

PROOF

xrootd

- **Fons Rademakers**
- **Maarten Ballantjin**
- **Marek Biskup**
- **Derek Feichtinger (ARDA)**
- **Gerri Ganis**
- **Guenter Kickinger**
- **Andreas Peters (ARDA)**
BaBar

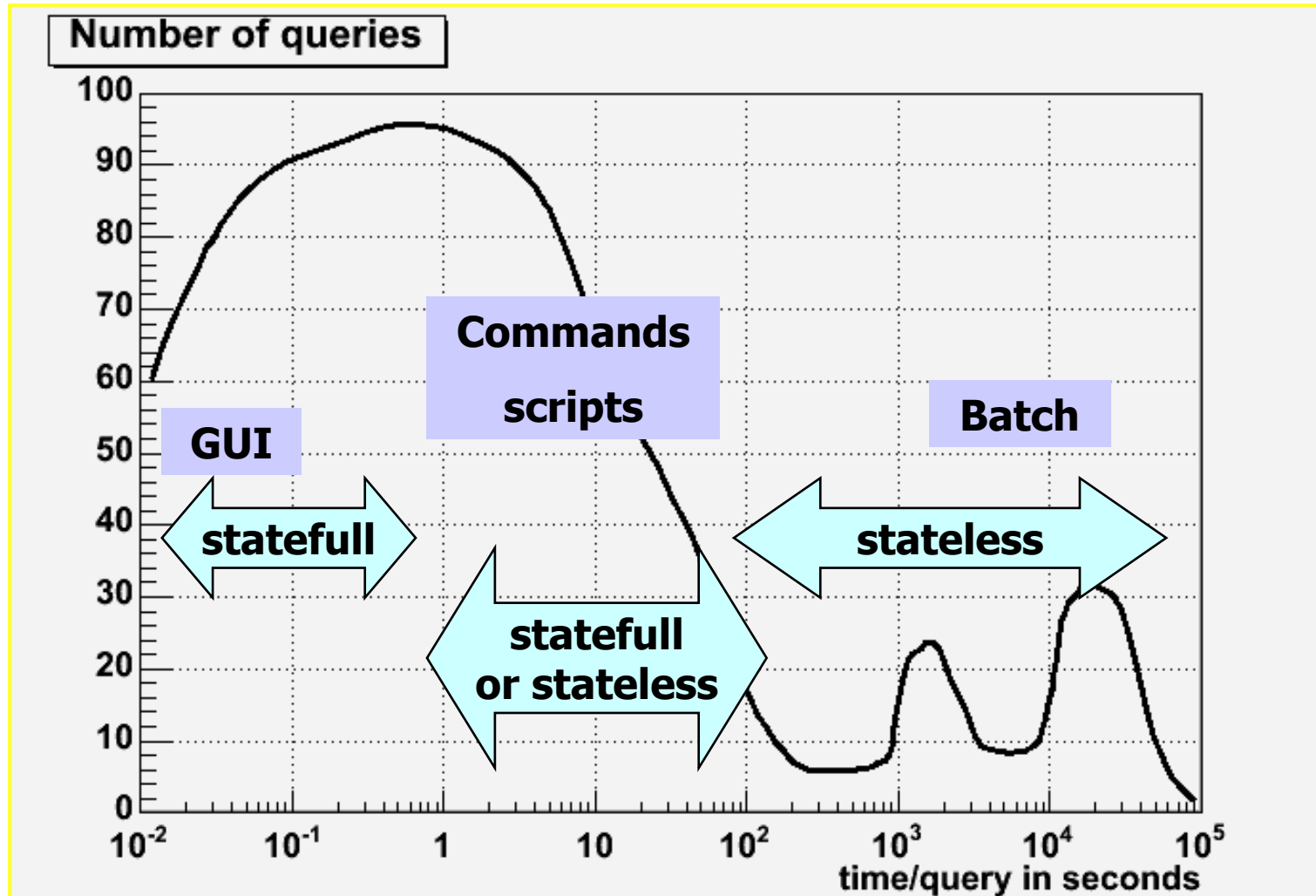
Original Goals

- Interactive parallel analysis on local cluster
- Transparency
 - same selectors, same chain Draw(), etc. on PROOF as in local session
- Scalability
 - quite good and well understood up to 1000 nodes (most extreme case)
 - extensive monitoring capabilities
 - MLM (Multi-Level-Master) improves scalability on wide area clusters
- Adaptability
 - partly achieved, system handles varying load on cluster nodes
 - MLM allows much better latencies on wide area clusters
 - not yet coming and going of worker nodes

New Additional Goals

- Support for “interactive” batch mode
 - allow submission of long running queries
 - allow client/master disconnect and reconnect
- Support “hostile” grid environments
 - startup of agents via Grid job scheduler
 - agents calling out to master (firewalls, NAT)
 - dynamic master-worker setup

Interactive/ Batch queries



Analysis session snapshot

what we want to do

AQ1: 1s query produces a local histogram

AQ2: a 10mn query submitted to PROOF1

AQ3->AQ7: short queries

AQ8: a 10h query submitted to PROOF2

**Monday at 10h15
ROOT session
On my laptop**

BQ1: browse results of AQ2

BQ2: browse temporary results of AQ8

**BQ3->BQ6: submit 4 10mn queries to
PROOF1**

**Monday at 16h25
ROOT session
On my laptop**

