

Projet ANR- 18-CE25-0005

Pythia : décisions à l'exécution se basant sur des prédictions

Programme JCJC 2018

A IDENTIFICATION.....	2
B LIVRABLES ET JALONS.....	2
C RAPPORT D'AVANCEMENT.....	2
C.1 Objectifs initiaux du projet.....	2
C.2 Travaux effectués et résultats atteints sur la période concernée.....	3
C.3 Difficultés rencontrées et solutions.....	3
C.4 Faits et résultats marquants.....	3
C.5 Commentaires libres.....	4
D VALORISATION ET IMPACT DU PROJET DEPUIS LE DÉBUT.....	4
D.1 Publications et communications.....	4
D.2 Autres éléments de valorisation.....	4
D.3 Pôles de compétitivité (projet labellisés).....	5
D.4 Personnels recrutés en CDD (hors stagiaires).....	5
D.5 État financier.....	5

A IDENTIFICATION

Acronyme du projet	Pythia
Titre du projet	Pythia : décisions à l'exécution se basant sur des prédictions
Coordinateur du projet (société/organisme)	François Trahay (Télécom SudParis)
Date de début du projet	01/10/2018
Date de fin du projet	31/08/2022
Labels et correspondants des pôles de compétitivité (pôle, nom et courriel du corresp.)	SYSTEMATIC Paris région
Site web du projet, le cas échéant	

Rédacteur de ce rapport	
Civilité, prénom, nom	François Trahay
Téléphone	01 60 76 47 40
Courriel	francois.trahay@telecom-sudparis.eu
Date de rédaction	01/10/2020
Période faisant l'objet du rapport d'activité	01/03/2019 (T0 scientifique) 01/09/2020 (T0+18)

B LIVRABLES ET JALONS

N°	Intitulé	Nature	Date de fourniture			Partenaires (souligner le responsable)
			Prévue initiale ment	Replan ifiée	Livrée	
D1.1	A tracing library able to generate OTF2 trace files	Logiciel	T0+6		Public ¹	<u>Télécom SudParis</u>
D1.2	A trace analysis tool able to detect patterns of events	Logiciel	T0+15		Public ²	<u>Télécom SudParis</u>
D1.3	An Oracle library	Logiciel	T0+18	T0+24		<u>Télécom SudParis</u>
D2.1	A runtime system that proactively adapts the threads placement	Logiciel	T0+24			<u>Télécom SudParis</u>
D3.1	A runtime system that prefetches irregular IO requests	Logiciel	T0+30			<u>Télécom SudParis</u>
D4.1	Evaluation of Tasks 2 and 3 on applications	Rapport	T0+36			<u>Télécom SudParis</u>
D4.2	A website with a trace repository	Données	T0+36			<u>Télécom SudParis</u>

C RAPPORT D'AVANCEMENT

C.1 OBJECTIFS INITIAUX DU PROJET

Les supports d'exécution prennent des décisions critiques pour les performances d'applications parallèles. Malheureusement, ces décisions ne peuvent se baser que sur des heuristiques prenant en compte l'état actuel de l'application, et estimant son comportement probable dans le futur. Par conséquent, les supports d'exécution prennent parfois des décisions qui détériorent les performances au lieu de les améliorer.

Pythia vise à fournir aux supports d'exécution des moyens de prédire le futur de manière précise. Pour cela, Pythia se base sur la nature déterministe de la plupart des applications parallèles : généralement, un programme aura le même comportement d'une exécution à l'autre. Dans le cadre de Pythia, nous concevrons une chaîne d'outils qui analyse l'exécution d'un programme afin de fournir aux supports d'exécution des indications lors d'exécutions futures du même programme. Grâce à des indications, un support d'exécution pourrait prendre des décisions en se basant à la fois sur l'état actuel de l'application, et sur le comportement futur du programme.

¹ EZTrace is available as open-source here : <https://gitlab.com/eztrace/eztrace>

² EasyTraceAnalyzer is available as open-source here :

https://parallel-and-distributed-systems.gitlab.io/easytraceanalyzer/eta_web/

C.2 TRAVAUX EFFECTUÉS ET RÉSULTATS ATTEINTS SUR LA PÉRIODE CONCERNÉE

La première partie du projet ANR Pythia a consisté à concevoir l'infrastructure logicielle sur laquelle va reposer l'ensemble du projet. Le travail effectué s'est donc concentré sur la tâche 1.

Mise à jour d'EZTrace

Le logiciel de collecte de trace EZTrace³ a été étendu afin de pouvoir générer des traces dans le format standard OTF2. La base de code d'EZTrace étant vieillissante, cela a nécessité un fort effort de développement qui a abouti à EZTrace 2.0, qui est actuellement en phase de test avant une sortie « officielle » dans les prochaines semaines.

Conception d'un outil d'analyse de trace

La conception et l'implémentation de l'outil d'analyse de trace EasyTraceAnalyzer⁴ ont été repensées afin de permettre une plus grande généricité et pour en améliorer les performances. Le résultat est un outil permettant d'analyser des traces dans des formats variés (y compris le format OTF2), facilement extensible. Outre les analyses actuellement implémentées, l'outil est conçu pour permettre le développement de nouvelles analyses dans le futur.

Conception d'un oracle

Parmi les analyses implémentées dans EasyTraceAnalyzer, une attention particulière a été portée sur la détection de la structure d'une trace. Cette analyse permet, à partir d'un flux d'événements de l'application de repérer les motifs qui se répètent. Cette analyse servira de base à l'oracle : en repérant la structure du programme et en la comparant à la structure de la trace en entrée, l'oracle devrait pouvoir prédire le comportement futur de l'application. Cette analyse a nécessité une longue étude bibliographique par le doctorant. Les travaux antérieurs sur ce thème étant issus de domaines de recherche variés, l'étude bibliographique s'est avérée plus complexe que prévue.

