

SPIREC

**Supervision et prédiction multi-niveaux pour infrastructures
géo-distribuées et hétérogènes dans le continuum Cloud-Edge-IoT**

Mario Südholt (IMT Atlantique)

Réunion de lancement PEPR Cloud, 15 nov. 2023

IMT Atlantique, Nantes

Partenaires

NGN - UVSQ

- Nadjib Aitsaadi
- Dana Marinca
- Emile Abdel Ahad

CEA

- Aurelien Mayoue

STACK - IMT Atlantique

- Aris Koutsiamanis
- Kandaraj Piamrat
- Mario Südholt

IPSL - CNRS

- Karim Ramage

DIEGO - TSP

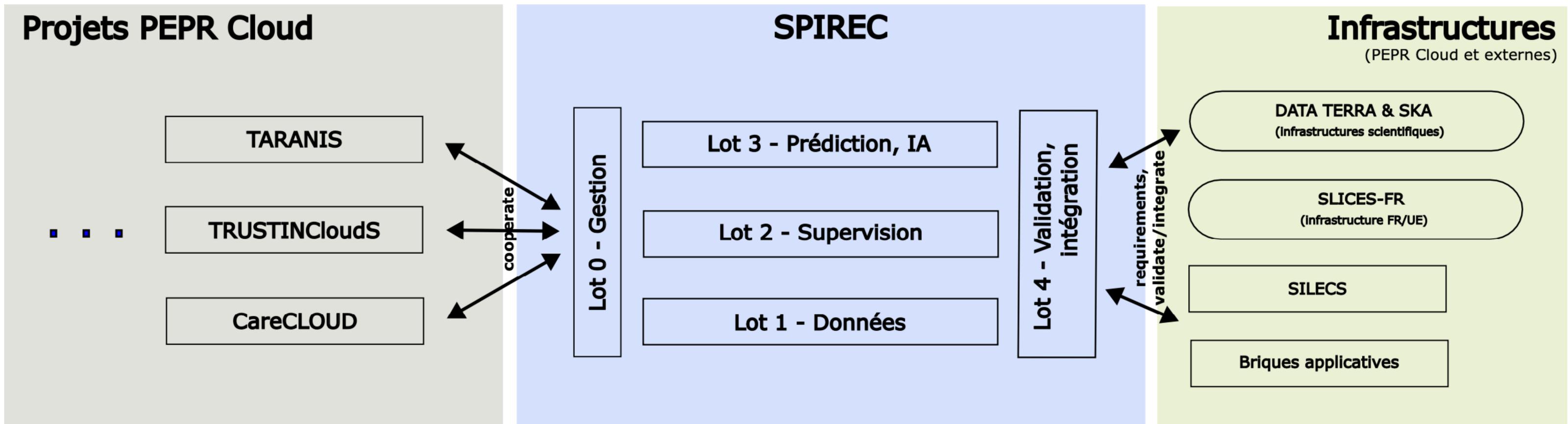
- Walid Gaaloul
- Mohamed Sellami

RESIST - Inria

- Isabelle Chrisment
- Abdelkader Lahmadi



Architecture globale



Enjeux scientifiques ciblés

Globalement :

Supervision et prédiction multi-niveaux pour infrastructures **géo-distribuées et hétérogènes** dans le continuum **Cloud-Edge-IoT**

Supervision horizontale et verticale sous contraintes fortes de ressources

Détection d'anomalies et prédiction de l'utilisation de ressources

Développement et utilisation de techniques d'**apprentissage automatique reparti** pour le CEI

Besoins, applications, validation/vérification : **Slices-FR, Data-Terra (CNRS)**

Enjeux 2

Supervision

- représentation de **données massives collectées en flux** dans le CEI
- architecture fondée sur **orchestrateurs** interagissant avec un réseau de **sondes légères réparties**
- nouvelles techniques de **réduction de dimensions**

Anomalies, prédiction

- **apprentissage automatique, autonome et incréments** pour classification
- application à des **besoins énergétiques** et de **sécurité/protection de données**

Apprentissage repart

- **hétérogénéité** des données et support de calcul
- apprentissage **hiérarchique**
- apprentissage **par renforcement** : autonomie...

