



INFN DCE/DFS WG Status report

Roberto Gomezel
HTASC Meeting
Geneve - October 2-3, 1997

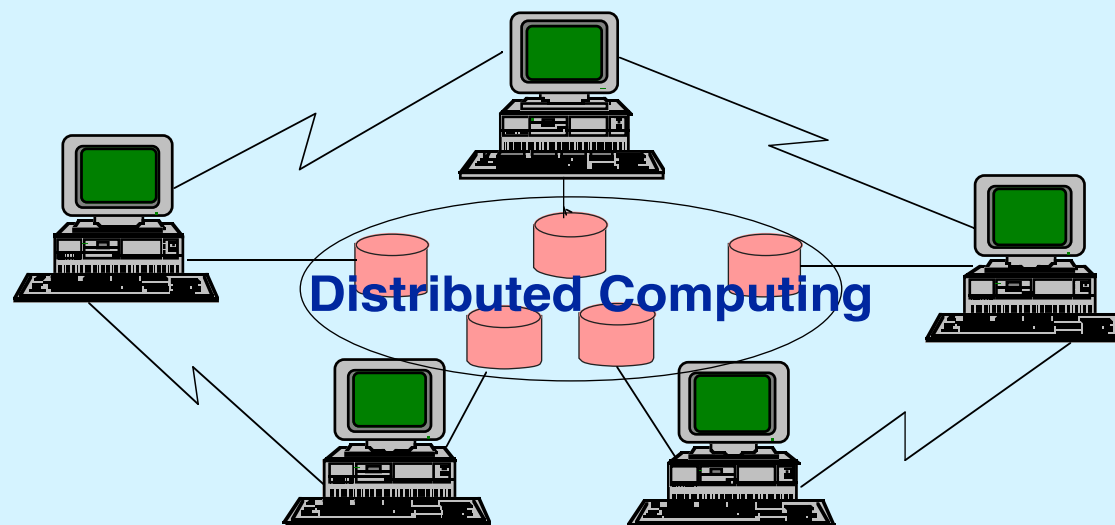


INFN DCE/DFS WG Status Report - R.Gomezel
HTASC meeting - CERN Geneva, October 2-3, 1997

gomezel@trieste.infn.it

What is DCE?

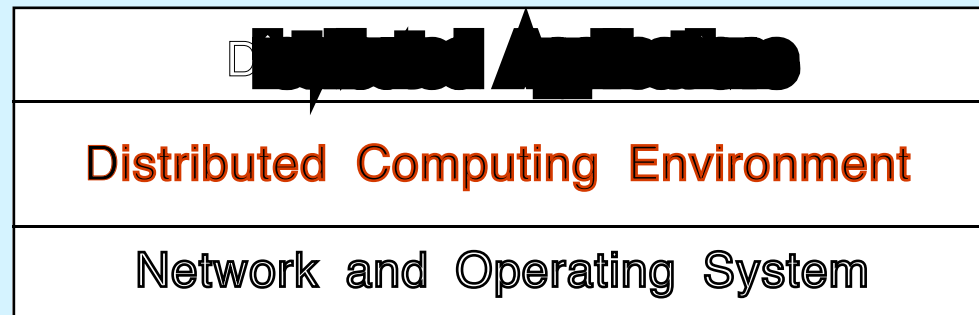
- DCE is an acronym for Distributed Computing Environment
- Provides services and tools that support the creation and use of distributed applications in a heterogeneous computing environment





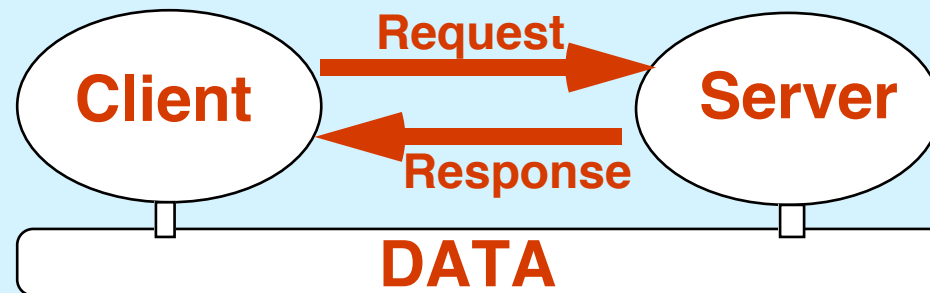
DCE architecture

- DCE is a layer between the operating system and network on the one hand, and the distributed applications on the other
- it allows a distributed application to communicate with a collection of heterogeneous computers as if they were a single system

