MadGraph Profiling

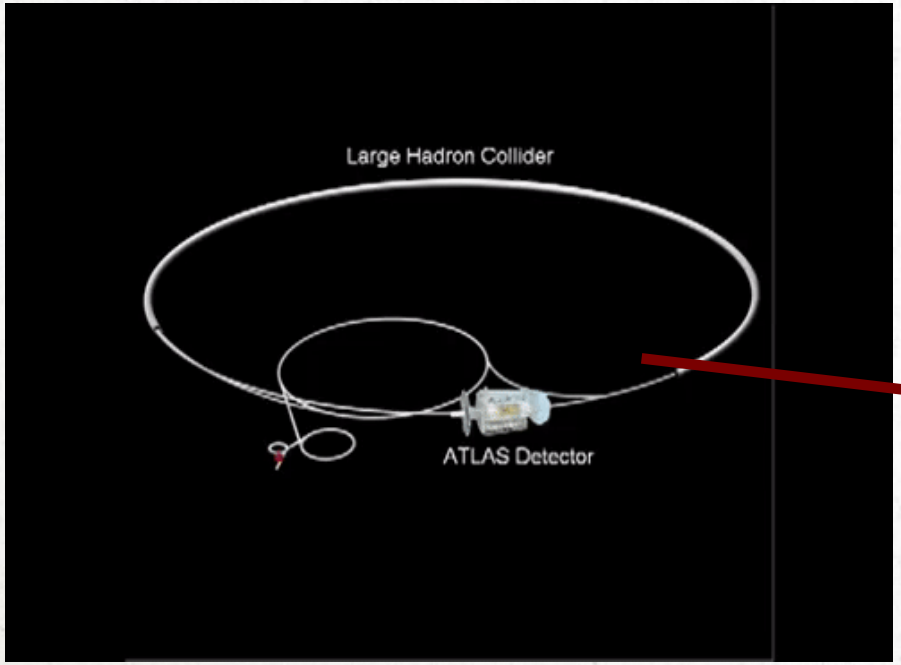
Smita Darmora, Walter Hopkins

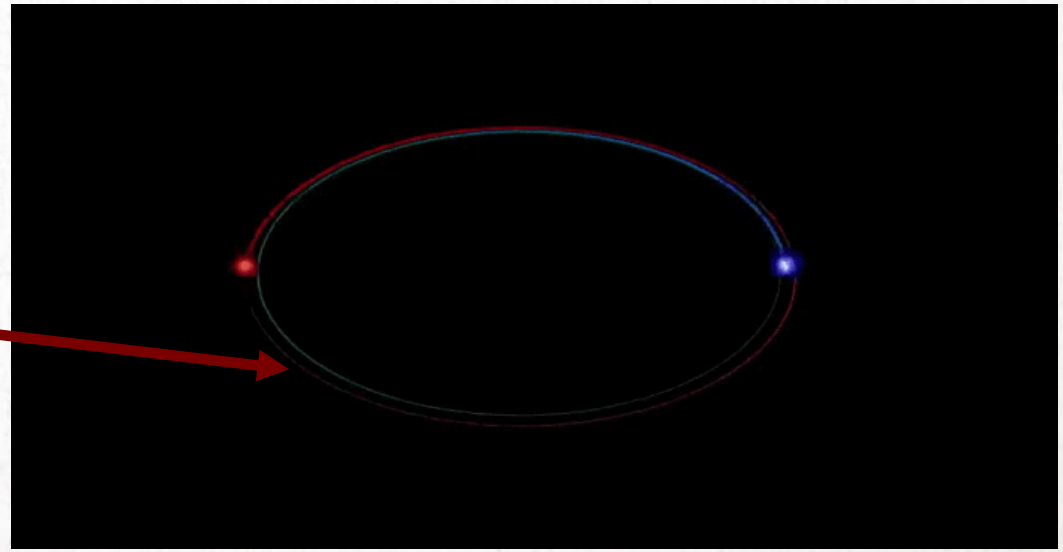
Young Scientist Symposium Series (YSSS)