CQ1: Browse results of AQ8, BQ3->BQ6

**Wednesday at
8h40
Carrot session
on any web
browser**

Disconnect / Reconnect

- Authentication, sessions token
- State issues
- Transparency issues
 - TSelector::Begin(), Terminate()
 - objects created in Terminate(), output lists, etc.
 - storage of intermediate results

Performance Issues

- Read ahead interface in (x)rootd
- Using and understanding xrootd
- Cache index on master, optimizes repeat queries
 - assign different sessions with same data set to same worker nodes, reuse in memory files
- Monitoring
- Proofbench
- Query estimator

Authentication, Authorization

- New xrootd authentication plugins
- Certificates (login and user name)
 - single experiment wide login
 - user name used for sandbox
- Authorization to sandbox and shared global space
 - not to other user's sandboxes under same account

Robustness

- Get rid of OOB
- Split in two processes
 - protocol, authentication processor (proofd)
 - proofserv
- Communicate via priority based message queue
- Threaded or forked?
- xrootd code reuse?
- Periodic output list reporting to the master

Usability

- Selector with dynamic variables
- Selector proxy support
 - (TTree::MakeProxy – better name)
- Tree friends
- Event lists
- Compressed bit slice indices
- All possible draw options of the Draw() interface

Usability

- Browser interface to:
 - proof sessions
 - on demand histograms (single shot, timer driven feedback)
 - monitoring histograms
 - other input/output list objects
- Stop / resume capability

