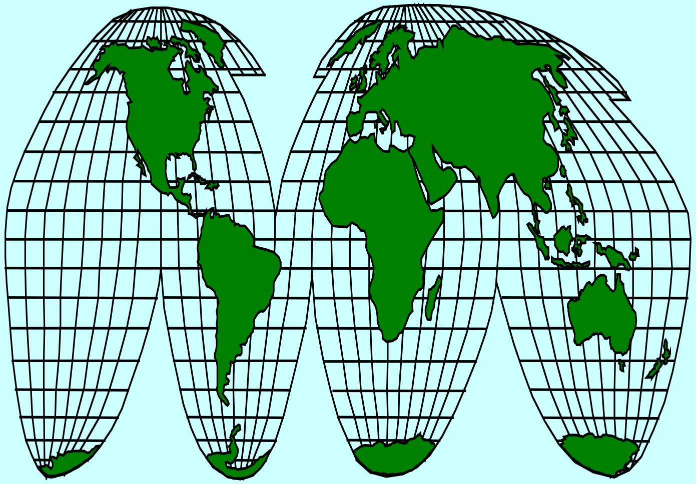
MONARC

Second Regional Centre Representatives Meeting



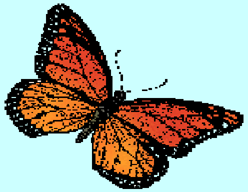
Harvey B. Newman (Caltech)

CERN

August 26, 1999

MONARC August 26 Agenda (1)

◆ Introduction	H. Newman		20’
◆ Reports from the Working Groups			70’
➨ Preliminary Document on Tier2 Centres and Architecture Open Issues	L. Barone	(7)	
➨ Analysis Process WG Status	P. Capiluppi	(7)	
➨ Regional Centre Activity Model	I. Gaines	(15)	
➨ Status of Simulation	I. Legrand	(10)	
➨ Status of Validation Activities	Y. Morita	(15)	
➨ Testbeds WG Status			
◆ Reports from Regional Centres			60’
➨ France (IN2P3)	D. Linglin		
➨ Japan (KEK)	K. Amako		
➨ Russia (JINR; Protvino)	V. Korenkov		
➨ US (ATLAS; BNL)	B. Gibbard		
➨ US (CMS; FNAL)	M. Kasemann		
➨ UK	S. O’Neale		
➨ Italy	E. Valente		
➨ Germany	A. Putzer		

