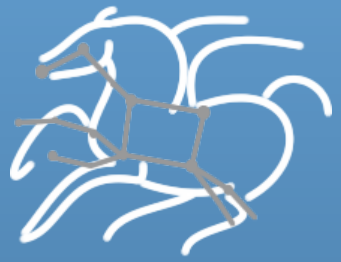


Toward Fine-Grained Online Task Characteristics Estimation in Scientific Workflows

Rafael Ferreira da Silva[1,2], Gideon Juve[2], Ewa Deelman[2], Tristan Glatard[1,3], Frédéric Desprez[4], Douglas Thain[5], Benjamín Tovar[5], Miron Livny[6]

[1]University of Lyon, CNRS, INSERM, CREATIS, Villeurbanne, France
[2]University of Southern California, Information Sciences Institute, Marina Del Rey, CA, USA
[3]McConnell Brain Imaging Centre, Montreal Neurological Institute, McGill University, Canada
[4]INRIA, University of Lyon, LIP, ENS Lyon, Lyon, France
[5]University of Notre Dame, Notre Dame, IN, USA
[6]University of Wisconsin Madison, Madison, WI, USA

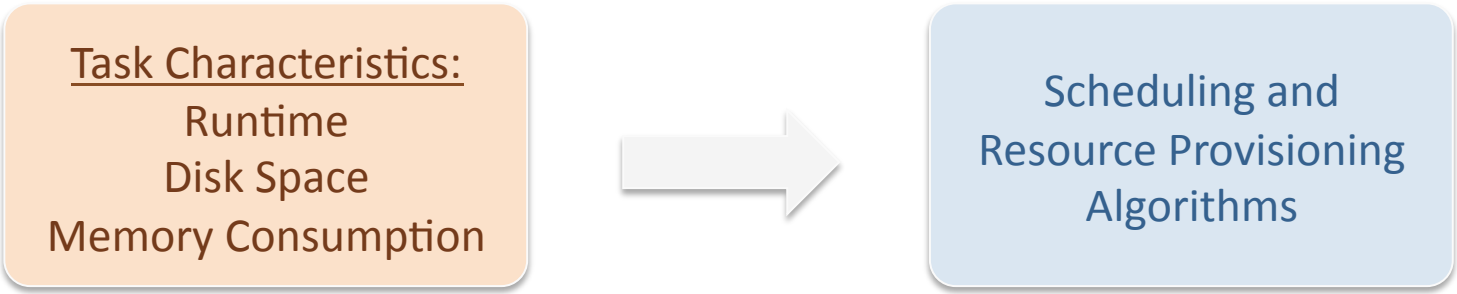
Outline

- Introduction
- Workflow Characterization
- Task Estimation Process
- Experiments and Evaluation
- Conclusions

Outline

- **Introduction**
- Workflow Characterization
- Task Estimation Process
- Experiments and Evaluation
- Conclusions

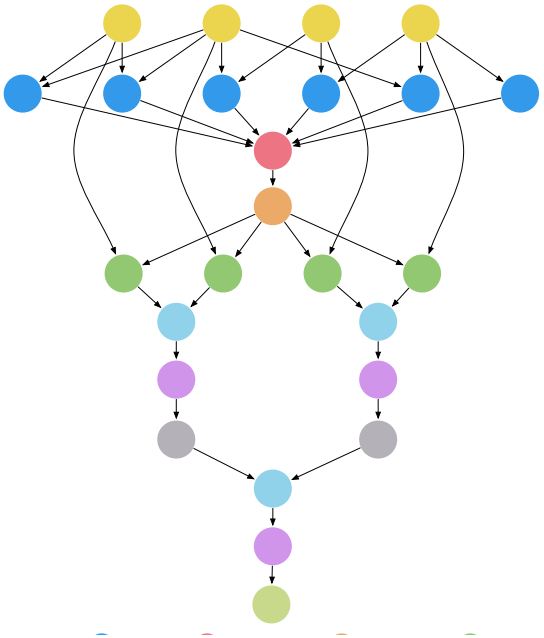
Context



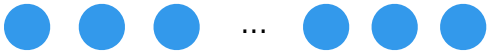
- Methods assume that accurate estimations are available
  - It is hard to compute accurate estimations in production systems
- A successful workflow execution mainly depends on how tasks are planned and executed
- We propose a method to online estimate fine-grained task characteristics

