WORKSHOP GARR 2023

NET MAKERS

Analisi sui flussi Netflow con sistemi OLAP, verso il ML

Alfredo Funicello

GARR

ETL moderno per l’analisi real-time dei flussi Netflow

Obbiettivo:

Evolvere il monitoraggio dei flussi Netflow in GARR-T

Perché:

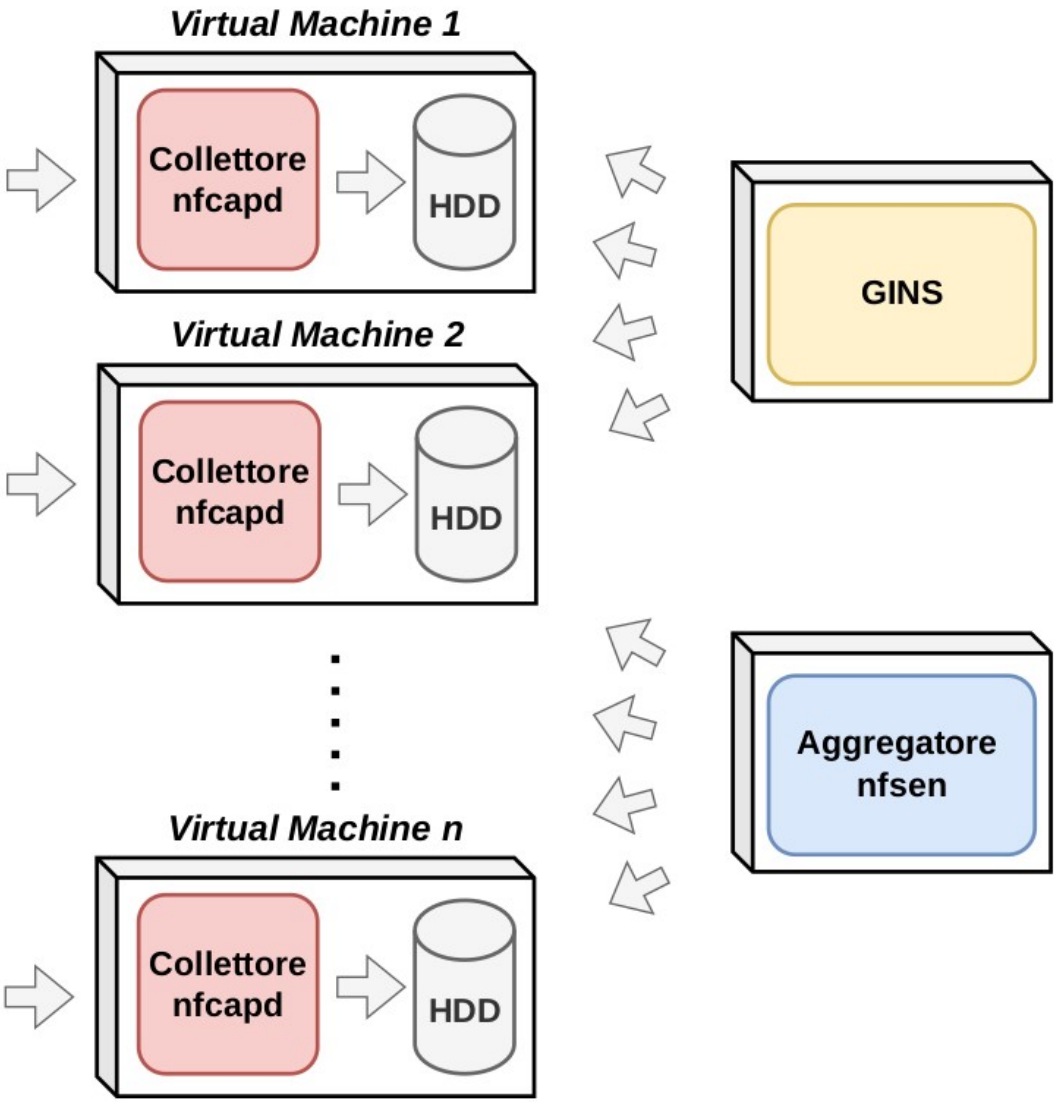
Una visione piu’ dettagliata del traffico

IP mittente	AS	Porta sorgente	...	IP destinatario	AS	Porta destinazione	Protocollo	...
216.58.205.x	15169	8092	...	193.206.1.x	137	8500	TCP	...

