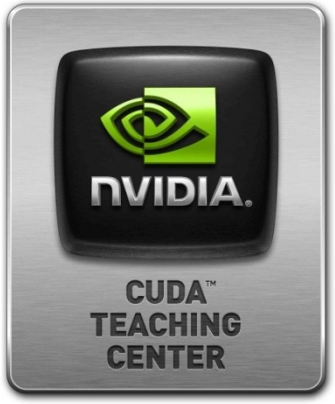
Computación de Propósito General en Unidades de Procesamiento Gráfico (GPGPU)

E. Dufrechou, P. Ezzatti y M. Pedemonte

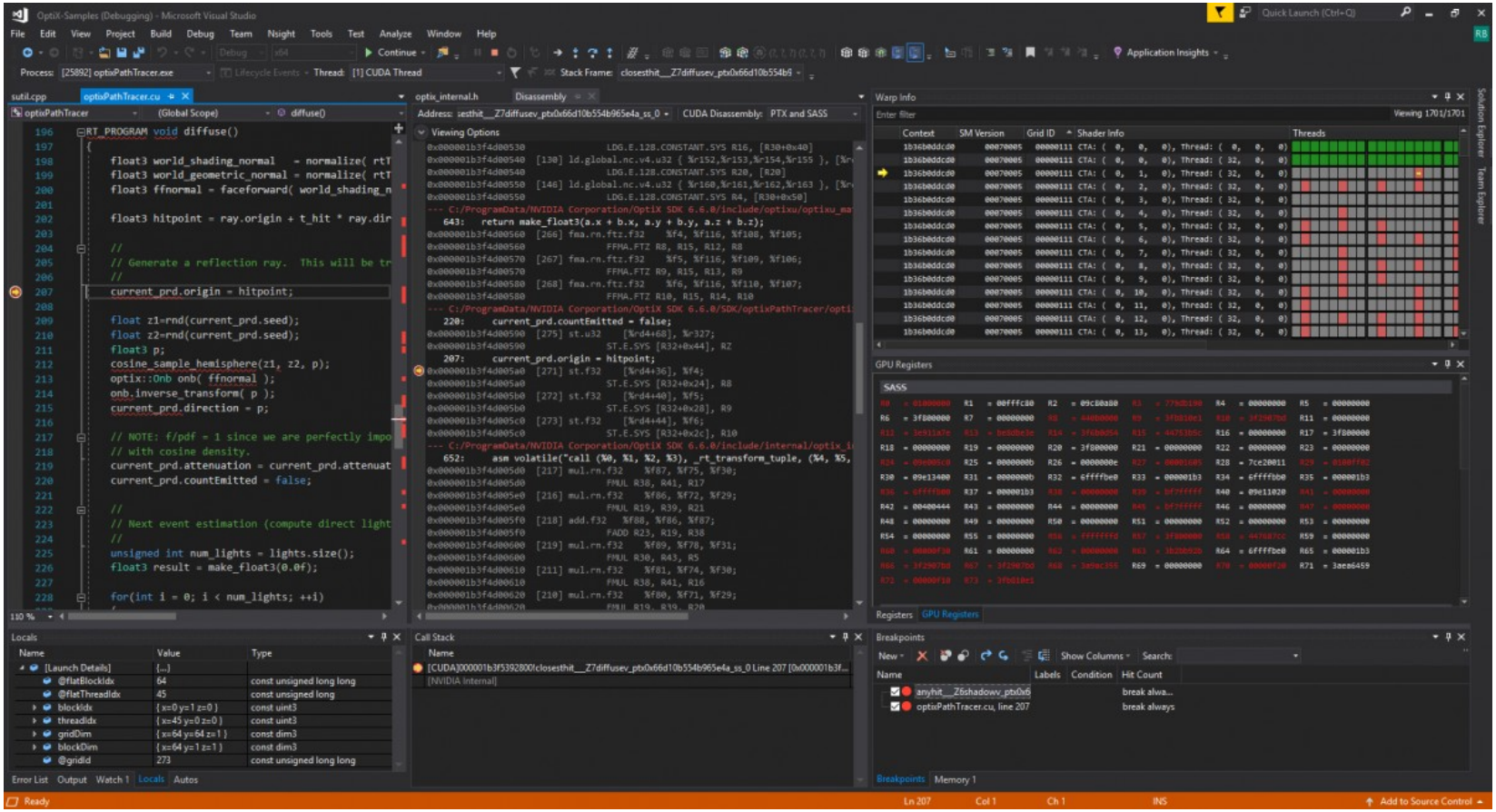


Clase 9:
Debugging y Profiling

Debugging

Distintas posibilidades para debugging de CUDA (https://developer.nvidia.com/tools-overview):

- **Plugins para Visual Studio, Eclipse, VS Code**
- **CUDA-GDB debugger de línea de comandos (Linux y Mac)**



CUDA-GDB

Funcionamiento similar a GDB:

- **Debugging simultáneo en CPU y GPU**
- **Breakpoints**
- **Inspección de memoria, registros, variables en memoria local/shared/global**
- **Soporta multiples GPUs, contextos, y kernels**
- **Debugging a nivel de código fuente y ensamblador (SASS)**
- **Es necesario compilar con las siguentes opciones: -g -G**

CUDA-GDB

Algunos comandos:

- run, continue, kill
- next (no entra en funciones), step (entra en funciones)
- break (pone un breakpoint)
