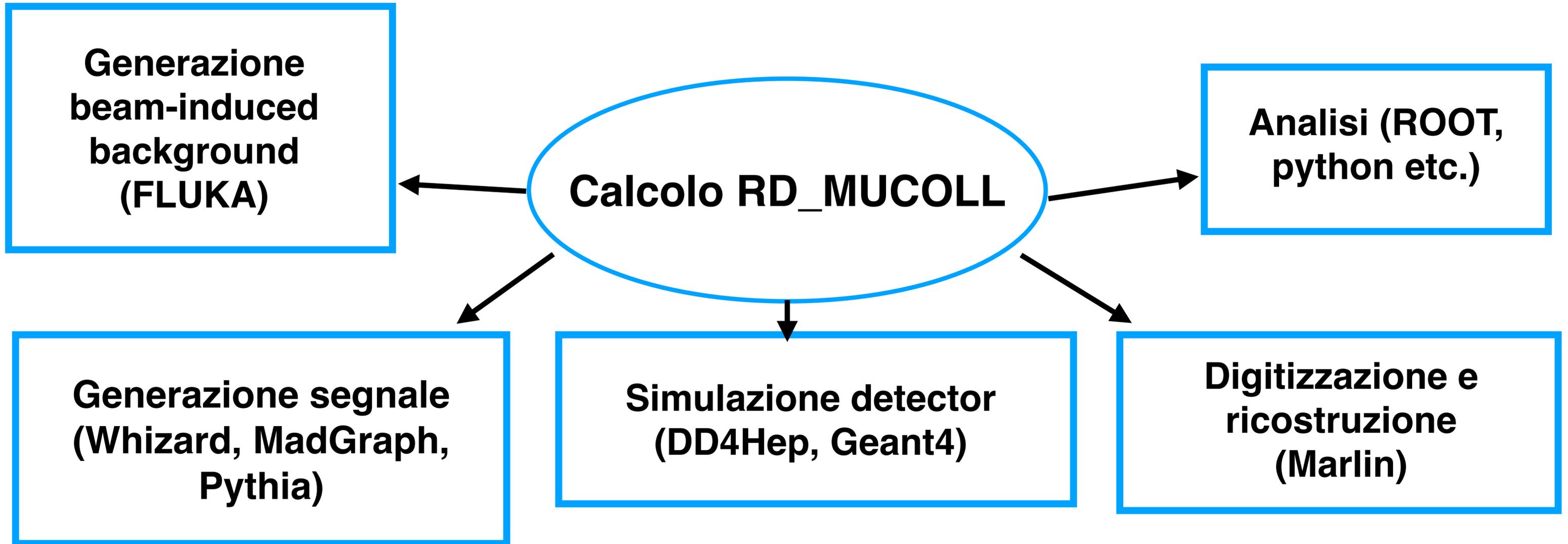




Calcolo RD_MUCOLL

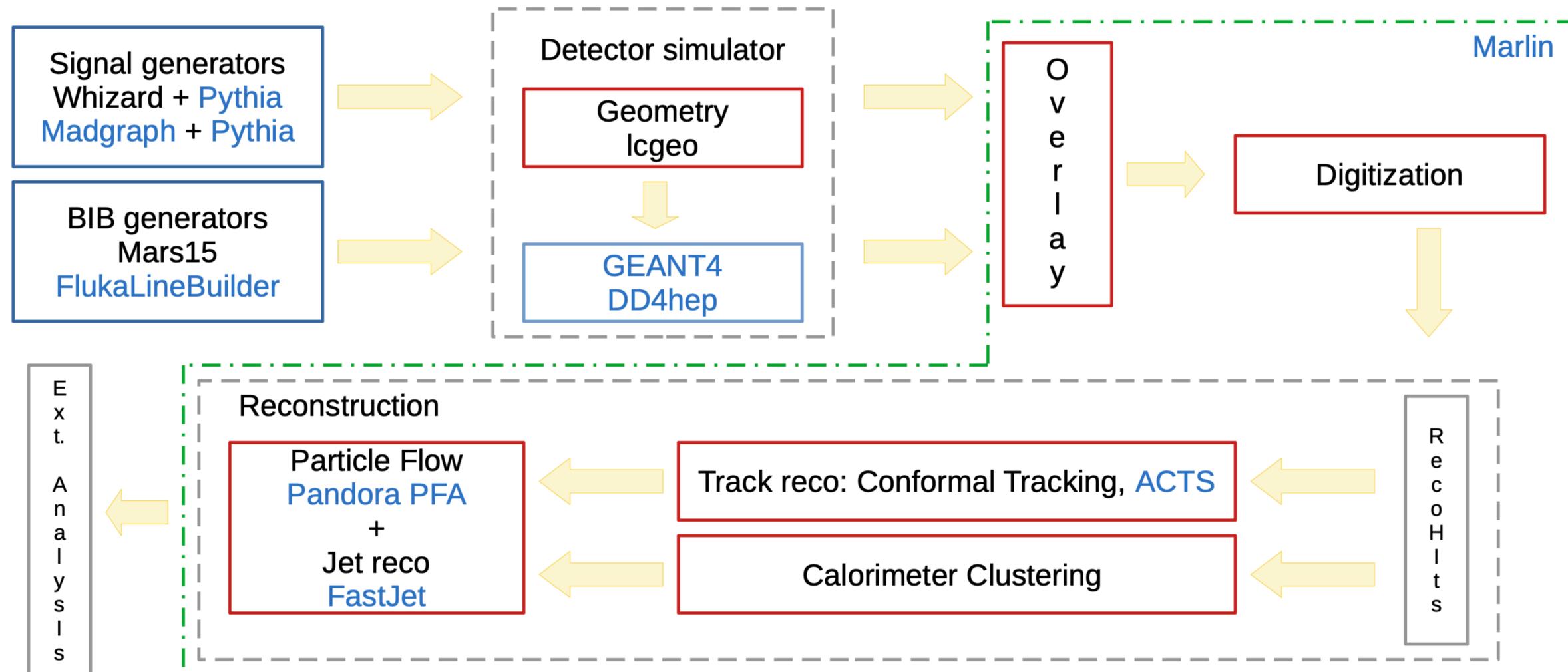
Lorenzo Sestini - INFN Padova

Meeting con i referee di RD_MUCOLL - 31/8/2022



Struttura software

Da presentazione Paolo Andreetto ad ICHEP 2022: https://agenda.infn.it/event/28874/contributions/169190/attachments/94243/128880/ICHEP_MC_Computing.pdf



- I processi più dispendiosi in termine di risorse di calcolo sono la simulazione del Beam-Induced Background (BIB, 30M di particelle per bunch-crossing) e la ricostruzione del segnale + BIB (combinatorio elevato)

Sviluppo e distribuzione software

Da presentazione Paolo Andretto ad ICHEP: https://agenda.infn.it/event/28874/contributions/169190/attachments/94243/128880/ICHEP_MC_Computing.pdf



Development and production previously based on CentOS 8
Temporary moved to CentOS 8 Stream

Waiting for a final solution from CERN/FermiLab



Development tools and guidelines from iLCSoft

The code is available in [Github](#)

A Github-based CI has been configured by BerkeleyLab people

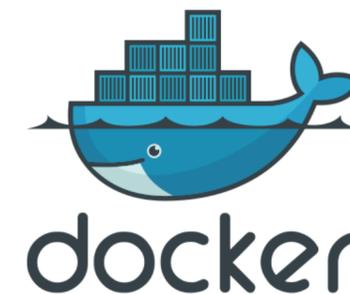
The reference artifact for the production is a docker container, published in [Docker Hub](#)

Conversion from docker image into singularity one is performed

The singularity container:

- is distributed via storage element for batch processing at INFN-CNAF
- will be available via CVMFS at CERN (**still work in progress**)

Stiamo inoltre seguendo lo sviluppo del framework comune per futuri esperimenti [Key4hep](#). Al momento non abbiamo forza-lavoro per contribuire ma valuteremo la migrazione nel futuro



Mantenimento del software e gestione risorse: [Alessio Gianelle](#) e [Paolo Andretto](#) (tecnologi INFN-PD)

Elenco siti risorse

- **Cloud-Veneto**: 200 VCPU, 740 GB di RAM, 90 TB di storage, accesso via INFN IdP
- **CNAF**: batch system basato su HTCondor, 150 TB di storage, 6 CE, accesso via VOMS
- **CERN**: batch system basato su HTCondor, 100 TB di storage su CERN EOS, in pre-produzione
- **IBISCO-Bari**: risorse condivise con altri progetti allocate al momento della richiesta, accesso con account a Bari
- **Risorse locali**: Farm Trieste (modalità parassitica), Pavia etc.