Risultato:

Un’infrastruttura di osservazione moderna

Architettura corrente



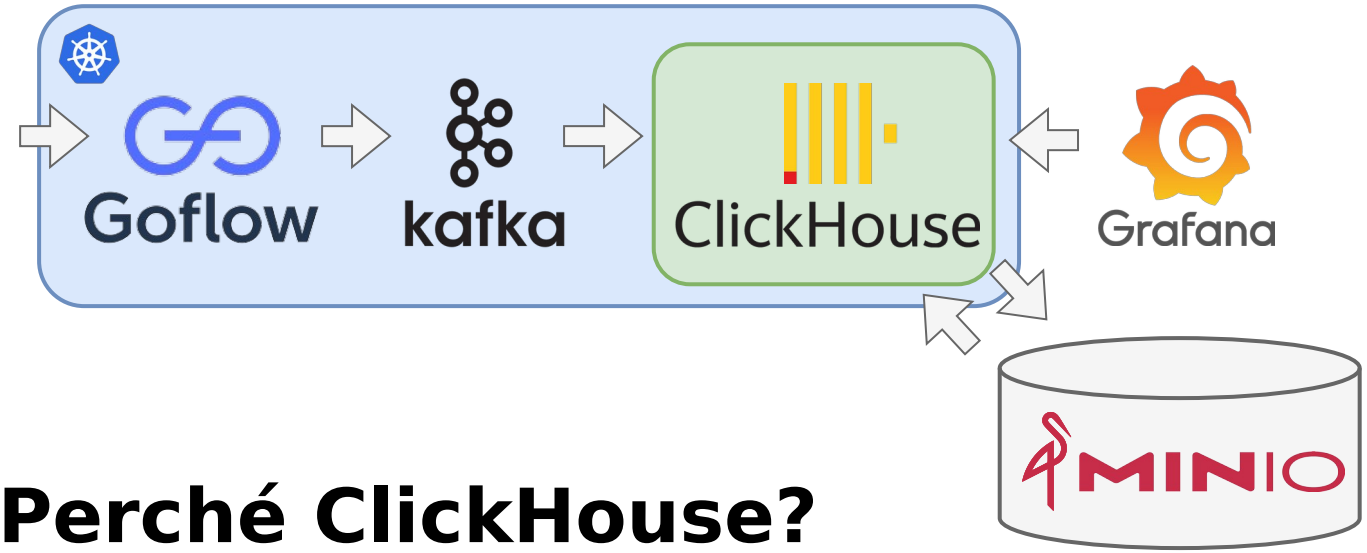
- VM con flussi in binario
- Nfsen per l'aggregazione e visualizzazione

