

# Distributed Computing at Penn: Observations From the Not-So-Bleeding Edge

Chris Shull, Open Systems Specialist  
Dan Updegrove, Associate Vice Provost  
Noam Arzt, Director, IT Architecture  
Information Systems and Computing  
University of Pennsylvania

# Background

- large, private, research university
- highly decentralized management structure
- participation and consensus required
- generally an early follower of IT advances
- long history of distributed computing

# Distributed Computing History

- 25 years of school-based computing
- 15 years of departmental, unit, and personal computing
- 10 years of NFS in engineering school, extending into other areas
- 4 years of PennNet Authentication System
- 5 year experiment with DECathena in Wharton School discontinued in 1994 because it wasn't distributed enough!

# In the Spring of 1994

## Heterogeneous Computing Environment:

- IBM: SAS, MIS, Library
- Sun: SEAS, Medicine, SAS, Library
- HP: Wharton, Medicine, Chemistry
- DEC: DCCS, Wharton, Medicine
- SGI: SEAS, SAS, Medicine
- 60/40 PC/Mac, plus Netware & AppleShare
- MVS, VM/CMS & VMS legacy systems

## PennNet Authentication System (PAS)

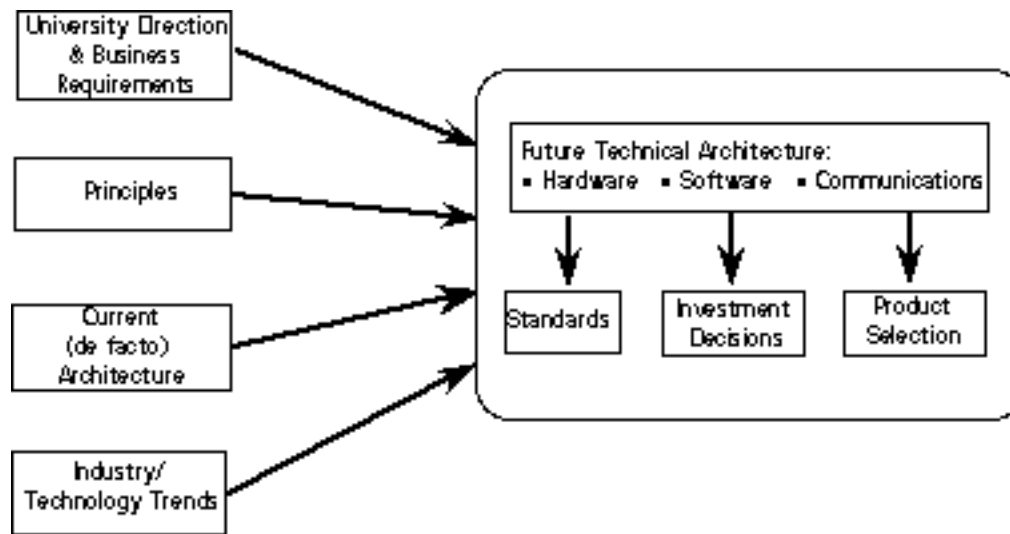
## And the Hospital and Medical System

# Major IT Initiatives

- Project Cornerstone
- Library Access 2000
- Electronic Mail Task Force
- Network Architecture Task Force
- ResNet
- CWIS / WWW
- New Media Center

# Task Force Methodology

## Technical Architecture Methodology



# Distributed Computing Task Force

- Spring 1994 launch
- campus-wide Task Force
- Three main objectives:
  - Learning, Learning and Learning
- Secondary objectives:
  - creating a campus-wide username space
  - testing, evaluating and piloting
    - . authentication services, and
    - . file services
  - writing system administrator guidelines

# Task Force Organization

Modeled after Penn's E-Mail Task Force

- working groups for separable, well-defined activities
- co-chairs for each working group intended to
  - distribute the leadership burden, and
  - provide complementary perspectives
- chairs meet on a regular basis to discuss interdependencies
- liaisons to other task forces and user groups

Advisory Committee:  
provides high level oversight  
helps assure consensus

Working Groups:

- Single-Signon Authentication Services
- Authorization and System Security
- File Services
- User Tools and Resources

Facilitator:

- responsible for "mothering" the Task Force

## Communications:

- Electronic Mail used extensively for meeting summaries and drafts
- World Wide Web for "full and prompt disclosure" of working documents

## *Project URL:*

<http://www.upenn.edu/open-systems/dctf.html>

# Task Force Tuning

- Scope broadened from DCE to Distributed Computing
- Facilitator joined by a Co-Facilitator
- Working Groups have discovered objectives

# Task Force Progress

Now available on the Web:

- University Direction and Business Needs
- Principles

This Summer:

- Current Architecture
- Industry and Vendor Trends

# On the Fast-Track

- Secure Access for Client/Server Administrative Systems
- Software Distribution
- Library Authorization
- Campus-Wide Usernames

# Unified PennNames

Methodology points to the need for unique, campus-wide usernames.

Users have too many usernames & passwords.

Solution: single sign-on authentication system.

Prerequisite to single sign-on authentication:  
unique, campus-wide PennNames

# Penn's Username Space Problem

- Many multi-user systems ( $>100$ )
- Many users ( $>25,000$ )
- Many username conflicts
  - Type 1 Conflicts: more than one person has a given username,  $\sim 2000$  conflicts effecting  $\sim 3500$  people
  - Type 2 Conflicts: a given person has more than one username,  $\sim 8000$  people
- More conflicts are being created constantly
- Chicken and Egg problem

# Chicken and Egg Problem

1. stop creating new conflicts
2. resolve existing conflicts

Maintain independence of

- unique username,
- single sign-on,
- electronic mail addressing,
- *et cetera*.

# 1. Stop Creating New Conflicts

- PennNames API and System
- SMTP-like protocol, maintains state information, but without a connection
- In startup mode, supports creation of PennNames Data Base by merging existing username spaces with Penn ID numbers allowing conflicts
- In operational mode, generates, reserves and adds non-conflicting PennNames
- Written in C using TCP/IP sockets

# PennNames System Requires

- input of existing username space, and
- modification of account generation procedures on multi-user systems
- first tools target Unix and the Modem Pool

## 2. Resolve Pre-existing Conflicts

A number of methods:

- replied and wants to keep name > didn't reply  
> volunteered to give up name
- faculty (by rank) > staff > graduate student >  
undergraduate student > guest
- older name > younger name
- more widely used > less widely used name
- has only one name > has multiple names
- system name > person name

# Conflict Resolution Implications

- The "rank" rule alone reduces the number of Type 1 Conflicts from 2000 to 600, a 70 percent gain!
- The 8000 Type 2 Conflicts are expected to be 90 percent spurious, as they arise from a historic quirk of PAS.
- Dummy accounts with `.forward` files can facilitate name changes.
- Privacy issues remain.
- Name changes can be deferred until forced.

# Conclusion

- It is not about "*DCE*" — it's about *Distributed Computing*.
- We can and must lay the foundation for 1 to 2 years out:
  - train someone to follow OSF/DCE,
  - manage a campus-wide namespace.
- Pilot and implement to your ability.
- Muddle through until you can develop clear objectives — it's too important to ignore!

# What Role Can CAUSE Play?