Proposal for a Collaboration Tool

PRELIMINARY draft version 4
07/02/2005

This document summarises recent discussions of the DESY ILC GANCOMM group. The primary goal is to recommend a possible infrastructure for the ILC global collaboration. A core part of this infrastructure is the need to make centralised collaboration software tools available to the community. The major part of this draft document details both the general requirements and possible implementations of such tools.

These requirements are far from complete. We invite comment and suggestions from the broader ILC community before finalising details and levels of support for these tools (both short- and long-term solutions).

1 The Value of Documentation

A project is exchanging information at different levels of quality, ranging from ad-hoc communication for discussion and coordination over minutes, notes and presentations to official documents like specifications, contracts, design models. For ad-hoc information, the focus lies on intuitive and easy-to-use communication channels, while official documents need to be reviewed, released and stored in way which guarantees their availability throughout the project lifecycle. A rule of thumb for deciding wether information is ad-hoc or of lasting relevance refers to “the magical fifteen minutes”: any information which takes less than fifteen minutes for authoring or for reading and understanding can be considered ad-hoc. Other information should be passed with metadata, e.g. instructions on how to use the information, keywords for information retrieval, status information.

2 General Description

The worldwide forming collaborations and working groups for the International Linear Collider (ILC) need a collaboration tool which allows simplifying the distributed work and helps foster communication inside and outside the ILC community.

The participants in the current ILC effort come from the Americas, Asia and Europe. Remote communication between the various actors works mainly via e-mail, websites, and telephone. Remote conferencing is usual via telephone systems or (less frequently) via video conference. In addition a number of physical conferences, workshops and less formal meetings have been held - or are currently planned - at various locations around the world. Most of these communications are currently organised on an ad hoc basis by the relevant interest groups. This is leading to a ever-growing number of web-sites and (more

important) *repositories* for the associated documents (minutes, copies of presentation etc.) Retrieving information across such a distributed and non-standardised system is cumbersome at best and potentially error prone at worst (finding the wrong version of a document, or not the latest version of a parameter list). In addition, there is a danger of loosing transparency across the evolving collaboration, if small groups of people unilaterally organise an ILC-related meeting but fail to openly inform the entire collaboration, primarily because centralised tools do not exist to support such information exchange. Centralising both the storage of information (or at least the *gateway* to the system) and an event calendar and meeting agenda system is a way around this problem. In addition a centrally managed set of tools for communication will become mandatory as the more formal structure of the GDE emerges, and the need for work-flow within the context of project management becomes unavoidable.

3 General Capabilities

The collaboration tool or tools should provide the following capabilities:

1. storage and management of documents
2. an agenda management system
3. an event calendar system
4. a discussion forum
5. a mailing lists system
6. a video conferencing system
7. a Webcast system

3.1 Document Management System (DMS)

The document management system should help the users in creating, updating, distributing, researching and navigating documents (reports, presentations, CAD drawings, pictures, technical specifications, parameter lists, etc.)

The system should support different *views* of the same subset of data. For example a technical view on a construction process defined in a set of documents might be different than a managerial view which concentrates rather on budget and scheduling than on technical details.

The document management system should help users to coordinate and automate pre-defined processes via a *workflow* mechanism. Examples for such processes include:

- the publication of an ILC paper from the initial submission of the author through several reviews to its final release by an appointed ILC editor;
- the post-processing of a superconducting cavity from its delivery to the receiving lab through different treatments and measurements to the installation in a cryomodule.

It should be possible to switch off workflow mechanisms when they are not required, e.g. presentations at meetings, memos, and other documents not necessarily requiring proof reading or some other sort of authorisation can be immediately uploaded to the document server. (We should note here that all documents *should* in principle require a minimal amount of workflow to make sure that required metadata is correctly attached.)

It is expected that the amount of data to be stored in such a system may eventually exceed several 100s of Gbyte.

General requirements for the DMS are:

- all required data formats should be supported;
- tools should be provided for automatically creating platform independent viewable documents from foreseen document formats (e.g. WORD, PowerPoint automatically converted to PDF).
- the system should foresee specially protected segments for sensitive data;
- the system should allow for re-organisation of the data and the data structure, which is important for a growing and developing project;
- it should be possible to reference data at different points of the data access tree
- the system should provide a multidimensional data categorization
    - topical (keywords)
    - emanation level
    - access level
- the system should support easy re-structuring (re-catagorising) of the data
- the system should support version control of the documents
- the system should support workflow mechanisms
- automatic back up of the data base
- authorisation of users should be organised based on encrypted passwords
- authorization should not be centrally administrated but de-centralized user management should be possible (local *administrators* for institutes, or work packages)
- the user should communicate via a web-interface with the system
- users should not be required to buy any licensed software
- the system should notify users or groups of users about changes of certain documents or groups of documents (user configurable)
- the system should support an efficient search engine