Cloud-Veneto

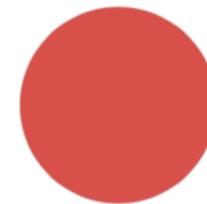
Report sull'utilizzo per periodo	2021-08-30	2022-08-30				
ID progetto	1d4bbed70b794917acaaaa69990873fb					
Istanze attive:	17					
Utilizzo totale VCPU (Ore):	922595	70				
RAM totale attiva (MB):	720896					
Utilizzo totale memoria (Ore):	4332019271	6				
Dimensione totale disco (GB):	425					
Utilizzo totale disco (Ore):	2707420	63				
Nome Istanza	VCPU	RAM (MB)	Disco (GB)	Utilizzo (Ore)	Age (Seconds)	Stato
MuonC_Laura	8	32768	25	5321,05	19155769	Attivo
MuonC_05	8	32768	25	2686,98	9673134	Attivo
MuonC_Lorenzo	8	16384	25	7074,38	25467760	Attivo
UI-1	8	8192	25	1198,62	4315032	Attivo
MuonC_UI	8	8192	25	4696,42	16907113	Attivo
MuonCServer	2	4096	25	8767,55	37299810	Attivo
MuonC_04	8	32768	25	4989,79	17963236	Attivo
Whizard_MPI	8	16384	25	5035,81	18128920	Attivo
MuonC_01	8	32768	25	6041,19	21748285	Attivo
ILC_Nazar	8	32768	25	8767,55	54504579	Attivo
MuonC_00	8	32768	25	7054,75	25397087	Attivo
MuonC_Monster	32	348160	25	5995,69	21584491	Attivo
MuonC_Stream	8	32768	25	2686,92	9672929	Attivo
Whizard2	4	8192	25	8767,55	56561194	Attivo
MuonC_Users	8	32768	25	5061,81	18222515	Attivo
MuonC_02	8	32768	25	6041,16	21748179	Attivo
MDI	8	16384	25	8767,55	35769995	Attivo

Istanze attive:	17
RAM attiva:	704GB
Ore-VCPU di questo perio...	922595,42
Ore-GB di questo periodo:	2707419,85
Ore-RAM di questo perio...	4332017959,86