- Utente principale: GARR CERT

Criticità

- Scalabilità
- Integrazione
- Flessibilità
- Storage

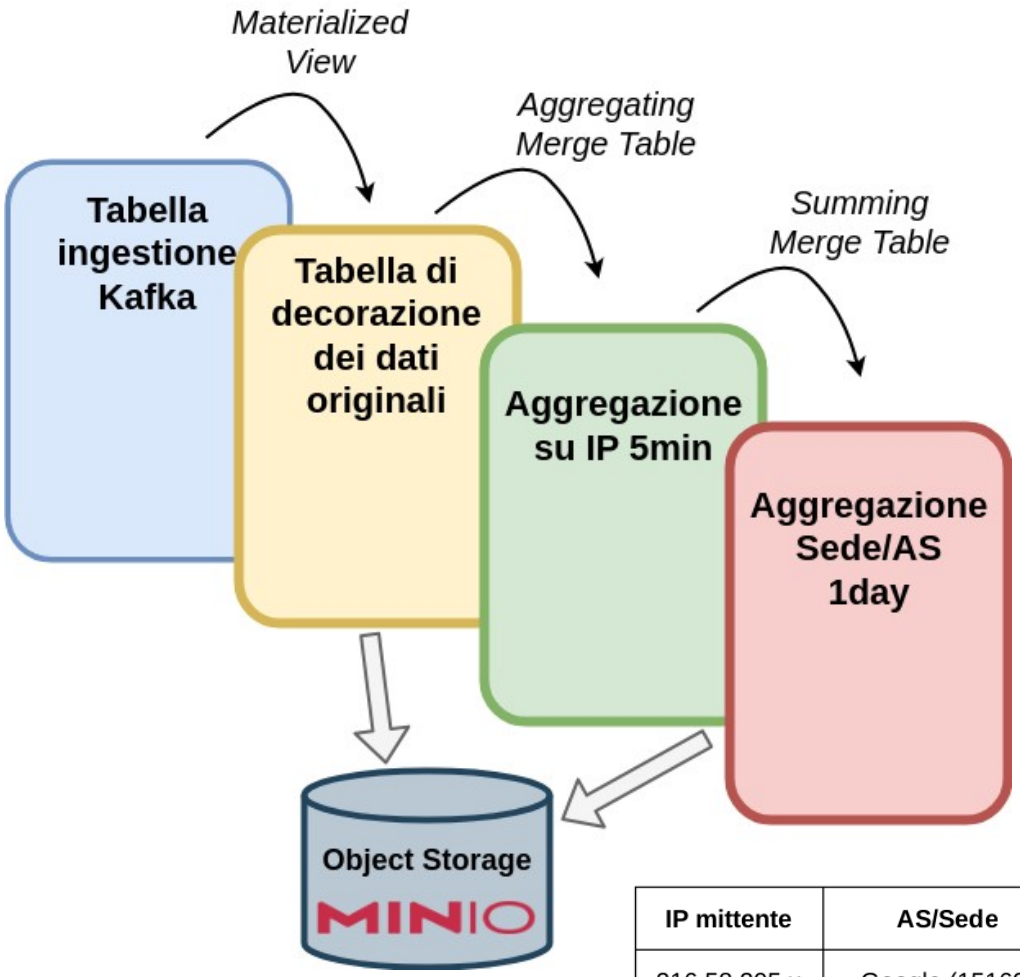
L’architettura della nuova piattaforma



Perché ClickHouse?

- Scalabilità: distribuito, Kubernetes
- Integrabilità: SQL, API
- Archiviazione: Tiering su Object storage
- Velocità: logica OLAP colonnare

Osserviamo un caso d'uso in logica OLAP



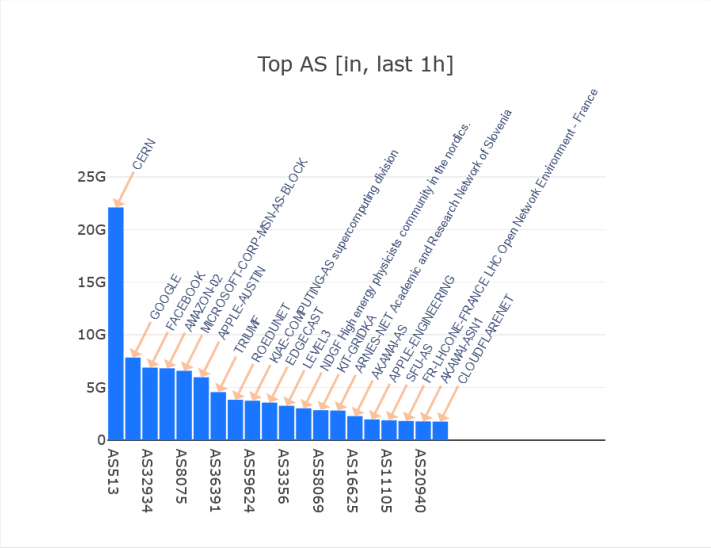
- Aggregazione a **5 min** per analisi di sicurezza
- Aggregazione a **1 giorno** per finalità di reportistica

