White Paper on Directory Services

July 10, 1989
DRAFT

M. I. T. Laboratory for Computer Science

Dr. Karen R. Sollins, sollins@lcs.mit.edu, (617)253-6006
Dr. David D. Clark, ddc@lcs.mit.edu, (617)253-6003

The MIT Laboratory for Computer Science believes that there is important work to be done in designing and building a *white pages* directory service for the National Research Network (NRN). For that reason, the Laboratory has begun an effort to understand and evaluate the alternatives and to participate in planning for this important service. The Laboratory proposes to continue this work in conjunction with the efforts of the Federal Research Internet Coordinating Committee (FRICC).

A *white pages* service is one in which one can look up people in order to learn information about them for finding them. In its simplest form, a white pages service provides what the white pages telephone book provides. Based on a name, one can find an address and a telephone number. In a network environment, there may be many other kinds of location information, such as electronic mailbox, electronic calendar, or file server, where one might leave a file for the recipient. In addition, the electronic white pages may support a much more sophisticated set of mechanisms for lookup. One might match on a more complex set of attributes than first and last name. In addition, the searching might span more than one local white pages service. There are a number of naming and directory service specifications and implementations in the field. They have differing functionality and mechanisms to address that functionality.

We propose to do this work in the context of a larger project, described initially in "A Plan for Internet Directory Services," NIC RFC 1107. That document is a report of a two day workshop held at NRI, in February, 1989, with the goal of defining a plan that would see the FRICC through to deployment of a white pages service in the NRN in three years.

The conclusions of that report can be stated as follows:

- There are a number of naming and directory service specifications and implementations available, having varying functionality and strengths and weaknesses.
- Because many of these standards and implementations are close enough to being complete, it is reasonable to undertake provision of an NRN white pages service.
- Although we are close, an effort is needed both in standards and code development to achieve the goal.
- An initial evaluation experiment is needed before making final detailed plans for a production version of the service.
- With suitable funding and encouragement, a production service is possible in three years.
- It is important to act now to provide a coherent solution. This means both having providing a unified, wide-spread solution before a plethora of differing solutions appear and an impact on the evolving standards.

NIC RFC 1107 suggests a three-stage general plan for achieving a white pages service. The first stage is a set of field trials of differing name services, concluding with an evaluation of the services and a proposal for a single form of service, based on the experience of the field trials. The second stage is the implementation. There will need to be some implementation in Stage 1 for at least some of the services to make them compatible with the environment and to build new user interfaces. Stage 2 will carry forward the implementation based on the conclusions of Stage 1, incorporating earlier code whenever possible. The third stage is deployment, including distribution of code,

placement of servers, delegation of authority, structuring of the global namespace (what the hierarchy will look like, if there is one), and other issues in bringing the service to the public.

In addition to the tasks noted in each of the stages identified above, there are a number of others that are necessary to the success of the plan. In each case, as they are identified and described below, we have also included an estimate of full time equivalents (FTEs). These estimates represent our judgment of task sizes, and were discussed as well at the February meeting. These are estimates for total funding from all sources, existing FRICC related funding, outside funding (such as the work at DEC and Steve Kille's funding through the Esprit project), and new FRICC support.

- Oversight and monitoring of the whole project (1/4 FTE): There will be a number of different activities ongoing at any given time. These must be coordinated, in addition to being monitored, to occur in a timely way. If parts of the project slip, others subsequently will be delayed. Completing this project in three years is cutting it close enough that it will be necessary to have a management and monitoring task assigned. We have an interest in activity here.

- Design and oversight of field trials (1/2 FTE): In order for field trials to be useful experiments, it must first be determined what is to be learned from them, and how that will be achieved. This will include placement of the services, exercising them in certain ways, and collection of information during the field trial period. The deployment of the field trials must be negotiated with potential user communities and the services put into place. If experimental data collection is not automated, this must be done manually. We have an interest in these activities.

