Consider that:

• Computing models have evolved
• Far better understanding of requirements now than 10 years ago
  – Even evolved since large scale challenges
• Experiments have developed (different!) workarounds to manage weaknesses in middleware
• Pilot jobs and central task queues (almost) ubiquitous
• Operational effort often too high; lots of services were not designed for redundancy, fail-over, etc.
• Technology evolves rapidly, rest of world also does (large scale) distributed computing – don’t need entirely home grown solutions
• Must be concerned about long term support and where it will come from
But remember:

• Whatever we do, we must evolve whilst not disrupting the ongoing operation
• We have a grid for a very good reason – we need to integrate the resources provided to us – but we can make use of other technologies
• We have the world’s largest (only?) worldwide trust federation and a single sign-on scheme covering both authentication and authorization
• We have also developed a very strong set of policies that are exemplars for all other communities trying to use distributed computing
• In parallel we have developed the operational security teams that have brought real benefit to the HEP community
• We have also developed the operational and support frameworks and tools that are able to manage this large infrastructure
WLCG must have an agreed, clear, and documented vision for the future; to:

• Better communicate needs to EMI/EGI, OSG,...
• Be able to improve our middleware stack to address the concerns
• Attempt to re-build common solutions where possible
  – Between experiments and between grids
• Take into account lessons learned (functional, operational, deployment, management...)
• Understand the long term support needs
• Focus our efforts where we *must* (e.g. data management), use off-the-shelf solutions where possible
• Must balance the needs of the experiments and the sites
To reassess the implementation of the grid infrastructures that we use in the light of the experience with LHC data, and technology evolution, but never forgetting the important successes and lessons, and ensuring that any evolution does not disrupt our successful operation.

The work should:

- Document a strategy for evolution of the technical implementation of the WLCG distributed computing infrastructure;
- This strategy should provide a clear statement of needs for WLCG which can also be used to provide input to any external middleware and infrastructure projects.

The work should, in each technical area, take into account the current understanding of:

- Experiment and site needs in the light of experience with real data, operational environment (effort, functionality, security, etc.), and constraints;
- Lessons learned over several years in terms of deployability of software;
- Evolution of technology over the last several years.

It should also consider issues of:

- Long term support and sustainability of the solutions;
- Achieving commonalities between experiments where possible;
- Achieving commonalities across all WLCG supporting infrastructures (EGI-related, OSG, NDGF, etc).

Deliverables

- Assessment of the current situation with middleware, operations, and support structure.
- Strategy document setting out a plan and needs for the next 2-5 years.
The MB manages the process directly, and appoints several TWGs, one for each area to be addressed, monitors the results of the technical analysis and takes final decisions on the proposals.

Each TWG should be co-chaired by 1 experiment and 1 site representative. The chairs and the group members to be proposed and agreed by the MB, perhaps with the assistance of a small steering group.

Steering group to define scope, mandate, and membership of each TWG, to be agreed by the MB, and to monitor progress and have editorial ownership of the process.
Some general comments

- Outcome should be a *documented* strategy
- Not everything needs to go back to square one!
  - Some things work!
  - Some work has already been (is being) done (e.g. Data management)
  - But the strategy needs to be documented
- The work must include consideration/models for long term support and sustainability
Data Management

• Distributed data management and tools
  – Follow on from work on Amsterdam demonstrators etc
  – Data placement, caching
  – FTS re-write/replace
  – “xrootd”
  – Transfer protocols
  – Data access security needs/access control
  – Continued use of POOL? (only ATLAS?, how long?)
  – Interaction between DM tools and ROOT, Proof?
• More to address site related services
  – Separation of archives and disk pools/caches
  – Future storage system needs: e.g. how should dcache, DPM, EOS, etc. evolve?
  – What are the storage system interfaces? SRM, S3, ...
  – Filesystems / protocols: NFS4.x, S3, ...
  – Security/access controls
  – These are site-run services: management interfaces? Monitoring? ...
Encompasses a broad range of topics
- Pilots and pilot frameworks
- What is needed from a CE now?
- Is a generic WMS required by WLCG?
- Use cases for clouds
- Use of virtualisation –
  - experiment expectations for using CERNVM, arbitrary images? (re-Hepix work)
  - Site expectations for use of virtualisation in managing sites
- Security model (MUPJs, etc.)
- Whole node scheduling
- How to ask for/use GPUs (is it required)
- Information services – what is needed, what tools?
• **Scope should be use of “grid” database services**
  - Not use of a db within an arbitrary application
  - LFC? What is still required? Deployment model
  - Frontier, squids, etc, vs 3D/streams, Goldengate, DataGuard, etc
  - What long term support is needed for Coral, COOL?
  - Work started in database workshop in June
Security model

- Should review risk analysis – what are the real threats now? Where should we focus
- Is the trust model still appropriate?
  - E.g. can we simplify the “glexec” issue?
- Do we still need open WN’s?
- X509/VOMS/IGTF have been essential in having a world-wide use of resources
  - But there are problems associated with proxies
- Can/should other federated ID management systems be integrated?
Need to document all the other pieces that we need supported

- Monitoring (SAM, Nagios, etc.); better monitoring needs? Ability to analyse monitoring data?
- Support tools (APEL, GGUS, etc.)
- Underlying services (ActiveMQ, etc.)
- Operational requirements on middleware
- Application software management (e.g. cernvmfs)
- Software management of middleware
- Deployment management
- Configuration management
Status

- Buy in to the process by experiments and sites – at WLCG collaboration meeting in July and over the summer details agreed in MBs
- Nominations for chairs received and approved
- Chairs being contacted now
- Work to start asap, initial reports hopefully by end 2011/early 2012
  - Frequent updates in MB/GDB to ensure broad dissemination of what is discussed