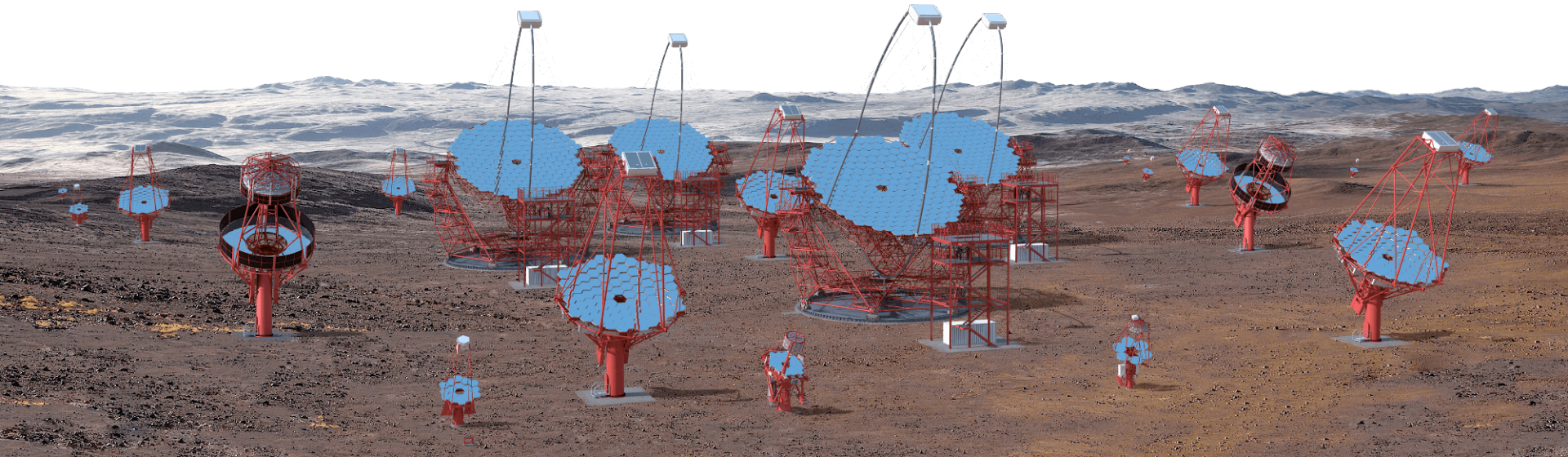




Intégration de la provenance dans CTADIRAC





Context



- Next Generation Gamma-Ray Astronomy (succeeding H.E.S.S., MAGIC and VERITAS)
- Two arrays of 99 (South) and 19 (North) Cherenkov telescopes (4, 12 et 24 m in diameter)
- Observatory open to the Astronomy community
- Timeline:



Pre-Construction phase:

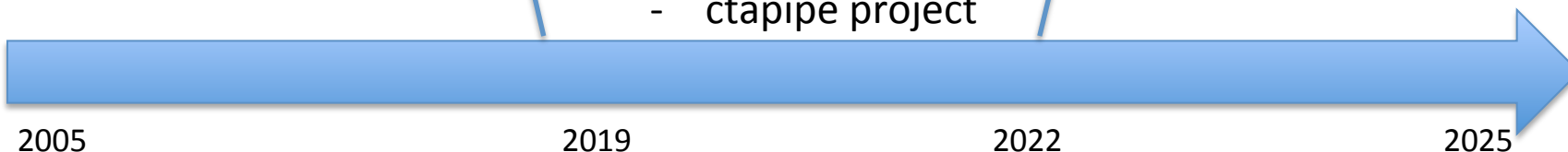
- Simulations

Pre & Production phase:

- Data model definition
- Software development
 - Prototype OPUS
 - ctape project

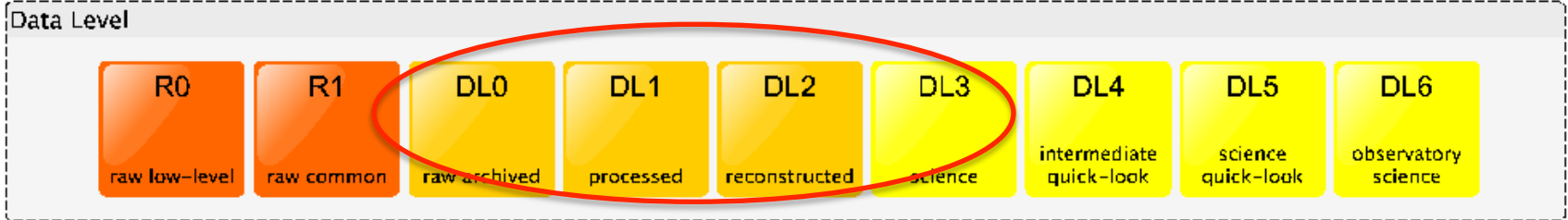
Operation phase:

