EGR-RG Session 2
Summary and Next steps

Ravi Subramaniam – Intel & EGR Co-Chair

Ellen Stokes - IBM

OGF IPR Policies Apply

- "I acknowledge that participation in this meeting is subject to the OGF Intellectual Property Policy."
- Intellectual Property Notices Note Well: All statements related to the activities of the OGF and addressed to the OGF are subject to all provisions of Appendix B of GFD-C.1, which grants to the OGF and its participants certain licenses and rights in such statements. Such statements include verbal statements in OGF meetings, as well as written and electronic communications made at any time or place, which are addressed to:
  - the OGF plenary session,
  - any OGF working group or portion thereof,
  - the OGF Board of Directors, the GFSG, or any member thereof on behalf of the OGF,
  - the ADCOM, or any member thereof on behalf of the ADCOM,
  - any OGF mailing list, including any group list, or any other list functioning under OGF auspices,
  - the OGF Editor or the document authoring and review process
- Statements made outside of a OGF meeting, mailing list or other function, that are clearly not intended to be input to an OGF activity, group or function, are not subject to these provisions.
- Excerpt from Appendix B of GFD-C.1: "Where the OGF knows of rights, or claimed rights, the OGF secretariat shall attempt to obtain from the claimant of such rights, a written assurance that upon approval by the GFSG of the relevant OGF document(s), any party will be able to obtain the right to implement, use and distribute the technology or works when implementing, using or distributing technology based upon the specific specification(s) under openly specified, reasonable, non-discriminatory terms. The working group or research group proposing the use of the technology with respect to which the proprietary rights are claimed may assist the OGF secretariat in this effort. The results of this procedure shall not affect advancement of document, except that the GFSG may defer approval where a delay may facilitate the obtaining of such assurances. The results will, however, be recorded by the OGF Secretariat, and made available. The GFSG may also direct that a summary of the results be included in any GFD published containing the specification."

- OGF Intellectual Property Policies are adapted from the IETF Intellectual Property Policies that support the Internet Standards Process.

Agenda

- Summarize the topics presented in the two workshop sessions
- Next steps
- Provide overview for an approach to analysis of the usages (use cases and scenarios)

Usages - Topics

- **Session 1**
  - "US Navy's Fleet Numerical Meteorology and Oceanography Center", Nick Werstiuk, (Platform Computing)
  - "Building a Campus Grid: Concepts & Technologies", Mary Fran Yafchak (SURA: Southeastern Universities Research Association)
- **Session 2**:
  - "Grid for Financial Services", Larry Ryan, (Hewlett-Packard)
  - "Requirement Analysis of Grid Scenarios", Mathias Dalheimer (Fraunhofer Institut fuer Techno- und Wirtschaftsmathematik: ITWM)

Summary – Weather forecasting

- **Enterprise/Corporate/Academic** - Fleet Numerical (Govt, Navy)
- **Domain**: Numerical weather prediction
- **Requirement types**: Same requirements as the EDA, financial and other HPC type of systems.
- **Number of machines**: 1000 CPUs
- **Model**: Cluster, local center, multiple applications, heterogeneous platforms (OS)
- **Challenges/Constraints**: Need high accuracy and detail, large volume, strict schedule, strong security (military grade)
- **Technology solution**: LSF Multicluster (compute oriented), data infrastructure "out of band" of job management system (requires explicit call out for data), queue oriented (single queue with farming of job)
- **Organization layout**: Single control oriented organization (controllable security and policy framework),
- **Motivation**: Link two geographically distant locations in a seamless fashion (Local sharing was already being done in a set of clustered machines), increased resource utilization, increased throughput to meet stringent time horizons, efficient use by pooling
- **Future Plans**: This case describes a test framework that will then migrate to a full production frame.
- **Financial Implication**: Deffered capital expense