Grâce à l'analyse développée, nous avons montré qu'EasyTraceAnalyzer peut traiter les événements de l'application en quelques centaines de nanosecondes, ce qui rend l'analyse utilisable pour la prédiction dans le cadre d'un support d'exécution.

Prévisions de travaux pour la prochaine période

La prochaine étape du projet consistera à finaliser l'oracle : à partir de la détection de la structure du programme, il va s'agir d'explorer l'arbre des événements futurs probables. Les étapes suivantes consisteront à évaluer l'utilisation de l'oracle pour améliorer les décisions d'un support d'exécution.

C.3 DIFFICULTÉS RENCONTRÉES ET SOLUTIONS

L'étude de l'état de l'art sur la détection de structure dans des traces s'est avérée plus complexe et longue que prévue. Le confinement et les mesures sanitaires à mettre en place pour l'enseignement à la rentrée 2020 ont également engendré une surcharge de travail qui a légèrement ralenti le projet. Le développement de l'oracle a ainsi pris un peu de retard. Ce retard ne devrait pas fondamentalement changer le calendrier du projet.

C.4 FAITS ET RÉSULTATS MARQUANTS

- Développement d'EZTrace 2.0 : une refonte du code source a permis de simplifier sa maintenance et de rendre le logiciel plus portable. Le logiciel est actuellement en phase de test et nous prévoyons une sortie « officielle » dans les prochaines semaines

³ Disponible en ligne ici: <https://gitlab.com/eztrace/eztrace>

⁴ Disponible en ligne ici: https://parallel-and-distributed-systems.gitlab.io/easytraceanalyzer/eta_web/

- Développement de l'outil d'analyse de trace EasyTraceAnalyzer : un outil extensible permettant l'implémentation d'analyses variés sur des traces dans des formats divers.

C.5 COMMENTAIRES LIBRES

Commentaires du coordinateur

Le projet se déroule conformément au planning prévu. Le confinement et les mesures sanitaires à mettre en place pour l'enseignement à la rentrée 2020 ont engendré une surcharge de travail qui a légèrement ralenti le projet, mais ce retard ne devrait pas fondamentalement changer le calendrier du projet.

Question(s) posée(s) à l'ANR

D VALORISATION ET IMPACT DU PROJET DEPUIS LE DÉBUT

D.1 PUBLICATIONS ET COMMUNICATIONS

Liste des publications monopartenaires (impliquant un seul partenaire)		
International	Revue à comité de lecture	1 M. S. M. Bouksiaa et al., "Using Differential Execution Analysis to Identify Thread Interference", in IEEE Transactions on Parallel and Distributed Systems, vol. 30, no. 12, pp. 2866-2878, 1 Dec. 2019, doi: 10.1109/TPDS.2019.2927481.
	Ouvrages ou chapitres d'ouvrage	
	Communications (conférence)	
France	Revue à comité de lecture	
	Ouvrages ou chapitres d'ouvrage	
	Communications (conférence)	1 A. Colin, F. Trahay, "Factorisation de traces d'exécutions de programmes pour l'analyse et la prédiction", in Compas 2020
Actions de diffusion	Articles de vulgarisation	
	Conférences de vulgarisation	
	Autres	

D.2 AUTRES ÉLÉMENTS DE VALORISATION

Liste des éléments. Préciser les titres, années et commentaires	
Brevets internationaux obtenus	1 2
Brevet internationaux en cours d'obtention	1 2
Brevets nationaux obtenus	1 2
Brevet nationaux en cours d'obtention	1 2

Licences d'exploitation (obtention / cession)	1	
	2	
Créations d'entreprises ou essaimage	1	
	2	
Nouveaux projets collaboratifs	1	
	2	
Colloques scientifiques	1	
	2	
Autres (préciser)	1	EZTrace 2.0, Logiciel libre, 2020
	2	EasyTraceAnalyzer, Logiciel libre, 2020

D.3 PÔLES DE COMPÉTITIVITÉ (PROJET LABELLISÉS)

Collaboration du projet avec le(s) pôle(s) ayant labellisé

--

Activités financées par le complément de pôle (laboratoires publics uniquement)

Montant du complément accordé par l'ANR (pour chaque labo public)	0 €
--	-----

Type d'action menée	Détails	Dépenses complément de pôle*
Actions contribuant à la réflexion stratégique et à la programmation scientifique du pôle		0 €
Actions de communication scientifique et publique bénéficiant à la notoriété du pôle		0 €
Développement de la recherche partenariale (recherche de partenaires, frais de gestion du partenariat, ingénierie de projets,...)		0 €
Valorisation de la recherche et transfert vers le monde industriel		0 €

* Estimation des dépenses imputées sur le complément de financement accordé au titre de la labellisation par un pôle de compétitivité, partenaires publics seulement.

D.4 PERSONNELS RECRUTÉS EN CDD (HORS STAGIAIRES)

Identification				Avant le recrutement sur le projet			Recrutement sur le projet			
Nom et prénom	Sexe H/F	Adresse email (1)	Date des dernières nouvelles	Dernier diplôme obtenu au moment du recrutement	Lieu d'études (France, UE, hors UE)	Expérience prof. antérieure (ans)	Partenaire ayant embauché la personne	Poste dans le projet (2)	Date de recrutement	Durée missions (mois) (3)
COLIN Alexis	H	alexis_colin@telecom-sudparis.eu	Récentes	Master	France	3 ans (ingénieur)	Télécom SudParis	Doctorant	01/05/2019	36

D.5 ÉTAT FINANCIER

Nom du partenaire	Crédits consommés (en %)	Commentaire éventuel
Télécom SudParis	40 %	