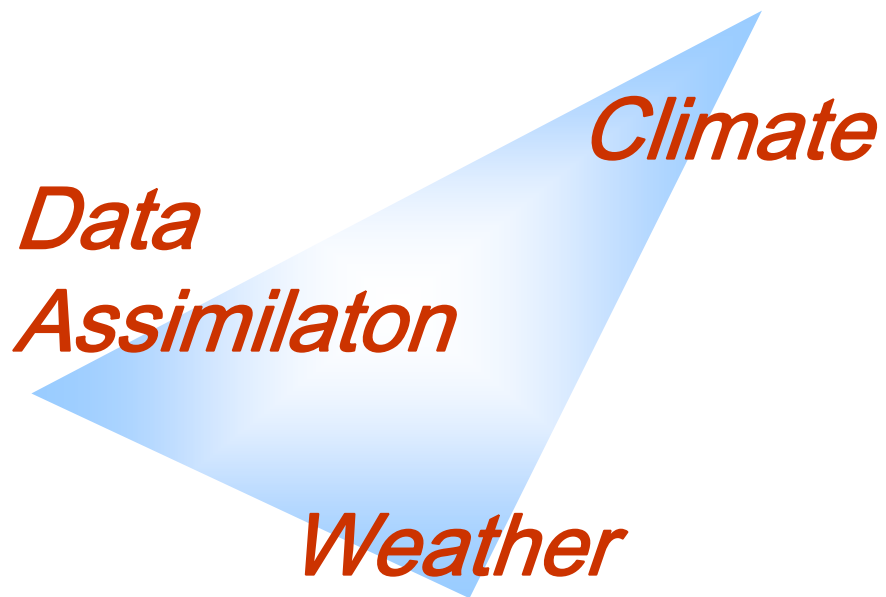


ESMF/Curator Status



Cecelia DeLuca
cdeluca@ucar.edu

CCSM Software Engineering Working Group
Boulder, CO
March 16, 2007

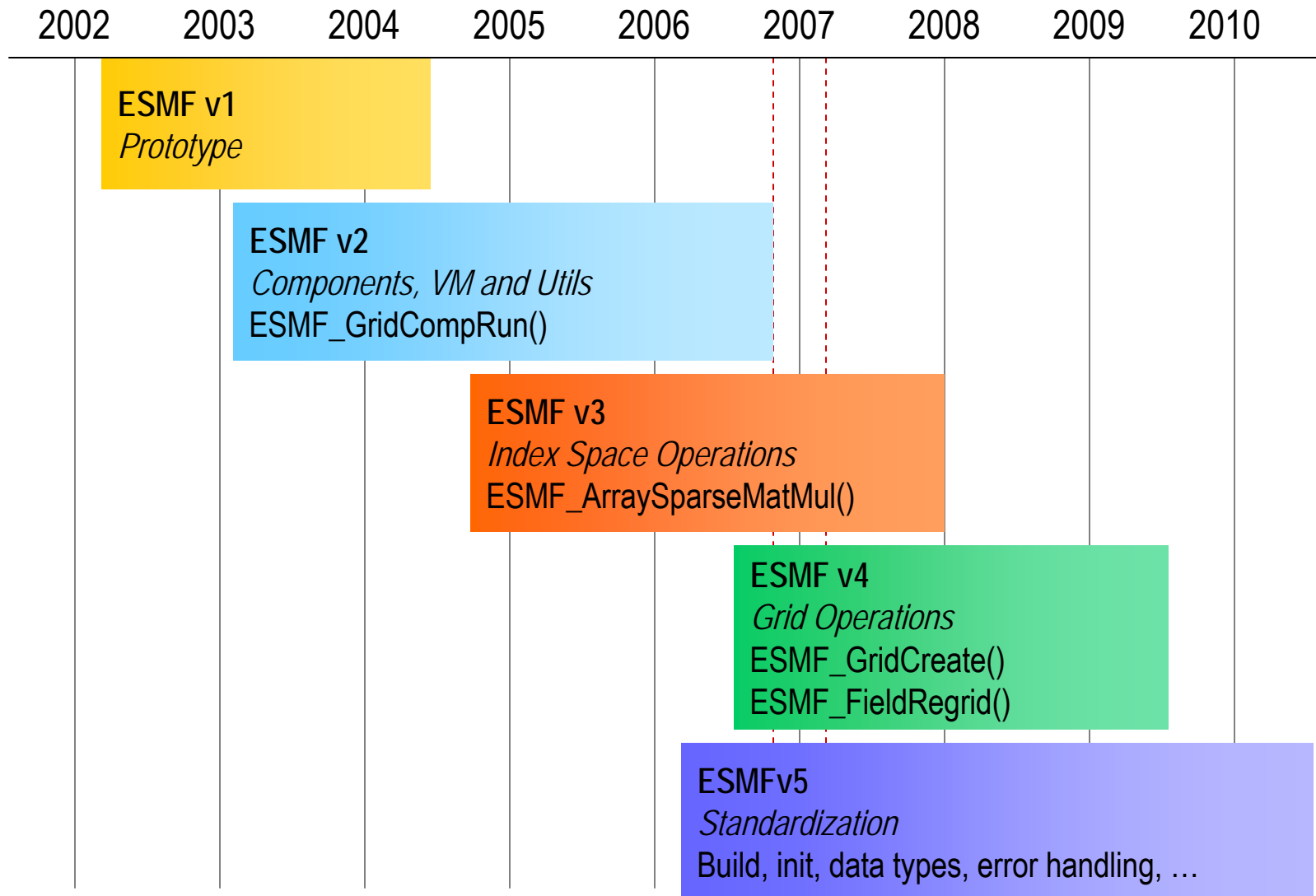


Outline

- Overview of the ESMF development plan
- Progress towards strategic goals
- Other ESMF activities
- Moving towards integrated systems
- Looking forward with Earth System Curator



ESMF Release Plan



ESMF v2.2.2r ESMF v3.0.1



GOAL:

Implementation of Arrays and Grids

Delivery of an ESMF release with a full implementation of the data block and distribution class interface, the ability to read in grid specifications for unstructured and multi-patch curvilinear coordinate grids, and the ability to generate interpolation weights for regridding such grids.



Index Space Operations

Current work (Theurich)

- TKR overloading of distributed array create (3.0.2)
- Topology information in dist grid (3.0.2)

Upcoming

- Further optimization of array sparse mat mul (3.0.3)
- Array bundles (3.0.3)
- Array halo (3.0.3)
- Array gather and scatter (3.0.4)



Grid Operations

- Structured grids (Oehmke)
 - Design targeting multi-patch, curvilinear grids
 - Single-patch curvilinear grids are being implemented first
 - Design review in progress (on 3rd iteration)
 - WRF-HYCOM is the use-test case for single patch
 - GFDL MOM4-cubed sphere atm is the use-test case for multi-patch
 - Grid representation with array extract, index space sparse mat mul in 3.0.3
- Unstructured grids (Neckels)
 - Representation of meshes for hydrology and related codes
 - Must relate to exchange grids (for NASA, DoD, GFDL) and observational data streams
 - Design reviews just beginning
- Test harness (Stark)
 - Grid operations span a huge parameter space and we need automation to specify tests and process output
 - Harness uses a shorthand syntax to describe and classify data, data decomposition and grid options, and resource files for specific configurations
 - Design review in progress (on 3rd iteration)
 - Scheduled for completion with 3.0.3