Summary – Campus Grid

- **Enterprise/Corporate/Academic** – University oriented enterprise - Several institutions – UAB, GSU, UMich, TACC, USC, UVA
- **Domain**: Campus Grid – multi-faceted work, multiple security/administrative domains
- **Requirement types**: cycle scavenging, dedicated, shared, arbitrated (not first come first served)
- **Number of machines:**
- **Model**: multiple administrative domains, geographical dispersed, local centers (as opposed to pure remote access), multiple applications, heterogeneous platforms (OS)
- **Challenges/Constraints**: Willingness to share, get by-in an multiple levels, slow progress to changes (not single authority to motivate?), develop policies for sharing (over come concerns), accounting (to allow complex policy enforcement, showcase usage)
- **Technology solution**: Primarily based on Globus, open source (versus proprietary), build versus buy outlook
- **Organization layout**: single security framework (centralized ID management),
- **Motivation**: Share resources, centralize ID management, increase computing cycles available to individuals or projects, visualization intensive usages, show unusual uses of grid and show that it is ubiquitous in various disciplines,. Collaboration, critical timeline for project, cycle scavenging (utilization)
- **Future Plans**: This case describes a test framework that will then migrate to a full production frame.
- **Financial Implication**: Deferred capital expense
- **User types**: Researchers, Educators, Administrative and other campus staff;
- **Motivators**: early adopters (visionary, collaborator) , top down and bottom up, funding agencies are driving collaborations.
- **Capabilities**: Accounting, Fault handling, error recovery, easy maintenance and support

Summary – Grid for Financial Services

(specific discussion Hartford Financial services)

- **Enterprise/Corporate/Academic** – Financial services (Hartford 30,000 employees)
- **Domain**: Financial positions, trading
- **Requirement types**: compute grids, data grids, memory grids, process grids, cycle scavenging
- **Number of machines:** tending to over 10,000 (capacity grown to above 8000 ;looking to expand to 25000)
- **Model**: multiple administrative domains migrate to single admin domain or federated, geographically dispersed, local centers (as opposed to pure remote access), multiple applications, heterogeneous platforms (OS) (federated model for architecture, local pools with technology agnostic interfaces that are integrated with meta-schedule; single OS image)
- **Challenges/Constraints**: Caching of information is critical as decision get more real-time (e.g. keeping indexes in memory to achieve competitive advantage)); ideally would like to keep most data in near "memory"; increase utilization of capital investments; security, standards maturity, organization politics, Business issues (charge back), expectations management (be realistic on what Grids can deliver; manage the hype) (funding is not an issue but need to curtail spending so that current capital will be used effectively)
- **Technology solution**: Globus, (Condor – scalable, flexible, mature and free; commercial vendor (?)), active directory, oracle databases
- **Organization layout:** Capacity determined by a minimum and grow as avallable (use policy structure in technology to enforce)
- **Motivation**: Reduce time to decision and minimize risk in decisions; real time calculation is a competitive advantage (higher profits; discover market opportunities; consolidating resources (scale computations), scalability, mitigate management overhead (hedging (modeling of exposure to variable annuities,, Stochastic modeling *monte carlo simulation) for annuities (e.g. mortality etc), convergence, current solution did not scale to handle number of workers that can simultaneous use and was difficult to use)
- **Future Plans**: .Grow the grid to incorporate desktops and reduce capital expenditure; create a shared file system to use local disk space.
- **Financial Implication**: Time to market and manage financial risk, reduce/defer capital costs, free up risk reserve.
- **User types**: Traders, modelers. actuators
- **Motivators**: Organization policy, demonstrated benefits
- **Capabilities**: Policies (fair use), Accounting, reporting, security, federation, automation (process)
- **Applications**: Master – worker model (centralized creating and workers are distributed and can be put on the grid); compute intensive and data light, parallelizable and partition able; task oriented, non-computational only (create reports on projections that will help manage the risks; modify the applications to run on the grid; mainly in-house developed applications; market scenarios generated on the fly in memory and so minimal data transfer)
- **Problems**: Around the management of the grid and get it operational and self-sustaining as a solution; how to maintain very high levels of uptime
- **Needs**: Grid architecture for resiliency, interoperability, simple to use – good user experience and transarency, management standards, reliable and secure infrastructure (SOX etc), chargeback models, transition/evolution models, capacity on demand

