Considerazioni su DataCloud (e alcune info di stato)

Davide Salomoni

C3SN, 19/4/2023

Principi *generali* per l'integrazione delle risorse hardware

- Risorse pledged Tier-2 restano accessibili via Grid
- Risorse pledged Tier-2 finanziate da PNRR saranno accessibili via Grid ai "detentori" delle pledged (ma integrate nel datalake per data management & compute, vedi dopo)
- Risorse non pledged finanziate da PNRR saranno integrate in Cloud ed accessibili a tutti gli utenti di ICSC
- Idem per le risorse Terabit: integrazione in Cloud e accesso *agli utenti Terabit* (=> utenti di HPC-BD-AI, formalmente questa è l'infrastruttura dell'INFN; plausibilmente, in qualche modo *integrati anche in ICSC*). Vorrei avere la vostra opinione su questo.

Principi generali per l'integrazione delle risorse umane

- WP1 sta lavorando ad un calendario per gli incontri con tutti i Tier-2 **e** con i siti dove è stato allocato personale. Il calendario sarà disponibile entro maggio 2023, con l'intenzione di terminare tutti gli incontri entro l'anno.
  - Si può velocizzare? Forse sì, ma come abbiamo visto in occasione della visita a Legnaro, serve tempo per capire le situazioni e le persone (e ci sono molti altri impegni in cantiere). Vedi anche dopo.
  - In parallelo, stiamo considerando se incontrare di persona i neo-assunti.
- È in preparazione (ETA: fine aprile) un file di mapping complessivo tra che copra tutti le assunzioni TDET PNRR, da discutere con PMB DataCloud e C3SN, che includa:

Persona TDET	Tipologia (Tecnico, Tecnologo)	Progetto PNRR di riferimento	WP DataCloud di riferimento	Sede

Alcune sfide (1)

- Stiamo assumendo persone per collocarle all'interno di un'organizzazione distribuita (DataCloud), senza che DataCloud sia una vera organizzazione riconosciuta formalmente dall'ente (vedi dopo ad esempio per una "non conformità" legata a questo rilevata dall'audit interno di INFN Cloud). Questo, come avevamo ampiamente previsto, crea diverse sfide.
- Alcune persone non hanno (a giudicare *almeno* dai CV che ho potuto vedere) competenze specifiche nelle tematiche di calcolo. In alcuni casi potrebbe anche esserci il dubbio che siano veramente *interessati* al calcolo.
- Dovremo gestire anche il caso di persone che rinunciano. In questi casi diventa dunque necessario scorrere la graduatoria. Io non ho informazioni su quando questo avverrà, sulle procedure previste e sulla loro attuazione.
- Fino ad ora, tutte le informazioni relative alle assunzioni sono state comunicate in modo strettamente partizionato sulla base dei progetti PNRR di afferenza, suggeriti dalle commissioni di concorso (e questo sulla base di comunicazioni forse con i direttori locali, certamente non con me). Ad esempio, a me è stato detto chi sono i tecnologi assunti a valere sul progetto DARE perché io sono referente DARE. Io non so chi sono i tecnologi di ICSC, e tanto meno i tecnici.

Alcune sfide (2)

- Poiché DataCloud deve fare da ombrello generale di integrazione per il calcolo dei progetti PNRR ICSC, Terabit, Ecosister, Itineris, credo sia indispensabile che ci sia un processo chiaro, comunicato e rapido per quanto riguarda assunzioni, CV, scorrimenti graduatorie, etc., anche condiviso con DataCloud.
- Un punto essenziale è che vogliamo che queste persone diventino parte integrante di DataCloud e che non “spariscano” nei meandri dei tanti impegni e di possibili spinte centrifughe, magari provenienti dalle loro stesse sezioni.
- Senza arrivare ad una formalizzazione che comunque non è prevista dall’ente, credo sia indispensabile che ogni sede a cui siano state assegnate risorse umane sul calcolo PNRR designi un «referente locale DataCloud», che dovrebbe avere il compito di fare da «liaison» (e anche un po’ da «verificatore») tra il personale neo-assunto in quella sede e DataCloud in generale.
  - Ricordiamo che le persone sono formalmente assegnate a una sede, non a DataCloud (che formalmente non è niente).
  - Vorrei avere la vostra opinione su questi punti.