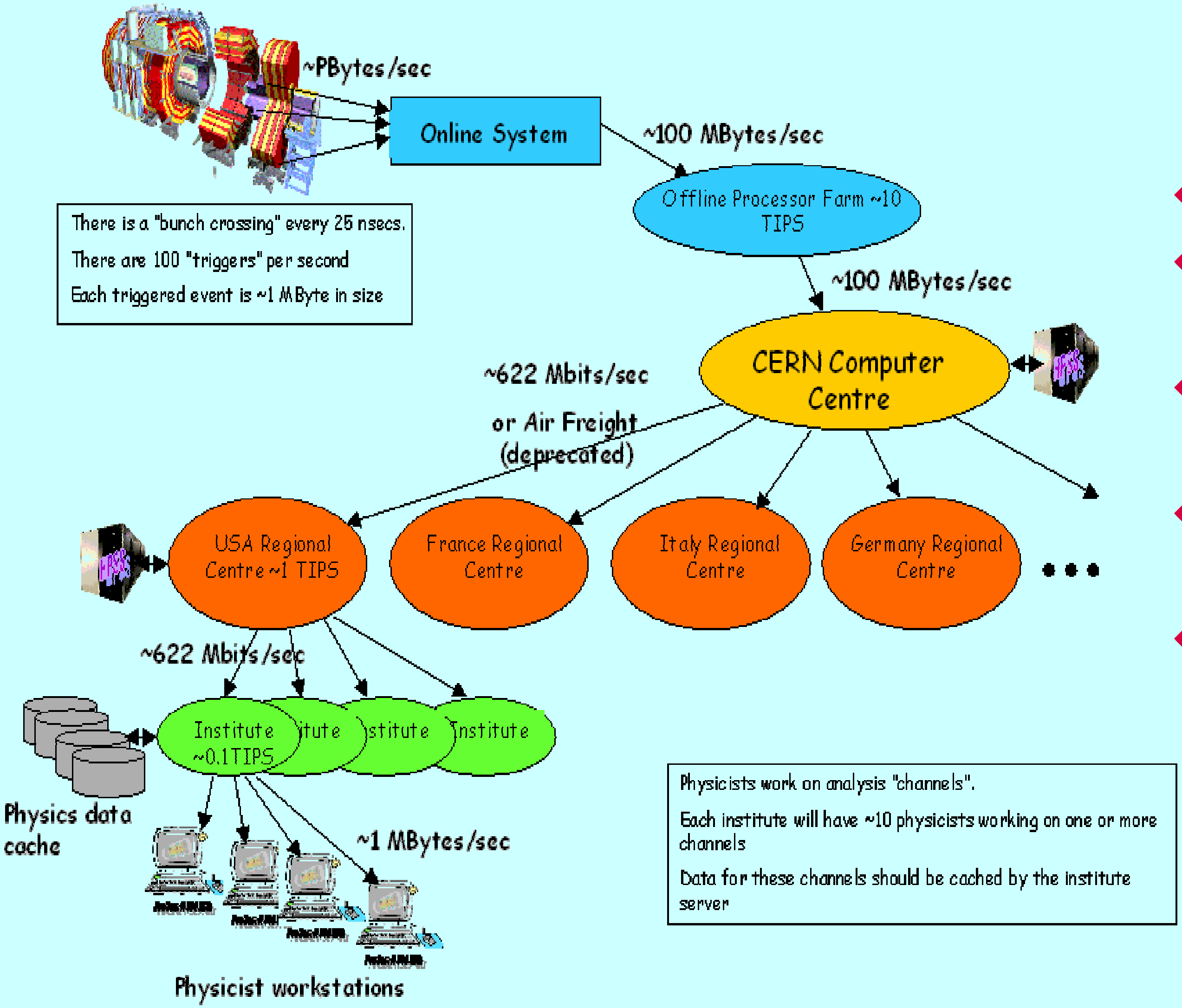


MONARC August 26 Agenda (2)

- Discussion and Preparations for LCB Marseilles Meeting 60’
  - Marseilles Meeting Purpose and Organization — M. Mazzucato
  - MONARC Deliverables by Marseilles — L. Perini
  - Remarks on Marseilles and CERN Review LHC Computing — H. Hoffmann
  - Maintenance and Operations costs for LHC Computing — L. Robertson
  - Role of R&D for LHC Computing — H. Newman
  - Discussion
- AOB; Next Meeting
- Adjourn Main Meeting by 19:30

- Steering Committee Meets for ~30 Minutes; immediately following the main meeting

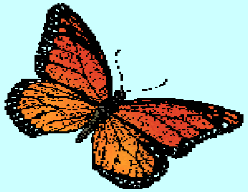
MONARC “Data Grid” Hierarchy Regional Centers Concept



LHC Grid Hierarchy Example

- Tier0: CERN
- Tier1: National “Regional” Center
- Tier2: Regional Center
- Tier3: Institute Workgroup Server
- Tier4: Individual Desktop

Total 5 Levels



Grid-Hierarchy Concept

Matched to the Worldwide-Distributed Collaboration Structure of LHC Experiments

Best Suited for the Multifaceted Balance Between

- **Proximity of the data to centralized processing resources**
- **Proximity to end-users for frequently accessed data**
- **Efficient use of limited network bandwidth (especially transoceanic; and most world regions)**
- **Appropriate use of (world-) regional and local computing and data handling resources**
- **Effective involvement of scientists and students in each world region, in the data analysis and the physics**

