IPP

Automated Plasma Reconstruction at ASDEX Upgrade

J.C. Fuchs, K. Behler, R. Fischer, R. Merkel, G. Neu, G. Raupp, W. Treutterer, and the ASDEX Upgrade Team

Max-Planck-Institut für Plasmaphysik, EURATOM Association, Boltzmannstr. 2, 85748 Garching, Germany

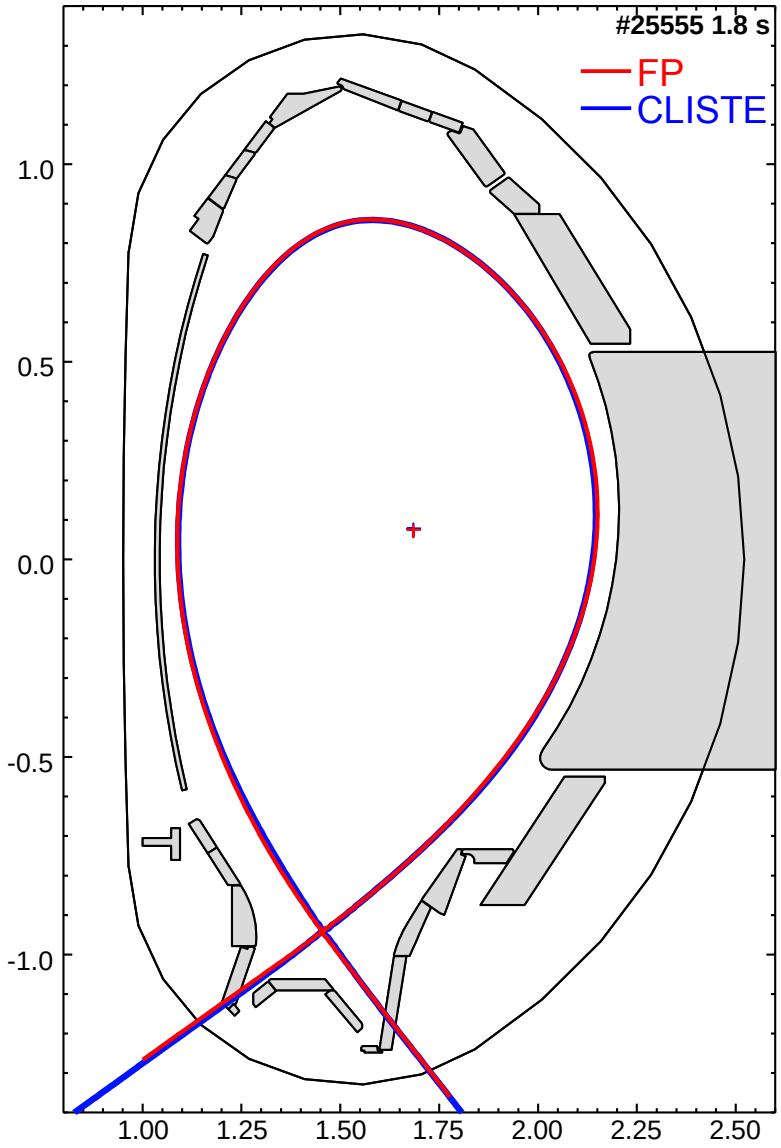
Integrated Modelling Technology workshop
Cadarache, 10 June 2011

Outline

- Algorithms of Plasma reconstruction at ASDEX Upgrade
- Why automated data analysis ?
- Flow chart of automatic data analysis:
  - Requirements for automated programs
  - Synchronization of different evaluations
  - Automated data validation
  - Dealing with invalid or suspect results
- Keeping evaluated data up to data: Data Dependence Base
- Summary

Algorithms of Plasma Reconstruction

- Equilibria reconstruction (poloidal flux matrix, plasma position, kinetic energy etc)
  - Function parametrization (FP):
    - linear or quadratic regression of several measurements (mainly magnetic loops)
    - necessary coefficients are determined using a database of several thousand ideal equilibria.
    - very fast, used for :
      - Real time reconstruction and control of plasma parameters and position
      - first offline check of the plasma equilibrium