- Acquisitions
- High level data available

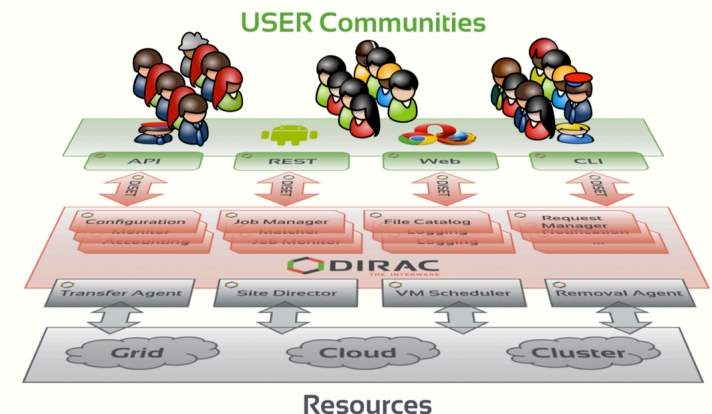




Pipeline DL0-DL3



- Software : Le pipeline utilise la bibliothèque **ctapipe** qui offre un ensemble d'outils permettant de calibrer, reconstruire, etc.
- Gestion des calculs distribués (simulations et traitement des données) : on utilise l'interware **DIRAC** (Distributed Infrastructure with Remote Agent Control).





ctapipe



- Ctapipe : Bibliothèque d'outils écrits en Python
- Outil **Provenance** (Karl Kosack) qui permet de **capturer** les informations de provenance et de les stocker dans un dictionnaire Python, enregistré à la fin de l'exécution de chaque outil, dans un fichier au format JSON.



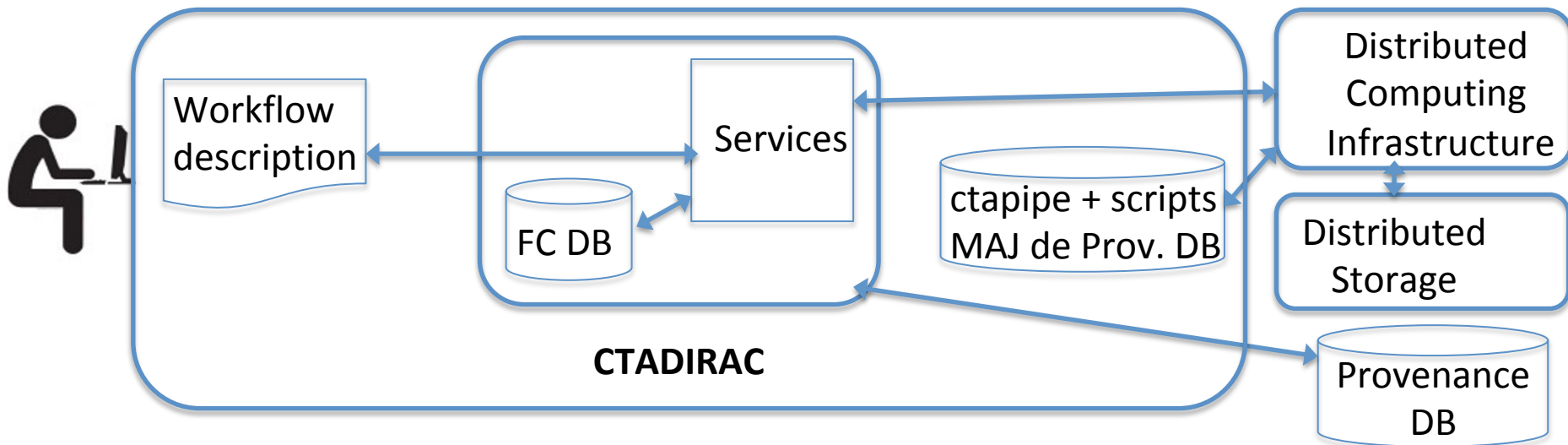
```
[{'activity_name': 'ctapipe-display-muons',  
  'activity_uuid': '93fc2206-59de-4852-868d-ec045ec25d1d',  
  'config': {'MuonDisplayerTool': {'events': 'proton_20deg_180deg_run22__cta-prod3-demo-2147m-LaPalma-baseline.simtel.gz'}},  
  'duration_min': 0.01649999999998819,  
  'input': [{'role': 'dl0.sub.evt',  
            'url': '/Users/bourgeat/Documents/CTA/Provenance/ctasoft/ctapipe/tests/proton_20deg_180deg_run22__cta-prod3-demo-2147m-LaPalma-baseline.simtel.gz'}],  
  'output': [{'role': 'dl1.tel.evt.muon',  
             'url': '/Users/bourgeat/Documents/CTA/Provenance/ctasoft/ctapipe/tests/muons.hdf5'}],  
  'start': {'time_utc': '2019-05-07T06:28:44.605'},  
  'status': 'completed',  
  'stop': {'time_utc': '2019-05-07T06:28:45.595'},  
  'system': {'arguments': ['/Users/bourgeat/anaconda3/lib/python3.6/site-packages/ipykernel_launcher.py',  
                           '-f',  
                           '/Users/bourgeat/Library/Jupyter/runtime/kernel-d793d0e4-abd3-4989-a71c-19b68194ea9e.json'],  
            'ctapipe_resources_version': '0.2.15',  
            'ctapipe_svc_path': None,  
            'ctapipe_version': '0.6.1',
```