RAM

Used 704GB of 740GB

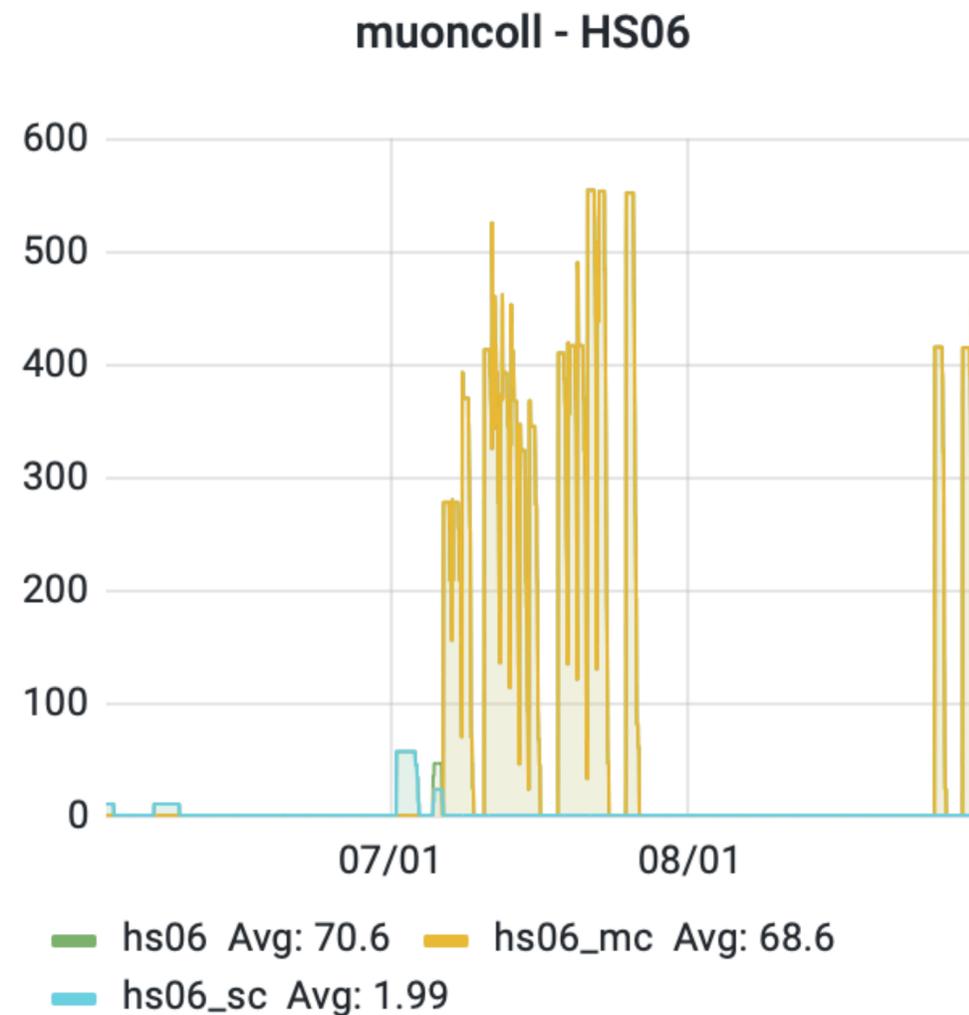
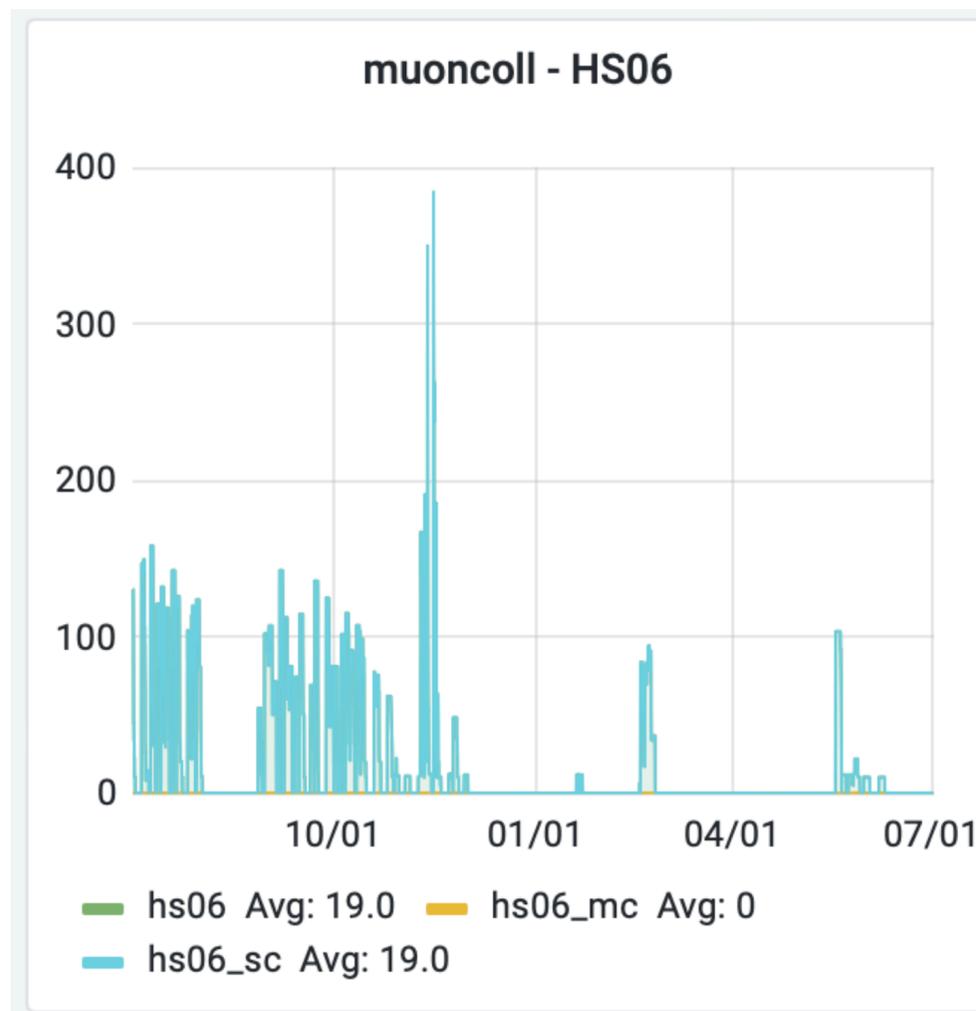


