Scalable, Autonomous Network Services Configuration

David L. Mills
University of Delaware
http://www.eecis.udel.edu/~mills
mills@udel.edu

DARPA Principle Investigators Meeting
6-7 March 1997

Sir John Tenniel; *Alice's Adventures in Wonderland,* Lewis Carroll

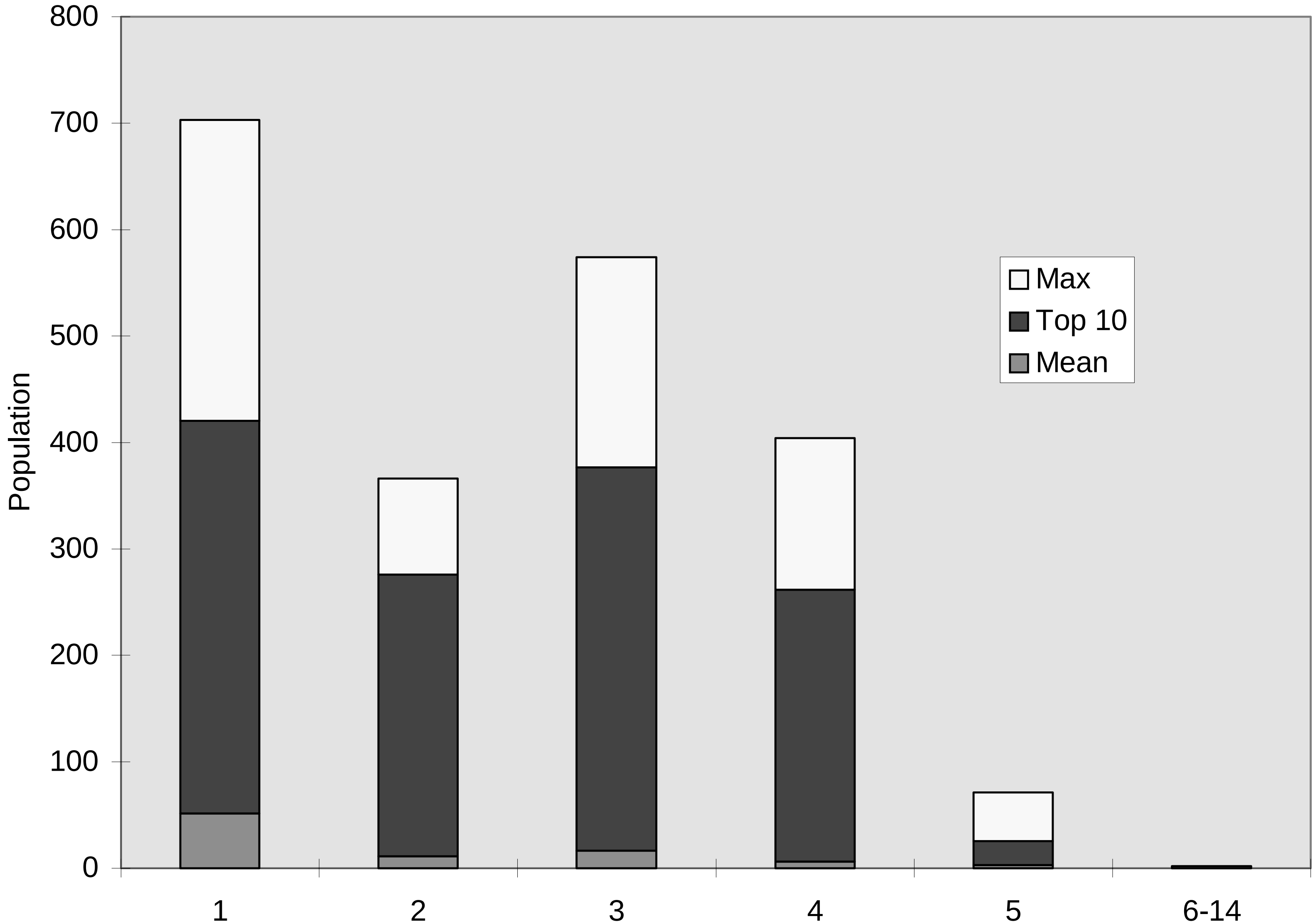
Introduction

- Autonomous configuration paradigm is designed to distribute cryptographically authenticated information from hundreds of servers to many thousands of clients
  - Uses a hierarchical network of primary servers, secondary servers and clients
  - Designed for ubiquitous, distributed services such as Web caching, time synchronization, news distribution
  - Avoids manual configuration witchcraft
- Potential peers discovered using directory services, service location agents and span-limited multicast messages
- Distributed algorithm determines subnet topology
  - Constructs shortest-path or minimum-weight spanning trees, subject to constraints of node degree and maximum distance
  - Operates continuously using low-overhead protocol
  - Engineered heuristic with provable complexity bounds
  - Design and implementation model using Network Time Protocol (NTP)