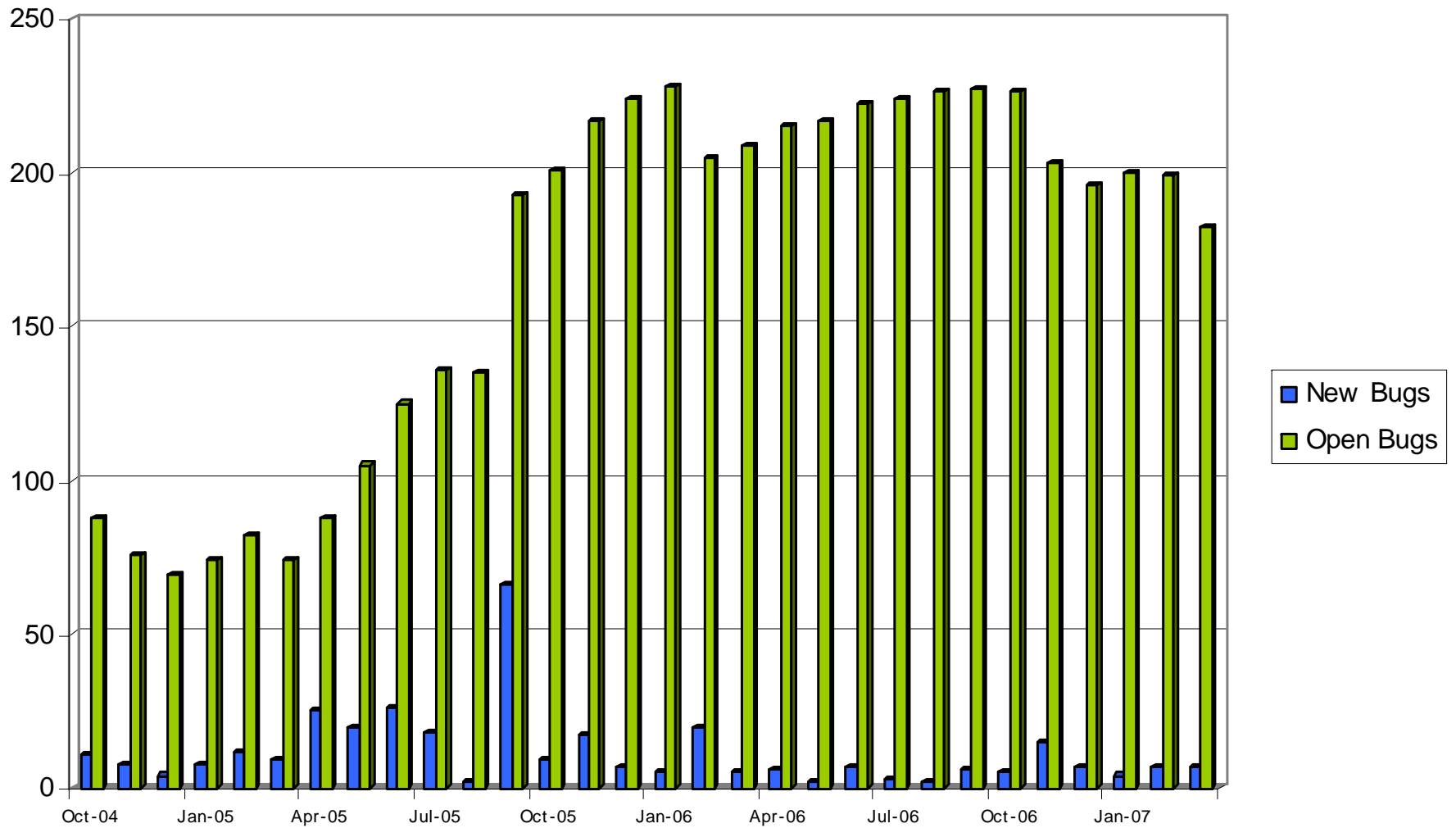


GOAL: Improve Usability

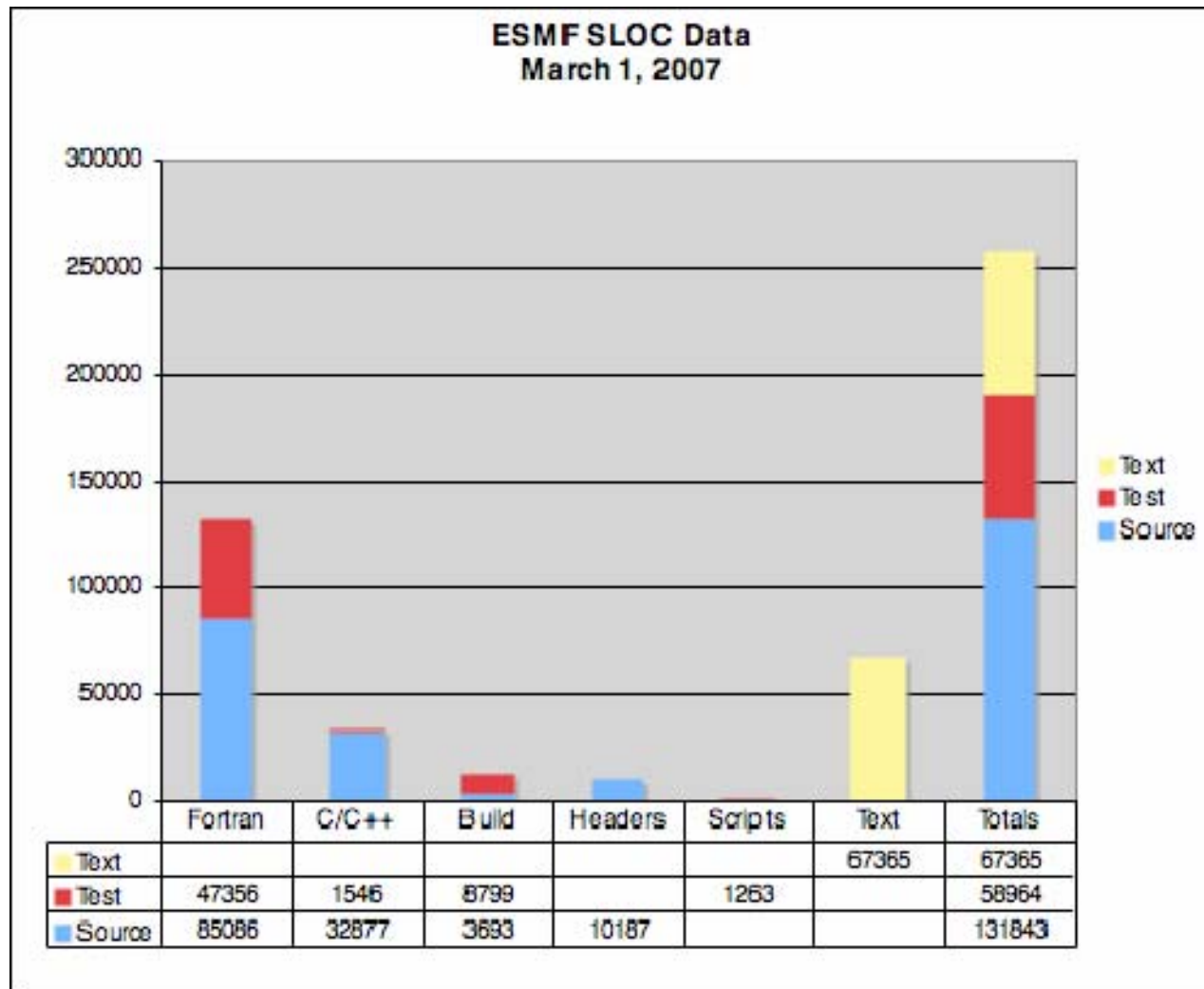
Substantial reduction in the number of open bug reports, support requests, and untested interfaces and features. Delivery of a release that includes improvements in consistency in the treatment of basic data types, method argument conventions, and inter-language interfaces.