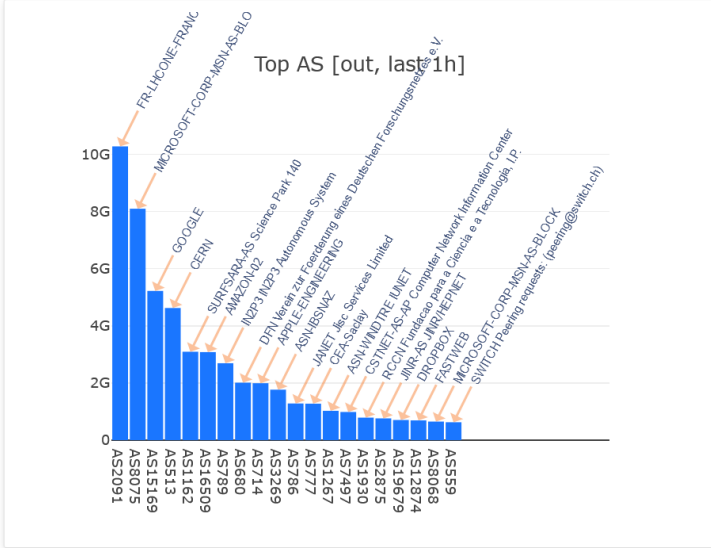
IP mittente	AS/Sede	Porta sorgente		
216.58.205.x	Google (15169)	8092		
IP destinatario	**AS/Sede**	**Porta destinazione**	**Protocollo**	...
193.206.1.x	ECMWF Bologna	8500	TCP	...