Audit interno di DataCloud (24/1/2023)

- Rilevate 4 “Non Conformità” (relativamente ad esempio a regolamenti nazionali o europei) e 8 “Azioni di Miglioramento”.
- Delle NC, quella essenziale è la numero 1 (NC01-INFN-DataCloud-22-23), che ha come base il Regolamento UE 2016/679 Art. 24, 25 (accountability nel GDPR).

Dalla relazione introduttiva e descrittiva della Struttura INFN-Data Cloud e le sue evoluzioni, risulta informalmente ben documentata la strutturazione e le responsabilità. Non esiste ancora però, un adeguato processo formale che costituisca un atto fondante dell'organizzazione e che definisca il contesto, gli stakeholder e gli altri aspetti necessari, così da poter comprovare all'esterno l'affidabilità e la capacità di essere garanzia di applicazione delle misure normative. Alla luce di tali misure, quali ad esempio quelle già dettate dall'Agenzia di Cybersecurity e della sua strategia nazionale, che modificano pesantemente le cosiddette "misure minime adeguate" presto in vigore e necessarie per una Cloud qualificata, e alla luce della Direttiva NIS2 e di tutte le normative a corredo che differenziano La Cloud e i Data Center locali, si rende necessario e strategico il coinvolgimento del management nel formulare l'ufficialità dell'organizzazione. Inoltre è da considerare che tale Cloud Nazionale ha al suo interno una sezione certificata (EPIC) che dovrà oltretutto rispettare ancor più le norme ISO e le sue necessarie verifiche. E infine che tale Cloud asservirà anche a parti esterne pubbliche e private.
Azione Correttiva: AC01-INFN-DataCloud-22-23 Predisporre atti formali costitutivi approvati a livello nazionale, necessari secondo le normative vigenti per dimostrare l'alta affidabilità dell'organizzazione, considerando il contesto, gli stakeholder anche esterni, la catena di responsabilità. A tali atti è necessario che vengano aggiunti i documenti, processi e procedure peraltro già predisposti, rivisti e revisionati, in modo da soddisfare i requisiti che nel complesso tutto il Data Cloud dovrà avere.

Policy di firma

- Caso di Torino (ma vale per molte altre sedi): «Essendo Torino un Tier-2 per ALICE, le attività del tecnologo [assunto a Torino su ICSC/Spoke 0] possono essere fatte valere per ALICE?» = può fare parte della collaborazione / firmare gli articoli di ALICE?
- Per me la sostanza è che se una persona è assegnata (al 100%) ad es. allo Spoke 0 (infrastruttura → DataCloud), la persona **deve** lavorare sulle attività dello spoke infrastruttura, che in questo caso non sono "di esperimento". Potrebbero essere attività di "supporto a un Tier-2 nel contesto del WG x di DataCloud" e magari quel Tier-2 supporta un solo esperimento (es. ALICE), ma il contesto di attività di spoke è il punto dirimente.
  - => essendo i TDET PNRR al 100%, in analogia a quanto abbiamo sempre fatto in casi simili nei progetti europei la risposta alla domanda di Torino a mio giudizio è «no, non può firmare per Alice», perché Alice non è in-scope con Spoke 0.
  - Questo 1) per non violare un principio base di rendicontazione e 2) anche per evitare che una persona, assegnata ad es. a spoke 0 proprio per svolgere attività di spoke 0, poi in realtà svolga attività che non sono in spoke 0 (e analogamente per qualunque altro spoke).
  - Vorrei avere la vostra opinione su questo.

Supporto agli spoke tematici di ICSC

(vale anche per altri progetti PNRR, prendo ICSC solo come esempio)

- Dobbiamo definire il modello di supporto legato alle richieste che verranno dagli spoke tematici 1-10 di ICSC.
- In concreto (ne ho parlato un po' oggi con Claudio, ma il processo va formalizzato, e per tutti i progetti, non solo per ICSC):
  - Gli spoke dovrebbero indicare – in modo standardizzato, ad es. sulla base di un portafoglio noto di componenti / soluzioni – i loro requirement a *qualcuno* (non so chi, ma certamente deve includere anche Spoke 0)
    - Parliamo qui dei flagship projects, non di attività gestite via ISCRA
  - Un *Resource Allocation Committee* di progetto valuta le richieste e approva, rifiuta, rimodula, etc.
  - Per richieste che richiedano risorse / soluzioni su INFN Cloud, dovremo avere un sistema di supporto = ticketing system, supporto di primo/secondo livello, implementazione, etc. Questo vuole dire probabilmente che dovremo anche avere un meccanismo dinamico per l'allocazione di risorse umane ai vari WP di DataCloud.
  - È essenziale continuare a seguire e anche direttamente *stimolare* queste discussioni nei progetti PNRR, per evitare di dover reagire poi all'improvviso.

