DynamicXcelerator (DX Fabric)

Key Benefits

- Access to hyperscaler architecture at an enterprise price point
- Better TCO by decoupling GPU resources from server upgrade path
- Deterministic topologies brings the right bandwidth to your workloads
- Lower latencies by using a direct connect all photonic fabric
- Easier to deploy and manage than traditional static, box by box solutions
- Grouping of valuable resources via direct connect resolves the stranded resource challenge
- Workload communication is restricted to defined topologies for secure and private workload slices
- Ride the optical innovation wave for years to come with advances in silicon photonics

Major shifts in the datacenter market led by Generative AI/ML are the driving force behind several challenges for traditional datacenter designs. Public cloud costs and security concerns have driven many companies to begin cloud repatriation to a colo or on-prem datacenter, while many new GPU cloud providers are building specialized cloud infrastructure for accelerator-based applications. Despite the emergence of the **#AcceleratorEra** for AI/ML datacenters, datacenter architects are left with long-established traditional interconnect architectures which are not well suited for the modern AI/ML datacenter.

The **#AcceleratorEra** presents the modern datacenter architects with unprecedented scale challenges. Specialized workflow needs, connectivity, power, cooling, and team skill set requirements hinder AI/ML infrastructure deployment. Traditional vendor solutions force forklift upgrades and fail to address root cause of the accelerator scale out challenge. Legacy vendor solutions dictate infrastructure design, forcing users to squeeze applications into it rather than vice versa. This results in networks built for general purpose compute or the **#HypervisorEra,** increasing cost, power, and complexity, promoting resource inefficiency.

A Better Interconnect Design

Drut's DynamicXcelerator provides customers the freedom to offer datacenter solutions that match their workload requirements as well as budgetary constraints, all within a multi-vendor dynamically reconfigurable computing infrastructure that can be tailored to their specific workload's requirements. Our key technology benefit is referred to as **"vPODs or virtual PODs"** which provides the ability to perform dynamic slicing of datacenter resources based on software workloads.

200 Innovative Way, Suite 1360
Nashua, NH 03062

Features

- GPU cloud at scale using standard datacenter devices
- Connects accelerator to accelerator, GPU to GPU, at small and large scale using RDMA over Photonics
- Scales up but with less cost than traditional electrical packet switching solutions
- Using software to define topologies
- Builds many dynamic GPU topologies
- Software integration with HPC and AI software tools
- Full disaggregated photonic solution for the data center
- Integrates with existing ethernet based networks
- Custom software integration available at the software layer for in house developed applications
- Uses open source software such as K8s, OpenStack, Ceph and many others
- Bare metal turn up and server OS deployment capability

Dynamic slicing is the ability to create sharded topologies from a larger environment and tuning the shards to the application's actual need (i.e. known patterns), versus traditional infrastructure designs that are built statically, without any awareness of flows or model traffic patterns.