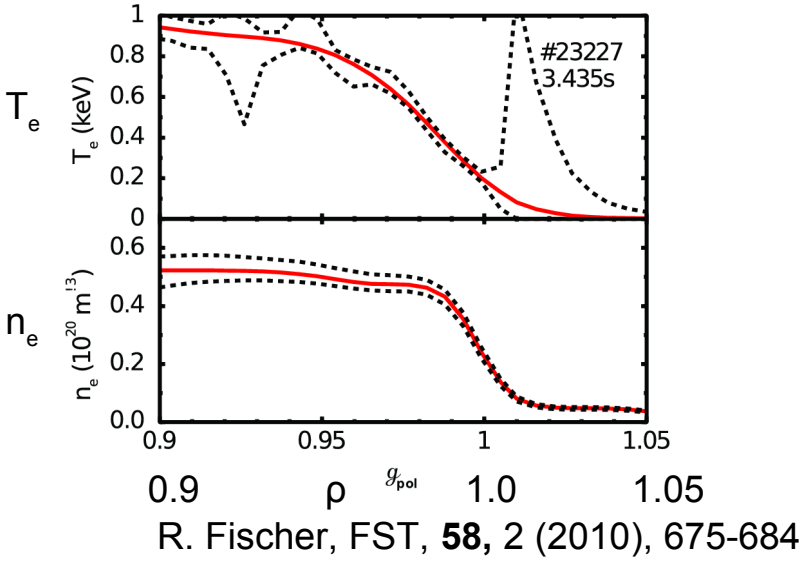
Algorithms of Plasma Reconstruction

- Equilibria reconstruction (cont'd):
  - CLISTE interpretative code:
    - numerically solves the Grad-Shafranov equation as a best fit to a set of experimental measurements:
      - magnetic probes, flux loops, edge currents (standard)
      - MSE, SXR, kinetic profiles, … (detailed analysis)
    - Various spatial and time resolution:
      - Coarse grid, 100 ms (standard, 1-2 minutes)
      - Fine grid, 1ms (now also standard, 1-2 hours)
      - Very fine grid and very high time resolution possible (depending on measurements, detailed analysis)
  - Results from FP and CLISTE agree very well
  - Ongoing projects:
    - Real-time grad-shafranov PDE solver using high-level graphical programming and COTS technology
    - Grad-Shafranov solver using Integrated Data Analysis

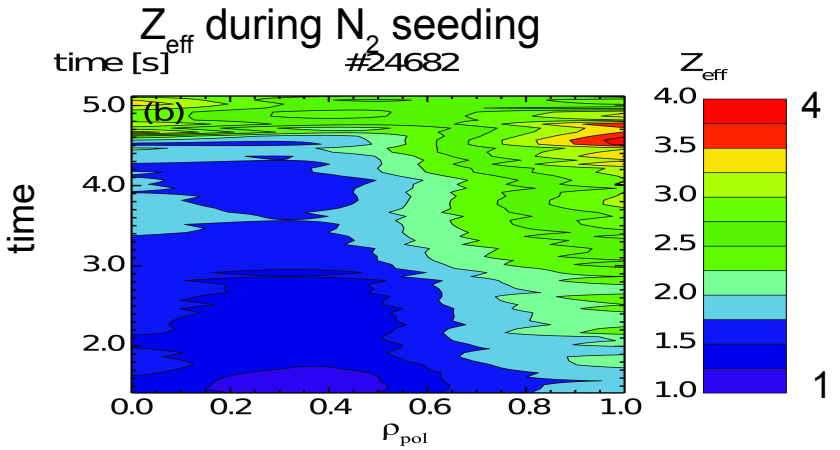
Algorithms of Plasma Reconstruction

- Profile reconstruction (e.g. n_e, T_e, Z_{eff}, …) :
  - integrated data analysis (IDA) within the framework of Bayesian probability theory
  - Combines measurements of several independent diagnostics
  - Gives profiles with confidence bands