Richieste “interne”

- Ci possono essere spoke che chiedono risorse / soluzioni per casi d’uso tipicamente INFN. In questo caso, assumendo che le risorse hardware PNRR non siano ancora arrivate / non siano ancora disponibili:
  1. Dobbiamo seguire correttamente tutte le procedure formali dei progetti in cui siamo coinvolti, senza shortcut. Questo è fondamentale ad esempio per evitare accuse di privilegiare l’INFN o problemi poi di rendicontazioni.
  2. Tuttavia, abbiamo riservato su DataCloud un pool di risorse per usi interni INFN. La mia proposta è che, in attesa che quanto detto sopra si concretizzi in quanto a Resource Allocation Committee etc, venga fatta internamente una richiesta a DataCloud, seguendo i canali attuali. Poi noi cerchiamo di soddisfarla (richieste piccole ovviamente). Queste risorse non le potremo rendicontare su PNRR. Quando avremo in funzione la procedura ufficiale (e il RAC avrà approvato una eventuale richiesta «INFN» per un progetto), faremo diventare le risorse già allocate «in-kind» come risorse ufficiali PNRR e dunque rendicontabili (dal momento dell’approvazione del RAC).
- Vorrei avere la vostra opinione su questo.

Servizi Cloud disponibili a INFN / ICSC / altri

Availability of INFN Cloud Services and Resources for ICSC

31 March 2023

INFN Cloud Hardware Resources that can be made available to ICSC

At the date of this document, INFN Cloud can provide the following resources to ICSC:

- 2000 vCPU
- 200 TB net disk space
- A marginal number of GPUs (NVIDIA V100 and A100)

INFN Cloud Services that can be made available to ICSC

At the date of this document, INFN Cloud can provide the following services to ICSC:

- Unified federated authentication via OpenID-Connect (OIDC) tokens, X.509 digital certificates, username/password
- Cloud storage, offered via an S3 interface
- Virtual machines, with different sizes and operating systems (1/2/4 vCPU, Ubuntu/CentOS)
- Jupyter Notebook as a Service, connected to persistent storage space
- Private Container Image Registry
- Automated container deployment directly via Dockerfile or Docker-compose files

- Kubernetes as a Service, with a configurable number of k8s nodes
- Sync-and-share as a Service ("Dropbox-like"), with a choice of either the Owncloud or Nextcloud backends, possibly associated to an encrypted filesystem.
- HTCondor-based Batch System as a Service, with a configurable number of worker nodes
- Elasticsearch & Kibana as a Service
- Spark & Grafana & Jupyter as a Service, deployed on top of a Kubernetes cluster, with a configurable number of k8s nodes
- Automated security scans on the deployed services
- A PaaS-level interface to external Cloud infrastructures such as OpenStack, Amazon Web Services, Google Compute Cloud, Microsoft Azure

Accesso a DataCloud

- Dobbiamo risolvere il problema legale dell'accesso alle risorse INFN in particolare da parte di esterni. Su questo abbiamo iniziato una discussione in ICSC/Spoke 0 ma questo vale per tutti i progetti calcolo PNRR. In DataCloud questo viene seguito da WP7. Mi pare però che ci siano ancora discussioni di principio all'interno del gruppo Harmony. Questo deve essere risolto o non potremo dare accesso ad utenti esterni all'INFN.
  - Notare che abbiamo potenzialmente un problema in WP7 legato alla possibile mancata conferma di uno dei 2 co-leader. Questo in qualche modo è un problema da affrontare e risolvere.
  - Dovranno necessariamente esserci accordi formali firmati dall'ente/impresa richiedente accesso, in cui siano presenti opportune clausole di designazione di amministratori di sistema, manleva, etc.
- Per quanto riguarda la nomina ad amministratore di sistema per utenti *interni*, segnalo che tuttora la procedura si basa su un processo per la raccolta / gestione delle nomine implementato da WP2. Questo però era sempre stato detto che doveva essere un processo temporaneo, perché le nomine avrebbero dovuto essere registrate su LDAP centrale da INFN AAI.
  - Se questo non avviene in tempi brevi dobbiamo saperlo, perché dobbiamo finalizzare processi potenzialmente complessi come la verifica periodica delle nomine, la loro revoca etc., e il modo di farlo dipende da come si risolve internamente la cosa.
- Ho saputo in modo indiretto di discussioni che avrebbero come oggetto la migrazione di dati gestionali INFN alla Cloud del PSN. Penso che DataCloud dovrebbe (o avrebbe dovuto) essere coinvolto in queste discussioni. Le discussioni avvengono invece mi pare solo in CCR. Chiedo la vostra opinione sul modo migliore di procedere.

