Consortium GARR

8° BORSISTI DAY
"ORIO CARLINI"
MERCOLEDÌ 14 DICEMBRE 2016 - ROMA

# Sistema di monitoraggio per datacenter distribuiti geograficamente basati su OpenStack

Dott. Ing. Gioacchino Vino

Tutor: Dott. Domenico Elia
Tutor: Dott. Giacinto Donvito

INFN
Istituto Nazionale
di Fisica Nucleare
Sezione di Bari

# INDICE

- Descrizione del progetto
- Architettura finale
- Confronto tra attività programmate e realizzate
- Proposta di attività per l'ulteriore anno di studio

# DESCRIZIONE DEL PROGETTO

Implementazione di un sistema di monitoraggio che supporti la gestione di datacenter distribuiti geograficamente basati su OpenStack

**Fasi chiave:**

- Acquisizione di metriche eterogenee provenienti da:
  - Servizi offerti
  - Livello IaaS basato su OpenStack
  - Livello Hardware
- Analisi sui dati raccolti:
  - Anomaly Detector
  - Root Cause Analysis

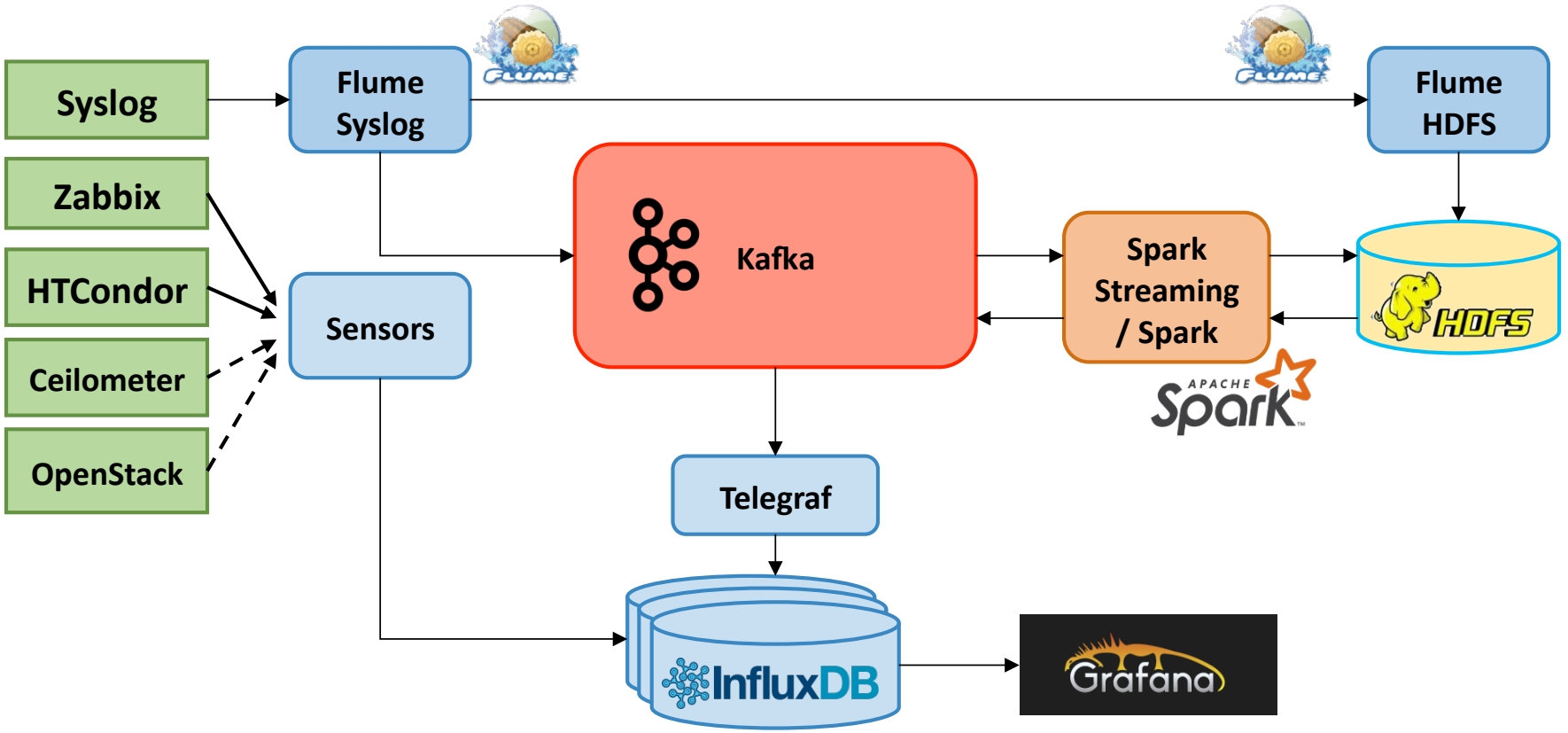
# DESCRIZIONE DEL PROGETTO

**Datacenter ReCaS di Bari**:

- 128 server con 8192 core complessivi
- 3.5 PB di spazio disco
- 2.5 PB di archivio su nastro
- Cluster HPC composto da 20 server con 800 core complessivi
- Connessa alla rete GARR con 2 link da 10Gbps
- Piattaforma cloud: OpenStack