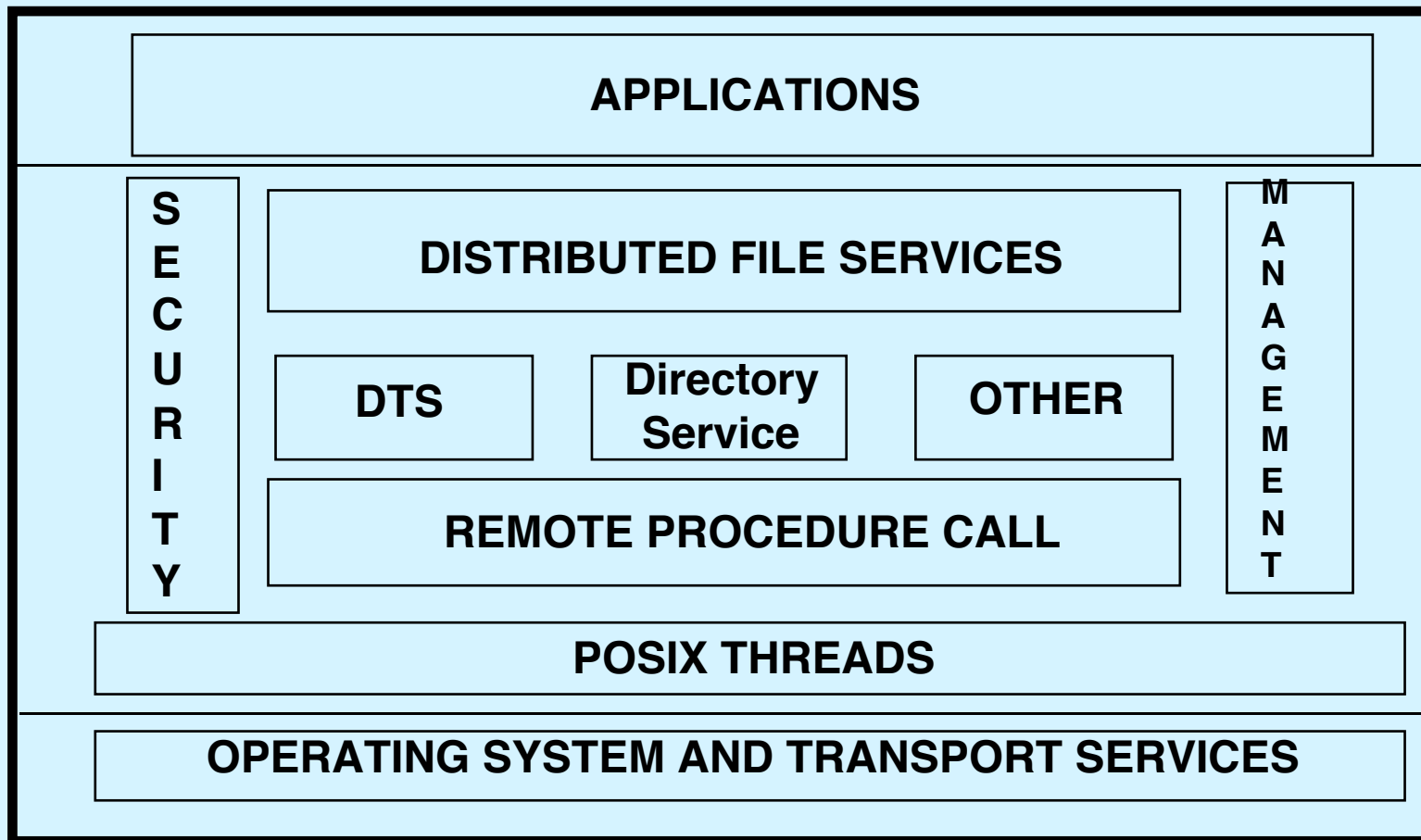


DCE: models of distributed computing

- DCE is based on three different distributed computing models
 - client/server model:
 - client requests a service/server executes
 - Remote Procedure Call
 - client makes what looks like a procedure call, this call is translated into network communication so the server can execute the procedure returning the results to the client
 - Data sharing model
 - data is shared by distributing it throughout the system



Architectural overview of DCE





Architectural overview of DCE

- **DCE threads:** support the creation, management and synchronization of multiple threads of control within a program
- **DCE RPC:** facility for calling a procedure on a remote machine as if it was a local procedure call, shielding the details of network communications
- **DCE Directory Service:** allows to keep track information about users, machines and other resources geographically dispersed