December 15, 2020

Experiment

ATLAS experiment at the Large Hadron Collider (LHC), a particle accelerator at CERN (the European Organization for Nuclear Research)

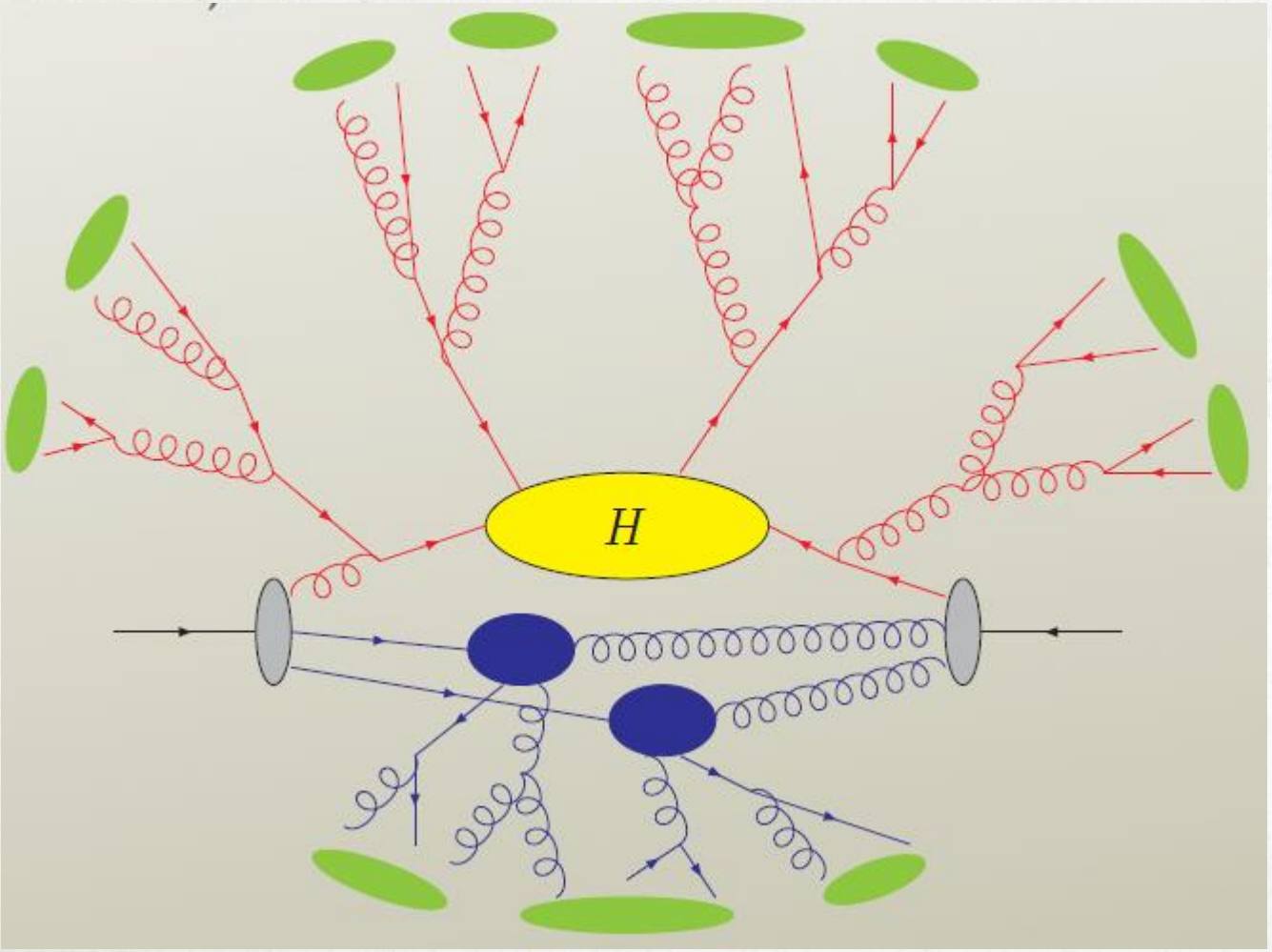




Introduction

- Simulation is a very useful, essential tool in modern particle physics for:
  - designing an experiment (e.g. ILC/CLIC, FCC)
  - analyzing the data (e.g. LHC experiments)
- For the LHC experiments, the simulation is made of two distinct steps:
  1. **Simulation of the p-p collision**
     - Monte Carlo event generators
  2. **Simulation of the passage of the produced particles through the experimental apparatus**
     - Detector simulation
     - The output of 1 is the input of 2.