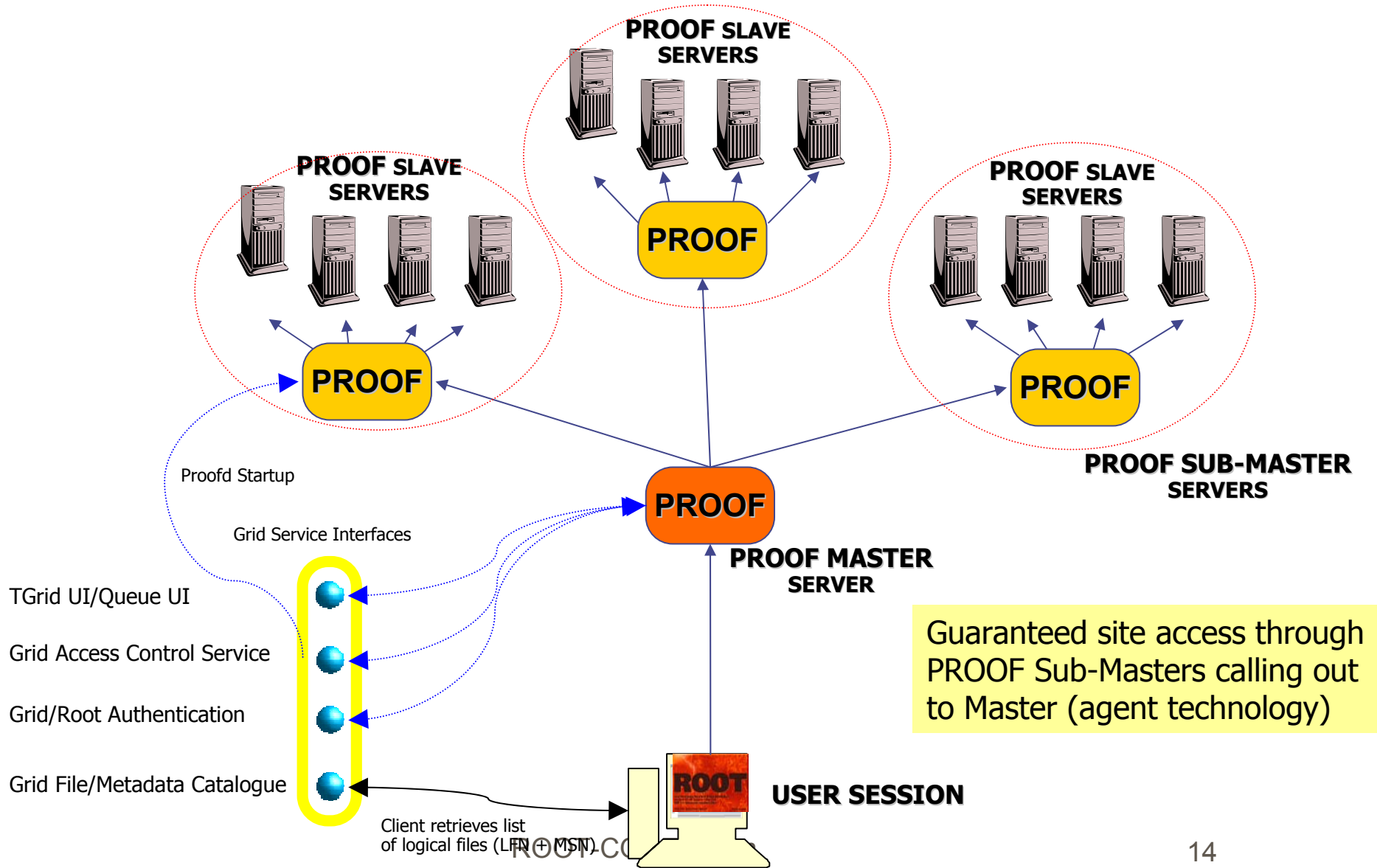
Usability

- Limiting processing to specific part of the data set
 - allow meta data in the TDSset (time, run, conditions, etc)
 - be independent from catalog
 - allows creation of new reduced data sets
- Remote display package, all canvases created in the master are send back to the client