Summary – Requirement analysis of Grid scenarios

- **Enterprise/Corporate/Academic** – Fraunhoffer Institute (academic; Institute for Industrial Mathematics)
- **Domain**: Campus Grid – multi-faceted work, multiple security/administrative domains
- **Requirement types**: collaborative grids, enterprise grids (globally dispersed, federated, single policy structure), cluster grids (same user id, share file system)
- **Number of machines:**
- **Model**: multiple administrative domains, geographical dispersed, local centers (as opposed to pure remote access), multiple applications, heterogeneous platforms (OS)
- **Challenges/Constraints**: Willingness to share, get by-in an multiple levels, slow progress to changes (not single authority to motivate?), develop policies for sharing (over come concerns), accounting (to allow complex policy enforcement, showcase usage)
- **Technology solution**: PHASTgrid (in-house developed) – user interface, grid server (manage the grid), job servers (abstraction of resource pool), compute nodes; security (firewall) is done in the grid server, Unicore front-end (for client processing)
- **Organization layout**:
- **Motivation**:
- **Future Plans**: Market driven costing by pricing based on the usage of the system (Calana broker to help implement)
- **Financial Implication**:
- **User types**: Researchers, Educators, Administrative and other campus staff;
- **Motivators**: early adopters (visionary, collaborator) , top down and bottom up, funding agencies are driving collaborations.
- **Capabilities**: Accounting, Fault handling, error recovery, easy maintenance and support
- **Applications**: Business sector focused (automotive, etc), ***application needs to be integrated*** and the jobs are referenced and requests/data sent and results returned; applications are parallelizable (this is done dynamically in their environment) – transform (uses heuristics, application knowledge), compute and aggregate; defines applications based on the Grid infrastructure rather than the other way around
  - seismic image processing (acoustic source with hydrophones; analyze the reflections and return single analysis (Kirchhoff-Migration); includes visualization; demanding as far as data sets that are 100 GB to 1TB and also compute intensive; requires 1 hours for 10 nodes; use MPI for communication between process – will eliminate MPI dependencies to minimize application crashes)
  - Life sciences – phylogenetic analysis (gene comparisons) – requires correlation analysis and visualization of the results; similar to montecarlo
  - Financial application: 24/7 on 3000 CPUs, trading system maps requests to PHASTgrid; investment opportunities and portfolio management; 50 jobs/sec; not as demanding on data movement (stored centrally in a database) but is very compute intensive
  - Robust engineering: Minimize material cost by maintain specification while maintaining manufacturing tolerances; define optimal structure and configurations; uses Finite Element analysis; model optimization and then design criteria check (this is an iterative loop); source code available internally; FEM is tightly coupled processes in the analysis and require more than one compute node at a time (uses Calana Broker to allocate co-located nodes for the FEM analysis); large models (high data movement requirements) and highly compute intensive in the FEM portion
  - Mystery applications: Starts with priming a database (DB size 5TB) – data can be updated incrementally by customer in update chunks; data used in subsequent calculations that are portions of the data; required high security for the data in database; need DB close to applications; execution guarantee is required; low data movement but highly compute intensive because of the large number of jobs that execute at a time; has job server that manages resources that come and go
- **Needs**: Understand application thoroughly (has implications to infrastructure; resources need to be matched appropriately); general tendency is that there is a high correlation between data movement and computation requirements.; need to understand the user requirements (reliability, meet user SLA and time schedules) ;need to guarantee SLA and user defined requirements.

Next Step

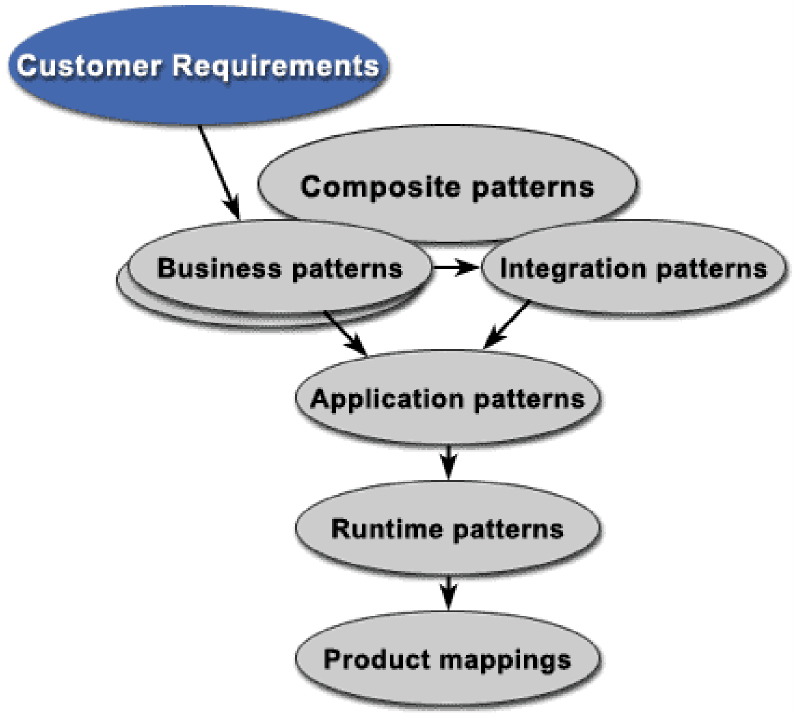
- Organize the workshop material as in the template so that we can have a consistent perspective for analysis
- Gather more such cases with individuals providing examples
  - Speakers for the next GGF – continue the workshop model?
- Define an approach for analysis towards finding tangible requirements for the standards development
- Engage and sync with other OGF activities in this space – e.g. VOC and VAF run by the OGF office and UC repository effort - underway
  - Please see Toshi's intro on the UC repository
  - Engagement with the VOC activities (session today at 4:00 in COMM31 in 147A) and VAF session (tommorrow at 9:30 in COMM34 in 147A)