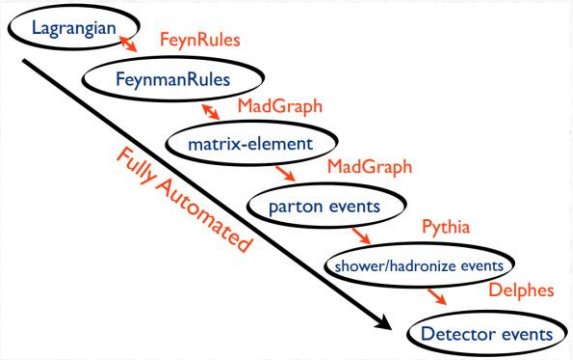
Description of an event



1. Incoming hadrons /particles (grey)
2. Hard part of the process (Yellow)
3. Parton shower (Red)
4. Underlying event (blue)
5. Hadronization (green)

MadGraph

- *Software that simulates collisions of particles, including the kind of very high-energy proton-proton collisions at the Large Hadron Collider (Python)*
- Integrated with an event generator, and with a program called Pythia for parton showering and hadronization, and with a choice of detector simulators
- *HEP Phenomenology framework*
  - *Feynman Diagrams*
  - *Cross Sections*
  - *Event Generation*



Users Requests:

MadGraph Returns:

- *List all contributing sub processes*
- *Feynman diagrams for each sub process*
- *Self-Contained Package:*
  - *Integrates cross section*
  - *Generates unweighted parton-level events*

Motivation

- *We want to be fast*

MadGraph Profiling

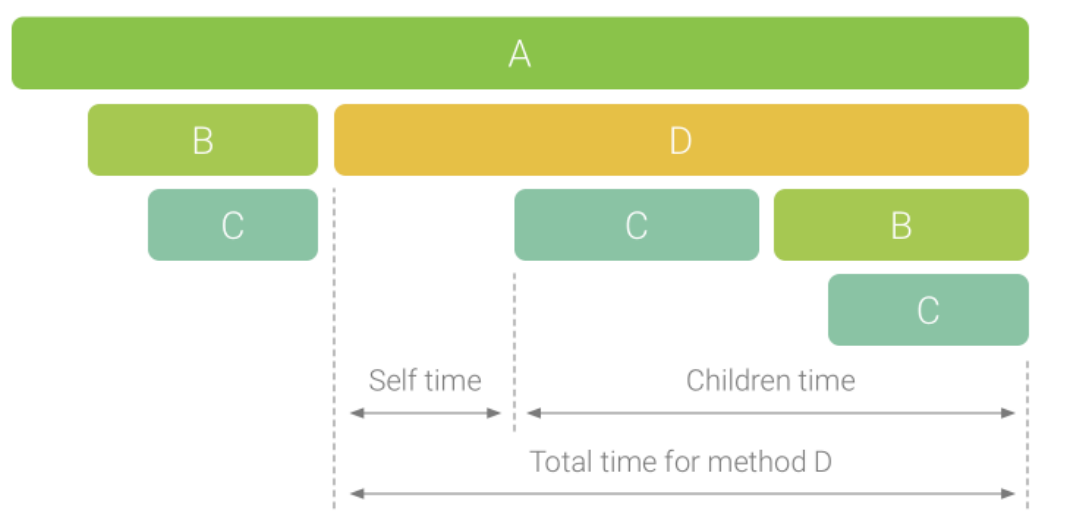
- *find the hotspot in the code and figure out what is using most of the CPU time.*
- *Intel VTune Amplifier, a profiling tool.*

Intel® VTune™ Profiler

- *Where do we spend our time?*
  - *In which function?*
  - *Called by whom?*
  - *Which instruction take time?*
  - *Where to parallelize?*

Profiling Terminology

- *Self-Time* is the time spent inside a function excluding the time taken by other functions called from within that function.
- *Total Time* is the accumulated time that a program unit incurs. For functions, total time includes self-time of the function itself and Self time of all functions that were called from that function.
- *Hotspot* is the most time-consuming parts of the code.



