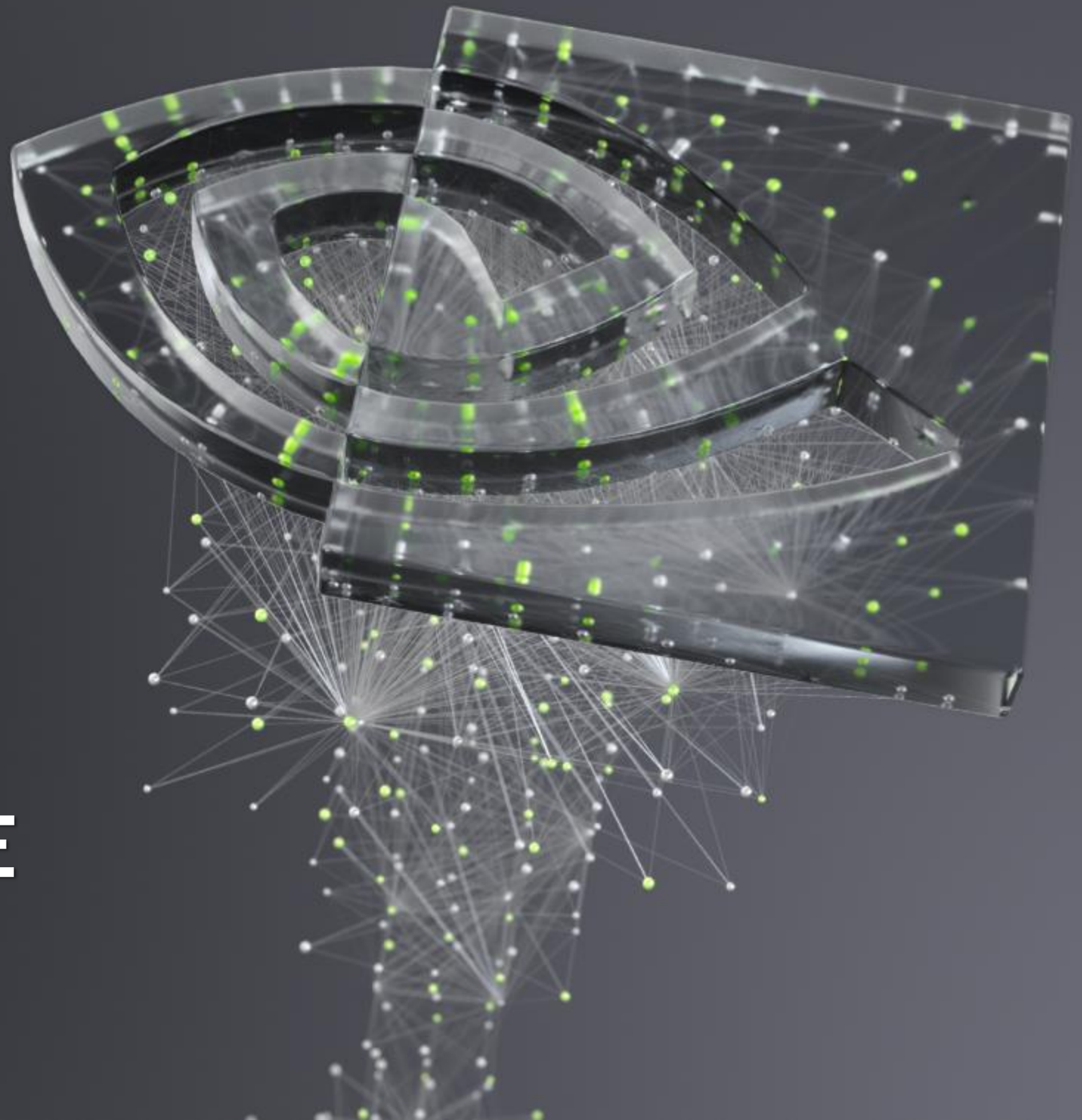


RDMA-CORE UPDATE

Jason Gunthorpe, Dec 1 2020

- Strong velocity in 2021:
  - rdma-core: 604 commits, 43kloc, 72 contributors
  - Linux kernel: 1109 commits, 54kloc, 156 contributors

WEAK ORDERING

- Relaxed ordering PCI support for user space was done in 2020
- On by default for Kernel ULPs now
- Better support for weakly ordered environments via ibv_query_qp_data_in_order()

“Last Data Polling” is becoming increasingly undesirable from a performance standpoint and more configurations are now outright incompatible with the scheme

DMABUF

For GPU Memory & PCI Peer to Peer

mlx5 and EFA can act as DMABUF importers:

- Able to issue DMAs to device memory supplied by the DMABUF exporter
- GPU drivers are the most common exporter
- Replaces the “ib_peer_memory” system previously used
- Mlx5 is fully featured, EFA can only work with “pinning” devices such as Habana Labs

To create a MR for a GPU memory object userspace will obtain a DMABUF file descriptor from the GPU driver then pass it to ibv_reg_mr_dmabuf() to create a MR.