Lots

1. Collecte et modélisation de **données** pour la **supervision** et la détection d'**anomalies**
2. **Surveillance** pour larges infrastructures et applications dans des environnements hétérogènes du **continuum CEI**
3. **Apprentissage** automatique multicritères répartie pour la **prédiction d'utilisation de ressources**
4. **Applications**, validation, implémentation et **intégration** (SILECS-FR...)

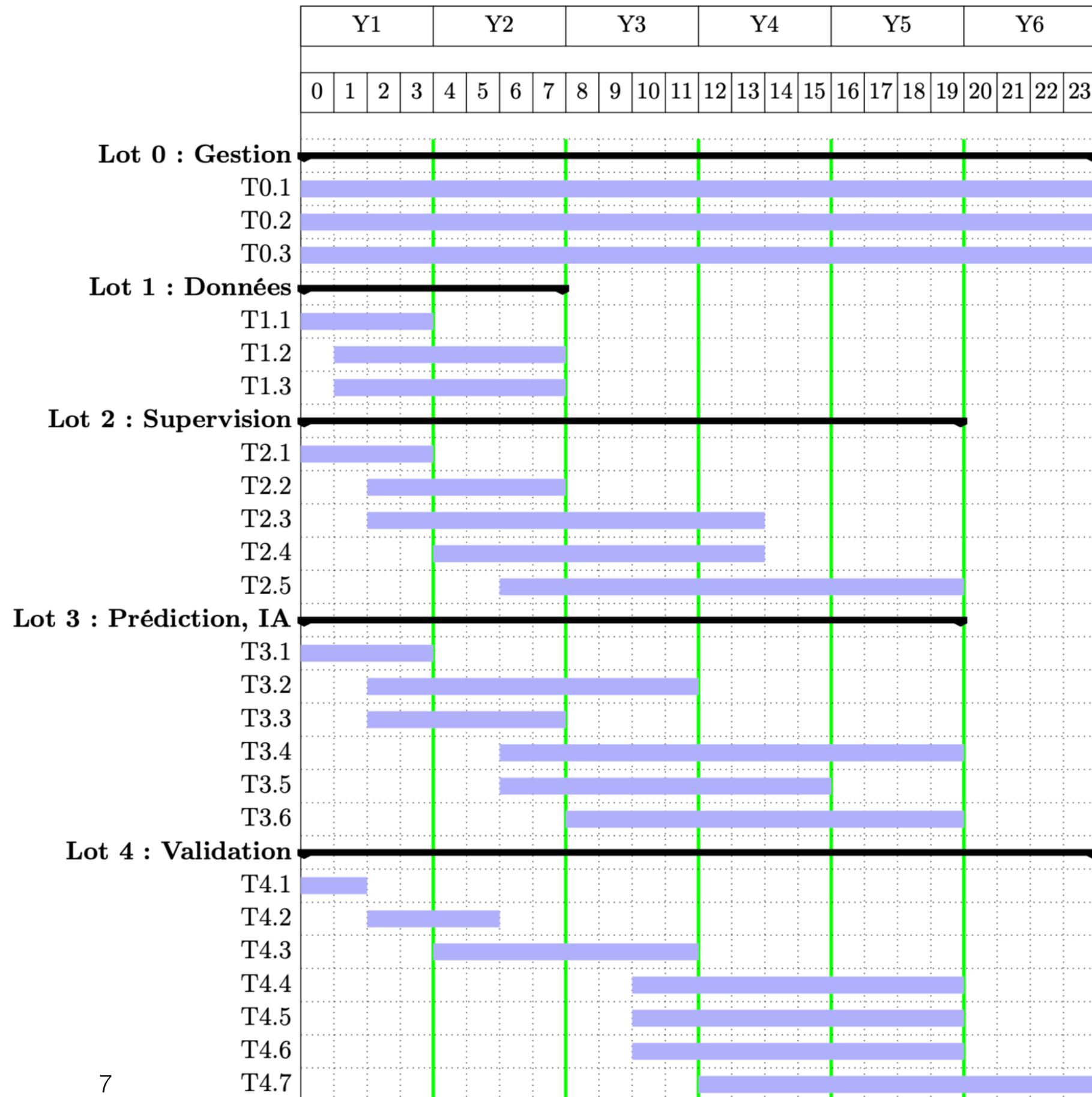
Gantt

5 ans

Périodes

- M1-6 : peu d'activité
- M7-A5 : activités principales
- A6 : peu d'activites

Problème : agencement avec personnels présents



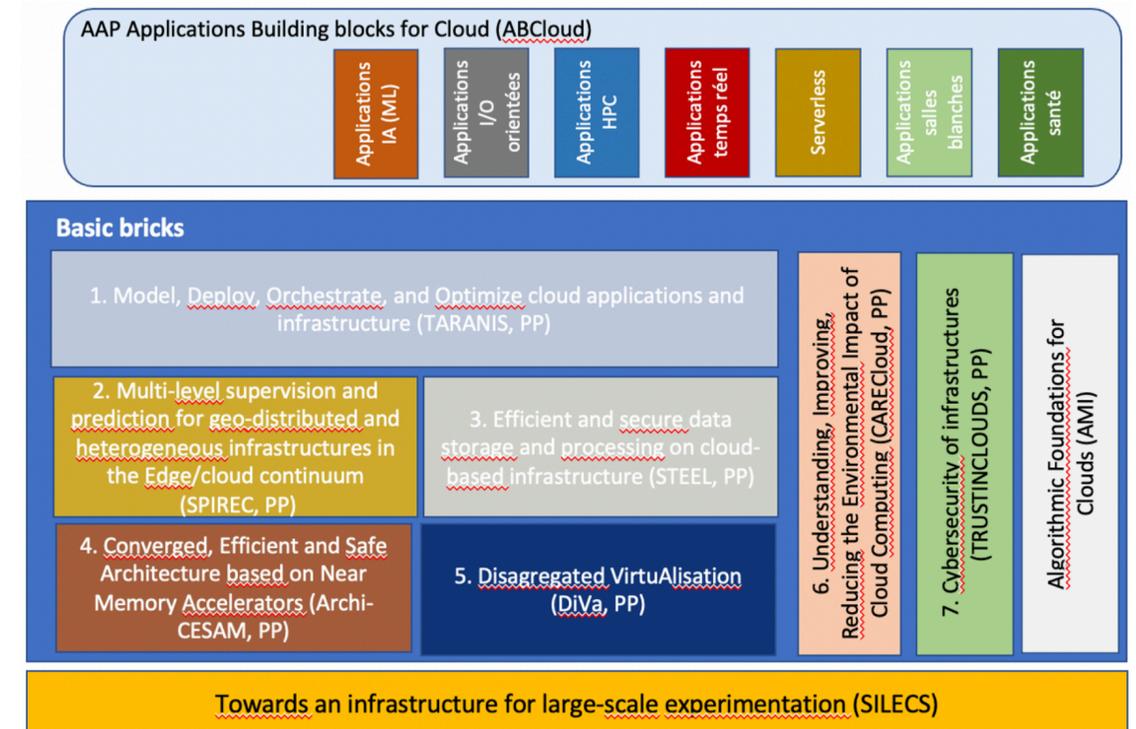
PEPR Cloud et autres PEPR

Coopérations inter-projets identifiées :

- **TARANIS** (gestion des infrastructures et applications)
- **TRUSTINCloudS** (sécurité)
- **CARECloud** (efficacité énergétique)
- ...

À concrétiser à court terme

Contacts avec d'autres PEPR dans le sillon du PEPR Cloud



Implication non permanents

Répartition relativement équilibrée entre lots / enjeux

Un peu plus de moyens sur application, validation, implémentation

	Lot 1 : Collecte et modélisation de données pour la supervision et détection d'anomalies			Lot 2 : Surveillance pour larges infrastructures et applications...			Lot 3 : Intelligence artificielle multicritères répartie pour la prédiction d'utilisation de ressources			Lot 4 : Applications, validation, implémentation et intégration		
	Thèse (nombre)	Postdoc (année)	CDD (années)	Thèse (nombre)	Postdoc (année)	CDD (années)	Thèse (nombre)	Postdoc (année)	CDD (années)	Thèse (nombre)	Postdoc (année)	CDD (années)
CEA					1		1	1				1
CNRS			1			0,5						2
IMTA	0,5			0,5			1	0,5			1	1
INRIA	0,5			1	1		0,5				1	
TSP	0,5	0,5		1			0,5				1	
UVSQ	0,5	0,5	0,5				0,5	0,5	1			1,5
TOTAL	2	1	1,5	2,5	2	0,5	3,5	2	1,5		3	5,5