Process:

Production of single electroweak bosons (Z) in association with jets in proton–proton collisions at center-of-mass energy of 13 TeV

$$\mathbf{Z \rightarrow \mu^{+}\mu^{-} + X,\ where\ X = 3\ jets}$$

- **Profilers:** Intel VTune and perf
- **Visualization:** Flame Graph and Call Graph
- **Steps:**
  - Compile without profiler on 1 core
  - Generate different number of events with Vtune and perf profiler.

Top hotspots (Total time and Self time)

nevents : 10000

Function	CPU Time: Total	CPU Time: Self	Module
__libc_start_main	99.8%	0.0%	libc.so.6
LHAPDF::mkPDF	33.1%	0.0%	libLHAPDF.so
LHAPDF::GridPDF::_loadData	32.0%	0.7%	libLHAPDF.so
____strtod_l_internal	22.4%	22.4%	libc.so.6
_start	11.2%	0.0%	madevent_mintMC
main	11.2%	0.0%	madevent_mintMC
MAIN__	11.2%	0.0%	madevent_mintMC
__multf3	10.9%	10.4%	libgcc_s.so.1
gen_	10.8%	0.0%	madevent_mintMC
sigintf_	10.8%	0.0%	madevent_mintMC
main	8.2%	0.0%	madevent_mintMC
_start	8.2%	0.0%	madevent_mintMC
MAIN__	8.2%	0.0%	madevent_mintMC
gen_	7.8%	0.0%	madevent_mintMC
sigintf_	7.8%	0.0%	madevent_mintMC
MAIN__	7.7%	0.0%	madevent_mintMC
_start	7.7%	0.0%	madevent_mintMC
main	7.7%	0.0%	madevent_mintMC
gen_	7.3%	0.0%	madevent_mintMC
sigintf_	7.3%	0.0%	madevent_mintMC
__addtf3	6.8%	6.6%	libgcc_s.so.1
compute_nbody_noborn	6.7%	0.0%	madevent_mintMC
bornsoftvirtual	6.7%	0.0%	madevent_mintMC
binothlha_	6.7%	0.0%	madevent_mintMC
sloopmatrix_	6.6%	0.0%	madevent_mintMC
MAIN__	5.2%	0.0%	madevent_mintMC
main	5.2%	0.0%	madevent_mintMC
_start	5.2%	0.0%	madevent_mintMC
gen_	5.0%	0.0%	madevent_mintMC
sigintf_	5.0%	0.0%	madevent_mintMC

__libc_start_main	99.9%	0.0%	libc.so.6
LHAPDF::mkPDF	19.1%	0.0%	libLHAPDF.so
LHAPDF::GridPDF::_loadData	18.5%	0.4%	libLHAPDF.so