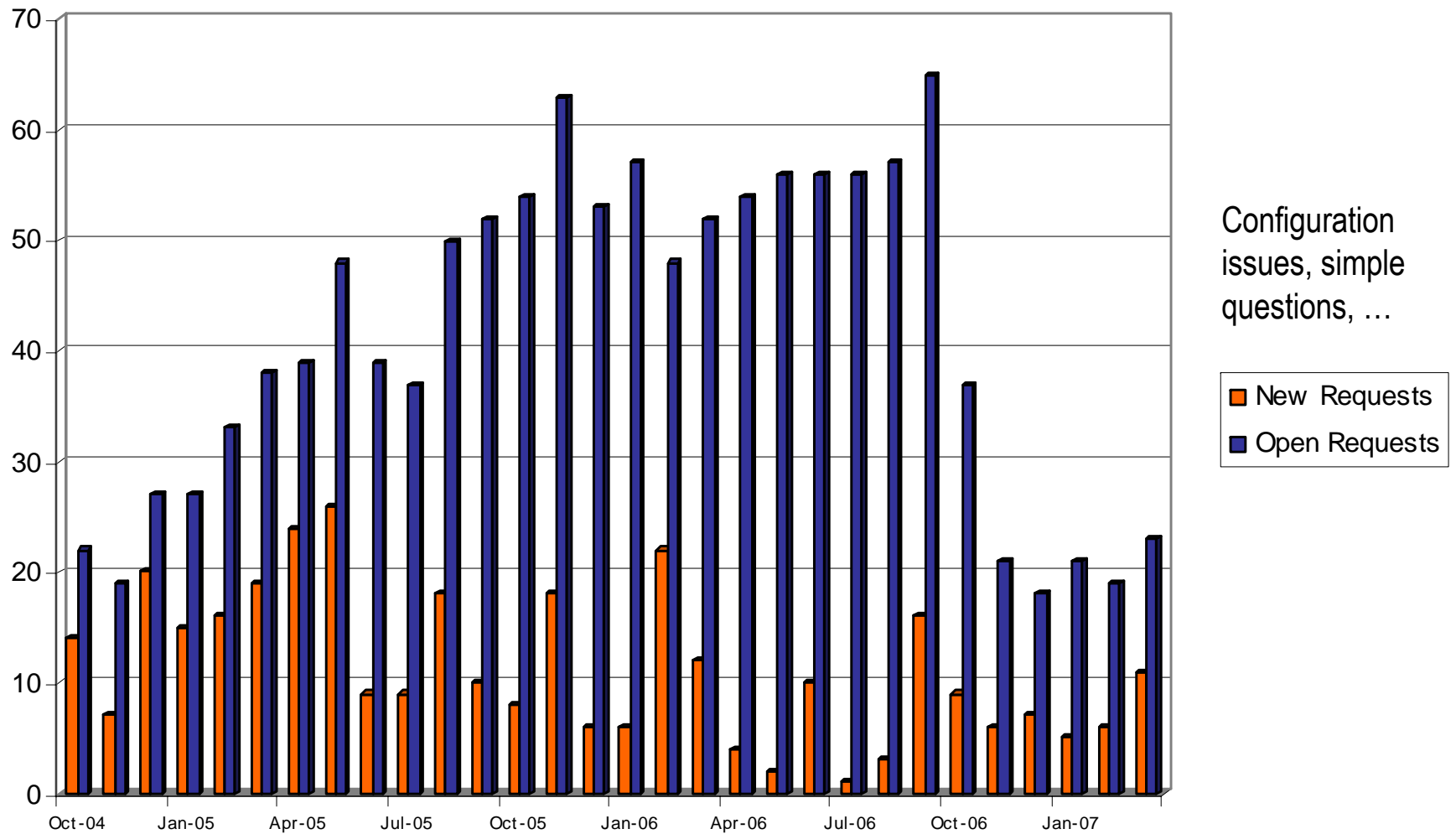
Metrics: Bugs



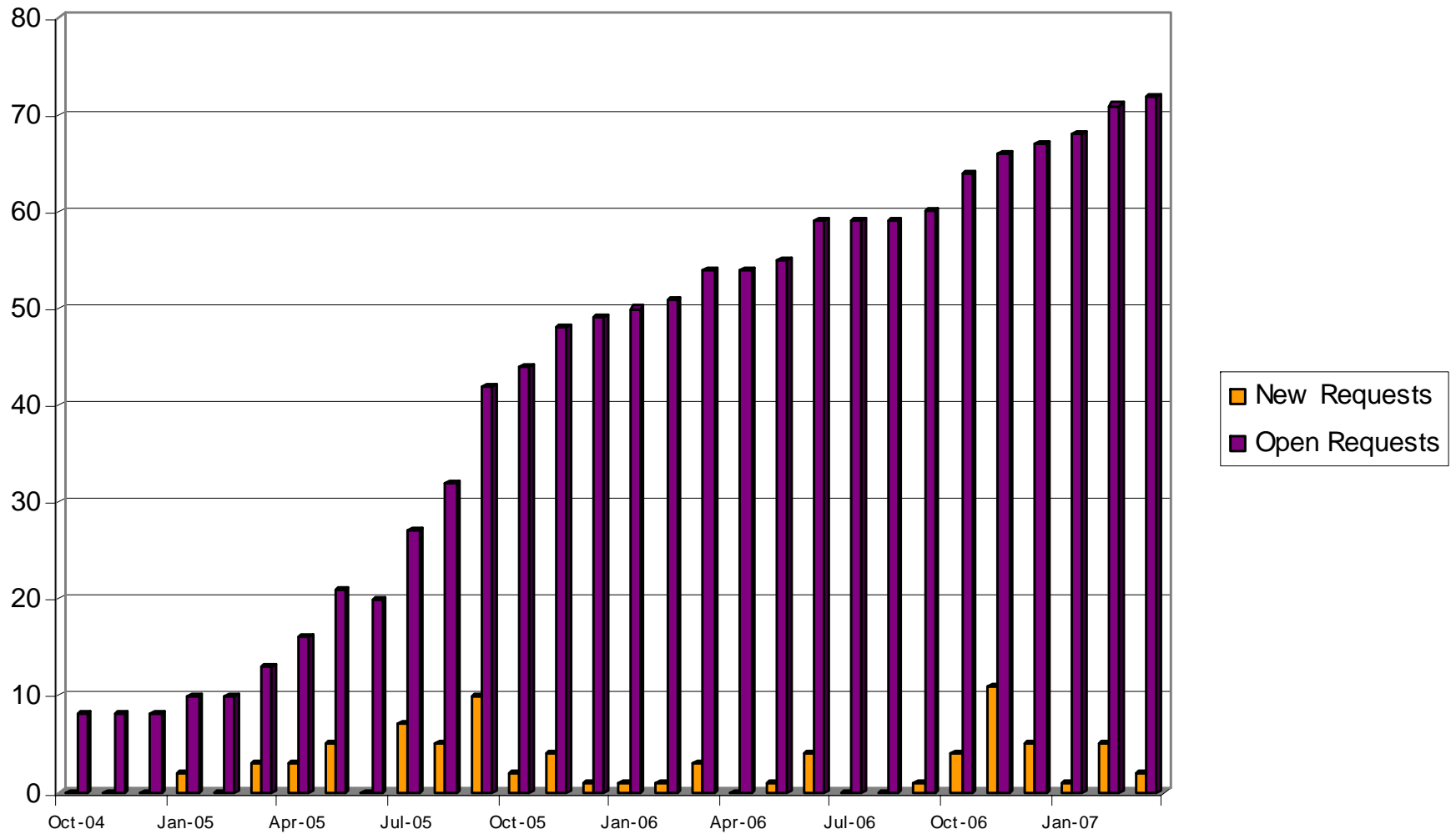
Metrics: Source Lines of Code



Metrics: Support Requests



Metrics: Feature Requests



Standardization and cleanup

Completed and current

- Build rework (2.2.2r)
- Comprehensive memory leak repair (2.2.2r)
- Initialization checks (3.0.1)
- Removal of extraneous files and documentation files shrank source to 1/10 its previous size (3.0.1)
- Auto-promotion policy (3.0.2)
- Real and integer data type standardization (3.0.2)
- Standard error handling (3.0.2)

Still to come (not all this year)

- Interlanguage interface consistency
- Standard handling of logicals and characters
- Standardization of method arguments
- Method-by-method reviews for adherence to conventions and proper behavior



GOAL: Utility Development

*Delivery of an I/O system to support complex grid implementation.
Delivery of a public ESMF release that includes new features for existing utilities. These will be prioritized in response to user input.*



Utility Improvements

Completed and current

- Config set implemented for NCEP
- Option to retrieve MPI communicator at the component interface for NCEP

Eventually

- I/O, which has been postponed again – but a petascale I/O project will be included