Pattern Oriented Analysis

Patterns for the Enterprise Grid

- Common designs that can be used / reused to develop solutions meeting an organization’s requirements

General Pattern Types

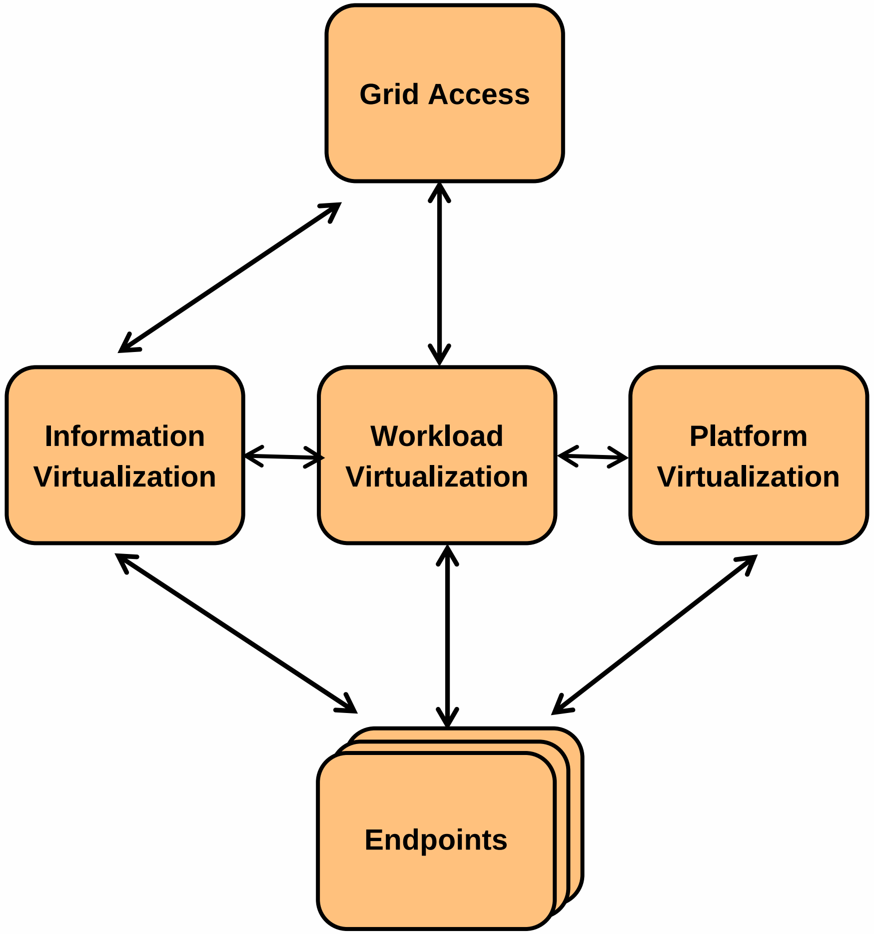


- Business patterns identify interaction between users, business, and data
- Integration patterns tie multiple business patterns together
- Composite patterns represent commonly occurring combinations of business and integration patterns
- Application patterns describe conceptually how application components and data within a business or integration pattern interact
- Runtime patterns define logical middleware supporting an application pattern

Proposal

- Define runtime patterns to describe how grid management service components interact to support a controlled and flexible execution of application workloads: depict major middleware components, their roles, and the interactions between those components.
- Composite pattern
  - Access (user access to grid: portal/authn/authz)
  - Workload virtualization (scheduling & workload mgmt)
  - Platform virtualization (resource mgmt & provisioning)
  - Information virtualization (local/remote data access)
- Apply to case study scenarios to validate / produce solutions

Composite Pattern



- Access patterns
  - Extended single sign-on
  - Directly integrated single channel
  - Store and retrieve
  - Replica location
- Workload virtualization patterns
  - Basic scheduler
  - Meta-scheduler
  - Peer-to-peer scheduler
- Platform virtualization patterns
  - Basic manage resources
  - Intelligent manage resources
  - Dynamic manage resources
- Information virtualization patterns
  - Local data access
  - Remote data access
  - Remote access to local data

Full Copyright Notice