Function	CPU Time: Total	CPU Time: Self	Module
____strtod_l_internal	22.4%	22.4%	libc.so.6
__multf3	10.9%	10.4%	libgcc_s.so.1
__addtf3	6.8%	6.6%	libgcc_s.so.1
getrusage	4.8%	4.8%	libc.so.6
__subtf3	3.4%	3.3%	libgcc_s.so.1
pow	2.2%	2.2%	libm.so.6
std::getline<char, std::char_trait	1.9%	1.9%	libstdc++.so.6
__ieee754_log_avx	1.7%	1.7%	libm.so.6
std::string::find	1.5%	1.5%	libstdc++.so.6
std::use_facet<std::ctype<char>	1.2%	1.2%	libstdc++.so.6
std::string::assign	1.0%	1.0%	libstdc++.so.6
ffv1_0_	0.9%	0.9%	madevent_mintMC
sread	1.4%	0.8%	libgfortran.so.3
std::vector<double, std::allocat	0.8%	0.8%	libLHAPDF.so
ffv1_0_	0.8%	0.8%	madevent_mintMC
__sfp_handle_exceptions	0.7%	0.7%	libgcc_s.so.1
LHAPDF::GridPDF::_loadData	32.0%	0.7%	libLHAPDF.so
ffv1_0_	0.6%	0.6%	madevent_mintMC
ffv1_0_	0.6%	0.6%	madevent_mintMC
ffv1_2_	0.6%	0.6%	madevent_mintMC
ffv1_1_	0.5%	0.5%	madevent_mintMC
ffv1_2_	0.5%	0.5%	madevent_mintMC
__memset_sse2	0.5%	0.5%	libc.so.6
read	0.5%	0.5%	libc.so.6
next_char	0.9%	0.5%	libgfortran.so.3
calloc	0.5%	0.5%	libc.so.6
__divtf3	0.5%	0.5%	libgcc_s.so.1
calliter_iternext	0.5%	0.5%	libpython2.7.so.1.0
ffv1_2_	0.4%	0.4%	madevent_mintMC
boost::multi_array<double, (uns	0.5%	0.4%	libLHAPDF.so

__multf3	10.6%	10.2%	libgcc_s.so.1
getrusage	7.3%	7.3%	libc.so.6
__addtf3	6.4%	6.2%	libgcc_s.so.1

Important functions:

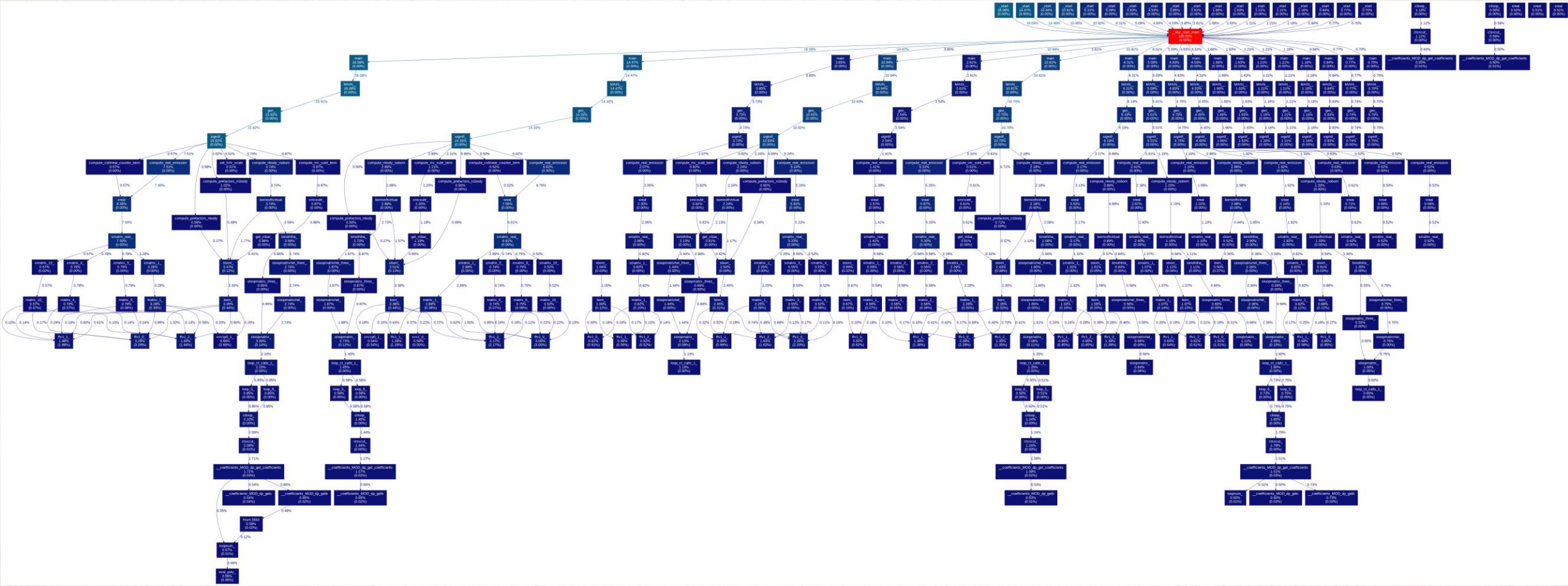
- *Particle Density function (PDF): a crucial input into cross-section calculations at hadron colliders*