- **DCE DTS:** Distributed Time Service provides a way to periodically synchronize the clock on the different hosts in a distributed system
- **DCE DFS:** Distributed File Service is an application built on the underlying DCE services that provides the ability to share data throughout a worldwide system





Overview of Directory Service Components

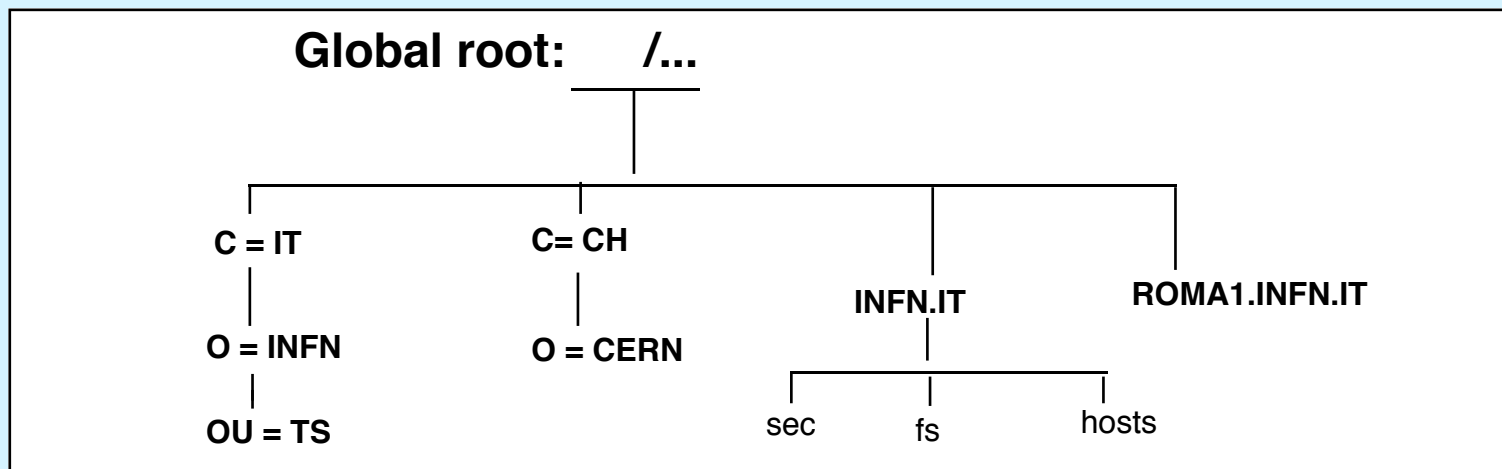
- it is a distributed and replicated database service and consists of a hierarchical set of names, the namespace
- there are three components that make up this service:
 - DCE Cell Directory (CDS)
 - DCE Global Directory (GDS)
 - based on X.500 international standard
 - it can participate in the worldwide X.500 directory service or can support the most widely used standard directory service, the DNS
 - DCE Global Agent (GDA)
 - intermediary between the cell's CDS and the rest of the world