*Sistema di monitoraggio per datacenter distribuiti geograficamente basati su OpenStack*

# DESCRIZIONE DEL PROGETTO



# DESCRIZIONE DEL PROGETTO

**Sorgenti dati**:

- Syslog:
  - Informazioni sui processi di sistema
  - Acquisiti tramite la sorgente Syslog di Flume
  - 5 - 6 milioni di righe di log al giorno -> 500 - 600MB/giorno
  - Raccolti 18 GB di storico dati dal 18 Novembre 2016

# DESCRIZIONE DEL PROGETTO

**Sorgenti dati**:

- Syslog
- Zabbix:
  - Risorse utilizzate dai nodi, monitoring delle componenti di OpenStack e dei servizi
  - Sonda scritta in Python
  - 42000 metriche campionate ogni 10 minuti -> circa 10MB/giorno
  - Raccolti 1.5 GB di storico dati dal 19 Luglio 2016

# DESCRIZIONE DEL PROGETTO

**Sorgenti dati**:

- Syslog
- Zabbix
- HTCondor:
  - Stato degli scheduler, dei job terminati e in esecuzione
  - Sonda scritta in Python
  - 750000 metriche campionate ogni 5 minuti -> circa 33 MB/giorno
  - Raccolti 5 GB di storico dati dal 18 Luglio 2016

# DESCRIZIONE DEL PROGETTO

**Analisi effettuate**:

- Batch:
  - Aggregazione dei file di log
- Real-time:
  - Estrazione di log contenenti parole chiave (come "err", "failed", "fatal", …)

# DESCRIZIONE DEL PROGETTO

**Risorse utilizzate**:

- 50 CPU
- 100 GB di RAM
- 2.3 TB di Volume Storage
  - 1.5 TB usati per **HDFS** in replica 3
  - 450 GB usati come buffer in **Flume** e **Kafka**
  - 150 GB usati per le tre istanze di **InfluxDB**
  - Unità disco persistente per ogni macchina

# DESCRIZIONE DEL PROGETTO

**Attività a supporto**:

- Configurazione centralizzata dei file di configurazione
- Monitoraggio di tutte le macchine virtuali utilizzate e delle componenti utilizzate:
  - Flume
  - HDFS
  - InfluxDBs
  - Sensors

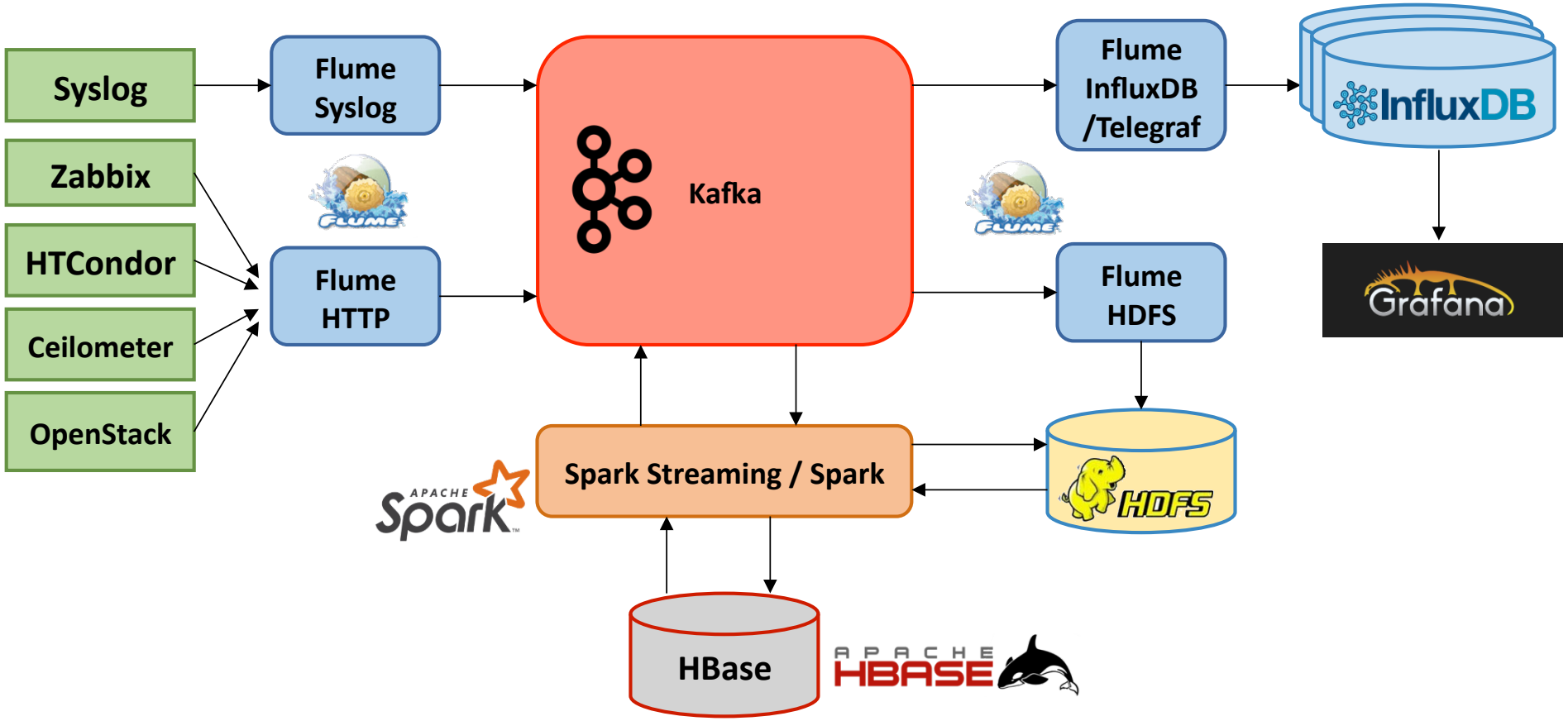
*Sistema di monitoraggio per datacenter distribuiti geograficamente basati su OpenStack*

