Computation and data intensive applications are increasingly prevalent, especially in research, as data acquisition and storage costs continue to drop towards zero.

**Composable Applications**

- **How to support these large scale applications?**
  - Efficient parallel/concurrent algorithms and implementation techniques
- **Key observations**
  - Most of these applications are:
    - A Single Program Multiple Data (SPMD) program or a collection of SPMDs
  - Exhibits the **composable** property
    - Processing can be split into small sub computations
    - The partial-results of these computations are merged after some post-processing
    - Loosely synchronized (Can withstand communication latencies typically experienced over wide area networks)
    - Distinct from the closely coupled parallel applications and totally decoupled applications
  - With large volumes of data and higher computation requirements, even closely coupled parallel applications can withstand higher communication latencies?

- **The computation is structured as a directed graph**
- A Dryad job is a graph generator which can synthesize any directed acyclic graph
- These graphs can even change during execution, in response to important events in the computation
- Dryad handles job creation and management, resource management, job monitoring and visualization, fault tolerance, re-execution, scheduling, and accounting

© 2009 Microsoft Corporation. All rights reserved.