Un paio di importanti attività di R&D, testbed, use case

- Queste attività avvengono in WP6. Concetto architetturale sottostante: DataCloud come infrastruttura “internamente eterogenea” ma “esternamente uniforme” in quanto a servizi offerti.
- Credits: Sgaravatto, Spiga

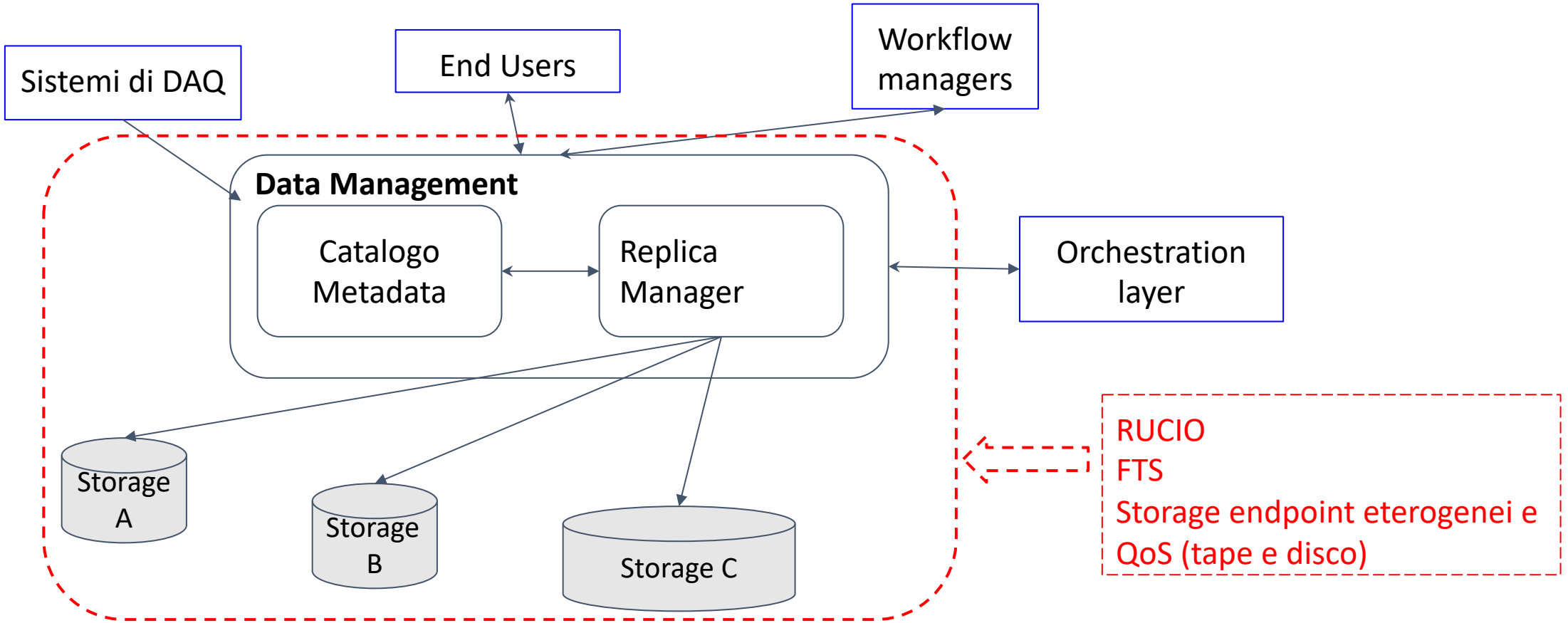
- Software Management (non ne parlo oltre qui)

Obiettivo	Permettere all’utente di gestire agevolmente ambienti di runtime (software in particolare) su risorse distribuite
Target	In particolare piccole comunità di utenti (o anche utenti singoli)
Strumenti identificati	CVMFS (ma senza che l’utente debba imparsarsi CVMFS e debba gestirsi una infrastruttura CVMFS)

