call. Information about calling into this line is available from the link above.
- ★ - Indicates who will be taking minutes

## Discussion Topic

[Planning Board](#)  
[Task Board](#)  
[DfR Planning and Estimating - 0.1](#)

Topic	Discussion Leader
SI Meeting	Jonathan
The DfR Value Proposition (Adv Grp Call)	Jonathan
Mini-Review of Iteration Planning - Key Tasks	Dan
Technology Choices for MEE - How to get there	Chris

## Actions from last meeting

[Last meeting](#)

Last Weeks Actions:

Action Item	Assignee	Status
Describe/script demonstrations to show	Dan and Jonathan	
Better conceptualize accounts for DfR	Team	Postponed
Conceptualizing/Design Synch Tooling	Dan and Bill with Team	
Better set the feasible scope for this year in automated metadata collection	Team	

### Related Information

[DfR Software Next Steps - Jan 2012](#)  
[2012-01-03 - Architecture Meeting, Temple University](#)

## Status

- Jonathan
  - Advisory Group meeting this afternoon on "what is the value proposition for DfR, given an increasingly crowded market?"
- Andrew
- Bill
- Brad
- Chris
- Dan
  - Still Working overview documentation (received some good input from Jonathan - not yet incorporated)
  - Still working on synch tool requirements
  - Still looking a SI VM (provided a copy)

- Prepared for SI meeting in DC
  - Creating cards from last weeks iteration meeting for critiques, pointing and prioritization
  - Reviewed Spring integration particularly for Apache Camel and Mule
  - Worked with the SI VM to look at integration points
  - Help Jonathan get the SI VM working in his desktop
- Danny

## Minutes

Coming up:

- Big House next week
- DfR travel tomorrow

Plans for getting cards up for iteration 2

- Dan already starting to put up cards, needs help from everyone to get cards put up
- Andrew: can put up security related cards
- Danny: want to run cards by Dan first
  - These are design deliverables
- Should plan to get all cards up by the end of the week (before BHR)
- Dan: need to determine how to break out sync tooling
- Jonathan: need to have a point person for SI integration - Dan to handle this
- Will be limiting the security work with regards to the SI integration
  - Goal is to be able to show "basic" set of data via SI UI

SI Meeting

- [Previous planning meeting](#)
- What are critical goals? Which ones match to 0.2 tasks?
- We need to "drive a spike" through the entire system
- 3 or 4 categories of discussion:
  - How will the technical integration actually work? How does our architecture mech with the SI application?
    - Includes authentication and access control
  - How will we keep everything working together? How will we collaborate? How do we keep communication open? How do we get DG support?
  - Review what each team has done so far
  - What is yet to be done in Thorny's requirements list?
- Andrew: There are several cases where we want the SI system to work in a certain way, that they don't care about
  - How do we get this work done, particularly the drupal work
  - Example: How to tie in a Shib module to the drupal system?
  - Brad: Practical question: Do we need to learn drupal or not?
  - Jonathan: Expectation is likely that DG will be paid for any work done specifically for DfR.
  - We need to be very precise about what we need to get done.
  - Should start with a small amount of contract time scheduled
  - How/Where can we best use DG as a resource?
  - We should not have to learn Drupal, would be much more efficient to use DG for that work
- Practical items for 0.2
  - Chris's work is to be able to take a file from DuraCloud, and create a basic Fedora object out of that
  - First "spike" is to be able to view objects in our Fedora via the SI interface
  - Need to determine what metadata needs to be in the object to make that object minimally visible in Islandora
  - Andrew: Concerns about getting there, we are in a position where we're blind to everything on the other side of the SI code
    - We should be comfortable spinning up the SI application
    - Brad: Should we be shooting for a shared integration build with DfR SI application
    - Dan: We may not need to go quite that far, we just need to have access to build the SI app
  - We need an expert on the SI application to talk to, to help with the learning curve. This will likely be Paul Pound.
  - Dan: Likely a strategy will be to create an AMI with the SI app installed, that we can spin up for testing
    - We'll need help from DG to help us get to the AMI built
    - This will be a separate AMI from DuraCloud
  - Andrew: The point is that we should be able to pull down the SI code, install it, and run it
  - Dan: Need to determine the first place in Islandora where we can view and download the file
  - Andrew will dial in on day 2 for discussion on authentication, others are welcome. Plan is for first thing on Thursday morning.
  - Meeting starts at 1pm tomorrow, goes through mid-afternoon on Wed

0.2 Jira Tasks

- DFR-63: Much of what is required for this was discussed in the previous discussion (what is needed from SI)
  - Andrew: It's important that we be able to operate within a security context. May want to be able to turn on/off security for development convenience.
- DFR-23: Need to outline the features of the sync tool
  - Under Dan's name, but will need discussion, work from many on the team
  - Danny/Peter will work on much of the user experience
  - Should start listing requirements and design for this work
  - Important part of the spike
- DFR-79
  - Need to think about when it makes sense to pull in some of the various technical pieces into the mix. First need to investigate and select tools.
  - Bill: Might be better to consider how we can meet user's needs

- Chris: One of these first needs is file type identification, as an example
- Dan: Want to be able to use a single execution environment to run all of the pieces
- Brad: System piece has a low-level piece for actually executing pieces, also have a need to provide a way to view/display this at the user level
  - My not need to couple these two pieces at this point in the project
  - On bottom-up should consider problems we need to solve for next step of development
  - Wait on trying to nail down how the user experience will occur
- Action: Keep DFR-79 around, but push it out, create a new, more specific jira

#### Value proposition of DfR

- We offer the ability for researchers to run processing close to their data (which minimizes cost)
- We can do preservation, archival, visualization, etc
- Preservation is the starting point for re-use
- Dan: The whole point is to promote re-use, particularly across data sets
- Other preservation services: bit-integrity checking, replication, metadata creation capabilities
- Richer access control
- Policy management - we can manage policies for executing specific actions
- Platform that enables you to build usage models beyond the file system
- Provides an on-ramp for moving data into a platform where more collaboration is possible
- When data is backed up as files only, there is no understanding of arrangement
- Access capabilities - providing the ability to create new exhibitions, APIs, endpoints, displays, etc
- Many on-ramps and many off-ramps - as well as being able to manage what's happening on the freeway
- Each piece of the system is open source and has its own community
- An end goal is to virtualize the cloud completely, so the same code can sit on any infrastructure