3.2 Agenda management system

An agenda system is needed to support the growing number of upcoming workshops, conferences, telephone/video meetings. It should provide a web-based distributed management, i.e. allow the organisers and conveners of a given event to register that event, construct an agenda and (eventually) upload presentations etc. Files with talks or contributed papers should be uploaded to

the DMS via the agenda system interface, and cross-linked to the agenda so that they can easily be associated to the speaker. A search engine is needed to easily find talks, speakers, or other agenda entries.

3.3 Calendar

A web-based event calendar is needed to help organise future workshops, conferences, and other events. The database should allow distributed management so that (for example) work package coordinators can register events. The calendar would then act as the single point of reference for all ILC related events (actors should be strongly encouraged to register their meetings with this system). Events with agendas (the normal case) should be associated with the respective entry in the Agenda Management System.

It should be possible to search the calendar for events by name, description, time interval, etc. The event calendar should be exportable to personal calendar tools like Outlook, Corporate Time, etc.

3.4 Discussion forum

To foster discussion inside the community a web-based discussion forum tool is needed. It should be able to create discussion forums which allow discussing different topics via a web interface. Discussion items should be stored in a database and should be accessible for an infinite time period. It should be able to post not only text (plain or html), but also documents to this forum. It should be able to cross-reference documents stored in the DMS to discussion topics. A search engine with the ability to search for authors, titles, text, documents is required.

The forum should provide user administration facilities. Discussion forums can be moderated or non-moderated. User registration is needed to allow for restricted access to certain forums. An e-mail notification service for new posts should be provided.

3.5 Mailing lists system

Mailing lists are an important communication tool in distributed collaborations. The mailing list system should allow for the easy administration of mailing lists via a web interface. It should be possible to install and maintain mailing lists in a distributed manner. Mails submitted to a mailing list should be stored in a database which is accessible via the web (the DMS). The web interface should allow to add/remove members and should additionally allow to get the members of a certain list easily. If possible mailing lists should check for duplicate e-mails. This is especially important if mailing lists contain other mailing lists as members.

Mailing lists are easily open to abuse. Provisions for spam filtering are needed but this may not be enough. Careful thought on how these lists are used should be made. In general, the functionality of a 'mailing list' can often be replaced by a discussion forum or hyper-news system which is easier to protect from abuse.

3.6 Video conferencing system

A video conferencing system which allows connecting desktops to virtual conference rooms is needed. It should allow supporting spontaneous forming conferences without the need to reserve lines or virtual rooms well in advance.

3.7 Webcast System

The live transmission and storage of video streams from conferences over the internet is gaining importance with the growing numbers of workshops and conferences. The collaboration tool should include the possibility to transmit live video streams together with synchronised images of transparencies from talks. A virtual desktop system might be needed to transmit (for example) software demonstrations.

4 General Constraints

The system should be usable for users around the world using different operating systems like Linux, Windows, Mac, Solaris, etc.

There should be no licence fees for the users.

The language must be English.

The user interface has to be simple and intuitive.

5 Users

We foresee the potential users of this system to come from the following groups:

- The worldwide ILC collaboration
- The ILC project groups at the various labs
- The working groups and consortia inside the ILC collaboration, e.g. LCRD, EUROTeV, CARE, etc.

The users come from all the regions participating in the ILC: Asia, the Americas, and Europe.

6 Operational Environment

The requirements for the client side have to be lightweight. It should run on any desktop or laptop PC without requiring special hardware or software. Supplying the primary interfaces via the web would allow this (with some notable exceptions where additional client-side code must be installed, such as VRVS or Skype).

On the server side the system needs some 100GB of storage space, backup facilities and broadband networking. The system should be maintained and operated on a 24h basis with the support of experts on hand.

7 Suggestions for Realisation

The ILC@DESY GANCOMM working group makes the following suggestions to realise the collaboration tool:

7.1 Document management system

The DESY EDMS is a system which matches most of the required capabilities. The development of a lightweight web interface to EDMS is under study and can be realised on reasonable short timescales. DESY is willing to provide the necessary system requirements and human resources for support and management.

7.2 Agenda Management System / Calendar System

The CERN agenda tool CDA and its successor InDiCo provide the capabilities described above. The adjustment of InDiCo to serve as a front-end to the DESY EDMS seems possible and is under study.