Scientific Workflows

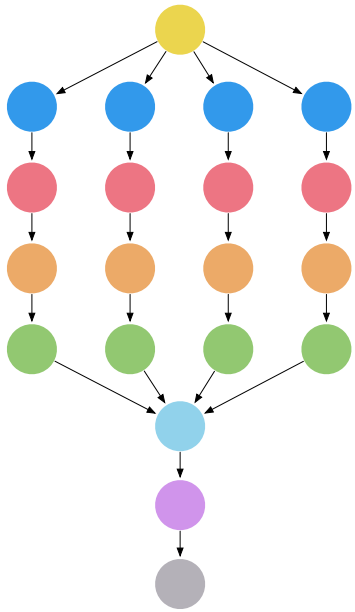
- Directed Acyclic Graph (DAG)
  - Nodes denote tasks
  - Edges denote task dependencies



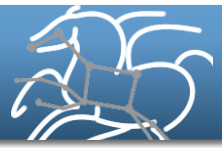
Small (20 node) Montage Workflow



Periodogram Workflow



Epigenomics Workflow



Outline

- Introduction
- **Workflow Characterization**
- Task Estimation Process
- Experiments and Evaluation
- Conclusions

Workflow Execution Profiling

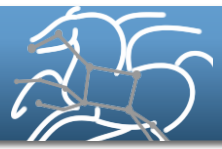
- Workflows were executed using Pegasus WMS

  and profiled using Kickstart profiling tool
  - Monitors and records fine-grained data
  - E.g. process I/O, runtime, memory usage, CPU utilization

- **3** runs of each workflow with different datasets
  - 16-core cluster

    5 Dual core MP Opteron™ Processor 250 2.4GHz / 8GB RAM

    3 Dual core MD AMD Opteron™ Processor 275 2.2 GHz / 8GB RAM

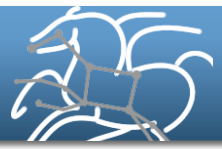


Execution Profile: Montage Workflow

Task estimation could be based on mean values

Task	Count	Runtime		I/O Read		I/O Write		Memory Peak	
		Mean (s)	Std. Dev.	Mean (MB)	Std. Dev.	Mean (MB)	Std. Dev.	Mean (MB)	Std. Dev.
mProjectPP	7965	2.59	0.69	4.24	0.19	16.20	0.80	9.96	0.40
mDiffFit	23733	1.25	0.92	24.08	5.76	1.35	1.11	5.32	0.90
mConcatFit	3	122.04	5.27	2.70	0.01	3.15	0.01	7.26	0.01
mBgModel	3	2008.08	88.50	4.14	0.04	0.27	0.00	14.41	0.01
mBackground	7965	2.14	1.68	13.67	6.78	13.05	6.44	11.75	5.78
mImgtbl	51	4.65	2.04	22.64	4.61	0.25	0.05	6.37	0.13
mAdd	51	47.69	14.03	2191.76	560.39	1574.22	383.86	21.66	3.40
mShrink	48	11.53	2.25	835.57	0.31	1.00	0.00	3.05	0.01
mJPEG	3	1.03	0.07	46.18	0.02	0.78	0.00	2.66	0.01

Task estimation based on average may lead to significant estimation errors



Automatic Workflow Characterization

- Characterize tasks based on their estimation capability
  - **Runtime, I/O write, memory peak ➔ estimated from I/O read**

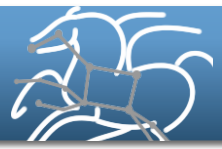
- Enforces correlation statistics to identify statistical relationships between parameters
  - High correlation values yield accurate estimations

Task	Runtime		I/O Write		Memory Peak	
	ρ	σ	ρ	σ	ρ	σ
fastqSplit	0.98	9.00	1.00	297.15	0.00	0.01
filterContams	-0.03	0.27	0.99	1.46	0.00	0.01
sol2sanger	0.21	0.41	0.90	1.49	0.00	0.01
fast2bfq	0.18	0.27	0.56	0.87	0.00	0.01
map	0.02	18.96	0.06	0.70	0.01	1.43
mapMerge	0.98	13.33	0.99	189.81	-0.36	2.15
pileup	0.99	4.73	0.17	249.78	0.87	25.70

