

ALMA MATER STUDIORUM · UNIVERSITÀ DI BOLOGNA

SCUOLA DI SCIENZE

Corso di Laurea in Informatica per il Management

TITOLO DELLA TESI

Relatore:

Chiar.mo Prof.
Moreno Marzolla

Presentata da:

Alessio Arcara

Correlatore:

Stefano Dal Pra

Seconda Sessione di Laurea
Anno Accademico 2022 - 2023

*Questa è la dedica:
ognuno può scrivere quello che vuole,
anche nulla . . .*

Sommario

Indice

Sommario	i
1 Introduzione	1
2 Capitolo di contesto	3
3 Estrapolazione dei dati	7
4 Analisi dei risultati	9
5 Conclusioni	11
A Prima Appendice	13
Bibliografia	15

Elenco delle figure

2.1	Struttura gerarchica del WLCG [4]	4
2.2	legenda elenco figure	5

Elenco delle tabelle

Capitolo 1

Introduzione

Questa è l'introduzione.

Capitolo 2

Capitolo di contesto

Il **grid computing** è un'architettura di calcolo distribuito che collega computer sparsi geograficamente allo scopo di condividere risorse e potenza di calcolo per raggiungere uno scopo condiviso. Attualmente, il più grande sistema grid al mondo è Worldwide LHC Computing Grid (WLCG). Questa è una collaborazione internazionale che coinvolge oltre 170 centri di calcolo sparsi in più di 40 nazioni. Lo scopo del WLCG è fornire l'infrastruttura computazionale necessaria per gestire i dati generati dagli esperimenti effettuati con il Large Hadron Collider (LHC) [2].

Come mostrato nella figura 2.1, i centri di calcolo all'interno del WLCG sono strutturati secondo il modello MONARC, che li organizza in un sistema gerarchico di livelli, noti come Tier, ciascuno dei quali ha funzioni e responsabilità ben definite. In questo contesto si colloca il centro nazionale delle tecnologie informatiche e telematiche (CNAF), che ospita il Tier-1 per tutti e quattro gli esperimenti del LHC. Oltre agli esperimenti LHC, vengono supportati presso il CNAF gli esperimenti non-LHC di astrofisica delle particelle e fisica dei neutrini [1].

Il CNAF offre più di 46000 core distribuiti su 960 host fisici per un totale di circa 630 kHS06¹ di potenza di calcolo [5]. L'allocazione di queste risorse segue il paradigma del **High-Throughput Computing** (HTC), dove a differenza dell'High-Performance Computing, che mira a eseguire calcoli ad alta velocità, l'obiettivo dell'HTC è massimizzare il

¹metrica per misurare le prestazioni della CPU, sviluppata dal gruppo di lavoro HEPiX. È utilizzata per confrontare le risorse di calcolo in ambito scientifico.

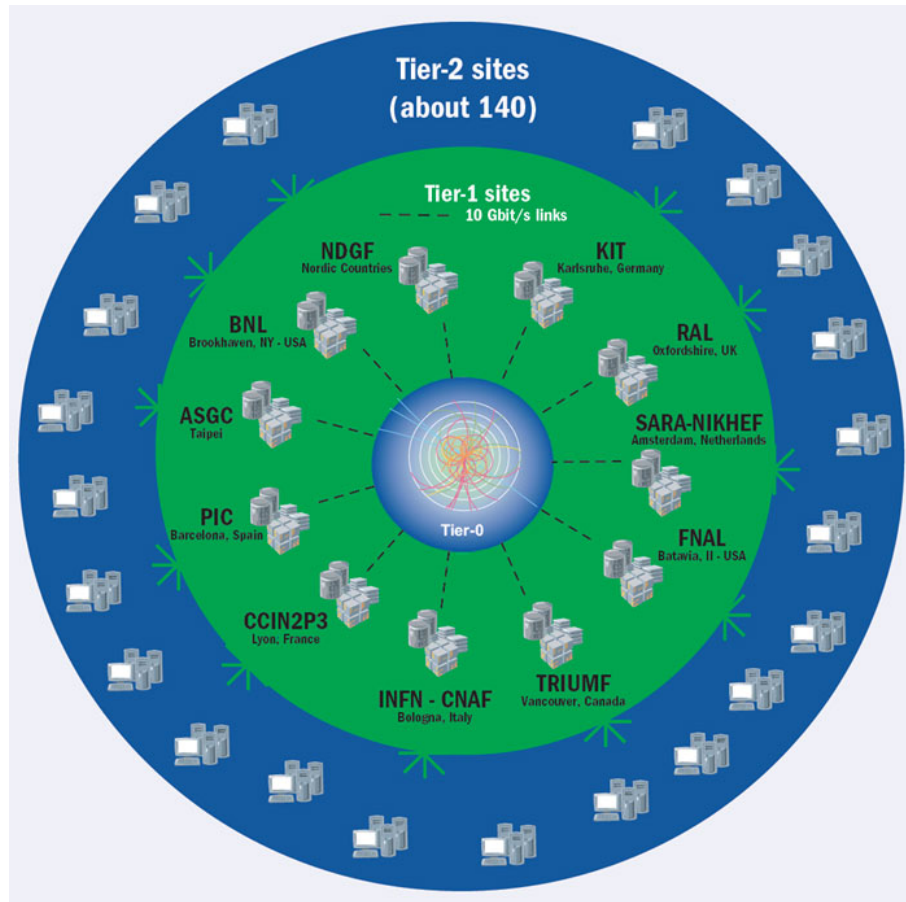


Figura 2.1: Struttura gerarchica del WLCG [4]

numero di operazioni compiute su un periodo di tempo prolungato.