A dedicated CDA/InDiCo server has been installed at DESY. This effort is mainly driven by DESY internal needs but it will be open to the ILC public. DESY offers to store ILC related meetings, workshops and conferences on this server with the idea to serve as a test bed for the central storage of the related documents. With this respect the connection to the DESY EDMS is planned.

The server will be operated and supported by DESY. This includes regular backups, 24h support and failsafe operations.

CDA includes a calendar which fulfils the requirements.

7.3 Discussion forum

There are existing freeware tools which provide the required capabilities. The discussion forum accessible at http://forum.linearcollider.org is a very good example. It should be studied whether the existing forum can be expanded to the needs of the broader community

7.4 Mailing lists

The ILC-BDIR working group uses a mailing list organised and hosted by KEK: http://ilcphys.kek.jp/mail/bds
This mailing list has the capabilities described above. It should be studied whether this system can be expanded to the needs of the broader community.

7.5 Video Conferencing System

VRVS (http://www.vrvs.org) is a tool which has the capabilities described above. Contacts with the VRVS organisers have been established. A set of private virtual rooms (ILC) has already been created in VRVS (actors will be required to register for the private room, and we encourage them to do so). Membership to the "ILC community" can be registered for on the VRVS site.

Four Virtual Rooms have been created:

- ILC-Cafe - Always on with no password
- ILC-Casino - Permanent VR with password (geneva)
- ILC-Vroom1 and ILC-Vroom2 that can be booked. (can be booked by members of the ILC community community)

6.6 Webcast System

Webcast systems exist and are in use at several institutes. A connection to the document storage system should be implemented to provide the ability to store video streams and associated documents in the central storage area. We are currently evaluating the InDiCo project is this respect.

8 Creating an Efficient World Collaboration

In the previous sections we have described a possible set of software tools to support a global collaboration formed by geographically remote sites. Having such centralised tools available will (we believe) greatly increase the efficiency of 'collaboration over distance' and help to keep the activities within the various institutes, working groups and regions transparent to the whole community.

It should be stressed that these tools only work if they are uniformly adopted by the community.

To further streamline the collaborative process (and again as a measure towards transparency), we propose two concepts that we believe - if adopted - will reduces the number of meetings and allow higher level of participation and information exchange between ILC related groups.

8.1 The Weekly ILC Day

The need for remote conferencing (be it telephone, video conference or vrvs) is clear. Such telephone meetings will become the mainstay of the global collaboration. With the number of work-groups and packages likely to increase, these meetings will become frequent. Attempting to attend even some fraction of all the possible meetings for those whose involvement spans more than one WP may well become difficult at least (impossible at worst).

We would like to propose that a three-hour slot be set aside each week which can be used by any ILC-related group for a remote meeting. The slots will always be on the same day each week, and can be rotated to provide equal 'discomfort' to each of the time zones involved. These slots should be established (and agreed upon) for the next six months. A central reservation system would allow an ILC related group to 'reserve' a given slot. The reservations would (naturally) be made public on the central calendar and agenda system.

Such a system has the following advantages:

- actors know in advance to try and avoid planning other activities during a regular time slot; this significantly eases the problem of finding convenient times for such meetings;
- the scheme removes the problems of parallel meetings, thus allowing a broader participation;
- actors will form the habit of checking what this weeks meeting is, and may choose to attend just for general interest (these meetings should allow unlimited participation and be generally open to any ILC collaborator).

The disadvantages relate to the need for a central reservation system and administration (which is arguably required in any case), the problems that may arise during a critical week when two or more meetings are needed (easily handled on a case by case basis if the need is not too frequent), and perhaps most importantly the need for the ILC community to embrace the concept to make it work.

8.1.1 Proposal for the 'Weekly ILC Day'

We propose the Thursday of each week to be the 'Weekly ILC Day'. Meetings should be held at the time slots shown in Table 1.

Table 1: Proposed time slots (starting time) for the weekly ILC days.

DATE	GMT-7 (SNOWMASS TIME)	GMT+1 MET	GMT+9 KEK-TIME
10. Feb 2005	00:00	08:00	16:00
17. Feb 2005	16:00	00:00	08:00
24. Feb 2005	08:00	16:00	00:00
03. Mar 2005	00:00	08:00	16:00
(…)	Cyclic permutations		

8.2 The Monthly ILC Week

Although the majority of meetings will presumably be remote (and hopefully within the formal structure set out in 7.1), physical meetings ranging from one hour meetings, through one-day mini-workshops to full-blown workshops will still

be needed. Again there is a need to formalise these events with a goal to maximising their benefit and usefulness for the community.