Constant values

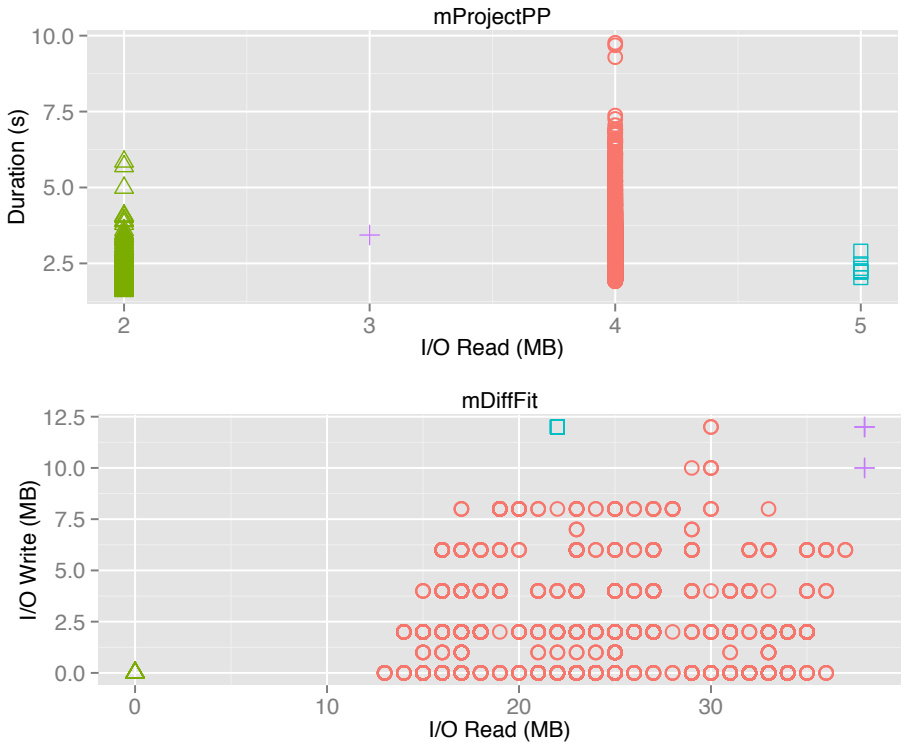
Correlated if $\rho > 0.8$

Epigenomics workflow



Density-Based Clustering

Identifies groups of high density areas where no correlation is found



Algorithm 1 DBSCAN algorithm.

inputs: D dataset, eps, $minPts$
cluster $C = 0$
for $p \in D$ **and** p is unvisited **do**
  mark p as visited
  neighborPts = regionQuery(p, eps, D)
  if neighborPts.size $<$ $minPts$ **then**
    mark p as noise
  else
    C = next cluster
    expandCluster(p, neighborPts, C, eps, $minPts$)
  end if
end for

expandCluster(p, neighborPts, C, eps, $minPts$)
add p to C
for p' $\in$ neighborPts **do**
  if p' is unvisited **then**
    mark p' as visited
    neighborPts' = regionQuery(p', eps, D)
    if neighborPts'.size $\geq$ $minPts$ **then**
      neighborPts = neighborPts $\cup$ neighborPts'
    end if
  end if
  if p' $\notin$ any cluster **then**
    add p' to C
  end if
end for

regionQuery (p, eps, D)
  return $D' \subseteq D$, where $distance(p, q) \leq eps$, $q \in D'$