  - break my_kernel, break _Z6kernelIfiEvPT_PT0, break acos.cu:380, break *0x3e840a8, break *$pc, set cuda break_on_launch application
- cuda (cambia el foco a un kernel/bloque/thread específico)
  - cuda kernel 2 block 1,0,0 thread 3,0,0
- info cuda devices/kernels/threads (muestra devices, kernels, threads activos)
  - info cuda threads kernel 2
- print (leer o escribir variables/direcciones de memoria)
  - print var, print &var, print var = 3

compute-sanitizer

Varias herramientas en una:

- Memcheck – accesos fuera de rango y desalineados
- Racecheck – condiciones de carrera en memoria compartida
- Initcheck – acceso a memoria global no inicializada
- Synccheck – uso incorrecto de sincronización

- No necesita compilar con flags especiales pero…
- para mostrar nombres de funciones en la salida:
  - nvcc -Xcompiler -rdynamic -lineinfo -o out in.cu

CUDA-MEMCHECK

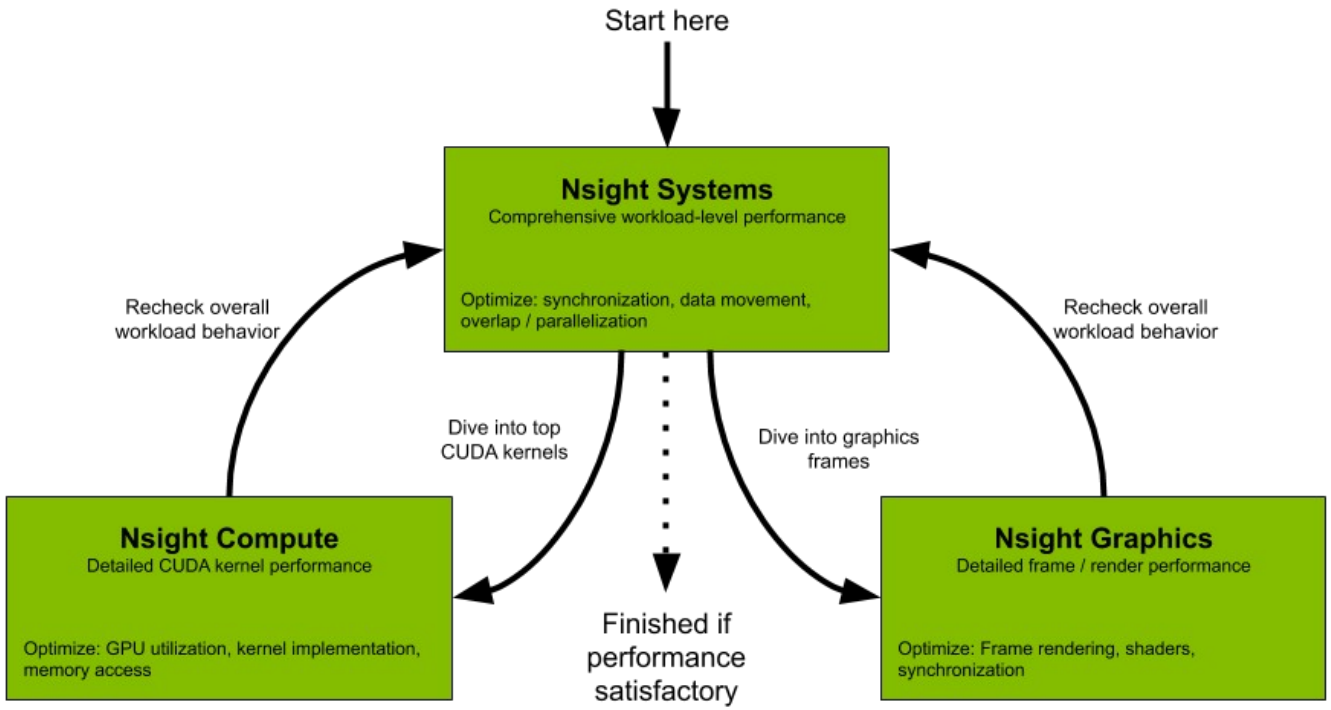
Dos formatos:

- **Standalone**
  - **cuda-memcheck [op] programa params**
- **Integrado con cuda-gdb (Linux y Mac)**
  - **set cuda memcheck on**

- **Util para chequear condiciones de carrera:**
  - **cuda-memcheck --tool racecheck programa params**

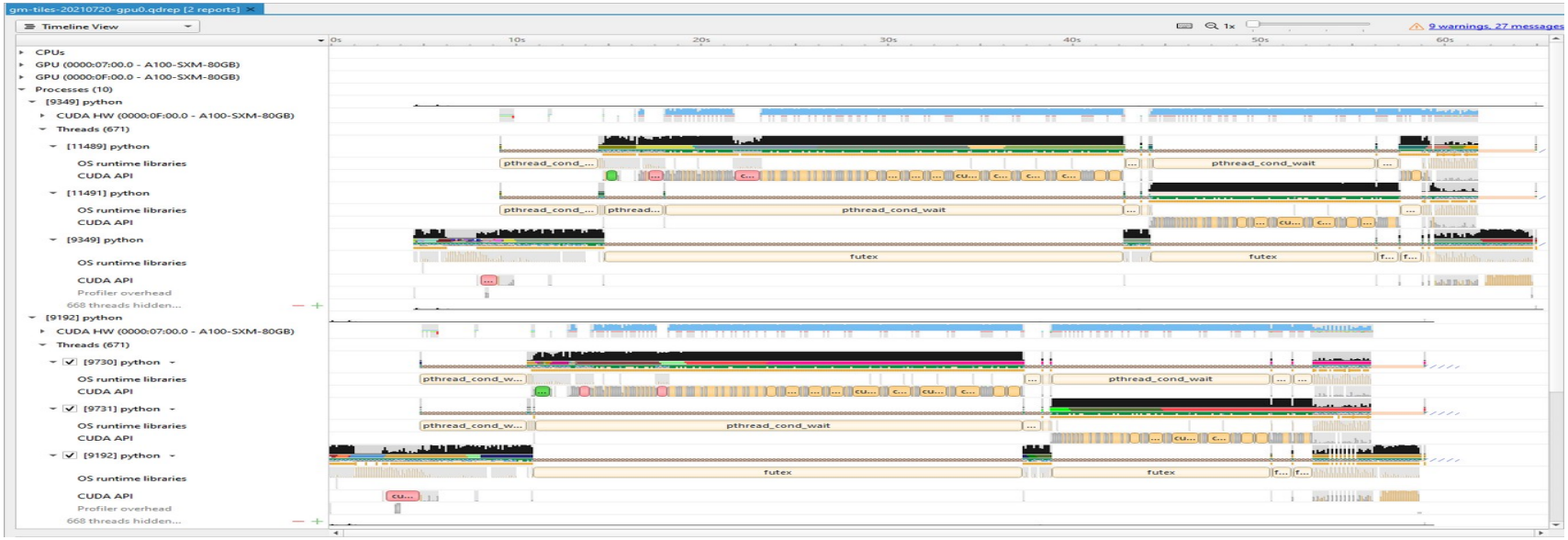
Profiling

- **Nuevas herraminetas: NSight Systems y NSight Compute**
  - **https://developer.nvidia.com/gameworksdownload#?tx=$gameworks,developer_tools**
  - **Ambas cuentan con GUIs y clientes de línea de comandos**



NSight Systems

- **Nos da una visión de todo el sistema:**
  - **CPUs, GPUs, hilos de CPU, streams en GPU**
  - **Ejecución de kernels, transferencias de memoria, etc.**
- **Nos muestra los kernels ejecutados, con los tamaños de grilla, recursos utilizados, etc.**
- **Puede generarse un archivo en línea de comandos y luego importarlo en la GUI**



NSight Systems

- Herramienta de línea de comandos: nsys
- Dos modos:
  - automático (profile)
  - interactivo (start,stop,cancel,launch,etc...)
- Ejemplo automático:

  nsys profile -o salida -f true ./prog params
  - Genera archivo salida.nsys-rep importable en NSight Systems
- Ejemplo interactivo:

  nsys start –stop-on-exit=false

