pap276 Review Details

pap276 Review Details

Title: P*: A Model of Pilot-Abstractions
Authors: Jha, Luckow, Maddineni, Merzky, Santcroos, Weidner

Key for the below column headings: hide

Review Categories (higher is better):

mer: **OVERALL MERIT.** In your estimation, how would you rank this paper with respect to other papers that have been submitted to HPDC. If you are unfamiliar with HPDC submissions, how would you rank this paper with respect to papers that are submitted to a range of high-performance conferences of the same caliber (e.g., SC, ICS, NSDI, USENIX ATC).
[Bottom 50% of submitted papers, Top 50% of submitted papers, but not in the top 25%, Top 25% of submitted papers, but not in the top 10%, Top 10% of submitted papers, but not in the top 5%, Top 5% of submitted papers. Implicitly recommended for the best paper award]

conf: **REVIEWER FAMILIARITY**. How familiar are you with the area.
[I know nothing or almost nothing about this topic, I am somewhat familiar with this area, but I can't claim I am an expert, or, my work in this area is out of date, I am well versed in this area, but it isn't my direct area of specialty, This is my area]

Summary of reviews of pap276s1

Reviewer	mer	conf
Reviewer 1	Top 25% of submitted papers, but not in the top 10%	I am well versed in this area, but it isn't my direct area of specialty
Reviewer 2	Top 10% of submitted papers, but not in the top 5%	I am somewhat familiar with this area, but I can't claim I am an expert, or, my work in this area is out of date
Reviewer 3	Top 50% of submitted papers, but not in the top 25%	I am well versed in this area, but it isn't my direct area of specialty
Reviewer 4	Top 50% of submitted papers, but not in the top 25%	This is my area
Reviewer 5	Top 25% of submitted papers, but not in the top 10%	I am well versed in this area, but it isn't my direct area of specialty
Averages:	2.8	3.0

Commitee Comments jump

Reviewer 1 Comments

top

Does this paper address an important issue?

Does this paper address an important issue? If the issue or problem addressed were (or now is) completely understood or solved, how important would that be, in terms of either fundamental concepts, or increased understanding, or practical relevance?

Pilot Job is an important mechanism widely used nowadays to access large scale distributed infrastructures. The main idea is to dissociate the resource acquisition from the job scheduling. In this paper, authors present P*: a Model for Pilot-Jobs. To my knowledge, this is the first attempt to provide a clear definition and abstraction for the Pilot Job concept. Such work may have a large impact on several categories of audience: developers of Grid software who seek interoperability, users who require to deploy their applications across multiple infrastructures as well as teaching of Grid technologies. It can also serve as a leverage for standardisation effort.

The P*-Model model is first presented descriptively, i.e. without formalism. P* is then discussed against several existing systems which implements directly the Pilot Job concept or similar mechanisms : SAGA BigJob, Diane, Condor, Swift, which illustrate its broad applicability. Next an API is derived from the model and experiments demonstrate that the Pilot-API effectively allows interoperability of various PJ frameworks across several DCIs. Overall, the article is very consistent and coherent and provides a more in terms of understanding a concept still not widely identified as a key mechanism of Grid Computing.

Does this paper present convincing results?

Does this paper present convincing results? Do the results provide worthwhile insight into the topic addressed? Are the results likely to be widely used by others? Does the work open up new areas, present new ideas, and/or serve as a foundation for new work?

The analyze of coordination in PJ framework is very focused and not exhaustive enough in terms of frameworks investigated. Thus it's difficult to draw general conclusion. It would have been perhaps better to have an analytical model of P* which identifies the major components of the performance and compares this analysis with experimental measurements.

Overall, the experiences of Sec 5.2 and 5.3 have a feeling of "déjà vu". Being able to span an application to several sites through the use of Pilot Job is already known: It's the essence of Pilot Job. However the experiences are technically well realized. They involve several international DCIs (FutureGrid, EGI, OSG, XSEDE) managed by a large variety of Grid middleware and illustrate perfectly the point of the paper.

Is this paper sufficiently well executed?