The DCE Namespace

- hierarchical set of names used by the Directory Service
- the names can be typed or untyped according to the formats supported by the two global directory services:
GDS and DNS
 - typed:: /.../C=IT/O=INFN/OU=TS
 - untyped : /.../ts.infn.it





DCE Security Service

- it provides a security system in order to control access to resources in a distributed environment
- comprises several parts like:
 - authentication service
 - privilege service
 - ACL facility
 - registry service
 - login facility



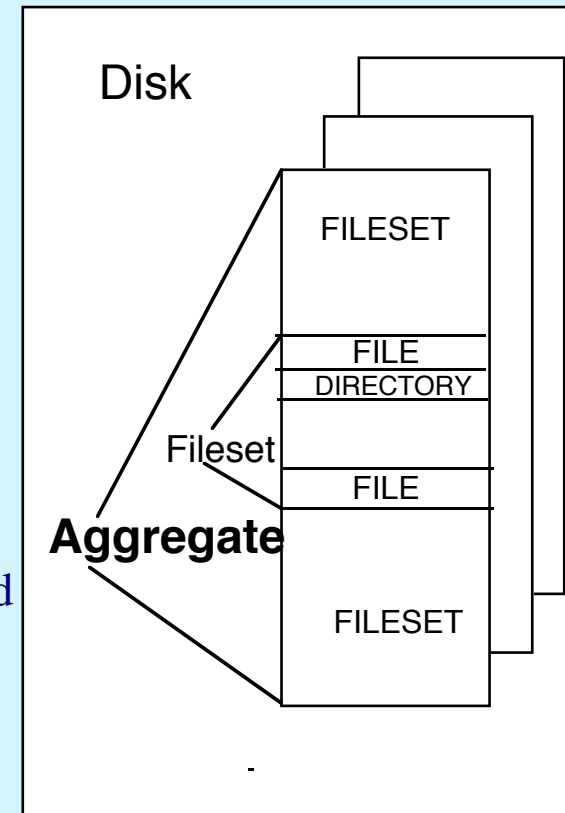
What is a DCE cell?

- it is a collection of machines, users and resources managed as a group
- a cell must have its own security service, directory service and time service
- optionally, it can have a distributed file service
- a Global Directory service will be necessary to enable the cell's directory server to communicate with other cell's directory servers



What is DFS?

- it provides the ability of sharing information in a form of a filesystem
- application client/server that takes full advantage of the lower-level DCE components such as RPC, the security service and the directory service
- DFS data organization
 - **Files and directories:** unit of user data
 - **Filesets:** unit of administration grouping of files and directories
 - **Aggregates:** unit of disk storage (similar to disk partition) and unit of fileset exporting that can contain one or more filesets



DFS components

- **Cache manager:** runs on any machine that is acting as a DFS client; it takes a user's file system request and manage the cache space on the local system
- **File exporter:** is the server side of DFS; it runs on a file server machine handling requests from clients for files that it manages
- **Token manager:** runs on file server machines to synchronize access to files by multiple clients
- **DCE Local File System:** the physical file system provided with DCE;it manages the store of files on the disk and allows more capabilities than the UFS such as the ability to replicate data,backup on-line, the use of ACLs and the possibility of moving parts of the file system without interruption to service
- *UFS systems are supported in DFS and can be exported as an alternative to the LFS, however you lose the most interesting functionalities offered by DCE LFS*



Advantages of DCE DFS

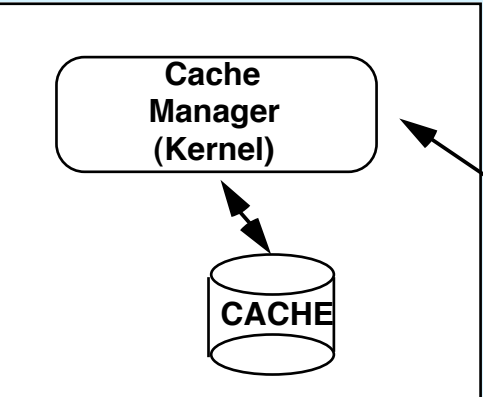
- **Uniform file access:** *based on a global namespace*
- **Intracell location transparency:** *data can be moved from one server to another within a cell*
- **High-performance file service:** *caching*
- **Availability:** *replication, caching and backup*
- **Support for distributed application programming**
- **Ease of administration and scalability:** *files grouped into filesets*
- **Interoperation:** *a UFS can be exported to users of DFS*