KPI

KPI	Cible
Nombre de thèses soutenues	Au moins 10
Nombre de composants logiciels et outils développés	Au moins 6
Nombre de composants logiciels et outils mature intégrés dans SILECS-FR	Au moins 1
Nombre de publications scientifiques avec comité de lecture	Au minimum 20
Nombre de plateformes utilisées pour valider les outils développés dans SPIREC	Au moins 2
Nombre de scénarios applicatifs validés grâce aux outils développés dans le projet SPIREC	Au moins 10
Gestion d'infrastructures et d'applications hétérogènes	Prise en charge de deux scénarios au moins avec un type d'infrastructure et des matériels différents à chaque niveau du continuum CEI
Capacité à superviser efficacement les appareils IoT de faible puissance.	Faible surcharge (consommation d'énergie, trafic réseau) de la supervision

Finances

Budget : 2,8 M€

- réparti essentiellement équitablement entre partenaires
- STACK un peu plus pour certaines (peu de !) tâches admin
- CNRS moins car implication moindre, essentiellement au niveau application
- « **provision** » **pour maturation logicielle/intégration SLICES finales**
(inclut actuellement dans budget STACK)

**SPIREC : très peu de support financier dédié à l'administration
=> production de livrables, rapports etc. essentiellement répartie**

Pilotage

Comité de pilotage

- mise en œuvre effective et efficace du projet
- responsables des lots et porteur

Comité scientifique

- 5-7 scientifiques renommés externes
(recherche en cours)

Réunions

- lots : mensuellement
- CT : bimestriellement
- CS : annuellement

Avancement

Mise en place SPIREC

- réunion lancement : 20 oct. 23
- responsables lots déterminés
- signature conventions financières avec l'ANR en cours
- installation comités SPIREC imminente
- installation Web, espaces partagé, git... en cours

Début de travaux SPIREC

- des lots, des tâches : lancement en cours par les responsables respectifs
- premiers travaux (état-de-l'art...) débutés
- **premiers recrutements printemps 2024**

Merci pour votre attention !

Des questions ?

Lot 1 : Collecte et modélisation de données pour la supervision et la détection d'anomalies

Porteur	TSP
Verrous	Inadaptation des modèles de données actuels pour la supervision au continuum CEI et les piles logicielles multi-niveaux
Contributeurs	CNRS, IMTA, INRIA, TSP, UVSQ
Objectifs	Définition d'un modèle de données hiérarchisées pour la supervision et la détection d'anomalies dans le continuum CEI
Tâches	<ul style="list-style-type: none">- T1.1 [CNRS, M0-M12] : Modélisation de données massives dans le continuum CEI et les piles logicielles multi-niveaux- T1.2 [TSP, M6-M24] : Collecte de données distribuée sous contraintes de qualité de service et à fréquence configurables- T1.3 [UVSQ, M6-M24] : Supervision de ressources dans le continuum CEI et observabilité des fonctions réseaux à définition logicielle

Lot 2 : Surveillance pour larges infrastructures et applications dans des environnements hétérogènes du continuum CEI