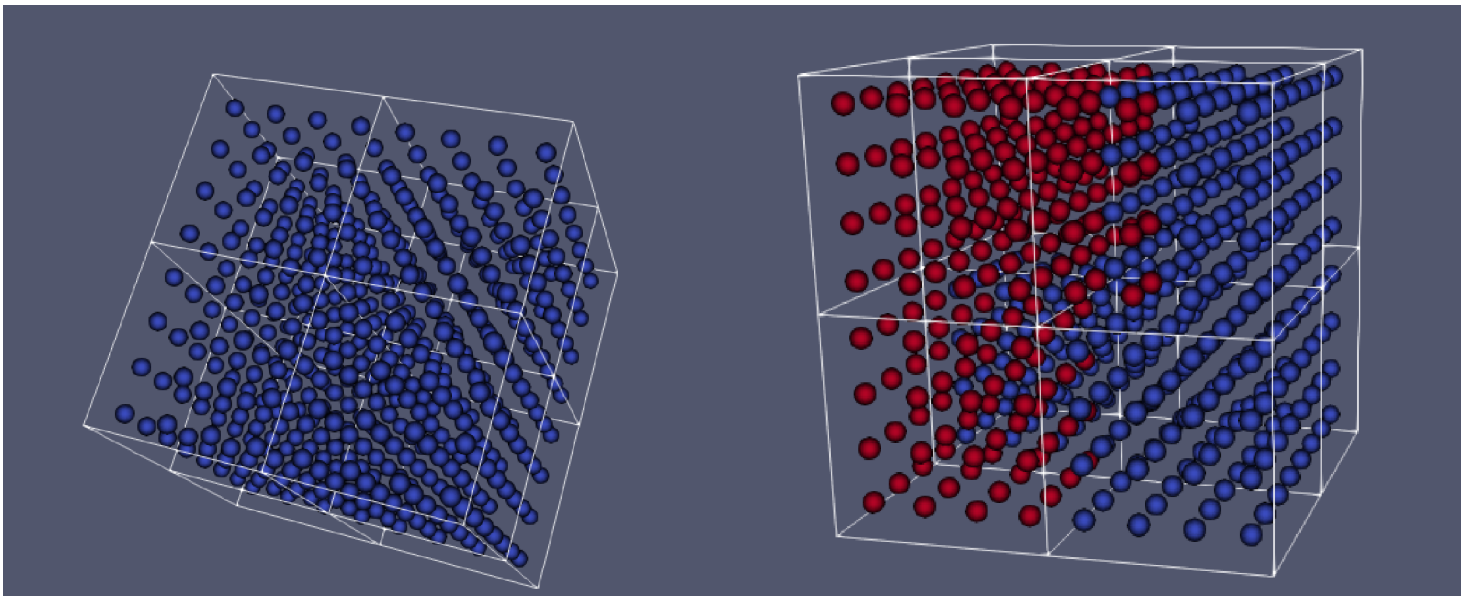
Bringing Brilliance to AI with the Power of Photonics

All this is made possible with the innovative use of Drut's photonic fabric. The DynamicXcelerator is a protocol agnostic connectivity solution, allowing dynamically reconfigurable low latency direct paths between various resource units inside a datacenter. Our industry has used optics in point-to-point links and talked about photonics for many years, but breakthrough advances in photonic technology now allow for it to be built at enterprise price and datacenter scale. There are many industry reasons why connectivity solutions continue to follow the model of "all-to-all" and "spray-and-pray" techniques, it is a combination of how it has always been done and incumbent suppliers protecting the legacy business model, but the AI revolution is changing the dynamics in such a way that this legacy model has now become a hinderance to the industry.

