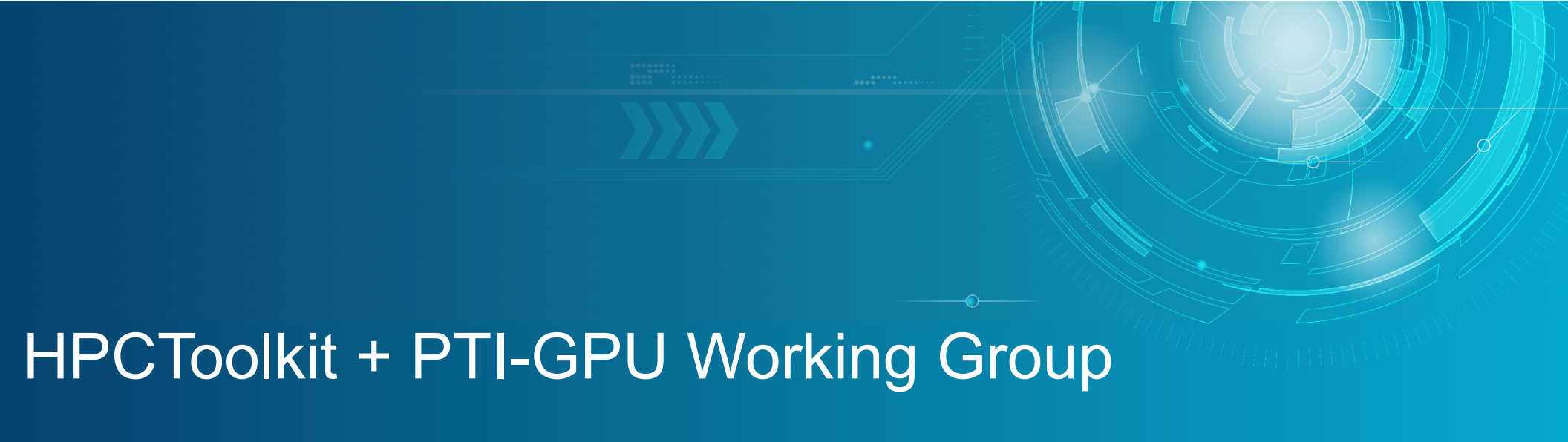


The background of the top half of the slide is a solid dark blue. It features several abstract, glowing light blue graphics. On the right side, there is a large, complex circular structure resembling a futuristic turbine or a multi-layered orbital ring, with various segments and a bright light source at its center. To the left of this, there are horizontal lines and a series of four chevron-like shapes pointing to the right. Small dots and faint circuit-like patterns are scattered across the background.

HPCToolkit + PTI-GPU Working Group

Scalable Tools Workshop 2026

Discussion topics

- **Associating external correlation records with GPU operations**
- **Tool initialization**
- **Correlation IDs for PC samples**

Associating correlation IDs with memory copy and kernel launch

- **Design and use of external correlations**

- PTI-GPU API supports push/pop of correlation IDs before/after invocation of GPU operations
- HPCToolkit uses correlation ID push/pop in synchronous PTI-GPU operation invocation callbacks
- PTI-GPU reports ExternalCorrelation records that contain a tuple (external id, internal id)

- **Problem**

- HPCToolkit wasn't seeing ExternalCorrelation records with expected information before GPU operations in activity trace

- **Diagnosis: two issues**

- HPCToolkit contained “hack” that pushed bogus ExternalCorrelation during tool initialization
 - added before PTI supported push/pop in PGI-GPU operation invocation callbacks
 - invocation callbacks pushed second ID on stack
- PTI-GPU reported outer correlation ID and ignored inner correlation

- **Moving forward**

- Removing push of bogus correlation made push/pop of ExternalCorrelation records deliver expected value in Activity Trace
- Intel needs to investigate why PTI-GPU is broken for nested correlations

PTI-GPU lacks a design for coordinating tool initialization

- **Wanted: analogue to rocprofiler_configure**
 - runtime looks for a tool-provided symbol at an appropriate time
 - tool provides initialization/finalization routines back to runtime
 - runtime calls these when appropriate
- **Wanted: analogue to rocprofiler_force_configure**
 - sometimes, a tool needs to trigger initialization before GPU naturally would
 - before main is entered
 - when a thread gets created
 - force runtime to initialize tool support before a runtime would naturally do so

PTI-GPU lacks support for attributing PC samples

- **Ponte Vecchio**

- HW accumulates (PC, stall reason, count) tuples in HW cache
- samples are reported when evicted from cache
- supports PC sampling during concurrent kernels

- **PC samples lack correlation IDs to kernel invocation contexts**

- Intel will investigate whether this can be fixed in software