Da così

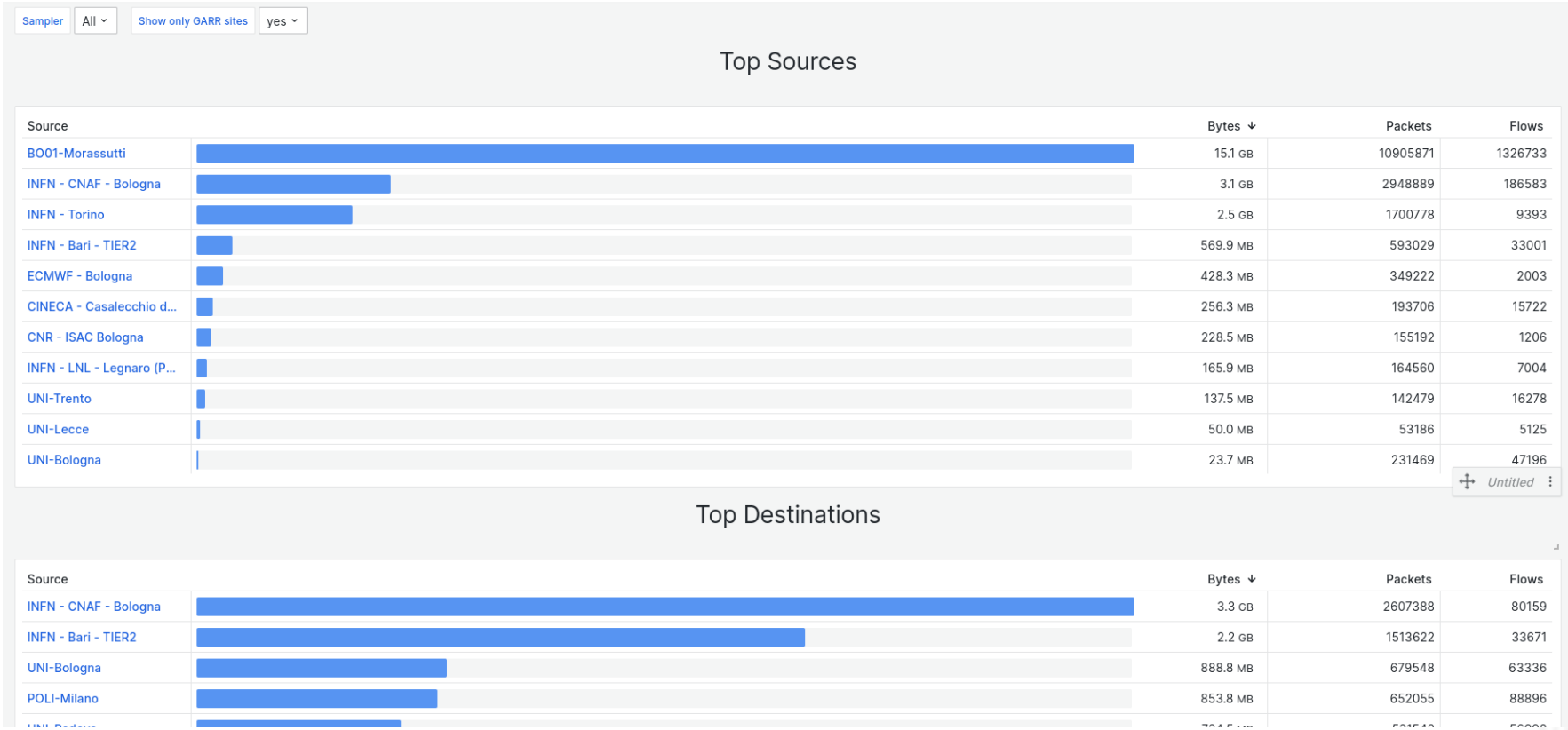
```
nfdump -m -M \
/data/nfsen/profiles-data/live/rl1mi02:rl1mi03:rl1mi08:rl1mi04:rl1ts01:rl2mi03:rl1mi01:re2ur:re2pz:rx2bo1:\
rl1fi01:rl1bo02:rl2fi01:rl2mi04:rl1rm13:rl2bo03:rl2rm04:rl1rm04:rl2rm02:rl1rm02:rl1rm01:rl1mb00:rl1pd01:\
rl1pd02:rl1pv00:rx3ba1:rx1me:rx1rc:rx1cs:rx1cz:rx2pa1:rx1ct1:rx2ct1:rx2mi1:rx1pa1:rx1na1:rx2na1:\
re1ss:rl1pi01:re1an:rl1fi03:rlabrm2:re2an:rl2fe00:rl1fi05:rl1fi02:rl1fi04:rl1co00:rl1fe00:rl1bo03:\
rl1bo01:rl1bg00:rx2ba1:rx1ca1:rl1ge01:rl2rm06:rl1rm06:rl1ge02:rl2ge01:rl1bo04:rl1bs00:rl1ve11:rl1ve04:\
rl1ve03:rl1ve02:rl2ve00:rl1ve00:re1pg:re2pg:re1aq1:re2aq1:rl1vr00:re2ss:rl1to01:rl1tn00:re2mt:rl1to02:\
rx1ca6:re1mt:rl1ud00:re1pz:rs1mi01:rs1bo01:rs1rm02:rs1mi02:re1rm02:re1mi02:re1mi01:re1ur:\
rl1tn01:rx1sa:rx1br:rx1na2:rx1na6:rx1le:rx1fg:rl2pv00:rl1pv01:rs1pi01:rs1to01:rs1pd01:rs1pd02:\
rl1pi02:rl2pi02 -R 2023/10/02/nfcapd.202310020300:2023/10/02/nfcapd.202310021500 \
-o "fmt:%ts,%te,%td,%pr,%sa,%sp,%da,%dp,%sas,%das,%pkt,%byt,%bps,%pps,%bpp,%fl" \
"(src port 161 or src port 162 or src port 389 or src port 53 or src port 123) and (src as 137 and not dst as 137)"
```

