NOW year 1 ARPA Review

David E. Culler

Motivation

- **Emerging "Killer Networks" represent a fundamental shift in how we design large-scale systems**
  - **serve a large number of users**
  - **solve large problems fast**
  - **accomplish a large number of tasks**

- **The hardware technology is evolving, but must**

  => rethink the communication abstraction

  => rethink system design in the new paradigm

Key Technological Change

- **Scalable, high-bandwidth, low-latency communication**

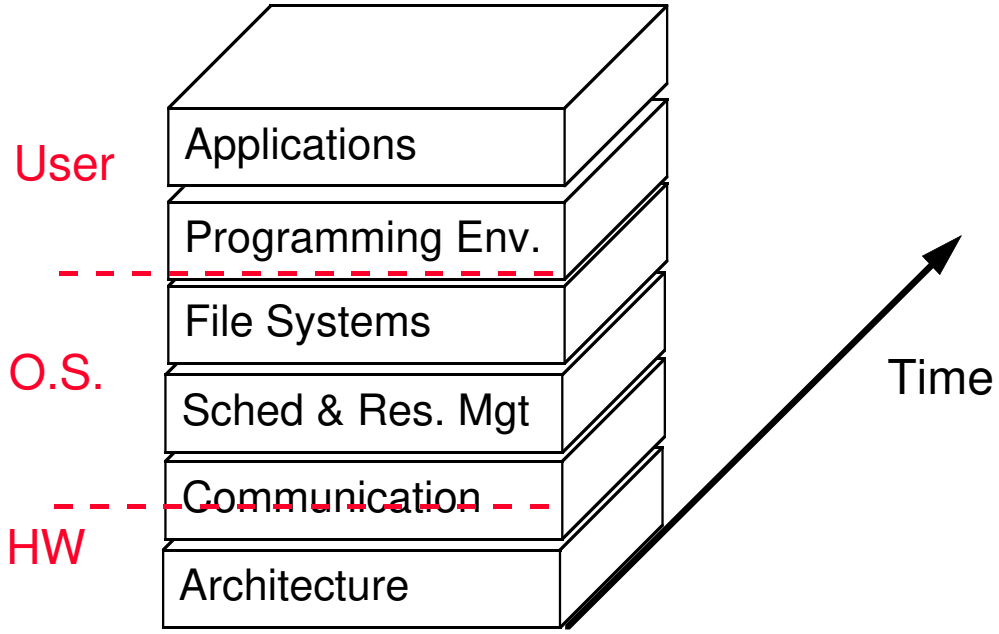
=> remote memory nearer than local disk

=> Processors can cooperate on shared tasks toward overall system good

=> inidividual nodes are powerful, general purpose, and available

- **Pool processing, storage, and bandwidth resources**
- **structure sub-system and appln design for *performance, scalability and availability***

NOW Vertical Slice



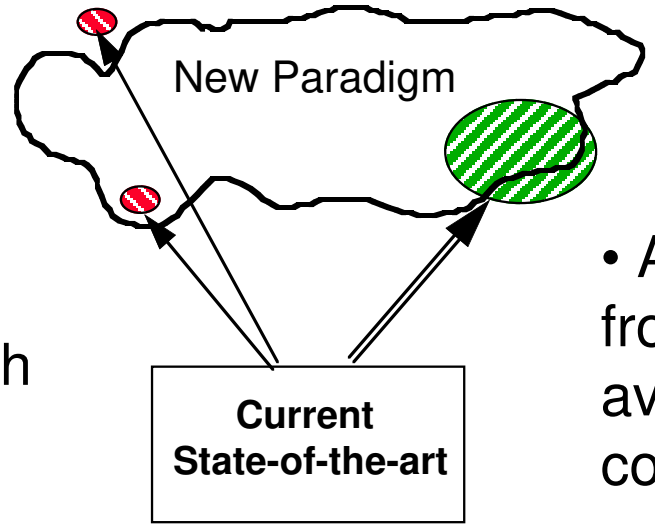
- Focus on innovation at each level
  - **acquire the rest of the infrastructure**
- Get into the regime, develop & explore core concepts
- Move the entire slice forward (rather than big bang)

NOW targets

- **Deliver low latency, full bandwidth communication to general purpose programs and the system in a scalable network.**
- **Deliver full aggregate disk bandwidth to *n* clients, full network bandwidth to one, while retaining low response time.**
- **Deliver full processing & storage resources to many sequential, few parallel or mix of jobs.**
- **Demonstrate fast, scalable, available applications (on demanding problems).**
- **Demonstrate effective maintenance, tuning, and diagnosis.**

