

r, ECS289F-W05, Topics in Scie

Step 1 Step 2 Step 3 Step 4



































Scientific "Workflows": Some Findings

- More <u>dataflow</u> than (business control-/) workflow
 DiscoveryNet, Kepler, SCIRun, Scitegic, Triana, Taverna, ...,
- Need for "programming eth (include)
 - Iterations over lists (foreach); filtering; functional composition; generic & higher-order operations (zip, map(f), ...)
- Need for <u>abstraction</u> and <u>nested workflows</u>
- Need for <u>data transformations</u> (WS1→DT→WS2)
- Need for rich <u>user interaction</u> & <u>workflow steering</u>:
 - pause / revise / resume
 - select & branch; e.g., web browser capability at specific steps as part of a coordinated SWF
- Need for <u>high-throughput</u> data transfers and CPU cyles: "(Data-)Grid-enabling", "streaming"
- Need for <u>persistence</u> of intermediate products and <u>provenance</u>

daescher, ECS289F-W05, Topics in Scientific Data Management

Scientific "Workflows" vs Business Workflows

- Scientific "Workflows"
 - Dataflow and data transformations
 - Data problems: volume, complexity, heterogeneity
 - Grid-aspects
 - Distributed computationDistributed data
 - User-interactions/WF steering
 - Data, tool, and analysis integration
 - → Data, tool, and analysis integration
 → Dataflow and control-flow are often *married*!
 - Business Workflows (BPEL4WS ...)
 - Task-orientation: travel reservations; credit approval; BPM; ...
 - Tasks, documents, etc. undergo modifications (e.g., fight reservation from reserved to ticketed), but modified WF objects still identifiable throughout
 - Complex control flow, complex process composition (danger of control flow/dataflow "spaghetti")
 - → Dataflow and control-flow are often *divorced*!





