MONARC Architectures WG

- Architectures WG: Lessons and Challenges for LHC Computing
  - SCALE: 60 Times more CPU and 10 Times more Data than CDF at Run2 (2000-2003)
  - DISTRIBUTION: Mostly Achieved in HEP Only for Simulation. For Analysis (and some re-Processing), it will not happen without advance planning and commitments
  - REGIONAL CENTRES: Require Coherent support, continuity, the ability to maintain the code base, calibrations and job parameters up-to-date
  - HETEROGENEITY: Of facility architecture and mode of use, and of operating systems, must be accommodated.
  - FINANCIAL PLANNING: Analysis of the early planning for the LEP era showed a definite tendency to underestimate the requirements (by more than an order of magnitude)
    - Partly due to budgetary considerations

Regional Centres Should Provide

- All technical and data services required to do physics analysis
- All Physics Objects, Tags and Calibration data
- Significant fraction of raw data
- Caching or mirroring calibration constants
- Excellent network connectivity to CERN and the region's users
- Manpower to share in the development of common maintenance, validation and production software
- A fair share of post- and re-reconstruction processing
- Manpower to share in the work on Common R&D Projects
- Service to members of other regions on a (?) best effort basis
- Excellent support services for training, documentation, troubleshooting at the Centre or remote sites served by it

Long Term Commitment for staffing, hardware evolution and support for R&D, as part of the distributed data analysis architecture

To Solve: the HENP "Data Problem"

While the proposed future computing and data handling facilities are large by present-day standards,
They will not support FREE access, transport or reconstruction for more than a Minute portion of the data.

- ➨ Need for effective global strategies to handle and prioritise requests, based on both policies and marginal utility
- ➨ Strategies must be studied and prototyped, to ensure *Viability*: acceptable turnaround times; efficient resource utilization

- **Problems to be Explored; How To**
  - ➨ Meet the demands of hundreds of users who need transparent access to local and remote data, in disk caches and tape stores
  - ➨ Prioritise thousands of requests from local and remote communities
  - ➨ Ensure that the system is dimensioned "optimally", for the aggregate demand