Is this paper sufficiently well executed? Are there flaws (e.g., technical mistakes, important uncited related work, poor assumptions, insufficient scope of evaluation, unsubstantiated conclusions, poor writing) in the paper? Are the flaws fundamental or superficial? That is, are the results likely to be true despite the flaws, or do the flaws fundamentally impact the results in the paper?

Two main issues with Pilot Job are not addressed correctly in this paper : security and data.

If the Pilot Job-API is used to provide PJ framework interoperability across infrastructures then the question of security and in particular accounting and traceability should be addressed. For instance, utilities such as GLexec which have been developed so that the security model is preserved between the execution of the SU and the CU should be part of P* and appear somewhere in Figure 1.

As the name suggest, the Pilot Job concept is very computation oriented and how to make the model evolve towards a solution which would be able to take into account application data requirements is an open question. Authors actually identifies this issue as Section 6 is devoted to an abstraction for Pilot-Data. However, the Pilot-Data API, which more or less mimics the PilotJob API is not very convincing in the sense that the interaction between the 2 APIs are not discussed at all. Maybe Pilot-Data should be presented as part of the P* model and section 6 should be dedicated to the implementation of the 2 APIs to data-intense computing.

Reviewer 2 Comments top

Does this paper address an important issue?

Does this paper address an important issue?

Does this paper address an important issue? If the issue or problem addressed were (or now is) completely understood or solved, how important would that be, in terms of either fundamental concepts, or increased understanding, or practical relevance?

This paper addresses an important issue: effective distributed resource utilization. In particular the paper is concerned with Pilot-Jobs, which are used by a diverse number of applications and which enjoy the support of a number of distributed cyberinfrastructures. The problem is that there are many different implementations, which are either slightly different or highly customized to take advantage of the underlying infrastructure. If the problem is completely understood and solved, it will have a positive effect in terms of portability, reusability and extensibility. The ultimate result will be efficient utilization of distributed cyberinfrastructures that in turn would mean fast, efficient execution of jobs.

Does this paper present convincing results?

Does this paper present convincing results? Do the results provide worthwhile insight into the topic addressed? Are the results likely to be widely used by others? Does the work open up new areas, present new ideas, and/or serve as a foundation for new work?

This paper presents relevant experiments with convincing results. It initially performs single experiments, where the authors examine performance and scalability of different PJ frameworks. The authors focus on coordination in PJ frameworks, which is (according to the authors) one of the primary barriers for performance and scalability. They also test sub-cases of BigJob in order to provide better insight. They measure runtime with different numbers of cores. Finally, they characterize PJ frameworks on different DCIs, in order to provide validation of the abstractions developed. They discuss runtime, queuing time, input files' transfer time etc in detail and then using insight gained from these experiments go on to talk about PJ framework interoperability. They validate their initial assumptions, that a Pilot-API is advantageous in terms of performance, and above all successful in enabling interoperability. The work is still in an initial stage and future research is suggested by the authors. Due to its nature this research can be of use to interested researchers to build upon and refine.

Is this paper sufficiently well executed?

Is this paper sufficiently well executed? Are there flaws (e.g., technical mistakes, important uncited related work, poor assumptions, insufficient scope of evaluation, unsubstantiated conclusions, poor writing) in the paper? Are the flaws fundamental or superficial? That is, are the results likely to be true despite the flaws, or do the flaws fundamentally impact the results in the paper?

The paper is very well executed. Its structure is very good and aids to a better understanding. The authors provide sufficient background backed by a good selection of relevant citations. They motivate their work in a very persuasive manner and make clear what their contribution is. As for the technical side, the paper is technically sound, without any ill-explained assumptions/restrictions. The authors explain current state of the art Pilot-Jobs frameworks, their inner workings and then introduce their P* model of pilot-abstractions. They prove that indeed such an analytical model can be built and provide a detailed description of it, as well as the correspondent functions in other PJ frameworks. They provide an API, which they explain briefly and provide a set of detailed relevant experiments (discussed above). They do a good job in analyzing the results and tying the analysis to their motivation, showing that they succeed in what they set out to achieve. Finally, they introduce an extension of the P* model for data, but integration in a concrete model is left for future work (TROY).