Density-Based Clustering: Epigenomics workflow

- Datasets with high correlation values are not clustered

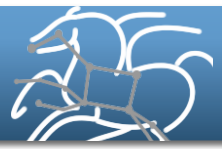
Higher correlation values

Task	Runtime			I/O Write			Memory Peak		
	c	ρ	σ	c	ρ	σ	c	ρ	σ
fastqSplit	1	0.98	9.00	1	1.00	297.15	1	0.00	0.01
filterContams	1	-0.03	0.27	1	0.99	1.46	1	0.00	0.01
	2	0.70	0.17						
sol2sanger	1	0.19	0.31	1	0.90	1.49	1	0.00	0.01
	2	0.39	0.31						
	3	0.17	0.08						
fast2bfq	1	0.12	0.21	1	0.24	0.73	1	0.00	0.01
	2	0.63	0.17	2	0.00	0.00			
map	1	-0.04	16.95	1	0.36	0.59	1	0.05	1.38
	2	0.41	14.10	2	0.37	0.55	2	0.54	0.89
mapMerge	1	0.98	13.33	1	0.99	189.81	1	0.55	1.98
							2	0.00	0.00
							3	0.00	0.00
pileup	1	0.99	4.73	1	0.17	249.78	1	0.87	25.70

Smaller standard deviation values

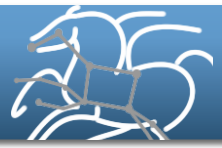
Constant values

c: clusters per task type
ρ: correlation value
σ: standard deviation



Outline

- Introduction
- Workflow Characterization
- **Task Estimation Process**
- Experiments and Evaluation
- Conclusions



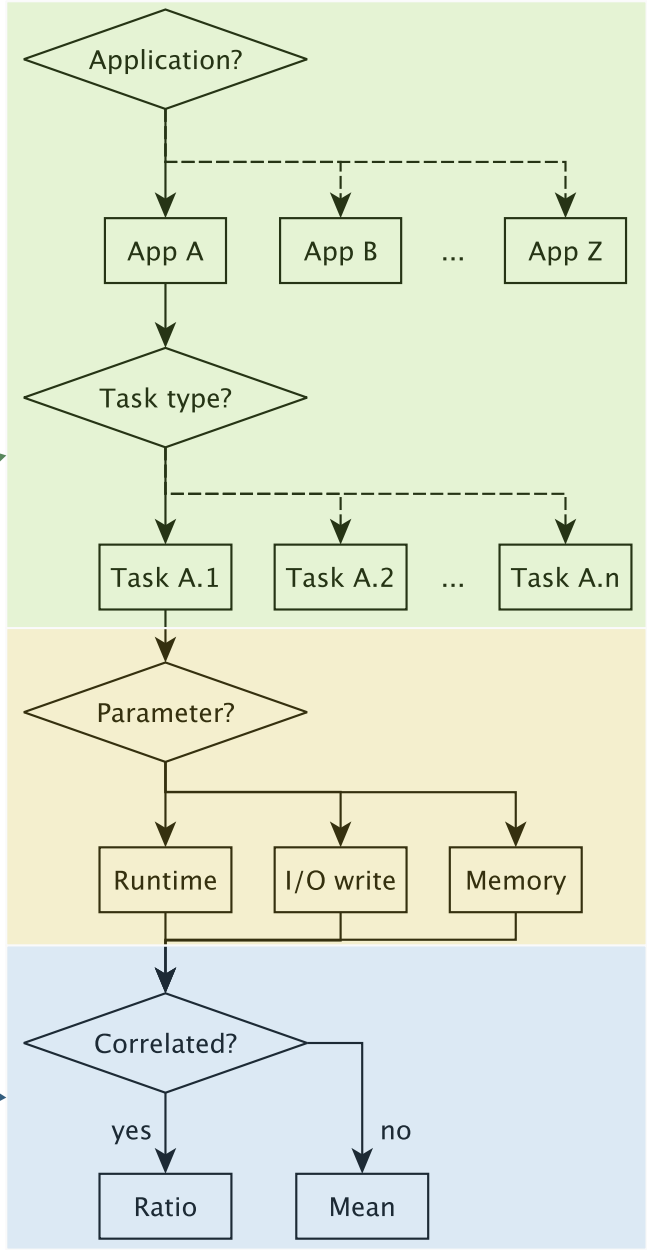
Task Estimation Process

- Based on Regression Trees
  - Built offline from historical data analyses

Tasks are classified by application, then task type

Estimation of runtime, I/O write, or memory peak

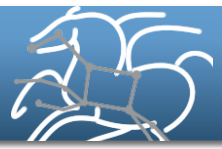
If strong correlated to the input data:
- Estimation based on the ratio parameter/input data size
- Otherwise, estimation based on the mean