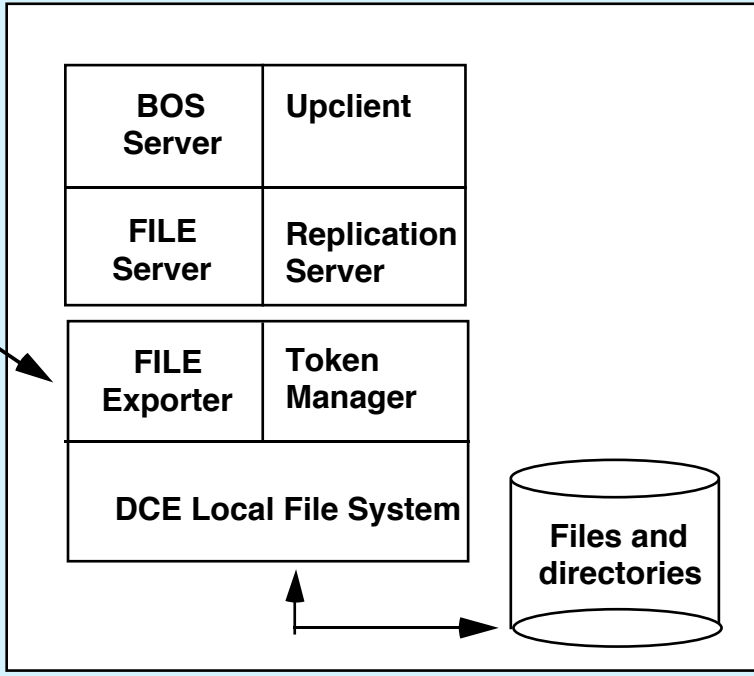


DFS Configuration

DFS Client Machine



DFS File Server Machine





INFN DCE/DFS Working group

- working group which began its activity in September 1996
- initial aim was to examine and improve knowledge about distributed applications and particularly about DFS
- this group exposed its first steps and initial testbed project during the last INFNet Workshop held in October 1996
- The Team:
 - **R. Gomezel (Trieste)**
 - **S. Angius (LNF)**
 - **S. Arezzini (Pisa)**
 - **C. Bulfon (Roma1)**
 - **R. Cecchini (Firenze)**
 - **E. Fasanelli (Lecce)**
 - **P.Mastroserio (Napoli)**
 - **D. Salomoni (CNAF)**
 - **S. Spanu (Roma1)**



DCE/DFS WG: goals

- at present, INFN sections utilize heavily AFS services to share data and ASIS/CERNlib collection regularly mirrored from CERN
- DCE WG was born to try out a new architecture providing not only data sharing services, but also a more wide and coherent computing environment for developing applications on a distributed system
- initial testbed consisted of bringing up DCE cells in 5 INFN sites (Firenze, Napoli, Pisa, Roma and Trieste) on different platforms in order to verify DCE/DFS in a very heterogeneous environment



DCE/DFS WG:Activity

- in order to report its steps this group has decided to draw up a WEB page where to collect meeting minutes, documents and links to other DCE WG throughout the world
 - <http://www.fi.infn.it/DFS>
- Meetings: 3 meeting held in December '96, March '97 and July '97
- a member attended the last DECORUM '97 in USA, entirely dedicated to tests and implementations of DCE cells in the world





What We Have Today

- we chose a multiple cells configuration to try out intercell communication between DCE servers running on different platforms