Roles of Projects for HENP Distributed Analysis

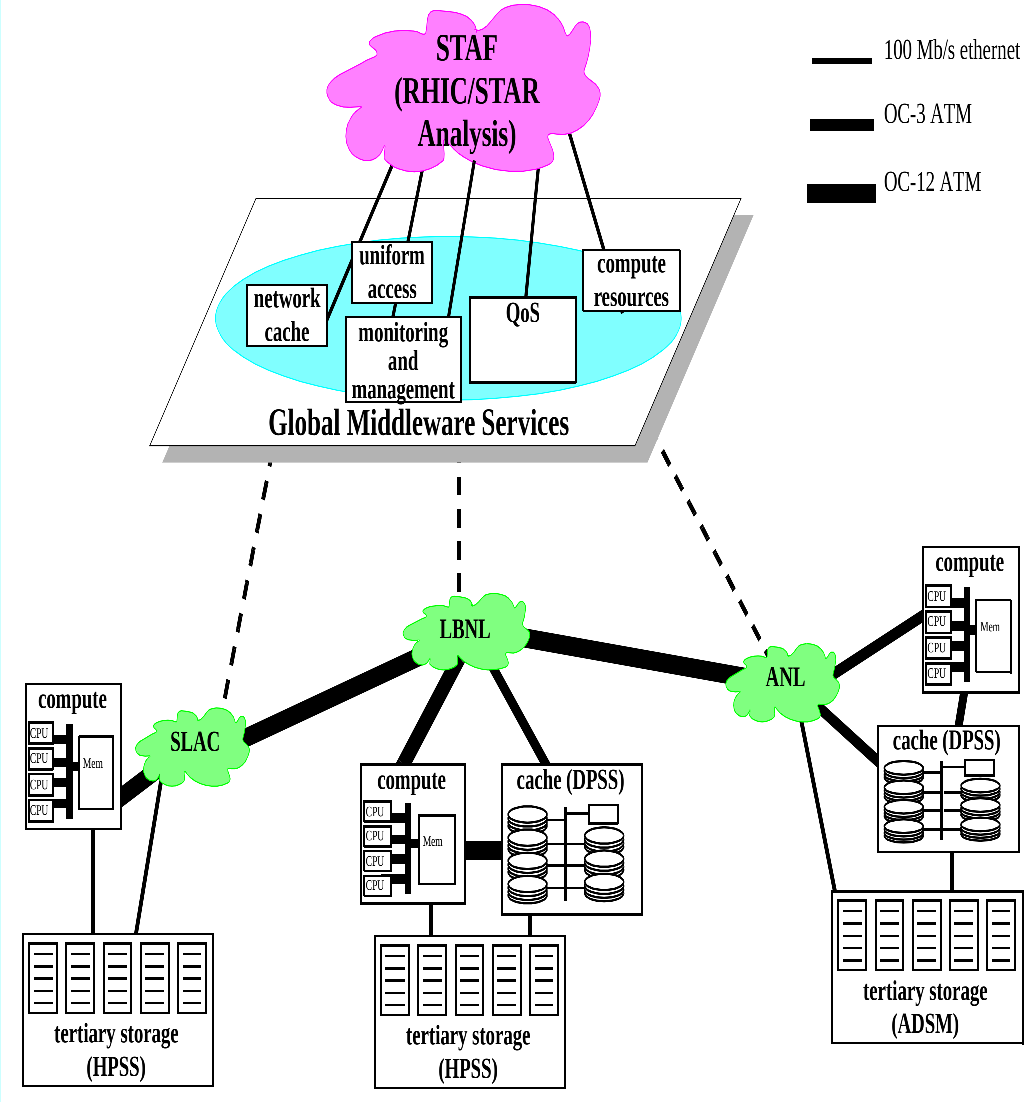
- RD45, GIOD: Networked Object Databases
- Clipper/GC; FNAL/SAM: High speed access to Objects or File data for processing and analysis
- SLAC/OOFS: Distributed File System + Objectivity Interface
- NILE, Condor: Fault Tolerant Distributed Computing with Heterogeneous CPU Resources
- MONARC: LHC Computing Models: Architecture, Simulation, Strategy
- PPDG: First Distributed Data Services and Grid System Prototype
- ALDAP: OO Database Structures and Access Methods for Astrophysics and HENP data

APOGEE: *Full-Scale Grid Design, Optimization, Prototyping*

- *Network Instrumentation Team*
- *Simulation/Modeling Team*
- *System Optimization/Evaluation Team*

- GriPhyN: Production Prototype Grid in Hardware and Software, then Production

HENP Grand Challenge/Clipper Testbed and Tasks



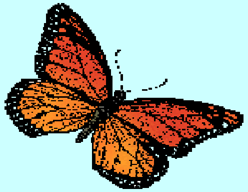
High-Speed Testbed
- Computing and networking (NTON, ESnet) infrastructure

Differentiated Network Services
- Traffic shaping on ESnet

End-to-end Monitoring Architecture (QE, QM, CM)
- Traffic analysis, event monitor agents to support traffic shaping and CPU scheduling

Transparent Data Management Architecture
- OOFS/HPSS, DPSS/ADSM

Application Demonstration
- Standard Analysis Framework (STAF)
- Access data at SLAC, LBNL, or ANL (net and data quality)



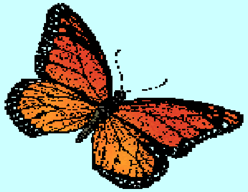
ALDAP (NSF/KDI) Project

ALDAP: *Accessing Large Data Archives in Astronomy and Particle Physics*

NSF Knowledge Discovery Initiative (KDI)

CALTECH, Johns Hopkins, FNAL(SDSS)

- Explore advanced adaptive database structures, physical data storage hierarchies for archival storage of next generation astronomy and particle physics data
- Develop spatial indexes, novel data organizations, distribution and delivery strategies, for *Efficient and transparent access to data across networks*
- Create prototype network-distributed data query execution systems using Autonomous Agent workers
- Explore commonalities and find effective *common solutions* for particle physics and astrophysics data



