

Indice generale

1 - Introduzione.....	1
2 – Tecnologie usate.....	1
3 – Analisi del problema.....	1
3.1 – network flows.....	1
3.2 – argus.....	1
3.3 - nProbe.....	2
3.4 - Stratosphere IPS.....	2
3.5 – problematiche dovute all’utilizzo di due formati diversi.....	2
3.6 – presentazione del problema.....	2
4 – Soluzione proposta.....	2
4.1 – installazione software.....	2
4.2 – conversione.....	2
4.3 – automatizzazione conversione.....	2
4.4 – rendere efficiente la conversione.....	2
4.4.1 – possibili soluzioni.....	2
4.4.2 – scelta effettuata.....	3
5 – Esperimenti.....	3
5.1 – dataset utilizzato.....	3
5.2 – benchmark single core.....	3
5.3 – benchmark multi core.....	3
5.4 – prestazioni.....	3
6 – Conclusioni.....	3

1 - Introduzione

Qui lascio quello che ho mandato, dove espondo tutti gli argomenti brevemente, il mio ruolo nella tesi e una descrizione per ogni capitolo.

2 – Tecnologie usate

Qui vado nel dettaglio sugli argomenti dell’introduzione,

3 – Analisi del problema

3.1 – network flows

descivo cosa sono i file di flow usati dai programmi che verranno usati dopo (argus e nprobe) e perchè si usano.

3.2 – argus

descivo cos'è argus

3.3 - nProbe

Come sopra

3.4 - Stratosphere IPS

tutto quello che c'è da dire su stratosphere. Metto anche una parte in cui descrivo i modelli comportamentali che utilizza e ci metto la tabella che è presente sul sito che spiega come leggere un modello

3.5 – problematiche dovute all'utilizzo di due formati diversi

Metto in evidenza le differenze tra i due formati

3.6 – presentazione del problema

Scrivo che i file sono in formato gzip, che c'è una struttura gerarchica fissa.

4 – Soluzione proposta

4.1 – installazione software

Descrivo che il tutto è stato fatto su VM, poi metto i passi che ho seguito per installare Stratosphere e argus. Dopodiché metto delle immagini che illustrano l'utilizzo di stf che prendo dal sito.

4.2 – conversione

Qui metto come ho scelto di convertire i campi tra i due formati

4.3 – automatizzazione conversione

La conversione viene fatta in modalità batch scrivendo un programma in python3. Metto lo pseudocodice del programma e commento le scelte che ho fatto

4.4 – rendere efficiente la conversione

Introduco il concetto di parallelismo applicato alla conversione

4.4.1 – possibili soluzioni

Metto i vari modi in cui è possibile parallelizzare l'esecuzione del programma.

- Scrittura su file 1:1
- lavorare su chunk di file
- meccanismi di lock

4.4.2 – scelta effettuata

Descrivo la scelta che ho effettuato e il perché con pseudocodice.

5 – Esperimenti

5.1 – dataset utilizzato

Descrivo le condizioni in cui si sono svolti gli esperimenti: il dataset utilizzato, l'hardware del pc

5.2 – benchmark single core

5.3 – benchmark multi core

5.4 – prestazioni

Introduco lo speedup lineare, il concetto di efficienza e applico le formule ai risultati ottenuti nei benchmark per far vedere quanto è efficiente il programma.

6 – Conclusioni

recap della tesi, ribadendo quale è l'obiettivo, cosa si è fatto, e che risultati sono stati ottenuti (il tutto in maniera discorsiva).