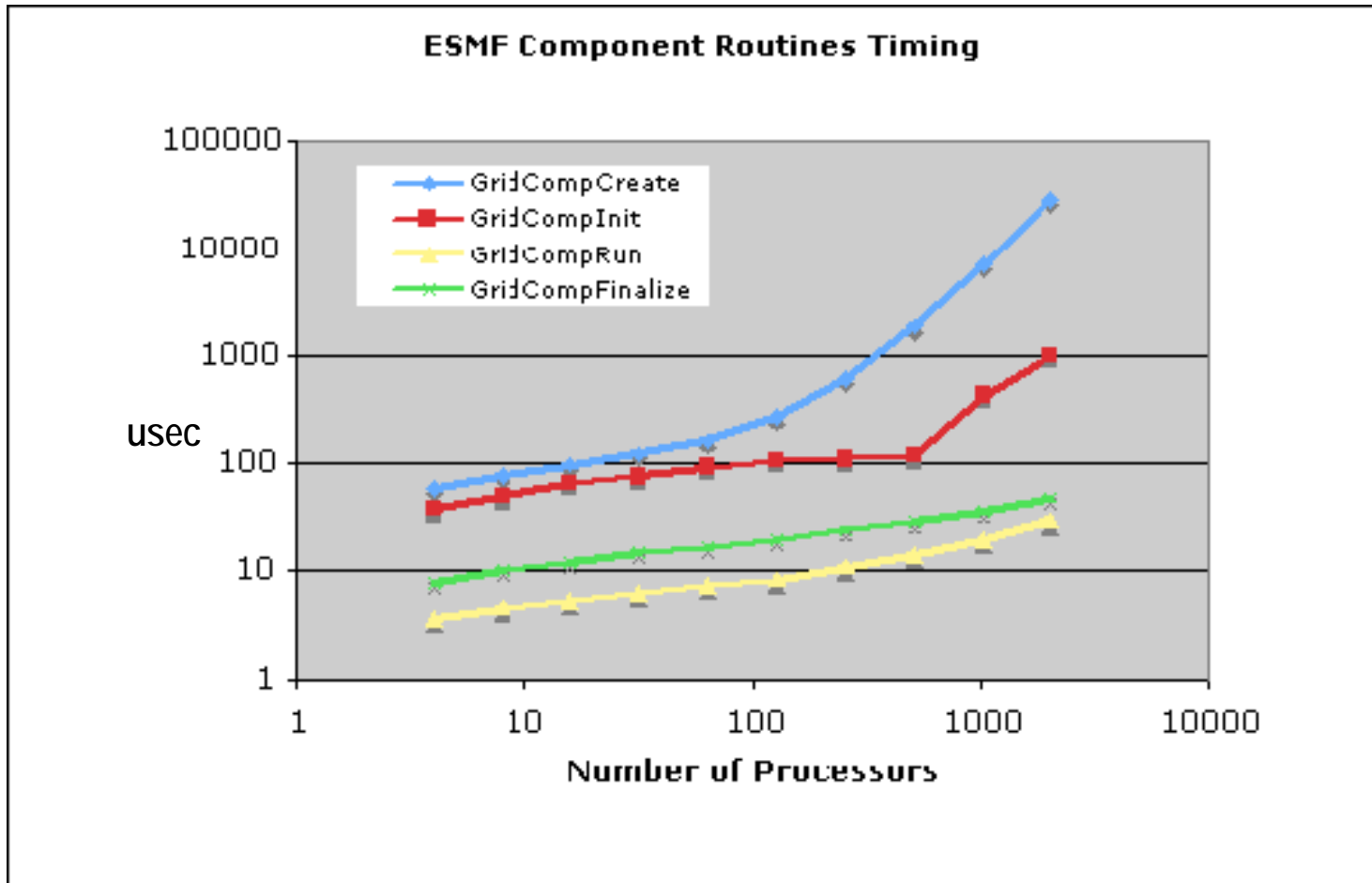
GOAL: Framework Optimization

Benchmark component overhead, regridding methods, redistribution, and middle- and low-level communications on a number of platforms. Include some evaluations at very high processor counts (1,000-5,000 processors).