- *LHAPDF: is a library by which PDFS are accessed for LHC experimental*

(general purpose C++ interpolator, used for evaluating PDFs from discretized data files)

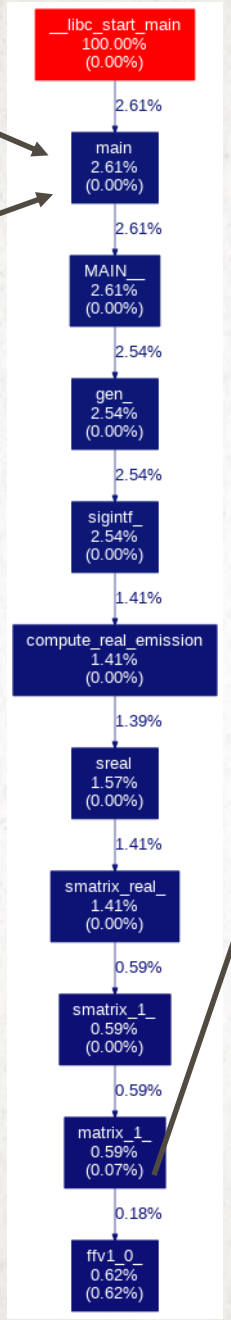
- *getrusage:*
  - *Monitoring function*
  - *used to examine the resource usage of a process*
  - *a function into a code to profile it.*