  nsys launch --trace=cuda --sample=none ./prog params

  nsys stop
- --stats true muestra un resumen de estadísticas en formato de texto plano

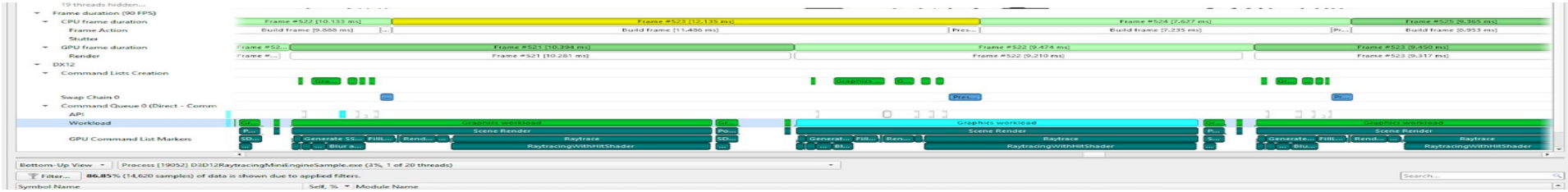
NSight Compute

- Herramienta de línea de comandos: ncu
- Tres modos:
  - Launch-and-attach (profile)
  - Launch / attach (attach permite hacer profiling remoto)
- Ejemplo lauch-and-attach (igual que nsys):

  ncu -o salida -f ./prog params
  - Genera archivo importable en NSight Compute
  - Por defecto extrae un conjunto pequeño de métricas
  - ncu --list-sets (lista los sets de métricas disponibles)
  - ncu -o prof_ncu -f --set detailed/full

NVTX

- Anotar y delimitar zonas del código
- Útil para poder visualizar la ejecución de un programa con varios kernels usando NsightSystems
- Se delimitan usando las llamadas a nvtxRangePush nvtxRangePop
- No genera overhead
- Los rangos o zonas pueden anidarse



Recursos

- Descargar herramientas (es necesario crear una cuenta en Nvida Dev Zone):
  - https://developer.nvidia.com/gameworksdownload
- Documentación:
  - https://docs.nvidia.com/cuda/cuda-gdb/
  - https://docs.nvidia.com/nsight-systems/
  - https://docs.nvidia.com/nsight-compute/
- Videos:
  - https://www.youtube.com/user/NVIDIADeveloper