Some issues to address if published:

Some of the choices and assumptions in the paper are not obvious. These could be elaborated upon.

What about other frameworks? Can they be incorporated? Why were these particular ones selected?

Is this really the first work of its kind? Are there no other related attempts?

Reviewer 3 Comments top

Does this paper address an important issue?

Does this paper address an important issue? If the issue or problem addressed were (or now is) completely understood or solved, how important would that be, in terms of either fundamental concepts, or increased understanding, or practical relevance?

This paper outlines the problems in the non-uniformity of pilot job systems, and works towards creating a generic framework enabling a unified system for submitting jobs that can utilize multiple different pilot job frameworks.

The main contributions of this work are practical, rather than fundamental concepts. Understanding how pilot frameworks differ, and how they map to each other is important---though the goal of a unified system to integrate all pilot systems is a rather elusive problem.

Does this paper present convincing results?

Does this paper present convincing results? Do the results provide worthwhile insight into the topic addressed? Are the results likely to be widely used by others? Does the work open up new areas, present new ideas, and/or serve as a foundation for new work?

One of the challenging aspects of this paper is the determining how to evaluate it. The authors present results that demonstrate that the tool they have built is working -- however, while some of the figures present performance numbers, the authors acknowledge that the evaluation should not be considered a performance analysis of their framework, as the performance is based largely on the computing infrastructure the pilot frameworks they consider are designed for.

I would say that this research does have the potential to open up new areas -- particularly the ability to build a scheduler that can match the characteristics of a job to the cluster/framework that best suits it, while being pilot job framework agnostic.

Perhaps the most interesting component of the research is the generalization of the model to data, beyond computing jobs. This portion of the research could probably benefit from being expanded.

Is this paper sufficiently well executed?

Is this paper sufficiently well executed? Are there flaws (e.g., technical mistakes, important uncited related work, poor assumptions, insufficient scope of evaluation, unsubstantiated conclusions, poor writing) in the paper? Are the flaws fundamental or superficial? That is, are the results likely to be true despite the flaws, or do the flaws fundamentally impact the results in the paper?

The paper seems to be reasonably well executed. It would however be a substantial improvement for the introduction to clearly define pilot jobs, and motivate the specific problem the authors are trying to solve. The primary motivation given is that in other frameworks that bear some similarity to pilot jobs, it has been hard to unify them.

Claims such as "Pilot-Jobs have been one of the most successful abstractions in dis- tributed computing", while arguably true, are hard to justify without being clear about what is meant by pilot jobs.

In the end, the paper seems like a pilot job system for pilot job systems. While the work seemed sound, I was unconvinced by the paper that another layer of indirection and abstraction was the right solution for connecting together pilot job systems. Perhaps another element which could strengthen the paper is a description, based on the authors analysis of pilot job systems, why pilot job systems cannot be unified in other ways.

Reviewer 4 Comments top

Does this paper address an important issue?

Does this paper address an important issue?

Does this paper address an important issue? If the issue or problem addressed were (or now is) completely understood or solved, how important would that be, in terms of either fundamental concepts, or increased understanding, or practical relevance?

This paper addresses the interoperability between Pilot-Job frameworks. By presenting a general abstraction, the authors analyze and compare the characteristics of some popular job scheduling frameworks. Based on the abstraction, a set of API is made for the coordination and communication of different frameworks. The primary intellectual contribution of this work is the development of the P* Model, the mapping of P* elements to PJ frameworks such as DIANE and Condor-G/Glide-in and the design and development of the Pilot-API – that reflects the P* elements and characteristics. The authors also validate the implementation of the Pilot-API by concurrently using multiple distinct pilot-job frameworks on distinct production distributed cyberinfrastructures. Finally, the authors apply the P* Model to Pilot-Data, but do not show how data-management can really be handled transparently across different pilot systems (especially given that many pilot job systems actually do not handle data explicitly).