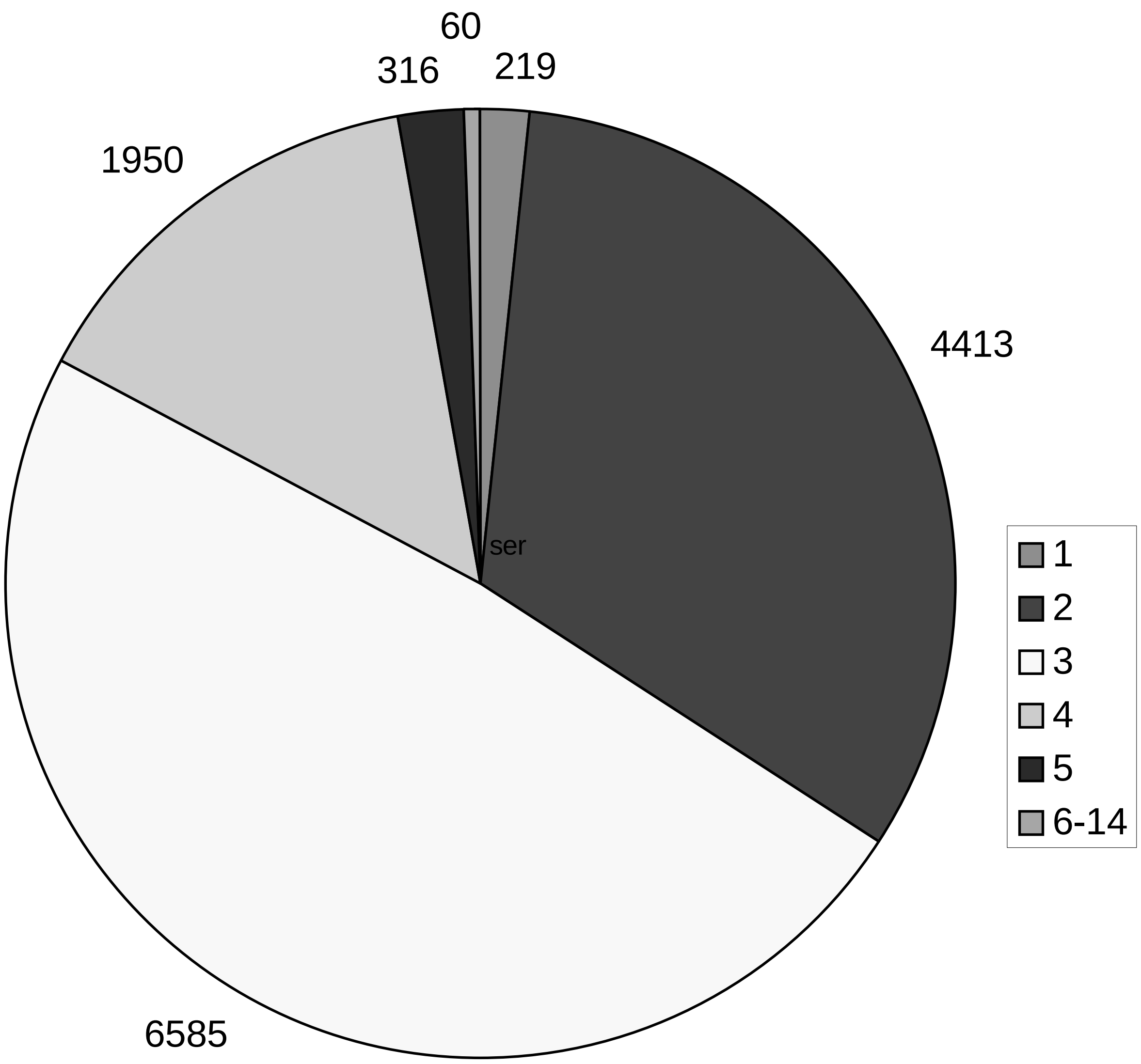
NTP capsule summary

- Network Time Protocol (NTP) synchronizes clocks of hosts and routers in the Internet
- Provides submillisecond accuracy on LANs, low tens of milliseconds on WANs
- Primary (stratum 1) servers synchronize to UTC via radio, satellite and modem; secondary (stratum 2, ...) servers and clients synchronize via hierarchical subnet
- Reliability assured by redundant servers and diverse network paths
- Engineered algorithms used to reduce jitter, mitigate multiple sources and avoid improperly operating servers
- Unix NTP daemon ported to almost every workstation and server platform available today - from PCs to Crays - Unix, Windows and VMS
- Well over 200 public NTP primary servers and 100,000 NTP peers deployed in the Internet and its tributaries all over the world