Minimal performance burden for working code (Target <5% overhead in component overhead and regridding methods).



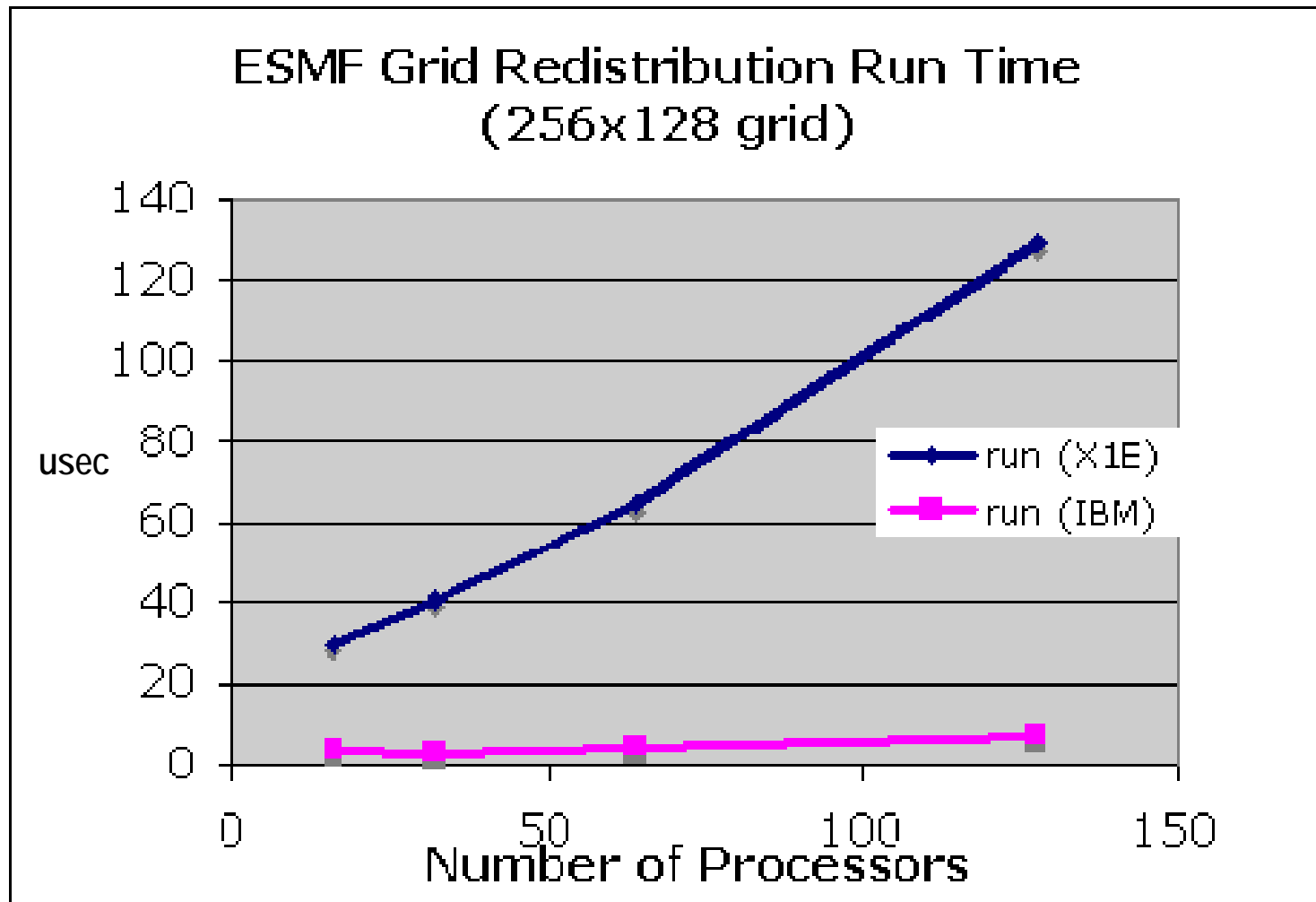
Metrics: Component Overhead



Measured by Peggy Li of NASA JPL on a Cray XT3 at Oak Ridge using ESMF 3.0.1.