- Reconstruction of various plasma quantities



R. Fischer, FST, **58,** 2 (2010), 675-684



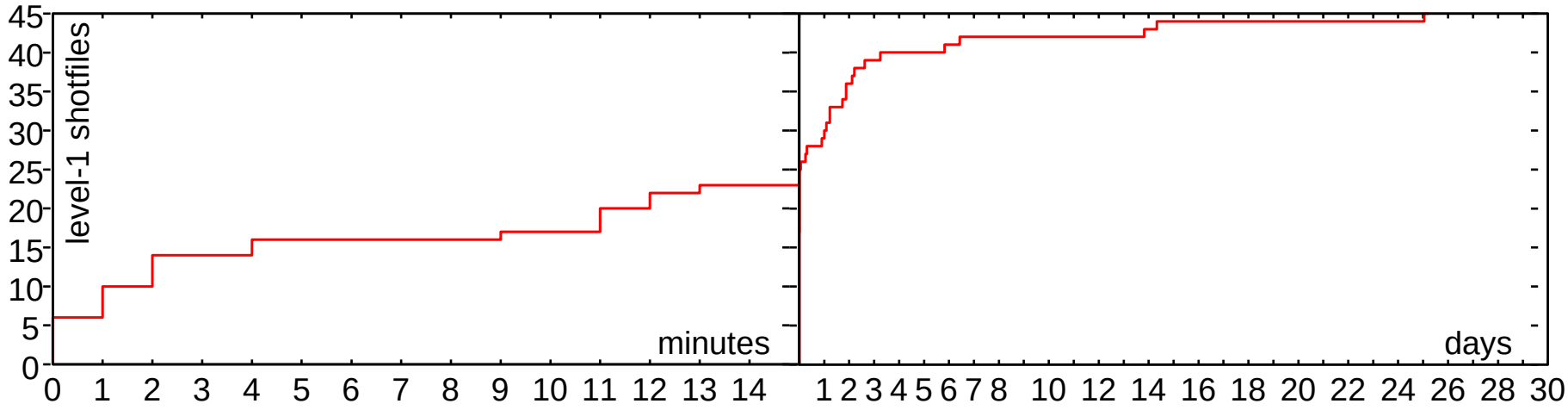
S.K. Rathgeber, PPCF **52**, 9 (2010), 095008

Automated data analysis: pros and cons

ASDEX Upgrade: environment of maximum scientific flexibility
fixed standard procedures should not prevent unforeseen tasks

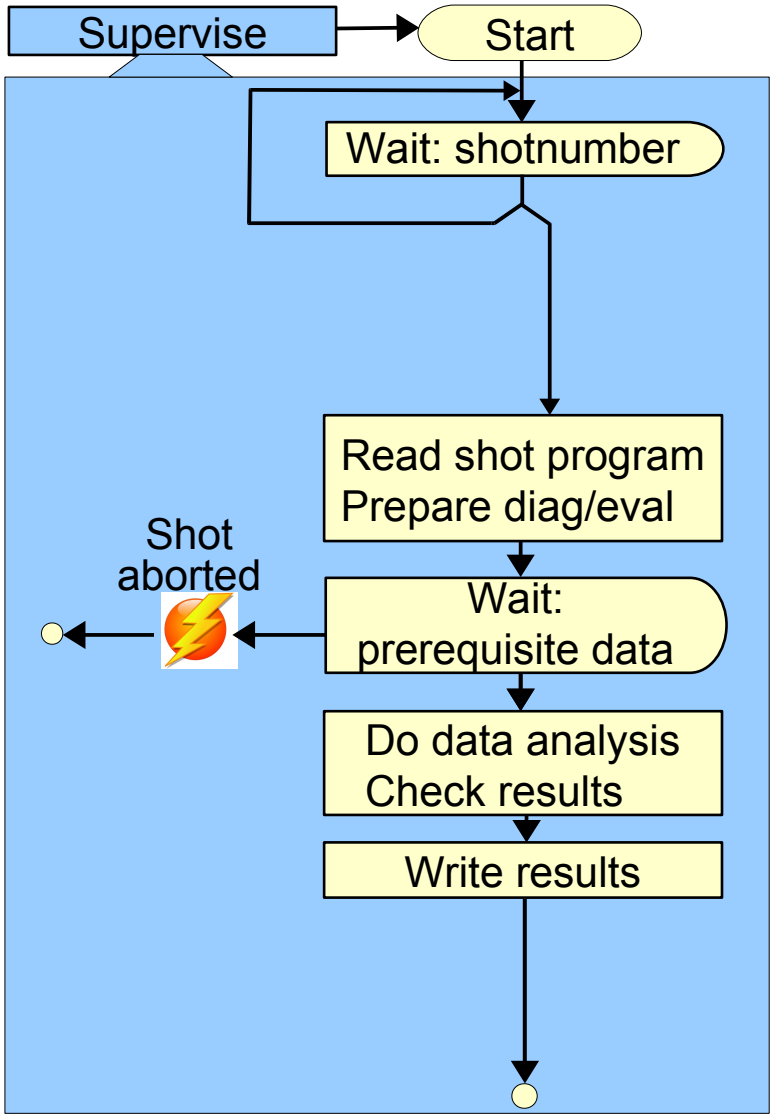
- Problems of automated tasks:
  - not as flexible as those with lots of human input
  - Overhead for
    - Implementation of automated tasks
    - Administration and organization
    - Supervision
  - Quality of automatic generated data is more difficult to judge
- Advantages:
  - **People are freed from time-consuming routine tasks**
  - **More time for creative scientific tasks**
  - Faster availability of results
  - Analysis is independent of the availability of specific people
- Requirement:
  - **Saved time > Overhead**