Clients per server population by stratum



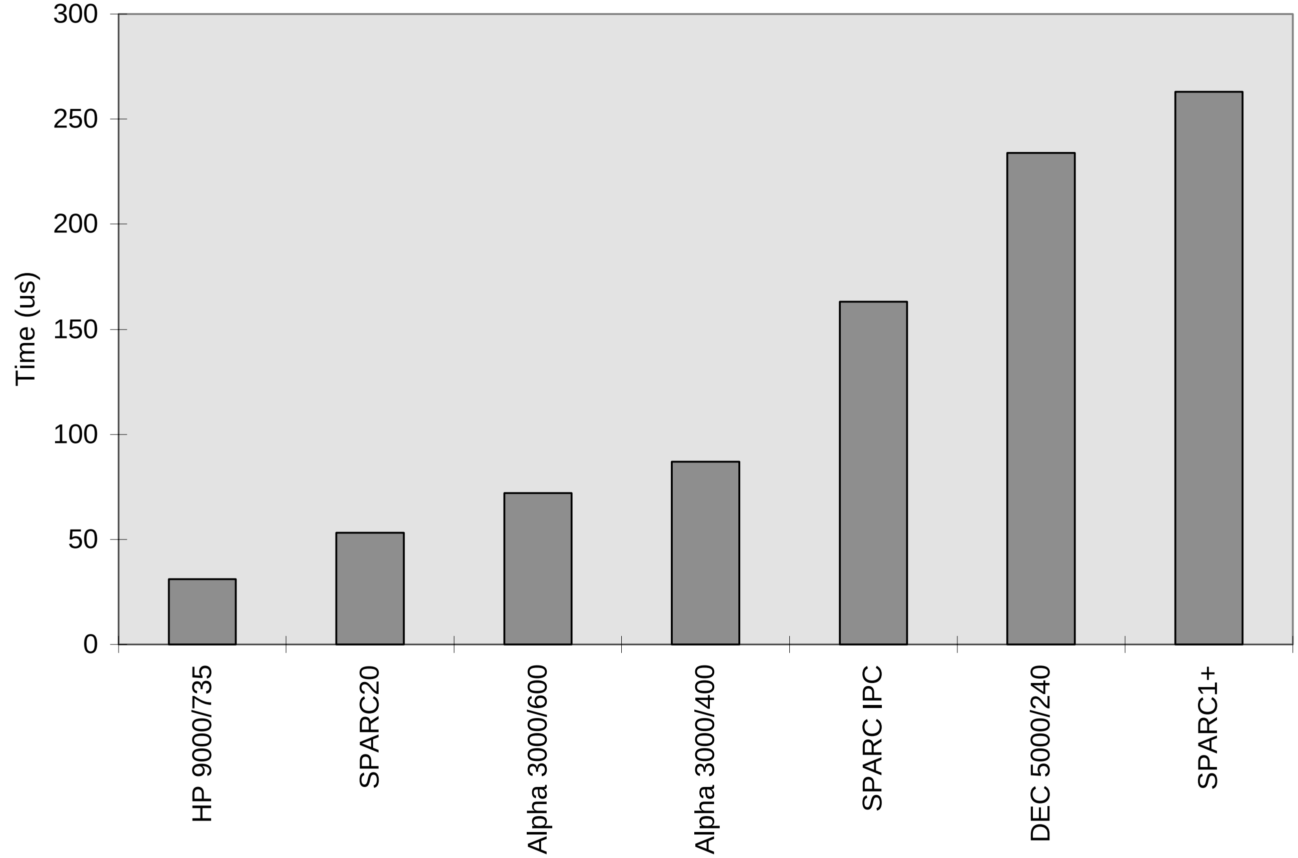
Server population by stratum



Autonomous configuration and authentication - issues

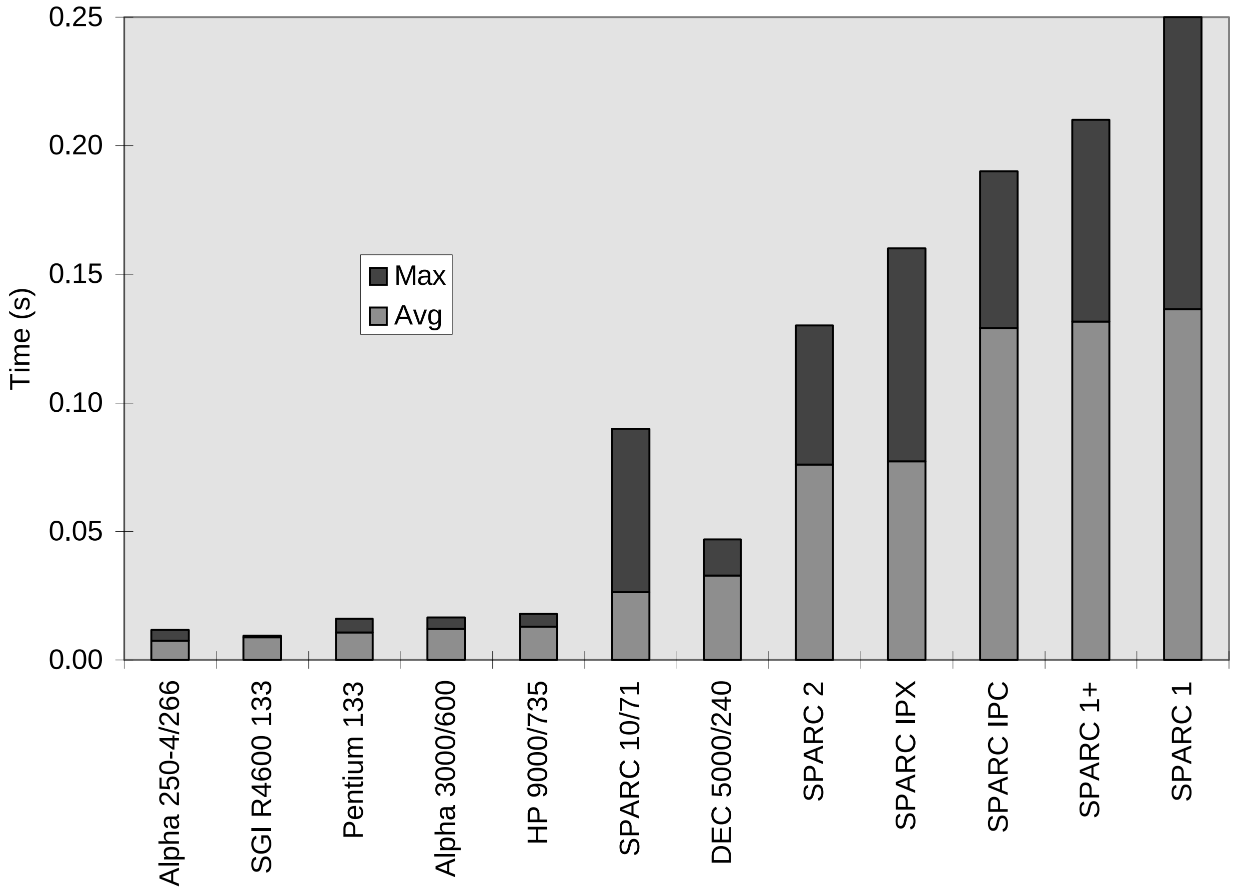
- Configuration and authentication and synchronization are inseperable
  - All three must be performed independently on a tentative basis
  - Only when all three “succeed” can a server or client claim to be authentic
  - This may require some redesign of current proposed IETF schemes
- Autonomous configuration
  - Discovery mechanisms based on DNS or service location protocols do not scale for large numbers of servers and clients
  - Finding optimal topologies in large subnet graphs under degree and distance constraints is NP-hard
  - Multicast techniques require engineered span controls to avoid congestion
- Cryptographic authentication
  - Centralized key management is incompativle with the Internet model
  - Symmetric-key cryptosystems do not scale for servers with large numbers of clients
  - Public-key cryptosystems are too slow and expensive for good timekeeping