# DESCRIZIONE DEL PROGETTO

- Flum
- HDF
- Influ
- Sens

*Sistema di monitoraggio per datacenter distribuiti geograficamente basati su OpenStack*

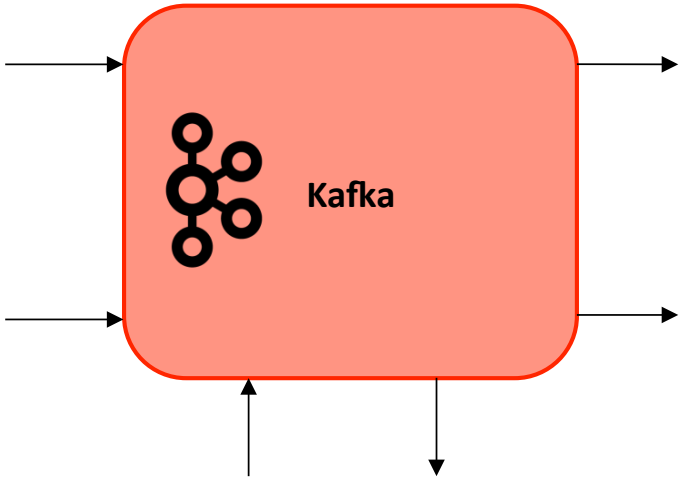
# ARCHITETTURA FINALE



# ARCHITETTURA FINALE

**Kafka:**

- Aumenta l'affidabilità dello strato di trasporto grazie alla possibilità di replicare i dati tra i nodi del cluster
- Spark Streaming interagisce unicamente con Kafka per la fase di lettura e scrittura dei dati
- Sono presenti componenti esterne (Flume, Telegraf) in grado di scrivere e leggere in/da Kafka



# CONFRONTO TRA ATTIVITÀ PROGRAMMATE E REALIZZATE

**Programma I trimestre**:

- Realizzazione dei sensori
- Configurazione dello strato di trasporto e degli archivi

**Attività realizzate**:

- Definizione dell'architettura del sistema e dei software da utilizzare
- Installazione e configurazione di **HDFS**
- Ceilometer*

# CONFRONTO TRA ATTIVITÀ PROGRAMMATE E REALIZZATE

**Programma II trimestre**:

- Installazione e configurazione del blocco di analisi
- Implementazione del rilevatore di anomalie

**Attività realizzate**:

- Configurazione di **HDFS** in High Availability
- Configurazione di **Flume**
- Scrittura dei sensori di **Zabbix** e **HTCondor**
- Monitoraggio delle macchine virtuali e delle componenti utilizzate
- Configurazione centralizzata dei file di configurazione

# CONFRONTO TRA ATTIVITÀ PROGRAMMATE E REALIZZATE

**Programma III trimestre**:

- Implementazione dell'estrattore di topologia e del Root Cause Analysis

**Attività realizzate**:

- Configurazione di **Spark** e analisi su storico dei dati
- Monitoring di **HDFS**

# CONFRONTO TRA ATTIVITÀ PROGRAMMATE E REALIZZATE

**Programma IV trimestre**:

- Fase di calibrazione e test di supporto con gli amministratori

**Attività realizzate**:

- Installazione e configurazione di **Kafka**
- Elaborazione dei dati in real-time con **Spark Streaming**
- Salvataggio dei risultati in **InfluxDB**
- Monitoraggio di **Zookeeper**

# PROPOSTA DI ATTIVITÀ PER L'ULTERIORE ANNO DI STUDIO

- Finalizzazione del progetto ordinario (stima un trimestre)
- Estensione e valorizzazione ulteriore del progetto e delle sue finalità:
  - Impiego di **Apache Mesos** come gestore delle risorse
  - Implementazione di uno **strumento rapido** e **intuitivo** a supporto della fase di **debugging**
  - Integrazione con il **sistema di ticketing** usato localmente
  - **Esecuzione automatica** di comandi semplici
  - Analisi di **colli di bottiglia** delle prestazioni
  - Test su siti distribuiti

# GRAZIE

# PER

# L'ATTENZIONE