Metrics: Redist Performance



Measured by Peggy Li of JPL on a Power5 IBM (bluevista) at NCAR and a Cray X1E at Oak Ridge National Laboratory using ESMF 2.2.2r.



GOAL: Support and Training

- Implementing an improved support and training program starting September 2006
 - New support lead position (Murphy)
 - Developers are each assigned a set of codes to support
 - All contacts are tracked
 - ESMF adoption level of codes is tracked
 - Established a customer relationship management database to store contact, code, and funding information



Other Activities

- CCSM Stage 1 Evaluation Plan
 - Considered a high priority by the ESMF Executive Board
 - Thoroughly reworked build system (Theurich)
 - Measured and optimized redistribution performance to equal existing CCSM capability (Li)
 - Overall this activity took about nine FTE-months
- New website out shortly, new support procedures, new metrics



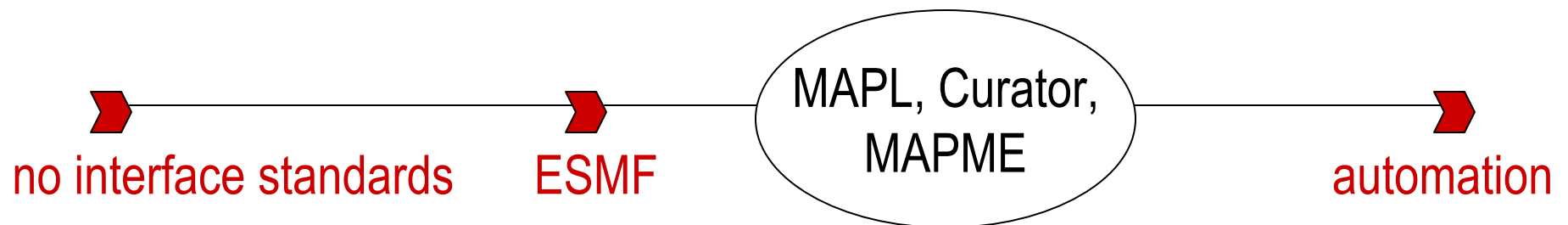
Metrics: Adoption

- Number of working ESMF components: 41
- Number of codes with ESMF in production: 6
 - NCEP GFS and Global ENSEMBLE (nested components)
 - NRL HYCOM-CICE (coupled components)
 - NASA GEOS-5 and variants (nested and coupled components)
 - MITgcm (nested and coupled components)
 - CSU (infrastructure)
 - LIS (infrastructure)
- More operational centers scheduled to transition codes this year (NAVO, Fleet Numeric, AFWA)
- Many codes in progress



Towards Integrated Systems

How do we increase model interoperability beyond what ESMF alone can do?
And how do we define and measure success?



NEED: Conventions and metrics for modeling component interoperability

- What level of interoperability is desired, and what is required to achieve it - time and sequencing rules, metadata, more built-in ESMF structure ... ? When do interoperability rules become too intrusive?
- Would ESMF compliance standards and “certification” be useful or threatening to modelers (or both?)
- How deep does ESMF need to go to be useful (high level wrappers? physics/dycores? parameterizations?)
- What metrics best measure interoperability?



Earth System Curator

- NSF-funded (leads NCAR ESMF/ESG, GFDL, MIT, GA Tech), just starting second year
- Developing a set of tools that allows climate modelers to:
 - archive and query models, experiments, model components, and model output
 - test the technical compatibility of model components
 - assemble and run multi-component models
- Prerequisite is metadata development – Curator works closely with European groups and ESG/PCMDI to develop international conventions for model architecture, coupling specifications, and grids



Curator Architecture

- The three pieces of the of the Curator architecture are:
 - CDP Curator, a catalog of models, components and data (based on existing NCAR/DOE *CDP/ESG BasePortal Technology*)
 - Satellite sites which house software for assembly and execution of modeling experiments (based on GFDL *FMS Runtime Environment*)
 - Provider sites that supply models and components to the CDP Curator via a web service interface
- Compatibility checking and automated assembly rely on *ESMF standard interfaces*
- For more information - <http://www.earthsystemcurator.org>



Curator Architecture