Archiviazione del volume

Used 87,9TB of 87,9TB

- Macchine acquistate principalmente da Università di Padova
- Grazie alla flessibilità della Cloud (ad es. nessun limite ad allocazione RAM o al tempo massimo di esecuzione), riusciamo a girare qualunque tipo di job
- Spesso le risorse Cloud-Veneto sono al limite, sia per l'esaurimento dello storage che per le CPU occupate

- Attuale occupazione Disco: 100 TB/150 TB



- Utilizzato per la simulazione di BIB e come storage
- Attualmente sono disponibili circa 560 bunch-crossing simulati di BIB
- Ricontrati problemi dovuti al limite della RAM per job, o ad errori di time-out

IBISCO-Bari

- **Le risorse di IBISCO-Bari sono state effettivamente disponibili dall'inizio di quest'anno (2022)**
- La richiesta del 2022 era stata calcolata sul progetto proposto dal gruppo di Bari
- Destinata alla produzione campioni di b, c e light jets (full simulation + ricostruzione con BIB) per studiare algoritmi di ricostruzione/identificazione basati su machine learning, e a campioni per lo sviluppo di HCAL-gas
- Richiesta 2022: 7k HS06 e 300 TB
- Le risorse sono attualmente condivise con gli altri progetti, vengono allocate al momento della loro richiesta
- L'accesso avviene tramite account a Bari, può essere aperto anche da utenti di altre sedi
- Per il 2023 si richiede il mantenimento delle risorse già allocate

Campioni simulati (Settembre 2021)

Campioni full simulation ricostruiti (segnale a 3 TeV+BIB):

- $H \rightarrow bb$ 3k eventi in totale (1.5k eventi in due configurazioni diverse)
- $\mu\mu \rightarrow bb$ 3k eventi in totale (1.5k eventi in due configurazioni diverse)
- $\mu\mu \rightarrow cc$ 1.5k eventi
- $\mu\mu \rightarrow$ light jets 1.5k eventi

Campioni per studi di tracking (particle guns):

- muons: 1800k particelle + 60k particelle ricostruite con il BIB
- pions: 240k particelle
- electrons: 120k particelle
- photons: 100k particelle

Campioni con solo segnale (full simulation + ricostruzione senza BIB):

- $HH \rightarrow (4b)$ 10k eventi per 7 valori diversi del coupling (70k in totale)
- 10k eventi di $\mu\mu \rightarrow 4b$ + 10k eventi di $\mu\mu \rightarrow H b b$ (fondi HH)
- 60k eventi di $\mu\mu \rightarrow bb$
- 60k eventi di $\mu\mu \rightarrow cc$
- 60k eventi di $\mu\mu \rightarrow$ light jets
- 100k eventi $\mu\mu \rightarrow \mu\mu$ (per studi misura luminosità)
- 100k eventi $\mu\mu \rightarrow \mu\mu$ Bhabha (per studi misura luminosità)
- 40k eventi di $H \rightarrow \mu\mu$ + 4M eventi fondo
- Dark Photons: 60k eventi di segnale + 420k eventi di fondo

Campioni simulati (Set. 2021 - Set. 2022)

Campioni full simulation ricostruiti (segnale a 3 TeV+BIB):

- H → WW: 10k eventi
- Fondi H → WW 40k eventi
- H → cc: 20k eventi
- Z → cc: 10k eventi
- H → bb: 10k eventi
- Z → bb: 10k eventi
- HH → bbbb: 80k eventi (10k per ogni ipotesi di trilinear coupling)
- Fondi HH: 20k eventi
- Campioni per misura Higgs width: 53k eventi (H → WW, H → ZZ on-shell ed off-shell, fondi)
- Campioni per studi su calorimetro Crilin: 20k eventi (jets, fotoni)

273k eventi ricostruiti di full simulation nell'ultimo anno
Anno precedente: 9k eventi di full simulation (+60k di particle gun)
(Ma nessuna richiesta di calcolo è stata fatta per il 2022)

Lista campioni: <https://confluence.infn.it/display/muoncollider/Monte+Carlo+Simulated+Samples>

Attività 2022 - 2023

- **Analisi in corso/pianificate**

- Misura della Higgs width (>100k eventi richiesti)
- $H \rightarrow \gamma\gamma$
- Studi di Flavour tagging
- Analisi ad 1 TeV