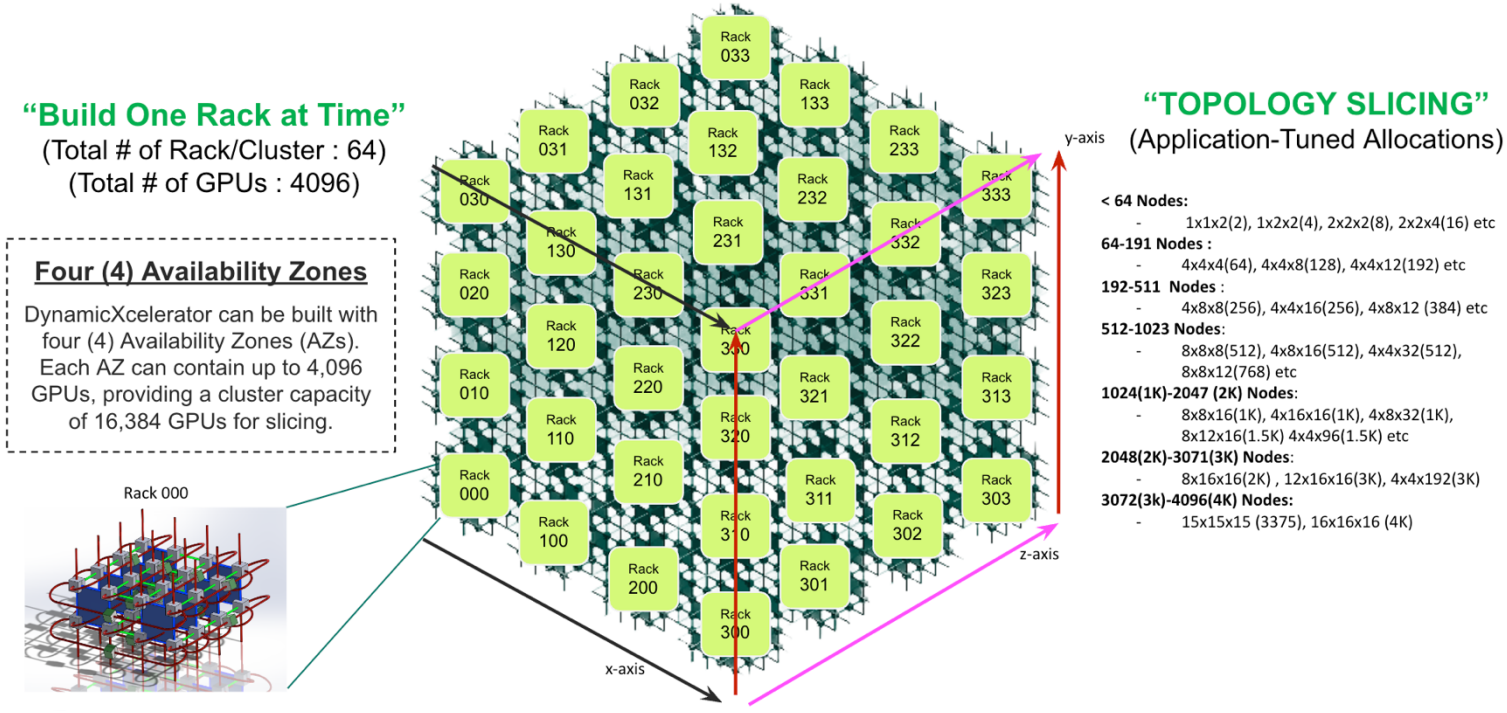
The characteristic that makes AI workloads so well suited to photonics is that the nature of the traffic patterns between GPUs are predictable within the model training and inference cycles. The result is that for shared environments you can build a number of smaller topologies, each of which matches these traffic patterns, and each of which is far simpler than a single shared topology.

Photonic fabrics provide several compelling advantages to AI deployments. Beyond scalability and availability, the modular nature of the fabric, the lower power consumption of the photonic switches, faster workload scheduling and better workload isolation are the key compelling advantages. The dynamic nature, the continued reuse and tuning of the photonic fabric delivers a better operating advantage.

The optimized variant of a modified 3D Torus design below on the left represents individual resources and the aggregate capacity in blue. The DynamicXcelerator can build slices of the topology shown in red based on the needs discussed above. You can program a wide variety of shapes and slices.



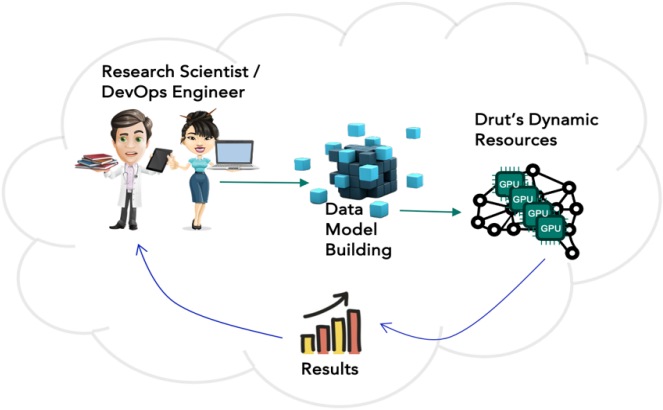
The diagram below shows the physical construct of an optimized 3D Torus (16x16x16), as well as a representative example of topology slices. A total of 64 cubes, with 64 nodes each, forms a 4k (4096) node cluster as a single availability zone (AZ) with three additional AZs possible, bringing the overall datacenter capacity to 16k nodes.