KERNEL

- New Intel irdma driver replaces the old i40iw driver
  - Supports new HW and RoCE
- NDR support
- > 255 device ports
- XRC support for HNS’s HIP09
- Broader support for optimizing HW DMA when contiguous memory is available
- Greatly improved ODP destruction performance

RDMA CORE

Device Memory import/export for sharing

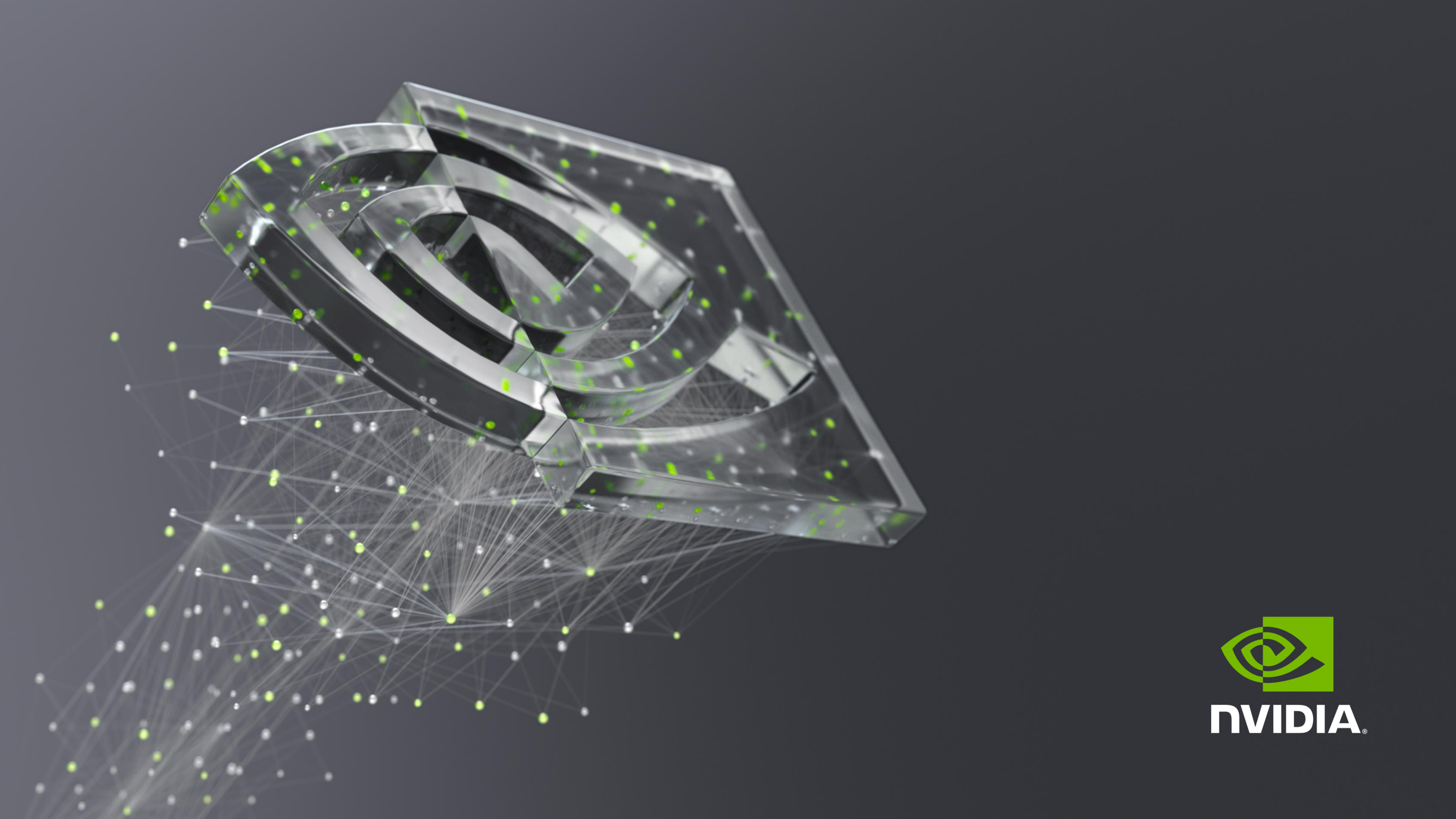
Automatic NOP of ibv_fork_init() when kernel support is present

Ex verbs support in more providers

Lots of pyverbs testing activity and increases in coverage

MLX5 FEATURES

- MEMIC operations
- “Shared DM” via import
- Sophisticated “memory windows” via userspace UMR
- DCT Streams feature
- mlx5 verbs provider directly over VFIO
- Trusted encryption key management for storage


NVIDIA