MD5 message digest



- Measured times to construct 128-bit hash of 48-octet NTP header using MD5 algortihm in RSAREF
- NTP reference implementation computes hash time and corrections

RSA/MD5 digital verify



- Measured times to construct digital verify using RSAREF
- Message authentication code decrypted with RSA 512-bit public key and compared with MD5 hash

Authentication scheme for client/server modes

- Scheme is based on public key (RSA) encryption and one-way hash function
  - Certificated public values for server provided by Secure DNS
  - Server computes session key as one-way hash (MD5 or DES-CBC) of server private value, server/client IP addresses and key identifier as each client request is received
  - On request, server sends session key to client using public-key cryptography
- Advantages
  - Simple ot implement; requires no protocol modifications
  - Server needs no persistent state variables for clients
- Disadvantages
  - Vulnerable to certain clogging and man-in-the-middle attacks
  - Not practical for multicasting

Authentication scheme for multicast modes

- Scheme is based on public key (RSA) encryption and S-KEY scheme
  - Certificated public values for server provided by Secure DNS
  - Server generates list of session keys, where each key is a one-way hash (keyed MD5 or DES-CBC) of the previous key
  - Server uses keys in reverse order and generates a new list when the current one is exhausted;
  - Clients verify the hash of the current key equals the previous key
  - On request, the server signs the current key and sends to client
- Advantages
  - Requires few protocol changes; backwards compatible
  - Requires only one additional hash and infrequent signature verification
- Disadvantages
  - Lost packets require clients to perform repeated hashes
  - Servers are vulnerable to clogging attacks
  - Clients are vulnerable to spoofing and man-in-the-middle attacks

Autonomous configuration - approach

- Dynamic peer discovery schemes
  - Primary discovery vehicle using NTP multicast and anycast modes
  - Augmented by DNS, web and service location schemes
  - Augmented by NTP subnet search using standard monitoring facilities
- Achieving optimal configurations
  - Distance metric designed to maximize accuracy and reliability
  - Constraints due to resource limitations and maximum distance
  - Multicast scope constraints to avoid network congestion
- Candidate optimization algorithms
  - Multicast with or without initial propagation delay calibration
  - Anycast mode with administratively and/or TTL delimited scope
  - Span-limited, add-drop greedy (SLAG) heuristic
- Proof of concept based on simulation and implementation with NTP

Disconvery and configuration strategy

- NTP multicast and/or anycast modes used to discover servers within the same hierarchical group
  - Multicast scope determined by group address and/or TTL
  - Anycast beacon intervals and number of servers determined as a function of prespecified timekeeping quality and reliability parameters
  - Expanding-ring search with self-equalizing peer renewal
  - All servers authenticated using Secure DNS and certificates
- Ancestors of hierarchical group discovered from NTP peer data, augmented by NTP monitoring data
  - Database is synthesized from all available data and distributed to "interested" servers and clients
  - SLAG used to optimize the inter-group peer paths and respond to failures
  - Service location schemes used where available

Span-limited, add/drop, greedy (SLAG) heuristic

- Algorithm finds shortest path trees constrained by degree and distance
  - Distance is computed from roundtrip delay to the root (primary server) and estimated error
  - Degree constraint controls load on the servers and network
  - Distance constraint discards subtrees unsuitable for synchronization
- Span limit controls size of server/client multicast group at each hierarchical level
- Greedy characteristic minimizes the computing load on the server and client
- Add/drop feature grows and shrinks subtrees according to metric and reachability
- Health of subtrees continually assessed by NTP protocol and monitoring tools, which alert configuration algorithms if a server or path fails

New clock discipline algorithm

- Goal is to reduce residual errors a factor of ten better than now -below 100 us on fast LANs (CAIRN) 1 ms on T1 WANs and 10 ms on others
- Secondary goal is to greatly increase peer poll interval to reduce cost (telephone modems) and detectability (radio), but minimize loss in accuracy
- Scheme uses redesigned clock discipline loop based on a hybrid phase/frequency-lock loop
  - Phase-lock loop (PLL) and frequency-lock loop (FLL) independently estimate frequency
  - Prediction errors measured and used to control the weight assigned the FLL and PLL predictions
  - Design exhaustively simulated under widely varying conditions of network jitter and clock oscillator wander
- Results confirm accuracy improves a factor of ten throughout the poll range from 16 s to 1.5 days