The Particle Physics Data Grid (PPDG)

DoE/NGI Next Generation Internet Program

ANL, BNL, Caltech, FNAL, JLAB, LBNL, SDSC, SLAC, U.Wisc/CS

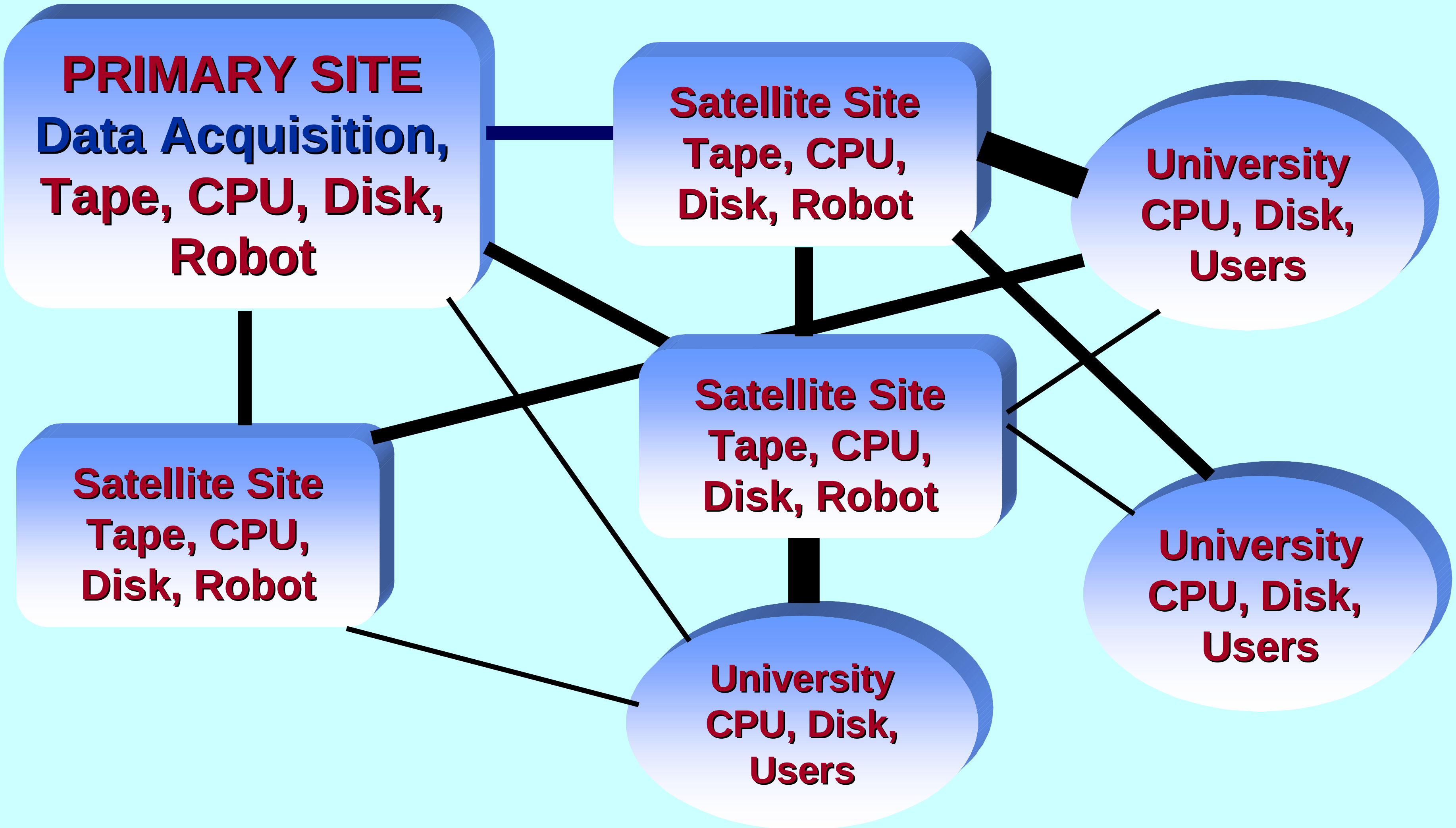
- Drive progress in the development of the necessary middleware, networks and distributed systems.
- Deliver some of the infrastructure for widely distributed data analysis at multi-PetaByte scales by 100s to 1000s of physicists