In many ways the concept of the ILC Week is exactly analogous to the ILC Week: A slot each month (this time a number of days - say four) is set aside which groups should use to hold meetings and workshops. Specifically:

- An ILC week is planned each month (for example the last week of the month) for the next six months.
- A Hosting Institute is identified willing to host the ILC week. The institutes should be in all three regions and cycled through.
- As with the ILC Day, specific groups (WG for example) can reserve a period of time within the ILC week for a workshop or a meeting.
- Parallel meetings can be reserved up to the capacity of the hosting institute (this will maximise overlap and potential synergy between groups).

The role of the Hosting Institute is pivotal, but their responsibilities should be no more than: guaranteeing that a certain number of seminar and meeting rooms are available; that a list of suitable accommodation is available; and that enough coffee is provided (and naturally that there is some minimum administrative support should somebody require local help). The responsibility for organising the meetings remains fully with the respective ILC group. Again one could imagine a central reservation administration where the Host Institute would register the available rooms (including capacities), and the rest would be up to the central administrator and the meeting organisers. The concept here is to reduce to a minimum the overhead for the Host Institute.

For 2005, several meetings and workshops (c.f. Table 2) have already been arranged which could be used in the interim as ILC weeks.

Table 2: Upcoming ILC related meetings, workshops, and conferences.

Start	End	Meeting	Venue
10.02.2005	12.02.2005	GAN-MVL Meeting	INFN-LNF, Frascati, I
15.02.2005	15.02.2005	Collimation for the ILC Meeting	CCLRC Daresbury Lab, UK
21.02.2005	22.02.2005	Wiggle 2005	INFN-LNF, Frascati, I
16.03.2005	17.03.2005	ILC Simulation Workshop	SLAC, Menlo Park, USA
18.03.2005	22.03.2005	LCWS 05	Stanford University, Palo Alto, USA
30.03.2005	01.04.2005	TESLA Tech. Collaboration Meeting	DESY, Hamburg, D
11.04.2005	13.04.2005	Workshop on Positron Sources for the ILC	CCLRC Daresbury Lab.

Start	End	Meeting	Venue
06.2005		EUROTeV workshop	TBC
20.06.2005	23.06.2005	BDIR/ELAN Workshop	Royal Holloway, London, UK
10.07.2005	15.07.2005	SRF 2005	SNS, Knoxville, USA
14.08.2005	27.08.2005	ALCPG Workshop	Snowmass, USA
14.08.2005	27.08.2005	2nd ILC Workshop	Snowmass, USA
17.10.2005	21.10.2005	Nanobeams 2005	Kyoto, J
14.11.2005	17.11.2005	ECFA ILC Workshop	Vienna, A

If this scheme is adopted, then the time between now and the Snowmass workshop can be used to put in place the necessary central administration, as well as identify the Host Institutes for the latter part of 2005 and into 2006.

8.3 Organisation and Administration

Both for the 'collaboration tools' outlined in section 6 and the ILC Day and Week concepts require some administration. Because of the obvious dangers of maintaining open and public write (i.e. upload) access to the various tools, it may be necessary to restrict certain access to specific designated individuals, which we will refer to here as *administrators*. Administrators have different roles depending on the scenario (and tool) being discussed.

As an example, we could image that each institute has a designated administrators, as does each Working Group (or eventually Work Package). A single individual can double as both a Work Package and institute administrator, thus reducing the total number of administrators required.

By way of example, we can take a scenario of a Work Package requesting an ILC Day slot for a vrvs meeting:

- At the request of the group, the group administrator requests a reservation for a two hour meeting for a specific ILC day meeting.
- (The reservation request could be automated within the calendar/agenda system, or it could be handled by a central administrator.)
- The group administrator now sets up a meeting and agenda within the Agenda system (this can only be done by an administrator).
- The administrator sets a modify password which is then sent to the speakers (the speakers will later use this password to upload their presentations to the agenda system).
- Note that the agenda is publicly visible to all, but only the administrator can modify the agenda.

In this scenario, the administrator plays the role of 'scientific secretary' for his group.

This is the lowest level of administration required. Direct access to EDMS and the use of workflow (which will require EDMS user accounts) will need to be reviewed to find acceptable and safe procedures. An example of such an administrator role is the updating of important (versioned) documents such as drawings, lattice decks or parameter sets. Figure 1 shows the interaction of the user, the collaboration tool and the administrator.

Figure 1: Interaction of user, collaboration tool, and administrator.