AI / Machine Learning Software Life Cycle

The Drut approach is to consider that as application workloads require ever changing resources, it is time to deploy a dynamic open looped resource scheduling architecture for your applications. How do you carve out the available resources? If we follow the typical model used by many research institutions, the available resources are clustered by how and when they were deployed, so a user can schedule cluster 1 with 4x GPUs, or cluster 2 with 8x/16x GPUs, etc. The reservation of resources has little to do with the wants or needs of the user, and if the available clusters do not match the user needs, then a significant effort is required to modify the available clusters to match their needs. The result is a poor combination of underutilized and limited capacity resources, leaving everyone dissatisfied with the result.

The Drut solution introduces dynamic cluster creation into an open looped workflow, allowing the research team involved with the building of their data model to provide input to their needs, which enables the Drut software and system to adjust by expanding or shrinking the cluster resources just-in time for their operations to run. As this is often a multi-phase training operation the results of the operations can be fed back into the research team in order to adjust their input and allow Drut to modify the datacenter resources that they can use.

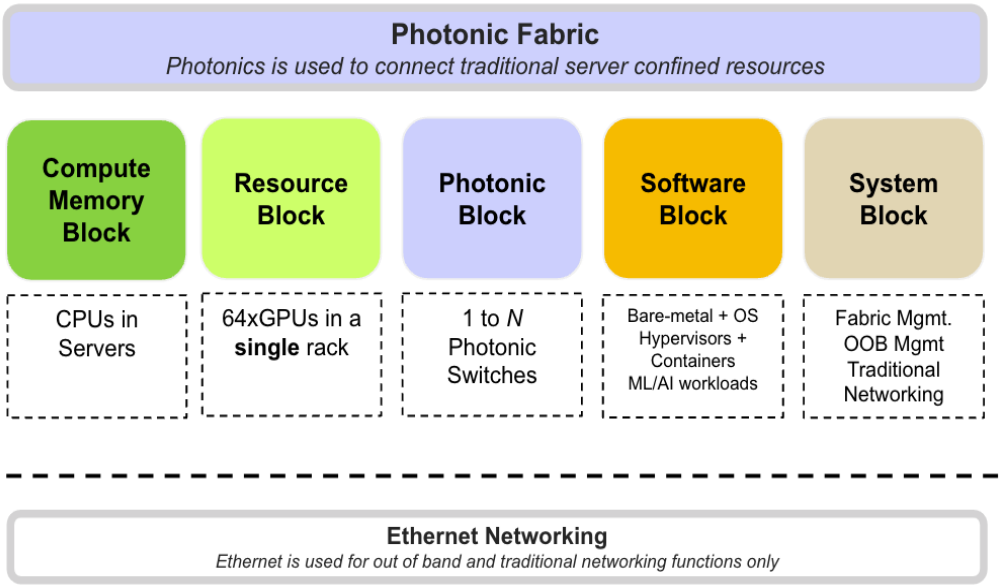


Research at various academic organizations has shown that the parallelism (data, tensor or pipeline) strategies can influence the traffic matrix, which can be used as one of the critical criteria in building the Dynamic Resources. This means that we are already at the apex where software can aid the researcher when they try to determine the resources needed.

Building Block Approach Matches Your Cloud Requirements

Drut's solution consists of a more rounded system approach to organizing constituent components into easily consumable blocks, that can be built to mirror the physical requirements of your datacenter and are flexible to map into your applications as needs change. These blocks can be seen as moveable parts, build the systems as you desire and expand and upgrade as needed. The experience of living with the DynamicXcelerator is much different than with static legacy box architectures. Systems can be composed, upgraded, and altered on demand. Resources can be taken out of service, new resources such as new GPUs added and then composed into nodes. Need more GPUs or more FPGAs or maybe you want to change GPU vendor because there is a new GPU vendor to try? Simply add the new GPUs to the Photonic Resource Unit (PRU) and put them into production. Need more bandwidth in the fabric five years from today? Well, most likely we have you covered because you will be upgrading the FIC 2500 to the FIC 4500 or something like that. The fabric is rate agnostic and new FICs will operate just fine.