The paper discusses different frameworks and tries to find common parts (to help it define P*). This work seems important in practical relevance, but it is not clear how much fundamental work there really is here.

Does this paper present convincing results?

Does this paper present convincing results? Do the results provide worthwhile insight into the topic addressed? Are the results likely to be widely used by others? Does the work open up new areas, present new ideas, and/or serve as a foundation for new work?

Unfortunately, the experimental results are not very convincing. Firstly, the overhead caused by interoperability (P*) is not provided. And the introduced overheads are critical in understanding if it is a worthwhile approach. Secondly, the demonstration of interoperability is overly simplified. For example, so of these pilot job systems were never meant to operate decoupled from their computational frameworks. For example, Swift and Swift-Coasters were meant to work together. It's not clear how one would interact with Swift-Coasters, without generating Swiftscript code (as opposed to simple job/task descriptions). The Swiftscript is very much a special instance of a way to implement some application logic. I would find it hard to believe that an application could adapt to such different requirements, where some application could be converted between some collection of jobs, to a Dagman/Condor workflow, to a Swift-Coaster workflow, and all this to be done automatically. If the complex workflow dependency management is taken out of the equation, then why bother to talk about the features of these workflow systems, such as the data from Table 2? More metrics like traffic, slowdown, and job speedup are needed to be measured to be convincing.

Is this paper sufficiently well executed?

Is this paper sufficiently well executed? Are there flaws (e.g., technical mistakes, important uncited related work, poor assumptions, insufficient scope of evaluation, unsubstantiated conclusions, poor writing) in the paper? Are the flaws fundamental or superficial? That is, are the results likely to be true despite the flaws, or do the flaws fundamentally impact the results in the paper?

The execution of the paper is ok, while there are two serious flaws with the paper: 1. The model seems unnecessary, straight forward, with not much substance. 2. Much of section 6 is repetitive

The experiments are by far the weakest part of this paper. For the experiments parts, the authors didn't explain very well how they test the interoperability between Pilot-Jobs and infrastructures exist. The results are also hard to follow, why certain experiments are done, and how to interpret the results.

Experiments are not clear enough to support the arguments. Does not say how to handle the case if the function in the framework can not be map to the P* Model. Also, I was expecting that more time would be spend on the Pilot-Data model, and how this translated to the various pilot-based systems.

Other trivial flaws: Section 4.1: "approach (ii) requires" should be approach (i) Figure 3: Physical Resource layer is confusing Section 5.1: Alamo and Sierra? Figure7: "The longer runtimes on EGI and OSG are mainly caused by the longer queuing times." is inconsistent with the figure

caused by the longer loading times" is inconsistent with the figure

Reviewer 5 Comments top

Does this paper address an important issue?

Does this paper address an important issue? If the issue or problem addressed were (or now is) completely understood or solved, how important would that be, in terms of either fundamental concepts, or increased understanding, or practical relevance?

The P* model is a model for reasoning about Pilot-Job types of systems. (The term "pilot job" is not referenced nor properly defined so the reviewer takes it to mean the union of functionality of the example systems identified in Section 1. The characteristics of the P* model are given, then the systems discussed. The paper includes performance comparisons of the different P* systems. The paper makes an important contribution in the framework it provides for reasoning about Pilot-Job systems. It is not a groundbreaking paper, however.

Does this paper present convincing results?

Does this paper present convincing results? Do the results provide worthwhile insight into the topic addressed? Are the results likely to be widely used by others? Does the work open up new areas, present new ideas, and/or serve as a foundation for new work?

The performance results are interesting and well done.

Is this paper sufficiently well executed?

Is this paper sufficiently well executed? Are there flaws (e.g., technical mistakes, important uncited related work, poor assumptions, insufficient scope of evaluation, unsubstantiated conclusions, poor writing) in the paper? Are the flaws fundamental or superficial? That is, are the results likely to be true despite the flaws, or do the flaws fundamentally impact the results in the paper?

Yes, the paper is sufficiently well executed.

Committee comments to authors: top

None