Current progress and status

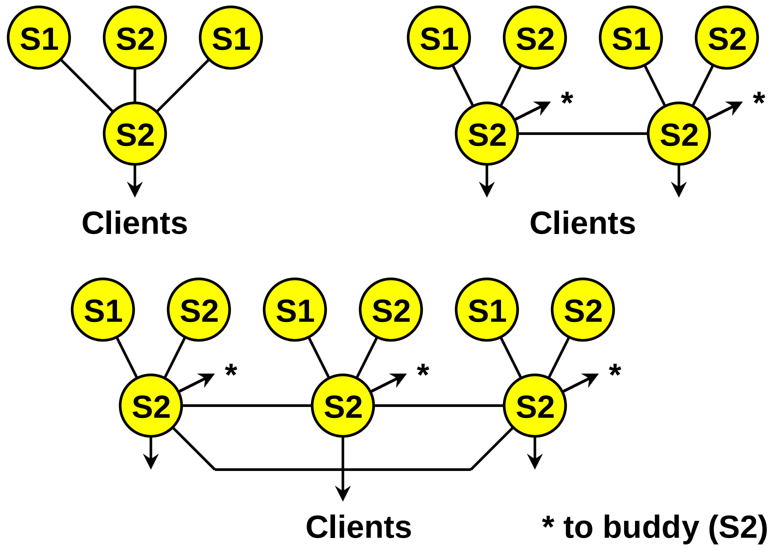
- NTP Version 4 architecture and algorithms
  - Backwards compatible with earlier versions
  - Documentation completely redone in HTML for browsing
  - New clock discipline algorithm designed and simulated; technical report in progress
  - Precision time kernel modifications for symmetric multiprocessors implemented, tested and deployed in current Digital Unix 4.0 kernel
- Multicast-based autoconfiguration schemes
  - Multicast mode with propagation calibration implemented and tested
  - Anycast mode with current multicast scoping implemented and in test
  - SLAG heuristic designed and simulated
  - Documentated in Ajit Thyagarajan's PhD dissertation
- Hybrid symmetric-key/public-key authentication schemes
  - Algorithms for client/server and multicast modes defined and documented
  - Implementation in progress
  - Documented in Electrical Engineering Report TR 96-10-3

Future plans

- Complete design and implementation of NTP Version 4 protocol model, state machine and supporting algorithms
  - Implement SLAG and integrate with other NTP Version 4 components
  - Implement new authentication scheme and integrate with other components
  - Implement new clock discipline algorithm
- Deploy, test and evaluate NTP Version 4 reference implementation
  - Test in SPARC, Alpha, RISC, HP and Intel architectures
  - Deploy and test in CAIRN/DARTnet testbed
  - Deploy and test at friendly sites in the US, Europe and Asia
- Prosecute standards agendae in IETF, ANSI, ITU, POSIX
  - Update the NTP formal specification and launch on standards track
  - Participate in deployment strategies with NIST, USNO, others
- Develop scenarios for other applications such as web caching, DNS servers and other multicast services

NTP online resources

- Internet (Draft) Standard RFC-1305 Version 3
  - Simple NTP (SNTP) Version 3 specification RFC-2030
  - Designated SAFEnet standard (Navy)
  - Under consideration in ANSI, ITU, POSIX
- NTP web page http://www.eecis.udel.edu/~ntp
  - NTP Version 3 release notes and HTML documentation
  - List of public NTP time servers (primary and secondary)
  - NTP newsgroup and FAQ compendium
  - Tutorials, hints and bibliographies
- NTP Version 3 implementation and documentation for Unix, VMS and Windows
  - Incorporated in stock kernels for Digital Unix, FreeBSD, Linux; planned for Solaris
  - Ported to over two dozen architectures and operating systems
  - Utility programs for remote monitoring, control and performance evaluation



New ideas

Dynamic peer discovery using engineered multicast schemes

Automatic, authenticated peer selection for ubiquitous distributed protocols

Multiple overlapping hierarchical subtree organization for redundancy and diversity

Maximize service performance relative to crafted metric and constraints

Impact

Minimize dependence on prior configuration data

Avoid intricate case-by-case analysis of failure/fallback/recovery scenarios

Provide automatic reconfiguration in case of network reconfiguration or failure

Schedule (second year)

Design NTP cryptographic authentication extensions for multicast (Aug 96)

Design NTP protocol enhancements for distributed mode (Aug 96)

Integrate with existing NTP (Feb 97)