- Data + Compute Management

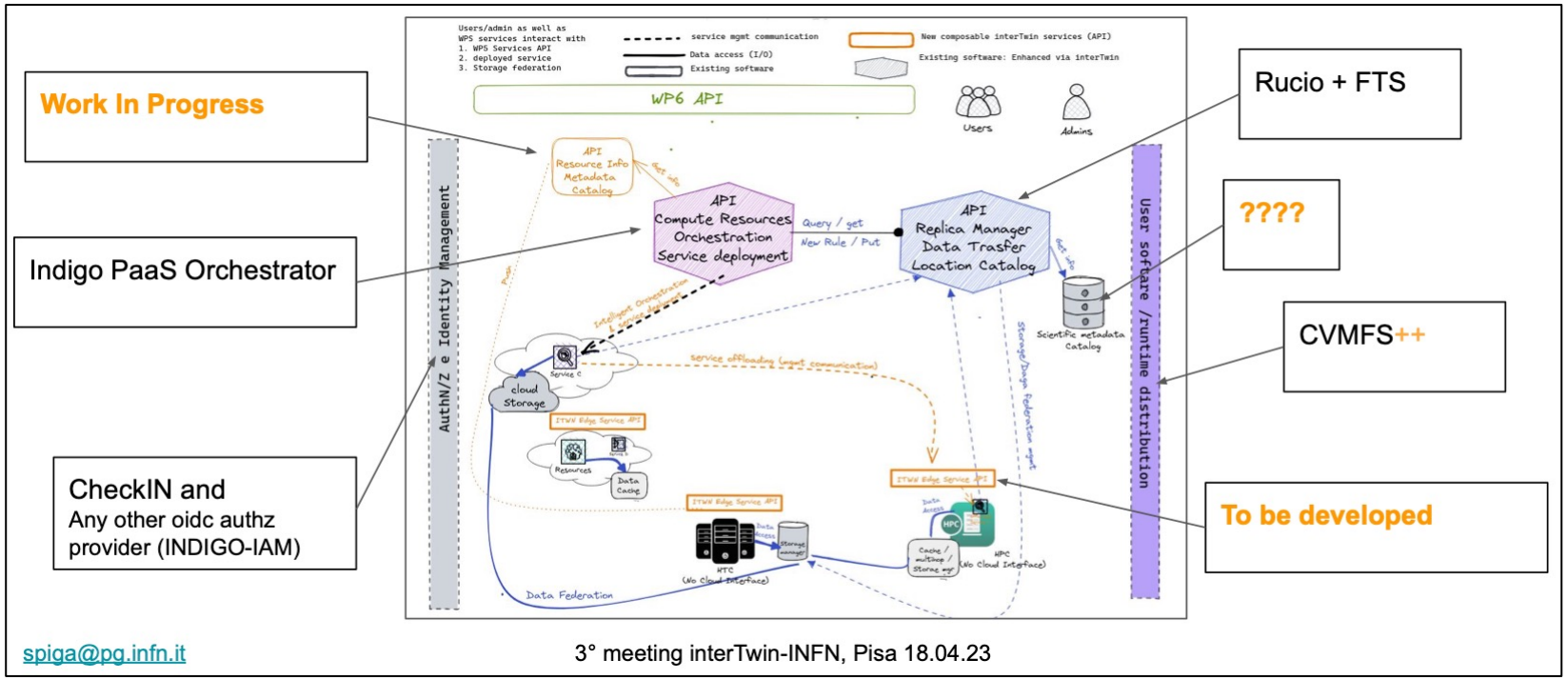
Obiettivo	Implementare un sistema di **data lake a livello di infrastruttura nazionale**
Target	In particolare i piccoli esperimenti (WLCG è già autonomo) L’infrastruttura nazionale distribuita
Strumenti identificati	Quelli a noi noti (quelli di WLCG in particolare) e quindi a servizi quali IAM, RUCIO e FTS

Data Management: abbiamo il prototipo



Compute integration

- Questo viene svolto in stretta collaborazione tra WP5/WP6 DataCloud e il progetto InterTwin. Credits: Spiga, Antonacci, Ciangottini.
- NB: questo funziona già ora in modalità PoC con risorse distribuite tra Cloud, Grid, centri HPC.



Corso di formazione per neo-assunti PNRR

- Previsto per settembre/ottobre 2023
- Programma provvisorio:
  - 16 h
  - Primo giorno: amministratori di risorse (cloud e non)
    - Policies, trattamento dati
    - Controlli di sicurezza, scansioni
  - Secondo giorno: INFN Cloud (user oriented)
    - Dashboard
    - VM, docker, jupyter, etc...
    - Esercitazioni