Example of Estimation Rules

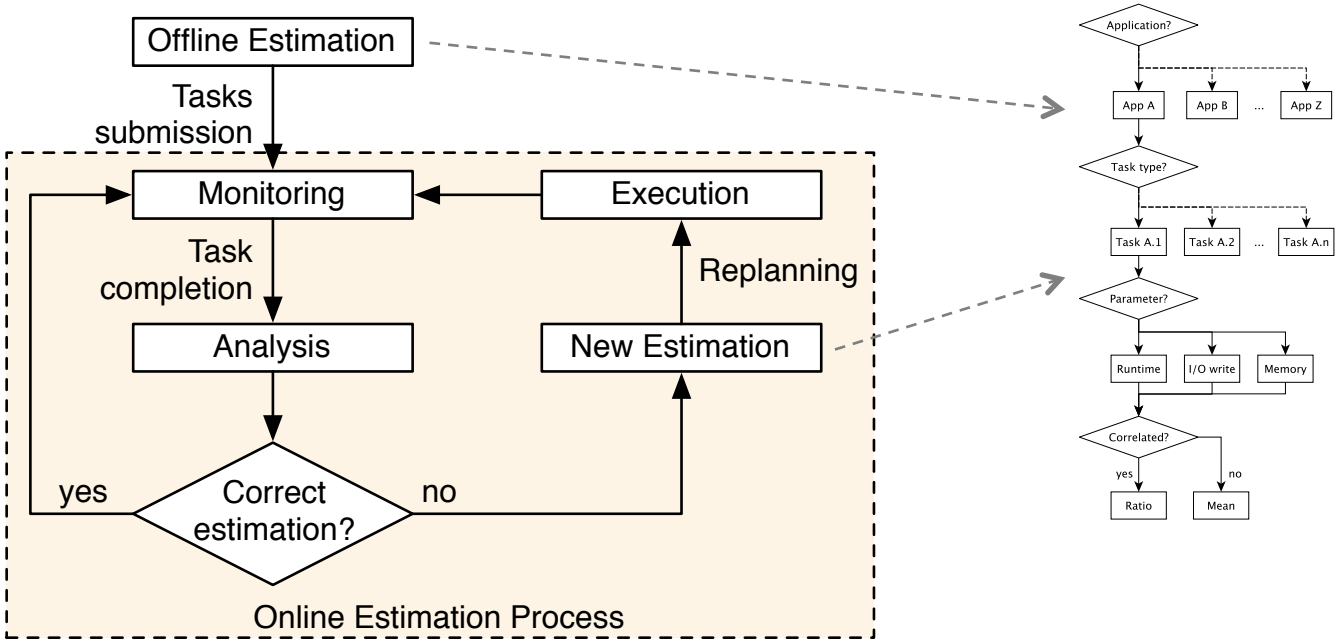
```
if workflow = 'Periodogram'
   and taskType = 'periodogram_wrapper'
   and parameter = 'write'
   and input_size ≤ 45088768 then
      return [7371489.28, mean] // mean value in bytes
end if

if workflow = 'Periodogram'
   and taskType = 'periodogram_wrapper'
   and parameter = 'write'
   and input_size > 45088768 then
      return [0.38, ratio] // ratio of output and input data
end if
```

Rules for I/O write estimation of the Periodogram workflow



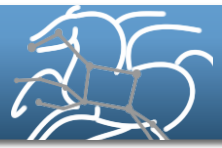
Online Estimation Process

- Based on the MAPE-K loop
  - Task executions are constantly monitored
  - Estimated values are updated, and a new prediction is done



Outline

- Introduction
- Workflow Characterization
- Task Estimation Process
- **Experiments and Evaluation**
- Conclusions



Experiment Conditions

- Trace analysis of 3 workflow applications
  - Montage
  - Epigenomics
  - Periodogram
- Leave-one-out cross-validation
  - Evaluate the accuracy of our online estimation process
  - 3 different workflow execution traces for each workflow
- Simulator
  - Simple DAG analyzer
  - Replays workflow executions

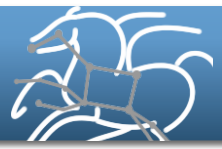
Results: Average Estimation Errors - Montage