Standard Data Analysis Chain: Status

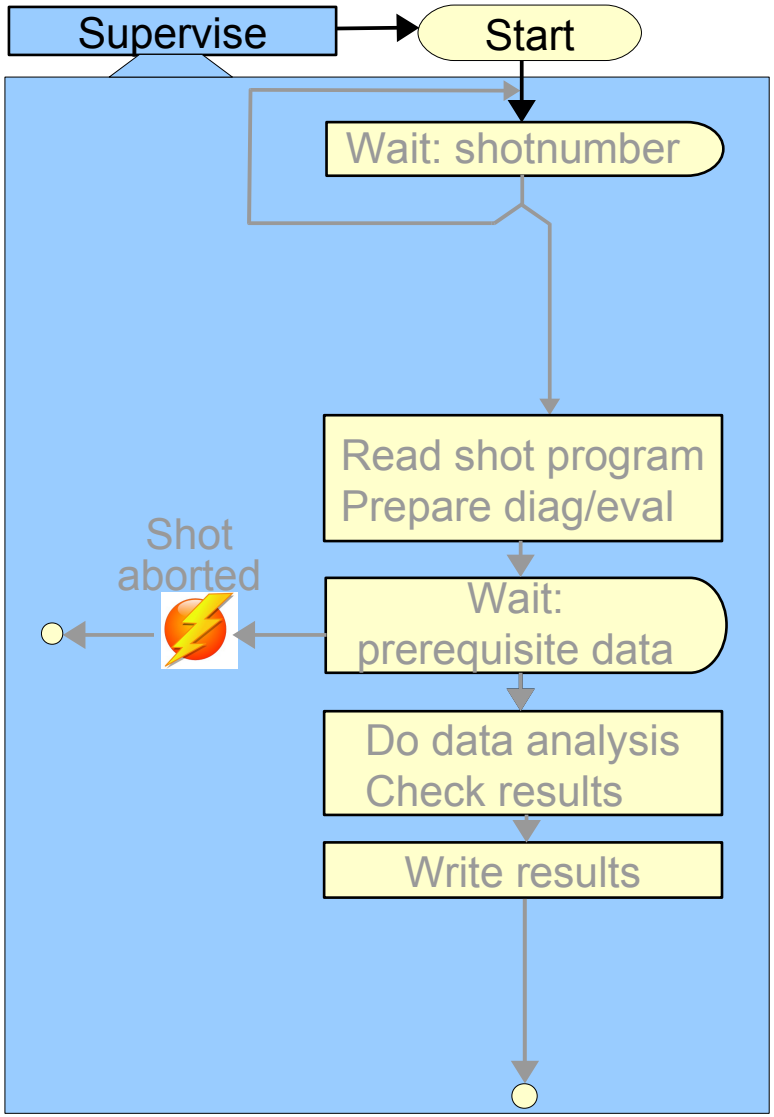


- Statistics over evaluated data:
  level-*n* shotfiles [≙ PPF at JET] #25001 - #25890
- 47 standard “diagnostics” [≙ DDAs at JET]
  (available for 75% of useful plasma shots)
- 15 diagnostics available on average 5 minutes after the discharge
  24 diagnostics available on average 15 minutes after the discharge
- Standard data analysis done:
  - IT group, using programs from diagnosticians (few)
  - Diagnosticians, using tools from IT group (many)