Implement and Run two services in support of the major physics experiments :

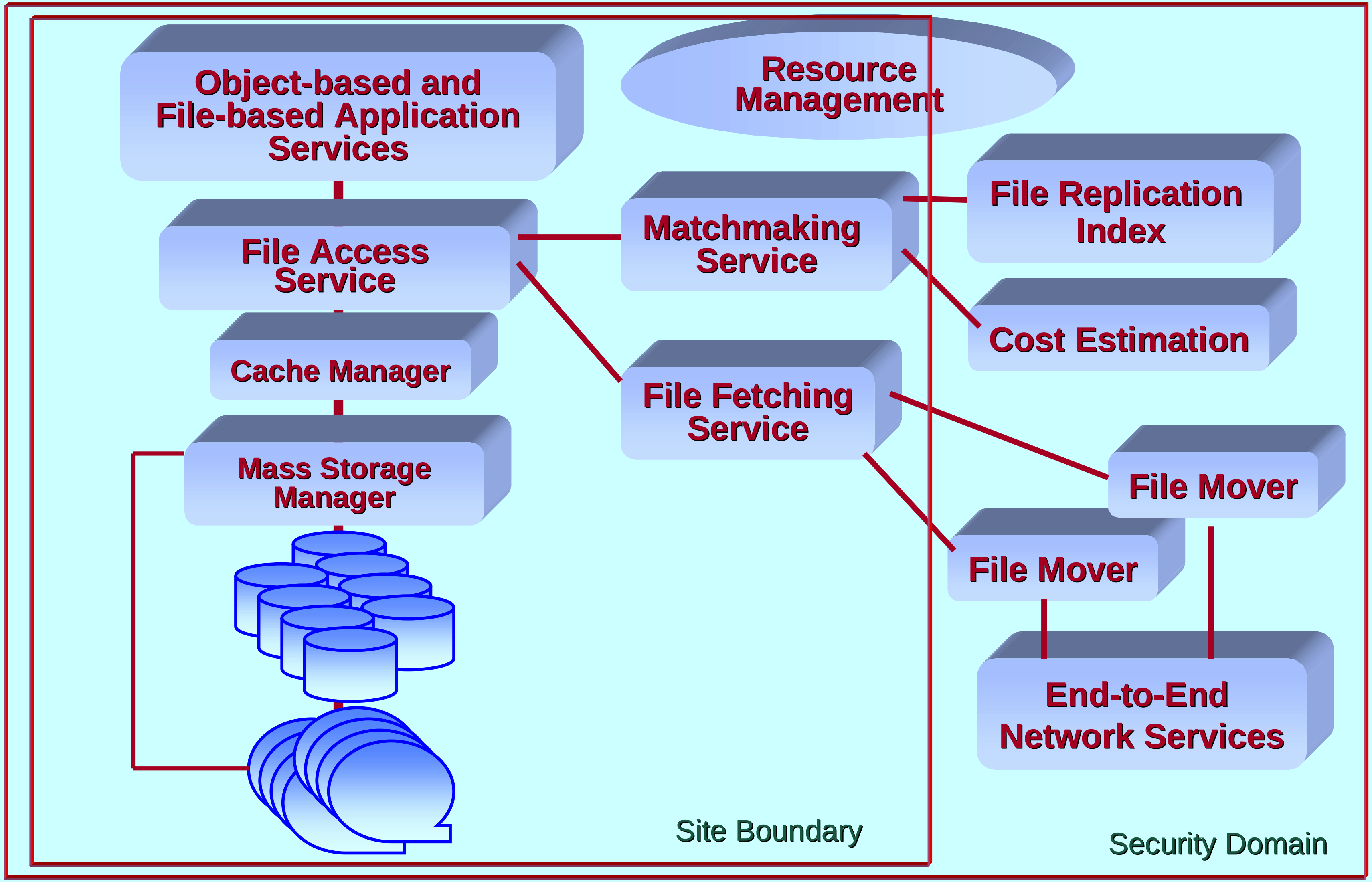
- *"High-Speed Site-to-Site File Replication Service";*
  Data replication up to 100 Mbytes/s
- *"Multi-Site Cached File Access Service":*
  Based on deployment of file-cataloging, and transparent cache-management and data movement middleware
- Deliverables for 2000: Optimized cached read access to file in the range of 1-10 Gbytes, from a total data set of order One Petabyte