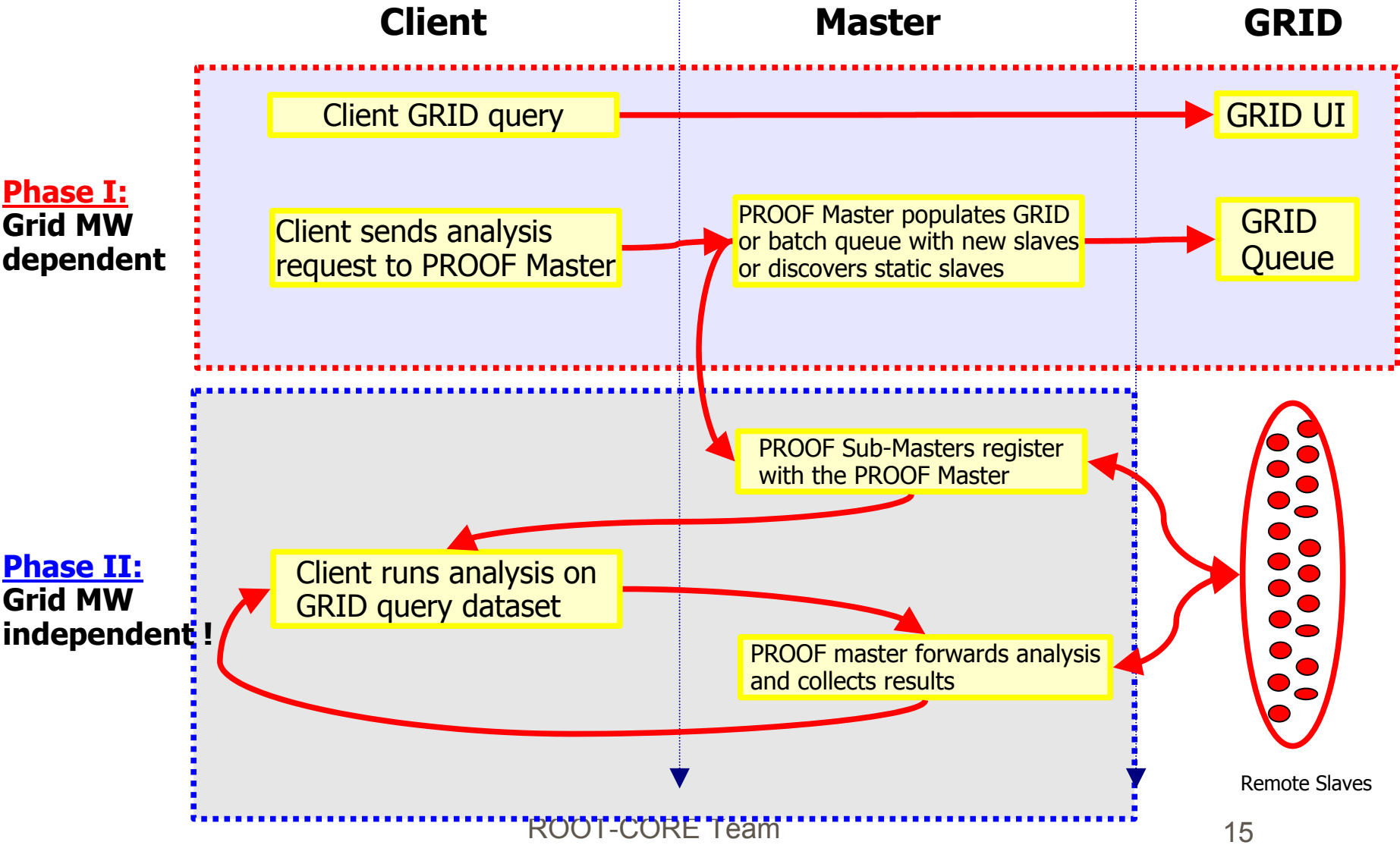
Grid Interfacing

- Grid catalog
 - data set creation
 - meta data, #events, time, run, etc.
- proofd agent creation
 - agents call out to (no-incoming connection)
- Config file generation / fully dynamic
- Coming and going of worker nodes
- Grid aware packetizer
- Scheduled execution
- Limiting processing to specific part of the data set

Interactive Analysis with PROOF on the Grid



PROOF Grid Session Diagram



Conclusion

- We have started an acceleration in the PROOF development towards achieving the new exciting goals that will hugely enhance the data analysis experience of very large data sets
- A first firm milestone is a demo at SC'05