Flow chart of automatic data evaluation



Flow chart of automatic data evaluation



- Supervision of evaluation loop:
  - Main task: restart if necessary
  - Separate process which must run very reliably and “survive” even reboots:
    - Cron job which runs every couple of minutes:
      - Easy to set up
      - Very flexible
      - No sysadmin required
      - May be too late
    - System services: Oracle Solaris “System Management Facility”:
      - Extremely fast restart
      - Higher privileges required
  - Monitor of progress

Monitor of data anlysis progress

ASDEX Upgrade

Monitorshotfiles

27031	BOL	BLV	BLK	BLB	BPD	FPC	FPG	FPP	GPI	EQI	GQI	EQH	GQH	UVD	MAG	MAI	MBI	MAX	MPC	MPG	MPP	FPR	OK
27032	BOL	BLV	BLK	BLB	BPD	FPC	FPG	FPP	GPI	EQI	GQI	EQH	GQH	UVD	MAG	MAI	MBI	MAX	MPC	MPG	MPP	FPR	OK
27033	BOL	BLV	BLK	BLB	BPD	FPC	FPG	FPP	GPI	EQI	GQI	EQH	GQH	UVD	MAG	MAI	MBI	MAX	MPC	MPG	MPP	FPR	OK
27036	BOL	BLV	BLK	BLB	BPD	FPC	FPG	FPP	GPI	EQI	GQI	EQH	GQH	UVD	MAG	MAI	MBI	MAX	MPC	MPG	MPP	FPR	OK
27041	BOL	BLV	BLK	BLB	BPD	FPC	FPG	FPP	GPI	EQI	GQI	EQH	GQH	UVD	MAG	MAI	MBI	MAX	MPC	MPG	MPP	FPR	OK
27042	BOL	BLV	BLK	BLB	BPD	FPC	FPG	FPP	GPI	EQI	GQI	EQH	GQH	UVD	MAG	MAI	MBI	MAX	MPC	MPG	MPP	FPR	OK
27043	BOL	BLV	BLK	BLB	BPD	FPC	FPG	FPP	GPI	EQI	GQI	EQH	GQH	UVD	MAG	MAI	MBI	MAX	MPC	MPG	MPP	FPR	OK
27044	BOL	BLV	BLK	BLB	BPD	FPC	FPG	FPP	GPI	EQI	GQI	EQH	GQH	UVD	MAG	MAI	MBI	MAX	MPC	MPG	MPP	FPR	OK
27045	BOL	BLV	BLK	BLB	BPD	FPC	FPG	FPP	GPI	EQI	GQI	EQH	GQH	UVD	MAG	MAI	MBI	MAX	MPC	MPG	MPP	FPR	OK
27046	BOL	BLV	BLK	BLB	BPD	FPC	FPG	FPP	GPI	EQI	GQI	EQH	GQH	UVD	MAG	MAI	MBI	MAX	MPC	MPG	MPP	FPR	OK
27047	BOL	BLV	BLK	BLB	BPD	FPC	FPG	FPP	GPI	EQI	GQI	EQH	GQH	UVD	MAG	MAI	MBI	MAX	MPC	MPG	MPP	FPR	OK
27048	BOL	BLV	BLK	BLB	BPD	FPC	FPG	FPP	GPI	EQI	GQI	EQH	GQH	UVD	MAG	MAI	MBI	MAX	MPC	MPG	MPP	FPR	OK

Exit

Status of EQH/GQH calculation - Mozilla Firefox

File Edit View History Bookmarks Tools Help

Status of EQH/GQH calcu...