- Design and development of DUAs (for each one, 1/4 FTE for toy, 1 FTE for production version): At present each directory service proposed for the field trials has a single fairly simple user interface. A variety of styles of interface should be available for different sorts of usage and levels of sophistication of users and well as different operating systems in the client workstations. The workstation operating systems that should be included are Unix, VMS, MSDOS, MacOS. The work in this area should be carried out both by the support teams of the various name services in the field trials (Larry Peterson at the University of Arizona, Steve Kille at University College London, Dave Oran at DEC, the implementers of another X.500 implementation if that is included) and others interested in this work. We would like to continue our ongoing work at the undergraduate level here.

- Additional implementation for services included in field trials (4 x 1/4 FTE): Some amount of programming is still required to bring the code of the name services to field trial quality. It also may be the case that they will need to be retrofitted with metering tools, in order to collect data from the field trials. The recipients of this funding will be the groups implementing the various name services used in the field trials, Larry Peterson, Steve Kille, Dave Oran, and another X.500 implementer if one is included in the field trials.

- Data collection and management tools (1 FTE for each of the first two years): All the services need to collect similar kinds of naming data, and of course it will all come from the same sources, the existing site files of user information and additional user input. Without this activity the directory services are useless. Each of the services included in the field trial provides some tools for data collection and management, but a consistent and complete set of tools will make everything run more smoothly, and make maintenance of up-to-date and correct information more likely. This work could be undertaken by the Network Information Centers (NICs) or the IETF Working Group on Network Information Services Infrastructure (NISI).

- Evaluation of field trials and report on an American specification of a directory service with conformance criteria (1/2 FTE): It is important both that the participants and funders understand the lessons learned from the field trials and that these conclusions take the form of information to be fed into the next stage of the plan. In this case, a detailed specification of what the service will be must be stated, including profiling and conformance criteria. There are two reasons for this careful work. The first is to allow multiple implementations. The second is to feed back into the standards activities (noted as a separate task below). We have an interest in participating in this activity.

- Implementation of production name services to be deployed (2 to 4 FTE): The amount of work here is difficult to predict without knowing the conclusions from the field trials. If an existing service is only modified a little from its present state, less work will need to be done. If the conclusion is that a mix of

the existing services would funding, provide the right balance, more work will be needed. There is a widely accepted belief in the networking world that two implementations are needed for any specification in order to validate it. This work will require production quality code and conformance testing.

- User assistance and training (1/4 FTE each year): Beginning with the field trials, network managers will be managing the servers and users will be using these services. Both will need help. To some extent this can be left to the purveyors of the code, but they will need some amount of support for this, and it may be both useful and educational to centralize this activity. Some of this money should be provided to support Larry Peterson and Steve Kille in deployment during the field trials. Later this funding might support either the NIST or the NICs in building user assistance tools for use in the final deployment of the white pages service.

- Planning for deployment - structuring the namespace, plan for distributed service (1/4 FTE): All of the proposed directory services subdivide the namespace into smaller namespaces. For example, in the Domain Name System, this is a combination of geographic and organizational partitioning, and reflects delegation of authority for name selection and management. This division will need to be done again, and planned carefully. Most of a year was expended in working on the Domain Name System in arguing about the structure of the namespace. There also must be a plan for distribution and replication of information for reliability and availability. In addition, a plan needs to be developed in detail for Stage 3, the deployment. Again, this work would best be performed by NIST or the NICs. We also have an interest in this activity, based on the interests expressed above.

- Deployment (1 3/4 FTE, each year after the first two): Deployment of the directory service will include delegation (or authorization) of subnamespaces, distribution and maintenance of code, placement of nodes in the distributed service, and monitoring and management of the distributed services. Other tasks may be identified during the planning for Stage 3. This should probably be performed by NIST. After the first two years, this activity combined with ongoing user assistance will comprise 2 FTEs.

- Standards participation (1/2 FTE, each year): It is critical that what is learned from the field trials and ongoing directory services activities as described above be channelled back into the standards activities. Because X.500 is being accepted as a standard in Europe, we must be able to interoperate with it. In addition, it is important that we influence the direction that the CCITT recommendations and later the ISO standards take, in order to enhance functionality and interoperability between Europe and the US. Again, NIST may be a good candidate for this activity. Debbie Deutsch at BBN has also expressed an interest in this activity.