Task	Estimation	Runtime Avg. Error (%)	I/O Write Avg.Error (%)	Memory Avg.Error (%)
mProjectPP	Offline	18.95	1.63	2.80
	Online	18.95	1.63	2.80
mDiffFit	Offline	191.02	159.46	91.07
	Online	46.52	69.14	73.72
mConcatFit	Offline	4.38	0.00	7.62
	Online	4.03	0.00	6.22
mBgModel	Offline	23.83	0.00	22.08
	Online	1.17	0.00	3.43
mBackground	Offline	65.13	102.80	104.62
	Online	44.90	1.23	1.84
mImgtbl	Offline	61.27	127.29	126.58
	Online	29.15	5.53	8.35
mAdd	Offline	9.67	113.14	110.20
	Online	9.31	3.43	9.06
mShrink	Offline	13.72	0.34	0.00
	Online	7.61	0.33	0.00
mJPEG	Offline	1.61	0.00	19.09
	Online	1.37	0.00	11.40

Online Process
Avg. Runtime Error: 18%
Avg. I/O Write Error: 9%
Avg. Memory Error: 13%

Offline Process
Avg. Runtime Error: 43%
Avg. I/O Write Error: 56%
Avg. Memory Error: 53%

Poor output data estimations leads to a chain of estimation errors in scientific workflows



Results: Average Estimation Errors - Epigenomics

Task	Estimation	Runtime Avg. Error (%)	I/O Write Avg.Error (%)	Memory Avg.Error (%)
fastqSplit	Offline	8.36	3.28	9.14
	Online	8.36	3.28	9.14
filterContams	Offline	59.31	109.81	102.83
	Online	29.13	5.35	8.15
sol2sanger	Offline	54.93	98.20	96.68
	Online	34.74	1.23	1.96
fast2bfq	Offline	27.13	128.18	99.98
	Online	17.09	15.11	10.65
map	Offline	23.62	0.00	21.07
	Online	1.39	0.00	3.33
mapMerge	Offline	53.74	93.34	1.01
	Online	10.22	9.39	1.00
pileup	Offline	6.00	4.17	49.42
	Online	5.11	3.87	19.31

Online Process
Avg. Runtime Error: 13%
Avg. I/O Write Error: 5%
Avg. Memory Error: 8%

Offline Process
Avg. Runtime Error: 29%
Avg. I/O Write Error: 57%
Avg. Memory Error: 48%

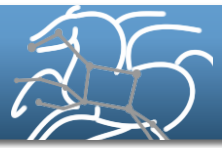
Poor output data estimations leads to a chain of estimation errors in scientific workflows

Results: Average Estimation Errors - Periodogram

Task	Estimation	Runtime Avg. Error (%)	I/O Write Avg.Error (%)	Memory Avg.Error (%)
periodogram_wrapper	Offline	45.13	16.72	1.02
	Online	45.13	16.72	1.02

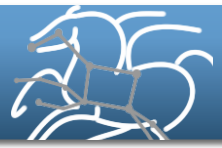
BoT: the online approach produces the same result as the offline

- **Experiment Conclusions**
  - Online strategy counterbalances the propagation of estimation errors
  - Estimation of first-level tasks have strong influence in subsequent estimations



Outline

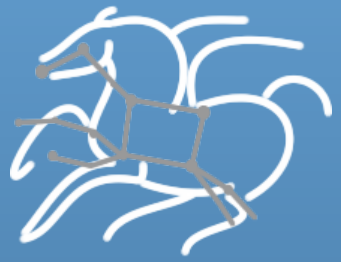
- Introduction
- Workflow Characterization
- Task Estimation Process
- Experiments and Evaluation
- **Conclusions**



Conclusions

- Summary of contributions
  - Automated method that characterizes scientific workflow executions
  - Fine-grained characterization of 3 real scientific workflows
  - An online estimation process to predict fine-grained task needs

- Future Work
  - Analysis of the impact of re-planning a workflow when using an online estimation strategy
  - Sensitivity analysis of the correlation value ρ
  - Increase the number of workflow samples
  - Compare the results with other monitoring tools



Toward Fine-Grained Online Task Characteristics Estimation in Scientific Workflows

Thank you.

Questions?

rafsilva@isi.edu

http://pegasus.isi.edu