- **MDI e generazione beam-induced background**

- Produzione BIB a 3 TeV con Fluka
- Produzione BIB a 10 TeV con Fluka
- Simulazione BIB a 3 e 10 TeV

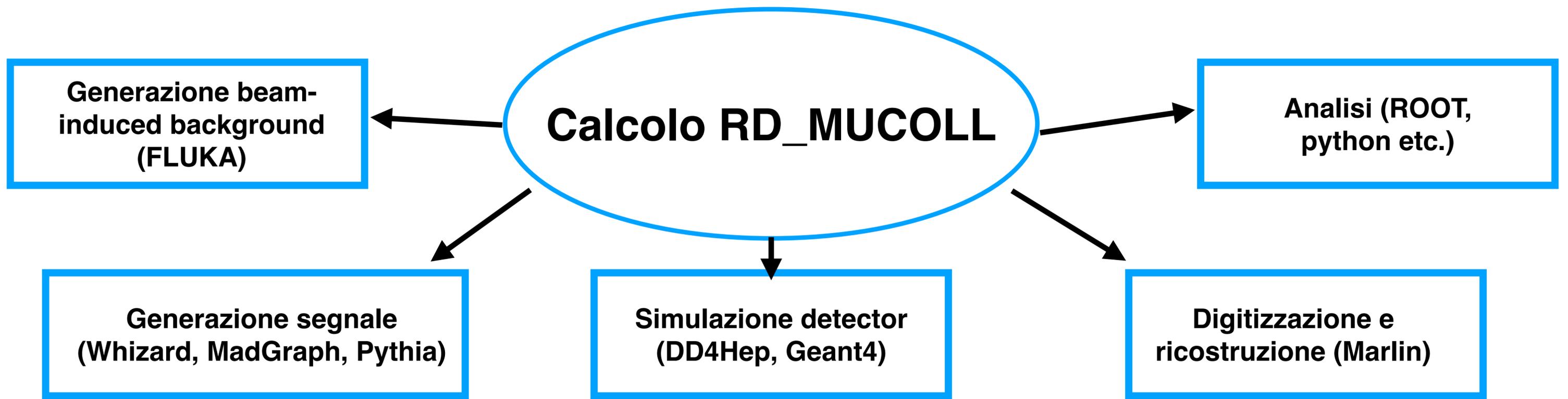
- **Sviluppo detector**

- Crilin ECAL
- HCAL gas
- Muon Detector (picosec)
- Tracking
- Studi per detector a 10 TeV

Con l'espandersi della collaborazione prevediamo un intensificarsi delle attività. Ci aspettiamo inoltre un periodo di sviluppo di codice, algoritmi e design detector (Software in continua evoluzione, difficile definire una versione di produzione)

Richieste 2023

- **Richiesta di calcolo per RD_MUCOLL è di un server per la Cloud-INFN**
- Confrontabile in termini di CPU e storage con quella già acquistata dall'Università di Padova per la Cloud-Veneto: **512GB di RAM e spazio disco di 300 TB** (circa 150 TB effettivamente disponibile per lo storage).
- **costo previsto di 50k euro** (possibile variazione del prezzo del 20% dovuta agli aumenti generali)
- **le risorse di tipo Cloud sono quelle che meglio si adattano al nostro modello di calcolo e le risorse Cloud-Veneto già a nostra disposizione sono al limite**
- ci aspettiamo un intenso periodo di sviluppo di codice, sia per algoritmi di ricostruzione che per studio di detector
- Avere nuove risorse sulla Cloud-INFN ci aiuterebbe anche a condividerle con i colleghi di tutte le sezioni
- Richiesta di 87k euro per IBISCO-Bari riguarda il mantenimento delle risorse già assegnate lo scorso anno (non gravano sul bilancio di commissione)



Risorse disponibili per calcolo RD_MUCOLL

- Cloud-Veneto: 200 VCPU, 740 GB di RAM, 90 TB di disco. Generazione beam-induced background, generazione segnale, simulazione detector, ricostruzione, analisi.
- CNAF: simulazione detector, storage (150 TB). 6 CE disponibili.
- CERN: nuove risorse in pre-produzione, storage su EOS (100 TB)
- IBISCO-Bari: studi HCAL, jet flavour tagging



Cloud-Veneto come risorsa principale, grazie alla sua flessibilità. Non più sufficiente con l'espandersi della collaborazione



Per il 2023 richiediamo l'acquisto di una macchina su Cloud-INFN