Status of EQH/GQH calculation

Tue May 31 13:38:49 2011

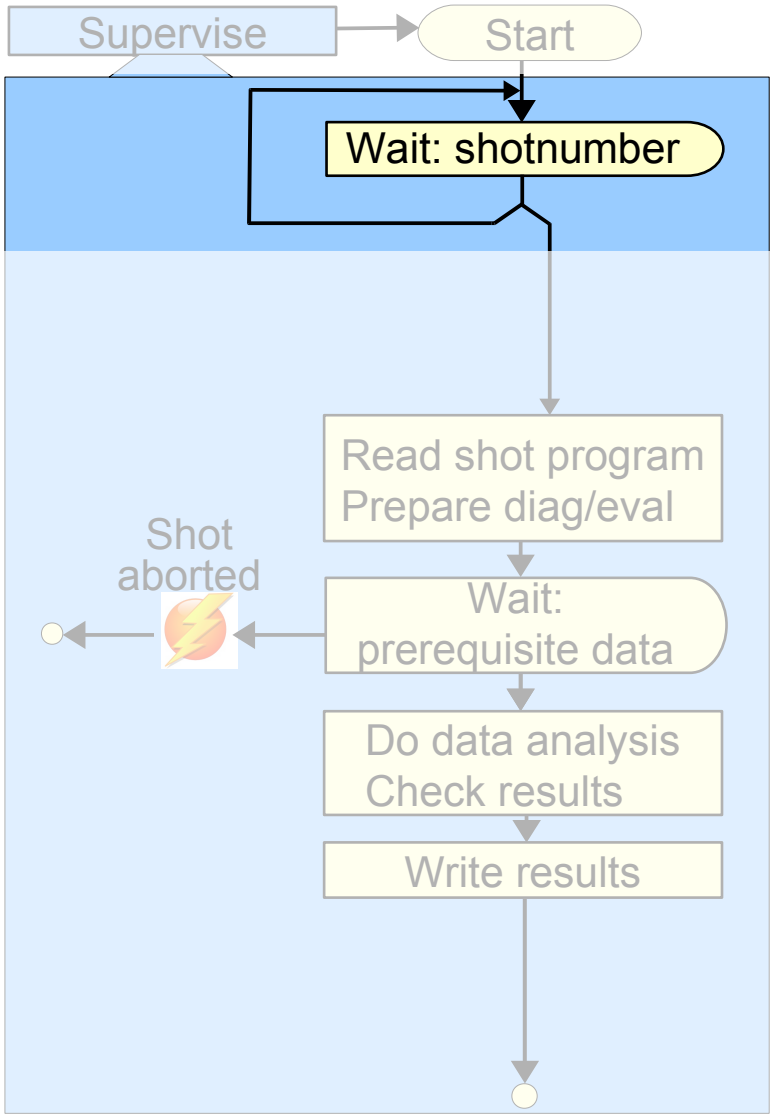
Shot	From	To	Nt	Status	Host	Done	T	Est	Speed	Shotfile	Date	Size
27050	0.0150	6.7460	6731	failed	sxaug20							
27052	0.0150	5.0510	5036	finished	sxaug20					EQH(1) (closed) GQH(1) (closed)	Tue May 31 12:03:42 2011 Tue May 31 12:08:38 2011	281600976 2064440
27053	0.0370	4.0000	3963	finished	sxaug20					EQH(1) (closed) GQH(1) (closed)	Tue May 31 12:36:42 2011 Tue May 31 12:39:42 2011	221630984 1257144
27055	0.0370	9.8070	9770	Running	sxaug20	4880	0:39	0:40	7507.7			
27056	0.0370	8.5420	8505	Waiting (1)								

(New requests should appear here after about 1-2 minutes)