Porteur	INRIA
Verrous	<ul style="list-style-type: none"> - Les techniques de surveillance actuelles ne couvrent pas le continuum, les piles logicielles et le réseau à définition logicielle - Les techniques d'apprentissages réparties ne sont pas adaptées à la supervision et la détection d'anomalies dans des environnements hétérogènes
Contributeurs	CEA, CNRS, IMTA, INRIA, TSP
Objectifs	<ul style="list-style-type: none"> - Développement de techniques de surveillance permettant de couvrir tout le continuum, les multiples niveaux d'infrastructures et de services ainsi que le réseau à définition logicielle - Définir d'une architecture de composants logiciels pour la supervision et la détection d'anomalies dans le continuum CEI
Tâches	<ul style="list-style-type: none"> - T2.1 [CNRS, M0-M12] : État-de-l'art de surveillance dans le Cloud, l'Edge et l'Internet des Objets - T2.2 [CEA, M6-M24] : Adaptation de la supervision et la détection des anomalies dans le continuum CEI - T2.3 [IMTA, M12-42] : Développement d'algorithmes d'apprentissage pour la supervision et la détection d'anomalies dans le continuum CEI - T2.4 [TSP, M12-42] : Développement d'algorithmes de traitement des données et de placement de la charge de travail, de réduction de la quantité de données transmises et d'optimisation l'utilisation de la batterie - T2.5 [INRIA, M18-60] : Infrastructure optimisée pour la supervision de larges infrastructures et applications

Lot 3 : Intelligence artificielle multicritères répartie pour la prédiction d'utilisation de ressources

Porteur	CEA
Verrous	<ul style="list-style-type: none"> - Les techniques d'apprentissage réparties actuelles ne sont pas appropriées à la du continuum CEI - Les techniques de prédiction actuelles ne permettent pas de prendre en compte l'hétérogénéité du continuum CEI
Contributeurs	CEA, IMTA, INRIA, TSP, UVSQ
Objectifs	<ul style="list-style-type: none"> - Définir un modèle pour l'apprentissage réparti pour infrastructures hétérogènes multi-niveaux - Définir des techniques pour la prédiction d'utilisation de ressources - Définir d'une architecture de composants logiciels pour la prédiction d'utilisation de ressources

Tâches	<ul style="list-style-type: none"> - T3.1 [INRIA, M0-M12] : État-de-l'art de la prédiction d'utilisation de ressources et de l'apprentissage automatique dans le continuum CEI - T3.2 [CEA, M6-M36] : Modèle de prédiction d'utilisation des ressources dans le continuum CEI - T3.3 [IMTA, M6-M24] : Modèle d'apprentissage réparti pour le continuum CEI - T3.4 [UVSQ, M18-M48] : Prédiction de ressources et mitigation automatique - T3.5 [TSP, M18-M36] : Adaptation et renégociation d'un modèle déclaratif et sémantique du SLO. - T3.6 [IMTA, M24-M60] : Infrastructure optimisée pour l'apprentissage réparti et la prédiction d'utilisation de ressources
--------	---

Lot 4 : Applications, validation, implémentation et intégration

Porteur	CNRS
Contributeurs	Tous les partenaires
Objectifs	<ul style="list-style-type: none">- Fournir des besoins d'applications réelles- Assurer la correction et l'efficacité des solutions fournies- Montrer l'applicabilité des solutions- Fournir des composants et des outils logiciels avec intégration partielle dans SLICES-FR

- **T4.1 [CNRS, M0-M6]** : Caractéristiques et besoins des infrastructures DATA-Terra et SKA pour la supervision, la prédiction et la mitigation automatique
- **T4.2 [TSP, M6-M18]** : Caractéristiques et besoins issus de SLICES-FR et des briques applicatives pour la supervision, la prédiction et la mitigation automatique
- **T4.3 [CNRS, M12-M36]** : Implémentation et validation à base de tests de composants d'infrastructure et d'outils pour les infrastructures Data-Terra et SKA
- **T4.4 [UVSQ, M30-M60]** : Déploiement et validation d'infrastructure logicielle réseaux (NaaS)
- **T4.5 [INRIA, M30-M60]** : Implémentation et validation de composants logiciels pour la supervision et la détection d'anomalies
- **T4.6 [CEA, M30-M60]** : Implémentation et validation de composants logiciels pour la prédiction de l'utilisation de ressources
- **T4.7 [IMTA, M36-M72]** : Intégration et validation de composants d'infrastructure et d'outils pour SLICES-FR

Implication permanents

	Lot 0	Lot 1	Lot 2	Lot 3	Lot 4	Total
CEA	1	0	4	6	3	14
CNRS	1	6	1	0	8	16
IMTA	5	3	8	8	6	30
INRIA	1	3	3	4	3	14
TSP	1	6	8	6	5	26
UVSQ	1	6	0	8	5	20
TOTAL	10	24	24	32	30	120