You can now think of your private cloud datacenter in terms of the aggregate resources available, knowing that you can carve out sections of resources that your workloads require. There is flexibility in terms of the placement of the blocks, but you can meld this with the physical requirements of your data center. For example, you can place the power hungry and cooling needed components (such as GPUs) within racks that have the appropriate power and cooling available, saving you time and money as most of your datacenter can remain untouched. If you have upgraded power on one side of your data center but not on the other, the DynamicXcelerator can stretch the distance allowing for better fidelity of power distribution amongst resources.



It's the Software!!

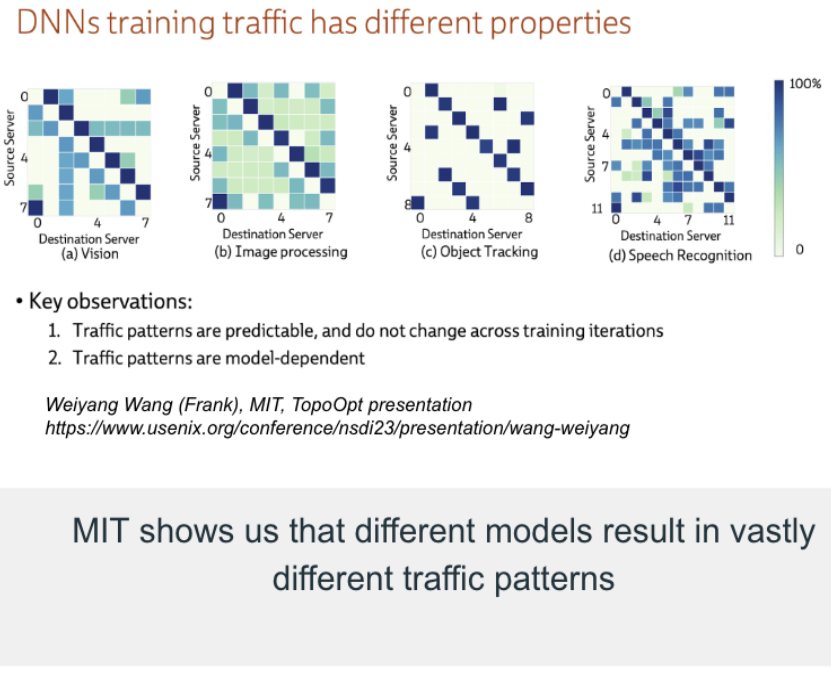
Combined with system level software stacks to drive workload level compute decisions Drut's photonic fabric allows for AI solutions to be deployed to a broader group of users by providing a dynamic rate of resource utilization. Moving away from traditional hierarchies of switches allows CPUs and accelerators to be directly connected and grouped by workload. This is how organizations will begin to deploy AI clusters. By decoupling all AI infrastructure resources from the traditional siloed box solutions will add the ability to dynamically create AI cluster nodes.

Drut offers complete system level solutions, from bare metal BIOS and OS all the way up the stack to AI/ML software stack integrations, and all the layers in between.

The Drut solution offers a componentized and layered software offering to fit the needs of every private cloud and service provider environment, both new and experienced customers can benefit from the Drut software offering.

This is just the beginning of the emergence of the software photonic revolution in the AI industry, and we can further enhance our slicing, with in-depth knowledge at the model level.

Research done at MIT has shown that model profiling allows us to predict traffic patterns. What they found is that not all communication is equal between Accelerators/GPUs, with this knowledge we can further optimize the topology offered, so that now the researcher who needs a 16 node 2x2x4 cube finds that he does not even need all-to-all capacity between this limited set of nodes, the power of the Drut solution is that we can map the capacity of the topology to the predicted traffic pattern, utilizing more or less links with more or less capacity.



MIT shows us that different models result in vastly different traffic patterns

Conclusion

Your next generation datacenter can be geared towards the ever-changing landscape driven by AI/ML workloads, you can now build architectures that will be useful many years into the future, at an immediate cost saving, which compounds into significant future savings in operational expenses in both personal and power costs. You can realize the advantages of using multi-vendor solutions and ride the hardware and software improvements without doing large forklift upgrades.

For more information on the Drut DynamicXcelerator please review our website https://drut.io .

For more detailed information contact info@drut.io