CELL name	DCE server OS	BIND	DCE patches
dceibm.le.infn.it	AIX 4.1.5	4.9.3*	yes
le.infn.it	Digital-Unix4.0a	4.9.3*	yes
dcecell.roma1.infn.it	HP-UX 10.10	multinet 3.2	PHSS_8344
fi.infn.it	HP-UX 10.10	multinet 4.0a	PHSS_8344
na.infn.it	DigitalUnix 4.0a	4.9.3*	yes
pi.infn.it	AIX 4.1.5	4.9.5	yes
ts.infn.it	DigitalUnix 4.0a	4.9.5	yes

Other cell information

- Cells were registered in DNS so as to be visible by DCE GDA of foreign cells
 - use of DNS resource record AFSDDB to define the hostname where the CDS server resides
- Use of NTP server as time provider to DCE DTS system
- Cross-cell authentication adopted to permit a trust relationship between cells
 - trust relationship established through cross-cell authentication should give DCE users of local cell authenticated access to resources in the foreign cell





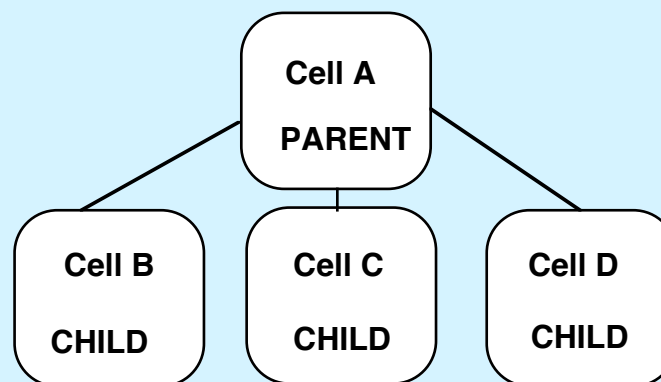
Problems encountered

- Lack of support of LFS (Local File System) in Digital and HP DFS releases (only UFS or proprietary file systems supported)
- Sometime, on Digital and HP platforms, intercell communication is unstable owing to the abrupt death of *dfsbind* daemon; this behaviour makes unreliable the data sharing between cells



Hierarchical cell configuration

- in this configuration one cell's CDS acts as a higher-level directory service to connect independent cells: this cell is known as a *parent* cell while the cells connected through the parent's CDS are known as *child* cells
- advantage of this configuration is the automatic trust relationship between the *parent cell* and the *child ones*
- unfortunately, this feature is only implemented in Digital DCE, so we cannot test it in a heterogeneous environment



Lessons learned

- DCE releases for IBM, HP and DEC are stable and seem to work fine within a cell; we cannot say the same about the intercell communication
- Access to DFS tree via NFS
- Only IBM and SUN provide a true DFS server
 - LFS, ACLs setting, backup and movable filesets
- Others can be used efficiently as DFS client
- SUN not yet tested because of lack of hardware



Features we are very interested in trying out

- Integrated login: home directories in DFS and authentication
- IBM will provide us with DCE and DFS client for Windows NT so as to test it
- Participate in upcoming field test of the new DCE release 3.0 (OSF 1.2.2) for Digital Unix - (scheduled to begin in October)
- Improve knowledge about developing DCE applications, exploiting the tools provided by DCE threads and RPC





Conclusions

- lack of vendor support
- IBM and SUN platforms are available only in a few INFN sites, so, at present, it is difficult to plan a migration from AFS to DFS
- If LFS is not available, use of DFS is not so attractive; it allows a caching access to a distributed file system, but you lose all the most interesting features you already have with AFS
- DCE manuals outline a wonderful set of facilities that are not equally implemented in the versions provided by different vendors, so it is very difficult to realize a world wide heterogeneous clustering
- there is a great interest in DCE architecture throughout the world, so we hope to have, as soon as possible, a common set of facilities equally implemented on main available platforms

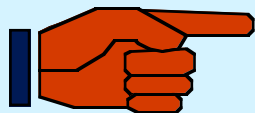


Final remarks

- Will the DCE architecture be a solution to our distributed computing needs?



It is still an unanswered question



<http://www.ts.infn.it/computing/htasc>

- This presentation