New request

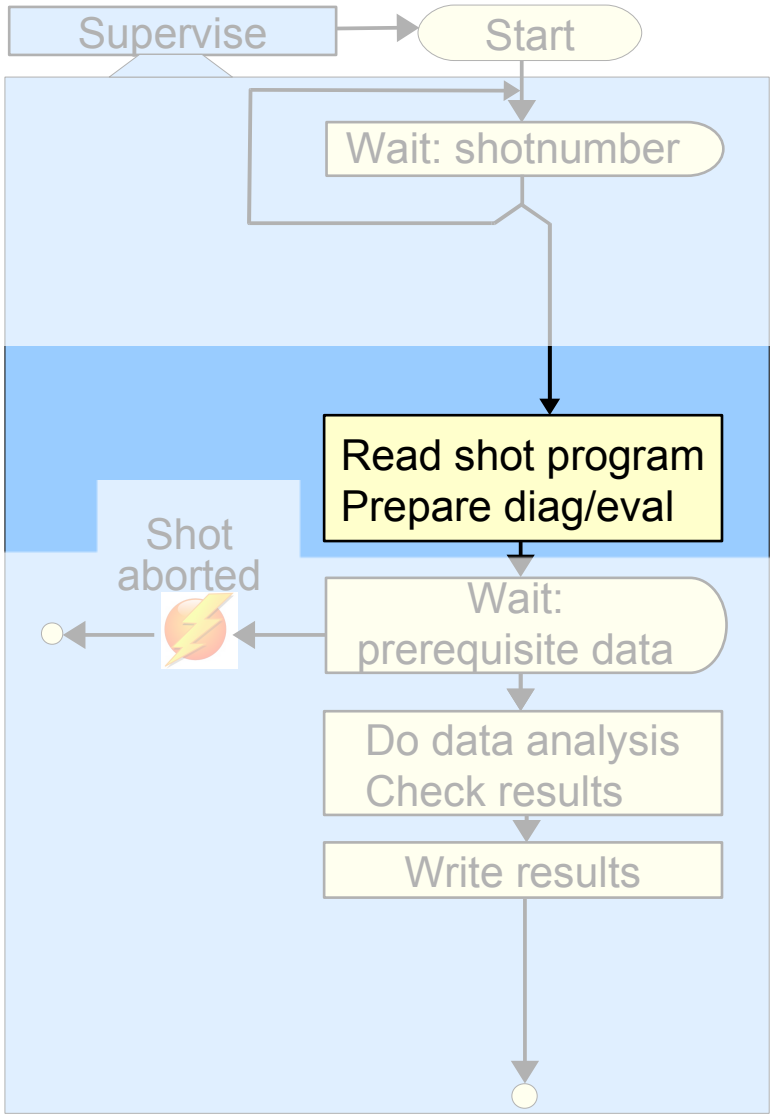
Done

Flow chart of automatic data evaluation



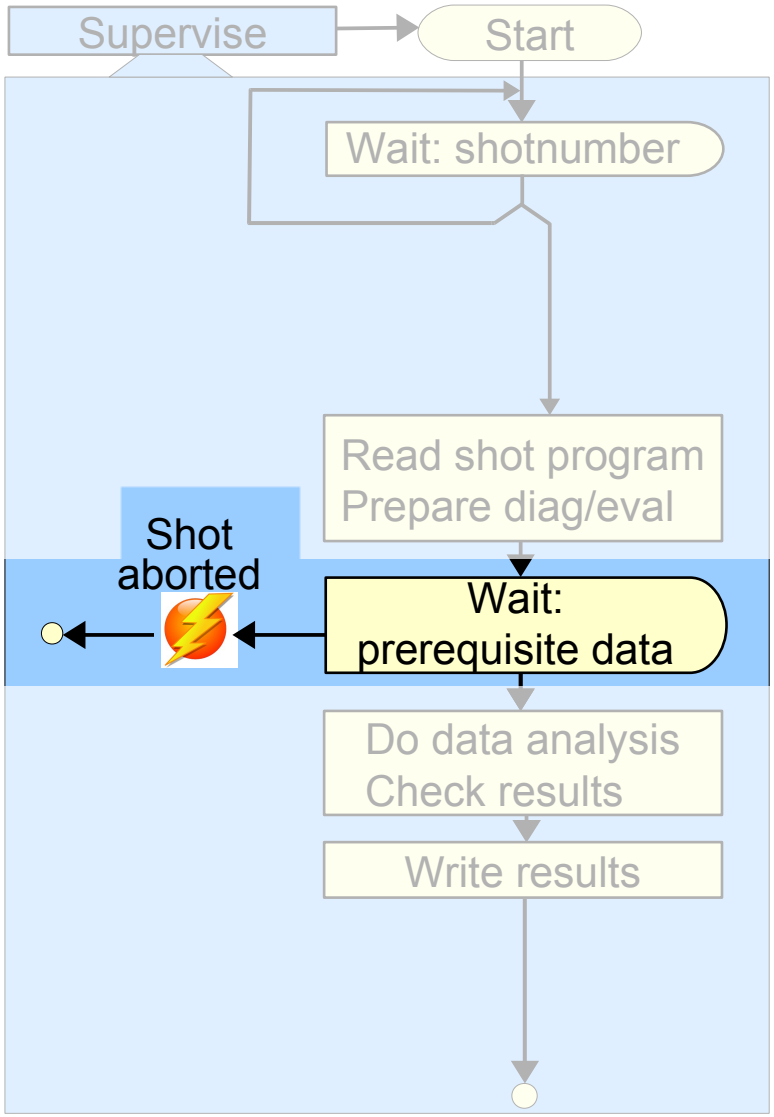
- Start of discharge cycle: "Shotnumber distributed"
- Several minutes (1.5 - >5) before plasma ignition
- User point of view:
  - function call which blocks until discharge cycle is started,
  - returns shotnumber
  - Provided in shared object library, callable from all major apps (Fortran, C, IDL, Matlab, Perl, ...)
- Implementation:
  - "shotnumber server" (RPC callback)
- For long running calculations:
  - Fork off separate process
  - Write request to batch queue