Ogni esperimento ha almeno una coda dedicata e le risorse di calcolo sono gestite centralmente da un unico batch system (HTCondor) [3]. In questo contesto, un **job** rappresenta un'unità di lavoro che un utente vuole eseguire. Questa può essere una qualsiasi operazione che richieda risorse computazionali. Una volta sottomesso, il job viene accodato da un batch system e attende la sua schedulazione in base ad algoritmi di fairshare. Questi algoritmi assicurano che le risorse computazionali siano distribuite equamente tra tutti gli utenti e le varie code, impedendo che un singolo utente o coda possa monopolizzare tutte le risorse o che una coda soffra di *starvation*².

²Si riferisce a una situazione in cui un job non viene mai eseguito perché sono presenti job con priorità più alta.

In media, ogni giorno vengono eseguiti 100000 jobs batch, e le risorse di calcolo sono utilizzate 24×7 dai vari esperimenti scientifici.

Una serie di dati è una sequenza ordinata di punti dati, ed esprime la dinamica di un certo fenomeno nel tempo. Quando questi dati sono ordinati in base al tempo, si parla di una **serie temporale**. Indipendentemente dal criterio utilizzato per ordinarli, i punti dati sono registrati seguendo intervalli di tempo equispaziati. Le serie temporali possono essere di due tipi: **univariate**, che coinvolgono una singola variabile misurata nel tempo, e **multivariate**, dove più variabili sono misurate contemporaneamente.

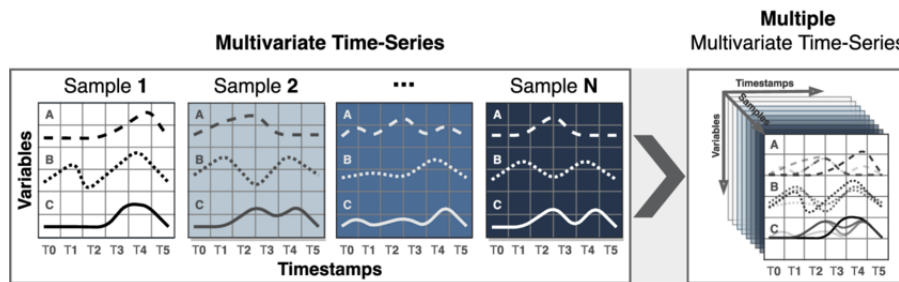


Figura 2.2: mettere riferimento

Capitolo 3

Estrapolazione dei dati

Capitolo 4

Analisi dei risultati

Capitolo 5

Conclusioni

Queste sono le conclusioni.

In queste conclusioni voglio fare un riferimento alla **[Bortolotti_2012]** bibliografia:
questo è il mio riferimento.

Appendice A

Prima Appendice

In questa Appendice non si è utilizzato il comando:
`\clearpage{\pagestyle{empty}\cleardoublepage}`, ed infatti l'ultima pagina 8 ha l'intestazione con il numero di pagina in alto.

Bibliografia

- [1] G Bortolotti et al. «The INFN Tier-1». In: *Journal of Physics: Conference Series* 396.4 (dic. 2012), p. 042016. DOI: 10.1088/1742-6596/396/4/042016. URL: <https://dx.doi.org/10.1088/1742-6596/396/4/042016>.
- [2] CERN. *Worldwide LHC Computing Grid*. 2023. URL: <https://wlcg.web.cern.ch> (visitato il 28/10/2023).
- [3] CNAF. *WLCG Tier-1 data center - Calcolo*. URL: <https://www.cnaf.infn.it/calcolo/> (visitato il 28/10/2023).
- [4] Stefano Dal Pra et al. «Evolution of monitoring, accounting and alerting services at INFN-CNAF Tier-1». In: *EPJ Web of Conferences* 214 (gen. 2019), p. 08033. DOI: 10.1051/epjconf/201921408033.
- [5] Andrea Rendina. *INFN-T1 site report*. https://indico.cern.ch/event/1200682/contributions/5087586/attachments/2538178/4368754/20221031_InfnT1_site_report.pdf. Accessed: 2023-10-28. 2022.

Ringraziamenti

Qui possiamo ringraziare il mondo intero!!!!!!!!!!
Ovviamente solo se uno vuole, non è obbligatorio.