CTADIRAC



- Prototype pour la gestion des workflow CTA basé sur DIRAC (Luisa Arrabito – LUPM).
- Ecrit en Python
- Fonctionne en mode client/serveur
- **Service de gestion de la Provenance intégré dans CTADIRAC**
- Base de données Provenance extérieure à CTADIRAC





Développements côté serveur



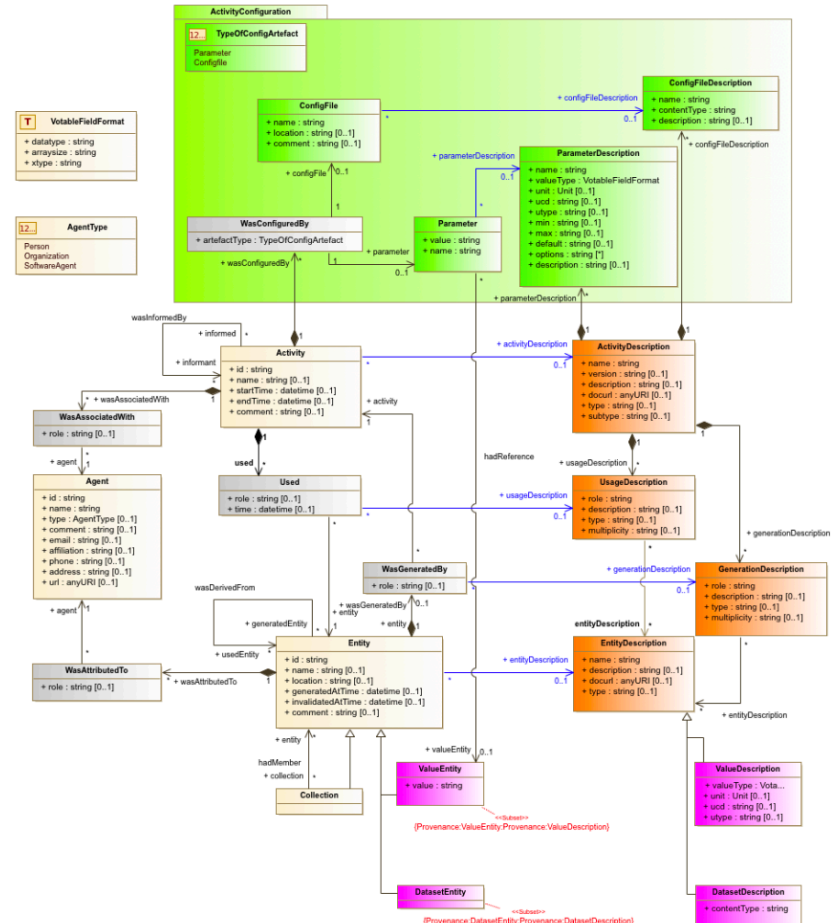
- Définition d'un nouveau service dans DataManagement qui écoute les requêtes venant du client : **ProvenanceManagerHandler.py** (Cf lancement sur machine virtuelle)
- Module de description de la base de données qui utilise le module sqlalchemy : **ProvenanceDB.py**. Description des tables et des méthodes d'accès (lecture, mise à jour) définies sur le serveur (Cf code sur machine virtuelle)



Base de données ProvenanceDB



- BD PostgreSQL
- Extérieure à CTADIRAC
- Cf module ProvenanceDB.py sous Pycharm





Développements côté client



- Script qui récupère le json produit par le tool de ctapipe qui contient le dictionnaire et lance l'ajout de la provenance : **cta-prod-add-prov.py**
- Script qui met à jour la partie descriptions (dans le futur à partir d'un fichier JSON) : : **cta-prod-init-prov.py**
- Programme qui définit les méthodes offertes au Client pour demander la mise à jour : **ProvClient.py**
- Programme qui définit les objets de Provenance qu'il peut manipuler : **ProvBase.py**



Côté utilisateur



- Programme Python qui permet de lancer dans DIRAC le workflow
- Programme Python qui décrit les différentes étapes du workflow



DEMO



- Avant utilisation :
 - Certificat GRID-FR demandé et installé
 - Inscription à la VO CTA
- BD Provenance vide au CC
- Sur machine virtuelle avec client CTADIRAC
 - dirac-proxy-init
 - Initialisation de la base avec la partie descriptions
- BD Provenance remplie des descriptions
- Sur machine virtuelle avec client CTADIRAC
 - Lancement d'un job
- Suivi sur <https://ccdcta-web.in2p3.fr/>
- BD Provenance remplie des informations de provenance du job



Je vous remercie pour votre attention