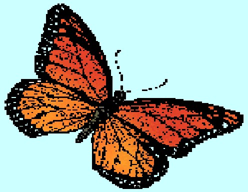
Using middleware components already developed by the Proponents

PPDG Multi-site Cached File Access System



PPDG Middleware Components





First Year PPDG "System" Components

Middleware Components (Initial Choice): See PPDG Proposal

File Access Service	Components of OOFS (SLAC)
Cache Manager	GC Cache Manager (LBNL)
Matchmaking Service	Condor (U. Wisconsin)
File Replication Index	MCAT (SDSC)
Transfer Cost Estimation Service	Globus (ANL)
File Fetching Service	Components of OOFS
File Movers(s)	SRB (SDSC); Site specific
End-to-end Network Services	Globus tools for QoS reservation
Security and authentication	Globus (ANL)

GriPhyN: First Production Scale "Grid Physics Network"

Develop a New Form of Integrated Distributed System, while Meeting Primary Goals of the US LIGO and LHC Programs

- Single Unified GRID System Concept; Hierarchical Structure
- Three (Sub-)Implementations, for LIGO, US CMS, US ATLAS:
  - 17 Centers: 5 Each in US for LIGO, CMS, ATLAS; 2 At CERN
- Aspects Complementary to Centralized DoE Funding
  - University-Based Regional Tier2 Centers, Partnering with the Tier1 Centers
  - Emphasis on Training, Mentoring and Remote Collaboration

Making the Process of Search and Discovery Accessible to Students

GriPhyN Web Site: http://www.phys.ufl.edu/~avery/mre/

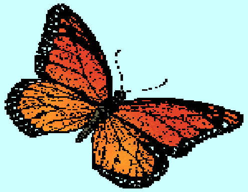
White Paper: http://www.phys.ufl.edu/~avery/mre/white_paper.html

APOGEE/GriPhyN Foundation

Build on the Distributed System Results of the MONARC, GIOD, Clipper/GC and PPDG Projects

- Long Term Vision in Three Phases
  - 1. Read/write access to high volume data and processing power
    - Condor/Globus/SRB + NetLogger components to manage jobs and resources
  - 2. WAN-distributed data-intensive Grid computing system
    - Tasks move automatically to the “most effective” Node in the Grid
    - Scalable implementation using mobile agent technology
  - 3. “Virtual Data” concept for multi-PB distributed data management, with large-scale Agent Hierarchies
    - Transparently match data to sites, manage data replication or transport, co-schedule data & compute resources