Globali

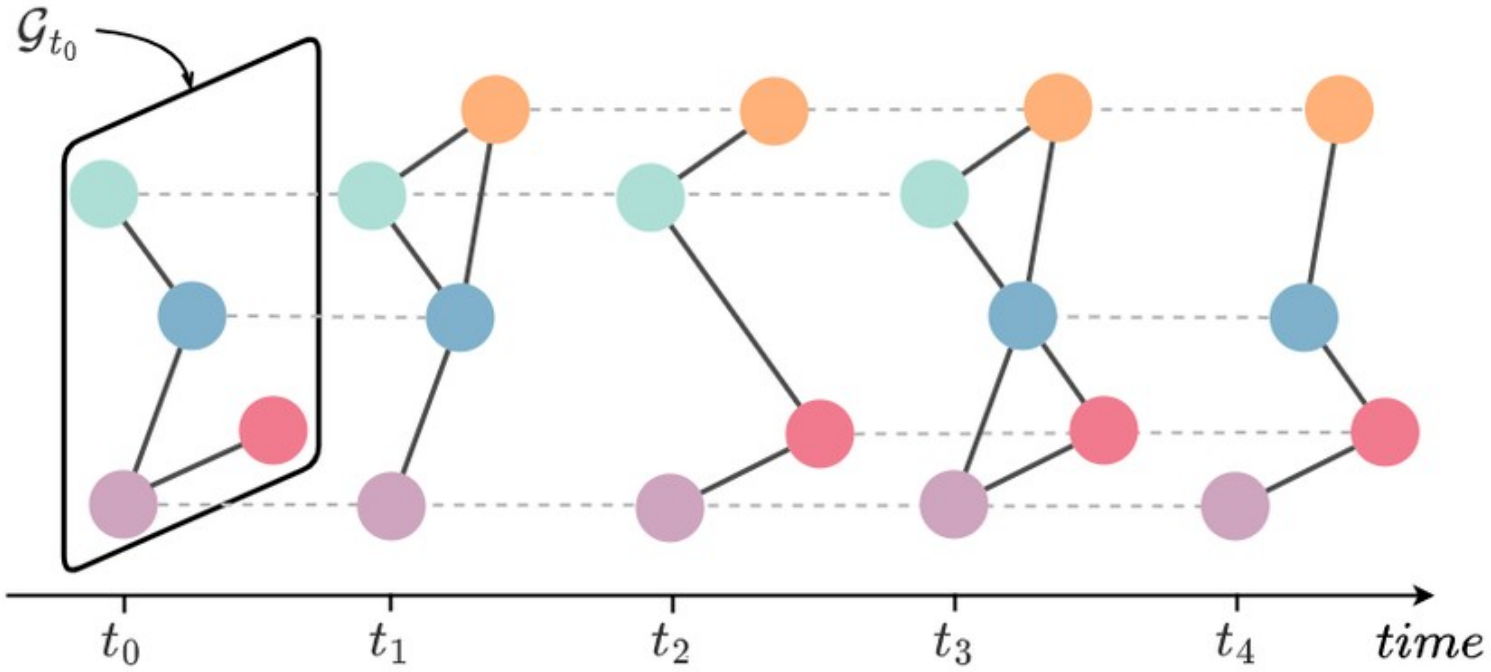




A così


Sampler
All
Show only GARR sites
yes
Top Sources
Source
Bytes
Packets
Flows
BO01-Morassutti
15.1 GB
10905871
1326733
INFN - CNAF - Bologna
3.1 GB
2948889
186583
INFN - Torino
2.5 GB
1700778
9393
INFN - Bari - TIER2
569.9 MB
593029
33001
ECMWF - Bologna
428.3 MB
349222
2003
CINECA - Casalecchio d...
256.3 MB
193706
15722
CNR - ISAC Bologna
228.5 MB
155192
1206
INFN - LNL - Legnaro (P...
165.9 MB
164560
7004
UNI-Trento
137.5 MB
142479
16278
UNI-Lecce
50.0 MB
53186
5125
UNI-Bologna
23.7 MB
231469
47196
Untitled
Top Destinations
Source
Bytes
Packets
Flows
INFN - CNAF - Bologna
3.3 GB
2607388
80159
INFN - Bari - TIER2
2.2 GB
1513622
33671
UNI-Bologna
888.8 MB
679548
63336
POLI-Milano
853.8 MB
652055
88896

Verso il ML per l'anomaly detection



"Le Graph Neural Networks (GNN) sono un metodo che ha ricevuto recente attenzione nella sfida al rilevamento efficiente e intuitivo delle anomalie in un grafo"

Structural Temporal Graph Neural Networks for Anomaly Detection in Dynamic Graphs (2020)
Lei Cai et. al

Grazie

www.garr.it/domande
codice: 2318 9129

Funicello Alfredo
alfredo.funicello@garr.it