As mentioned previously, there are three sorts of funding for the White Pages project as described here. The first is external funding that is already in place. This includes the support provided to projects such as the Quipu project (a British implementation of X.500), DEC's funding of the DNANS development and enhancement, and such smaller activities as participation in a task force or working group by industrial representatives. The second source of funding is currently being provided by members of the FRICC to do work in the naming and white pages service area. In some cases, this money is directly in line with the aims of the project and, in others, may need some redirection. This funding includes that provided to Larry Peterson for Profile and possibly Deborah Deutsch at BBN for work in internet naming activities and the ongoing work on X.500 at NIST.

The third category of funding is new money that is required from FRICC members in order to complete the project. This last category includes such work in Stage 1 as the planning, oversight and review of the field trials, in Stage 2, planning and oversight of the deployable implementations, also in Stage 2, support for the deployable implementations, and, in Stage 3, planning for deployment. A project review will be required to characterize the funding for actual deployment in the third year of the project. Such tasks as general management, standards work, and user services will require ongoing support as well. Table 1 summarizes FTE levels by year for each of the tasks listed above. This table represents total funding for tasks, based on all three types of funding as described above. Table 2 summarizes our best estimate of sources of funding for this effort.

It is important to understand one additional factor in the distribution of funding between FRICC and non-FRICC

organizations, the level of non-FRICC funding on which the FRICC effort will be building. There is probably at least 10 FTEs involved in the implementations of the three services currently under discussion as part of the field trials. In addition, a fourth service would bring in yet more work to base the FRICC effort on. Finally, a great deal of effort has gone into and will continue to be put into the specification X.500. FRICC funding will be directed in those areas where others are not addressing the detailed issues of specific importance to this project. The tables reflect new activity that is necessary for this specific project.

From the above list, the tasks in which we propose to be involved are: oversight and monitoring of the whole project, design and oversight of the field trials, design and development of DUAs, evaluation of the field trials and the report on the US specification, and planning for deployment. In order to participate in these activities, funding is sought from interested agencies within the FRICC.

The level of funding requested by MIT is for a base amount with extensions to permit greater participation in this project. At the base level, the MIT Laboratory for Computer Science is requesting funding to cover one half a researcher's salary. This would be directed at oversight and management of the project as a whole and the field trials specifically, and later to evaluation of the field trials. In addition, the base level funding would include one undergraduate to work on a DUA, and a small amount for travel and supplies. With our overhead, this comes to $85,000.

There are several items that would increase our effectiveness. The first is all or part of a technical staff member (programmer) at $110,000 with benefits and overhead. The staff member would work on a set of data collection and management tools and a DUA (production version). A second important addition is a graduate student at $33,000, to work on set up, data collection, and data analysis in the field trials. In conjunction with increased activity are increased computing, travel, and materials expenses. In summary, with increased funding, MIT can achieve more, thereby increasing the thoroughness and effectiveness of the FRICC white pages project.

Task \ Year	1	2	3
Oversight and management	*	*	*
Design & oversight of FT	**		
Design & Development of DUAs	**** **	**** **	**** **
Implementation for FT	****		
Data collection and mgmt tools	****	****	
Evaluation of FT and report	**		
Implementation of Name Service		**** **** ****	
User assistance & training	*	*	* Annually, henceforth
Planning for deployment		*	
Deployment			**** *** Annually, henceforth
Standards participation	**	**	**
TOTAL	5 1/2 FTEs	6 3/4 FTEs	4 1/4 FTEs

Note: Each * denotes 1/4 FTE

Table 1: Level of Effort in Tasks

Funding Source \ Year	1	2	3
Existing FRICC funding	****	****	**
Non-FRICC funding	****	****	****
New FRICC funding	**** **** **** **	**** **** **** **** ***	**** **** ***

Note: Each * denotes 1/4 FTE

Table 2: Estimated Sources of Funding

Comment: It is important to notice here that the levels of non-FRICC support is currently at at least the level of 10 FTEs, although that effort is not reflected in these numbers because it is part of other projects, on which this one will build.