Build on VRVS Developments for Remote Collaboration



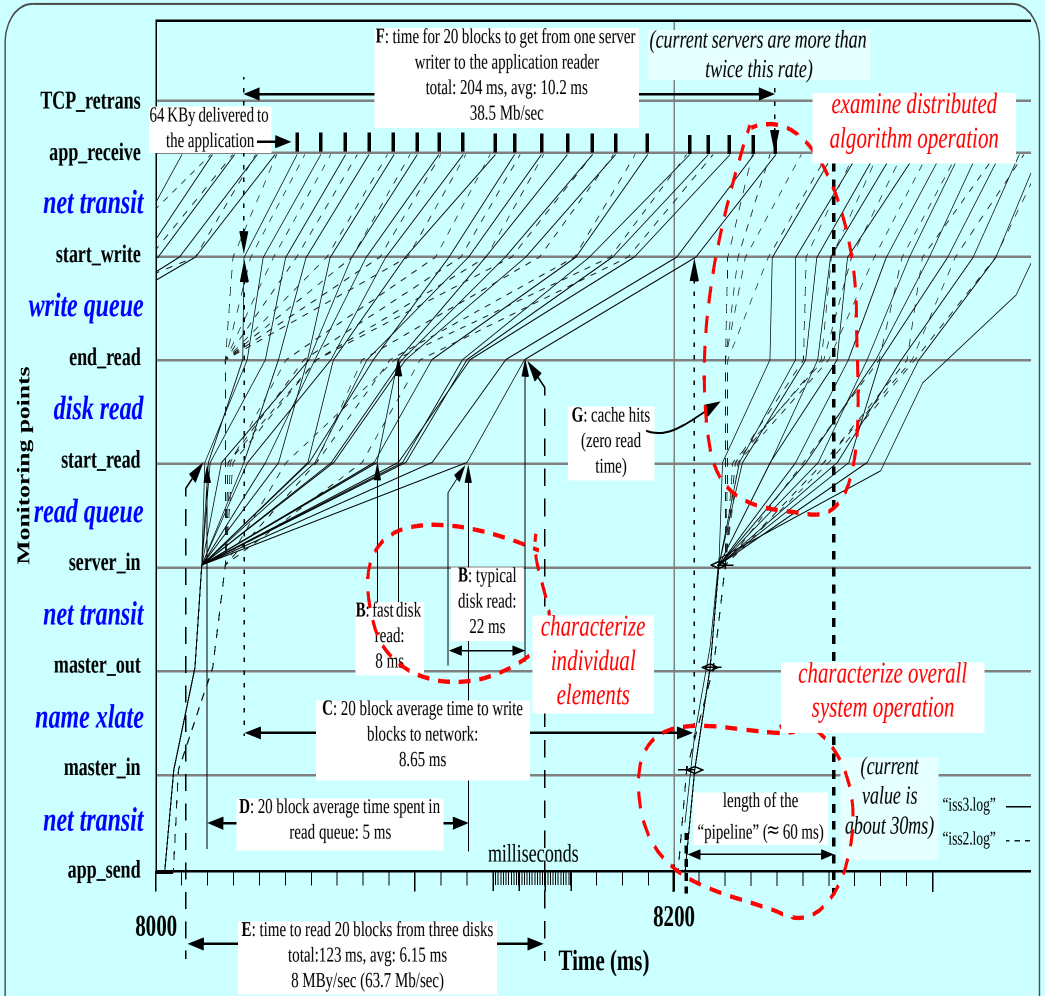
Monitoring Architecture: Use of NetLogger as in CLIPPER

End-to-end monitoring of grid assets is necessary to
- Resolve network throughput problems
- Dynamically schedule resources

Add precision-timed event monitor agents to:
- ATM switches
- Data Servers
- Testbed computational resources

Produce trend analysis modules for monitor agents

Make results available to applications



Netlogger: Two server, ATM *LAN* Event lifelines (detailed event analysis)

APOGEE: Production-Design of an HENP Data Analysis Grid

INSTRUMENTATION, SIMULATION, OPTIMIZATION, COORDINATION

- SIMULATION of a Production-Scale Grid Hierarchy
  - Provide a Toolset for HENP experiments to test and optimize their data analysis and resource usage strategies
- INSTRUMENTATION of Grid Prototypes
  - Characterize the Grid components' performance under load
  - Validate the Simulation
  - Monitor, Track and Report system state, trends and "Events"
- OPTIMIZATION of the Data Grid
  - Genetic algorithms, or other evolutionary methods
  - Deliver optimization package for HENP distributed systems
  - Applications to other experiments; accelerator and other control systems; other fields
- COORDINATE with Other Projects, MONARC, ALDAP, GriPhyN, and Experiment-Specific Projects: BaBar, Run2; RHIC, JLAB; LHC