

Software stack consolidation



Balázs Kónya, Lund University

3rd EMI all-hands, Padova, 17-19 October 2011

EMI is partially funded by the European Commission under Grant Agreement RI-261611

- Introduction: why, what, how and by when?
- Compute area consolidation plan (15 min)
- Data area consolidation plan (15 min)
- Security area consolidation plan (15 min)
- Infrastructure area consolidation plan (15 min)
- Discussion

- ***Project Objective #1: “Simplify and organize the different middleware services implementations by delivering a streamlined, coherent.. distribution”***
 - “major issue ... having to use multiple clients and APIs”
 - “struggling with different security models”
 - “struggling with incompatible libraries”
 - “existence of several different implementations of similar services or functionality leads to unnecessary confusion”
 - “reduction in the complexity of the middleware offering is necessary”

- (pre-EMI) portfolio contain too many products
- many of those are not functioning well together
- EMI proposed solution: consolidation

“the mergers or acquisitions of many smaller companies (products) into much larger ones”

- Harmonization: implementation of certain EMI agreement
 - Fixing interoperability, compatibility problems
 - Making things “coherent, integrated, interop”
 - Develop common interfaces, libraries
- Streamlining: removing products, cut code, reduce software stack

What else is in the contract (DoW)?



- KJRA1.3 Number of reduced lines of code
 - *“proves that the actual lines of codes that have to be maintained are actually reduced during the course of the project”*
 - *“indicates code re-use and the harmonization of products”*
 - SLOCS to be retrieved from ETICS
 - **Target: > 33% (1/3) reduction over the three-year activity**
- KJRA1.4 Number of reduced released products
 - *“a numeric value that indicates (the change of) the number of different products within an EMI product release”*
 - **Target: >= 2 products per year in average over three years**

- Remove duplications, confusing overlaps
 - are there duplications?
- Streamlining
 - make EMI more “focused”, limit scope
 - remove not used/not needed products
- Reduce maintenance
 - smaller code base = less maintenance (?)
 - do we know the maintenance cost of our products?

Consolidation: How?

- Identify duplicates, keep only one of them
 - Requires adjustments of other products
- Develop new common product AND remove old ones
 - Increases code for a transitional period
- Limit the scope of EMI and remove related product(s)
 - only removed from EMI, it can leave outside EMI
- Internal simplification of a product
 - PTs write more efficient code
 - Replace "old" with "new" (in-house consolidation)

Consolidation: How?

- Replace an EMI product with an external alternative
 - Look for a smaller, better community alternative
- “Replace proprietary technologies with off-the-shelf and community supported technologies”
 - Rewrite EMI products using off-the-shelf technologies
- Cristina forgets to add some product to the release
😊

Consolidation: What?



- EMI 1 Update 8:
 - 53 Products (104 Elements)
 - Total source lines of code: 2 943 939
 - *gLite: 1 688 670, dCache: 736 245, ARC: 306 374, UNICORE: 133 367*
 - Update 4 sloc: 2 730 359
- EMI 0:
 - Less than 53 Products
 - Did not contain everything
 - Majority of products were incomplete
 - Total lines of code: 2 542 762
- Unclear REFERENCE point
 - difficult to say what was pre-EMI, or even what was EMI-0

the "sloc table"

- Emi_products_sloc.xls attached to the indico
- Lists all the EMI products with their elements
- An attempt to specify a sloc number for every product for pre-EMI, EMI 0, EMI 1
- In principle, it should be easy to fill from SA2 metrics

- Area consolidation proposals
 - end of October
- Project Technical Board discussions
 - November/December
- Very little “consolidation” is expected for EMI 2
- Real effect:
 - EMI 3 workplan: April 2012
 - EMI 3 release: February 2013

- Phase out targets (incl. investigate):
 - ~~ARC CE/GM~~, ARC CE/gridftp jobplugin, ARC CE/JURA, ~~ARC CE/update erls~~, ARC CLI/ng*, ARC CLI/libarcdata2, arc-proxy, ARC GridFTP server, ARIS, EGIIS, Resource BDII, ~~gstat validation~~, Site BDII, CEMon, CREAM cli, FTS cli, GFAL, lcg_util, lcg-info-clients, RAL-SAGA-SD, Trustmanager, ~~Util~~ Java, U. client libs, U. ucc, U. HILA, U. Registry, U. Uvos, VOMS client, ~~SCAS~~, ~~SLCS~~

Consolidation “progress”



- Most of the EMI agreements in place
 - No visible “consolidation effect” on EMI software portfolio (yet)
- A couple of small things got dropped
 - see previous slide
- “New” products are still being added to EMI 1
 - Products that missed EMI 0: *too many to list here (e.g. CREAM LRMS modules)*
 - Products that missed EMI 1: FTS, StoRM, Hydra, Pseudonymity

Consolidation "progress"

- Most of the products are growing
 - New probes, new features, new integration
 - Needs to be monitored and understood
 - *e.g. glEXEC-wn: EMI 0: 49 295 EMI 1: 208 823*
- Latest EMI 1 contains much more code and products than EMI 0
- EMI 2 will contain even more products and code
 - The real "new products" are coming
 - EMIR, STS, EMI_datalib, EMI_authlib
- As of Year 2 EMI will have very little visible "consolidation result"
 - Considerable increase of code lines
 - Several new products

For the discussion:



- is the area plan already represents an agreement of the middleware consortia/ product teams?
- which are the products (product elements) that will/may/could be phased out?
- what are the main difficulties in the consolidation of your area?
- will there be a "common EMI cli/API" in your area? what would that be? who (which PT) would develop, maintain that?
- what will be the envisioned "harmonized" API/interface/ to access the services of your area?
- How will the EMI agreements/new product development in the area (e.g. emi_Authlib, emi_Datalib, EMIR) contribute to the consolidation? Which existing product/library/module will be replaced/substituted by these new developments?