Two ways to “sandbox the future”

- Invent it from scratch

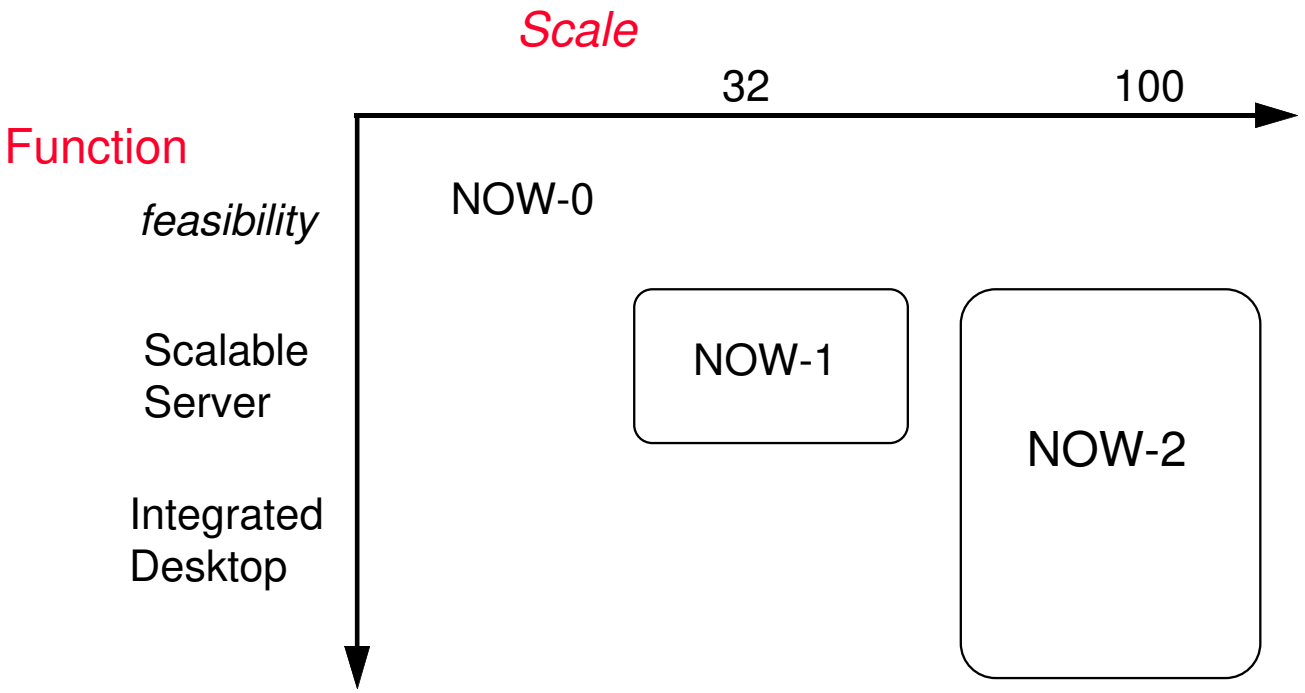


- Approximate it from best available components

Why the NOW approach

- **Build upon a substantial basis**
  - **conceptual foundations and technical infrastructure**
  - **MPPs, Distributed Systems, and Industry**
- **Seek broad-based solutions, rather than "silver bullets"**
  - **complete systems, I/O, VM, protection, availability, . . .**
- **Quantifiable performance, not just abstractions**
- **Synergy in the vertical integration**
- **Relevance doesn't hurt**
  - **exploit trends, leverage state-of-the-art**

Stages of Development



Today's A.M. Presentations

- **Communication (AM)**
  - build a "killer enough" network to enable novel subsystem and application design
- **File System (xFS)**
  - first truely distributed file system
- **Scheduling and Resource Management (Glunix)**
  - making sense of a huge resource pool
- **Scalable Web Service (Inktomi)**
  - world class performance on a NOW
  - scalable, available application

Other Areas to Note

- **Parallel Programming Environment**
  - **Split-C, Mantis**
  - **MPI**
- **Applications and Benchmarking**
  - **database**
- **Network RAM**
- **System Administration**
- **Integrating the desktop environment**
- **Tertiary Disk Storage**

Common Issues

- **Performance**
- **Scalability**
- **Availability**
  - **Graceful degradation on failure**
  - **Dynamic reconfiguration and tuning**

- ***"Federation"***