Call Graph



Total time

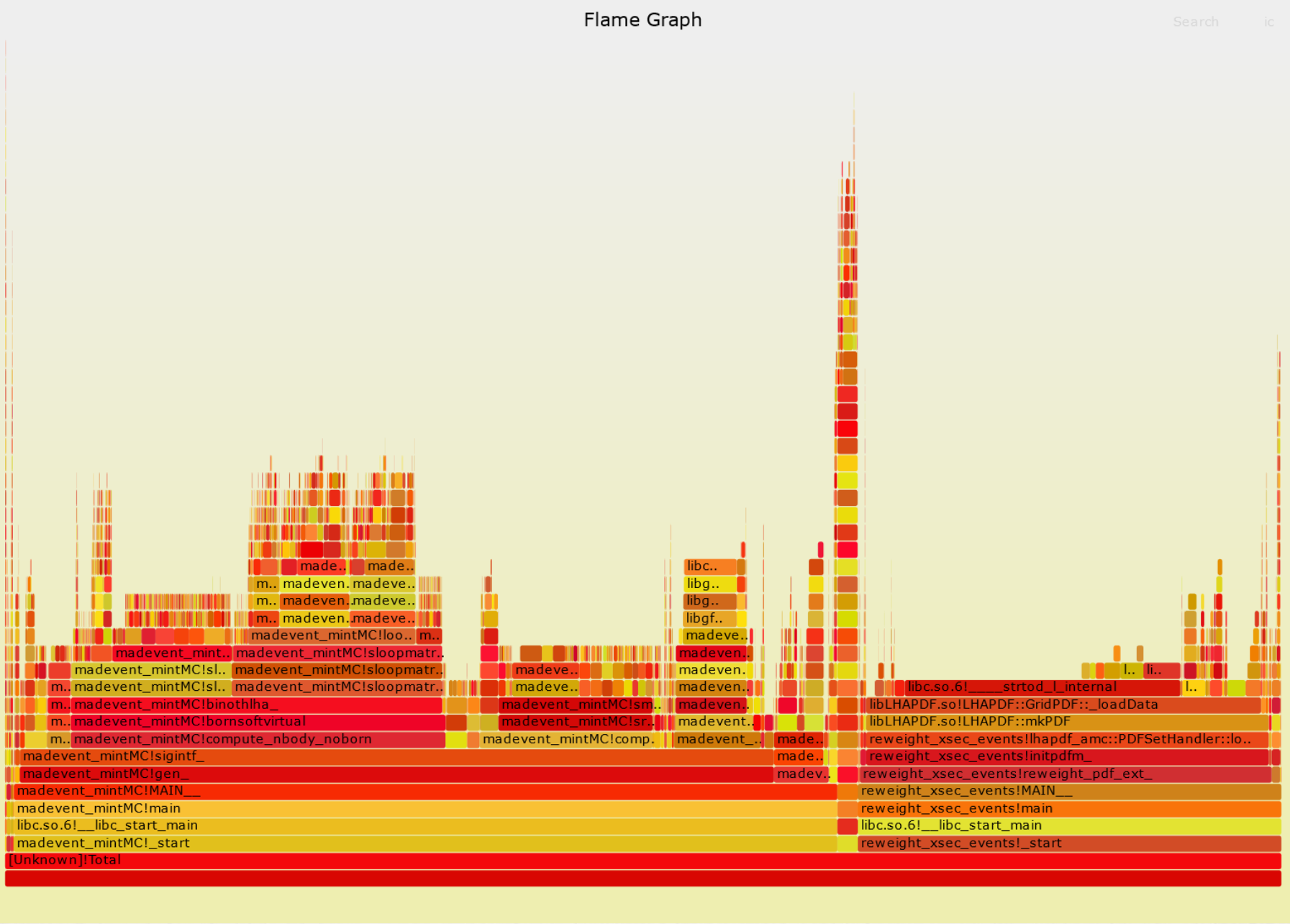
Self time



Function	CPU Time: Total ▼	CPU Time: Self	Module
smatrix_1_	0.6%	0.0%	madevent_mintMC
matrix_1_	0.6%	0.1%	madevent_mintMC
loop_5_	0.6%	0.0%	madevent_mintMC

Callers	CPU Time: Total	CPU Time: Self ▼
▼ matrix_1_	100.0%	12.5%
▶ smatrix_1_	100.0%	12.5%

Callees	CPU Time: Total
▼ matrix_1_	100.0%
▶ ffv1_0_	34.5%
▶ ffv1_2_	21.6%
▶ ffv1_1_	9.4%
▶ vvv1_0_	6.0%


Flame Graph
Search
ic
made..
made..
libc..
m..
madeven..
madeve..
libg..
libgf..
madevent_mintMC!loo..
madevent_mint..
madevent_mintMC!sloopmatr..
madeven.
l..
li..
madevent_mintMC!sl..
libc.so.6!____strtod_l_internal
madevent_mintMC!binothlha_
madevent_mintMC!sm..
libLHAPDF.so!LHAPDF::GridPDF::_loadData
madevent_mintMC!bornsoftvirtual
madevent_mintMC!sr..
madevent..
libLHAPDF.so!LHAPDF::mkPDF
madevent_mintMC!compute_nbody_noborn
madevent_mintMC!comp..
madevent_..
made..
reweight_xsec_events!lhapdf_amc::PDFSetHandler::lo..
madevent_mintMC!sigintf_
reweight_xsec_events!initpdfm_
madevent_mintMC!gen_
madev..
reweight_xsec_events!reweight_pdf_ext_
madevent_mintMC!MAIN__
reweight_xsec_events!MAIN__
madevent_mintMC!main
reweight_xsec_events!main
libc.so.6!__libc_start_main
madevent_mintMC!_start
reweight_xsec_events!_start
[Unknown]!Total

Summary

1. Analyzed different process with Vtune profiler.
2. Highlighted top hotspot MadGraph functions.
3. LHAPDF::makePDF has a high contribution.
4. getrusage, examine the resource usage of a process, takes a significant amount of time
5. MadGraph function visualization.

Thank you