

# Report delle Attività di Ricerca

Pierpaolo Guidotti

26 maggio 2025

## Obiettivo del Progetto

L'obiettivo principale della ricerca è la progettazione e realizzazione di un'infrastruttura IaaS per il **Cineca**, orientata alla gestione di grande quantità di dati. Il sistema dovrà essere:

- Reproducibile tramite strumenti di provisioning automatico (**Vagrant**, **Ansible**);
- Basato su **Kubernetes**, preferibilmente installato tramite **Kubespray**;
- In grado di supportare pipeline di machine learning e data science con **Kubeflow**;
- Eseguitibile in ambienti locali o virtualizzati tramite **VirtualBox** o simili.

## Attività Tecniche Svolte

### 1. Studio e Configurazione di Kubernetes

- Studio dei concetti fondamentali di Kubernetes: architettura, networking, gestione risorse.
- Installazione e test di K3s tramite K3d per ambienti locali leggeri.

### 2. Utilizzo di Vagrant e Ansible

- Progettazione di un **Vagrantfile** per la creazione automatica di una o più macchine virtuali.
- Integrazione con **Ansible** per l'installazione automatica di Kubernetes o altri componenti.
- Definizione di variabili configurabili (numero di nodi, RAM, CPU).

### 3. Deploy di Cluster con Kubespray

- Clonazione del repository ufficiale di **Kubespray** da GitHub.
- Configurazione dell'inventary Ansible per cluster a 1, 2 e 3 nodi.

- Personalizzazione dei parametri di installazione (network plugin, componenti opzionali).
- Esecuzione di playbook per il deploy completo di Kubernetes su macchine virtuali.
- Verifica dello stato del cluster con `kubectl` e strumenti di health check.
- Test di resilienza dei nodi e riconfigurazione di rete.

#### 4. Integrazione con Kubeflow

- Analisi dell'architettura di **Kubeflow** e dei suoi componenti (Pipelines, Katib, Notebooks, ecc.).
- Studio delle modalità di installazione: `kfctl` (deprecato), `kustomize`, manifest ufficiali.

### Decisioni Tecnologiche

- **K3d** scelto per testing rapido su ambienti locali.
- **Kubespray** per la creazione di cluster realistici e multi-node in ambienti simulati.
- Utilizzo di **Ansible** per il provisioning dichiarativo.
- **Vagrant** per isolare l'ambiente ed evitare interferenze con la macchina host.

### Prossimi Step

- Completare il `Vagrantfile` finale con opzioni multiple (K3s / Kubespray);
- Automatizzare l'installazione di Kubeflow via `kustomize`;
- Testare pipeline di esempio (notebook, training ML);
- Scrivere documentazione tecnica per la riproduzione completa.
- Spostare il cluster sul quale installare su kubeflow su VM fornite.

#### *Postilla*

Questo documento è stato redatto in forma generale per riassumere quanto svolto fino ad oggi. A partire da questo momento, esso sarà **aggiornato quotidianamente** con la descrizione puntuale delle attività svolte.

Tutti gli aggiornamenti saranno salvati in formato  $\text{\LaTeX}$  e pubblicati nella repository ufficiale del progetto, disponibile al seguente indirizzo:

<https://github.com/ppguidotti/ricerca>

La repository fungerà da riferimento pubblico e versionato per il tracciamento delle attività, in ottica di trasparenza e collaborazione.