Flow chart of automatic data evaluation



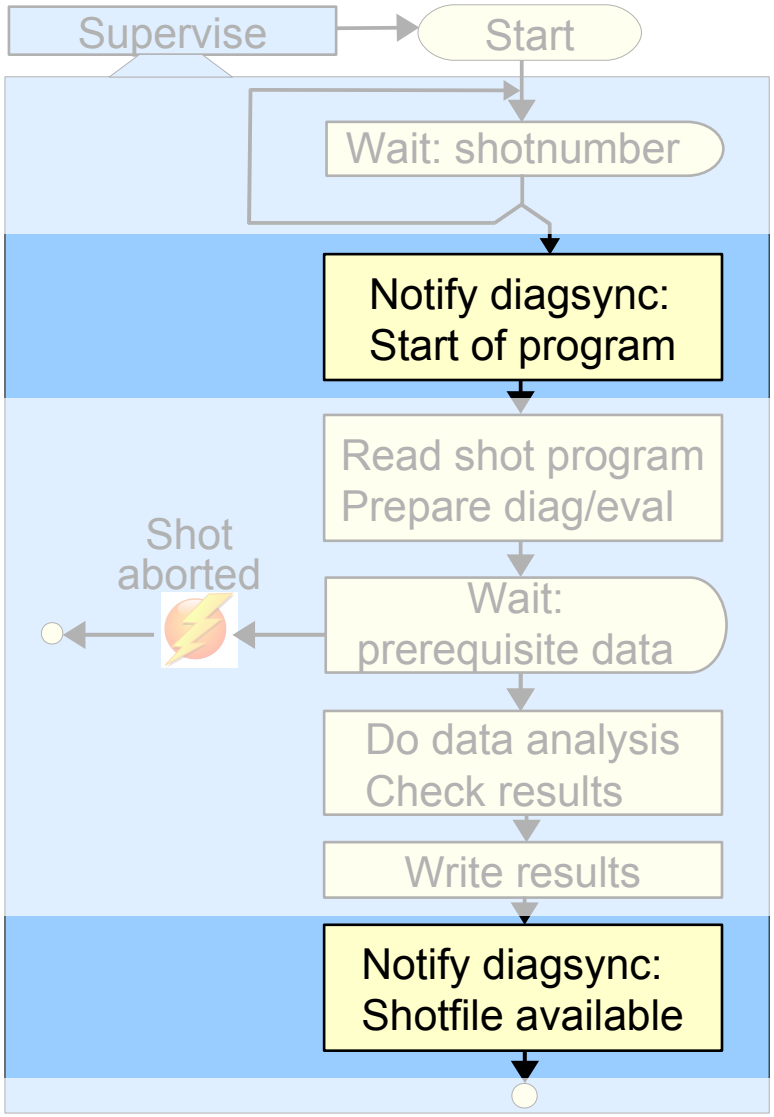
- All (planned) shot parameters are fixed
- Can be used to configure diagnostics or parameters for data evaluation

Flow chart of automatic data evaluation



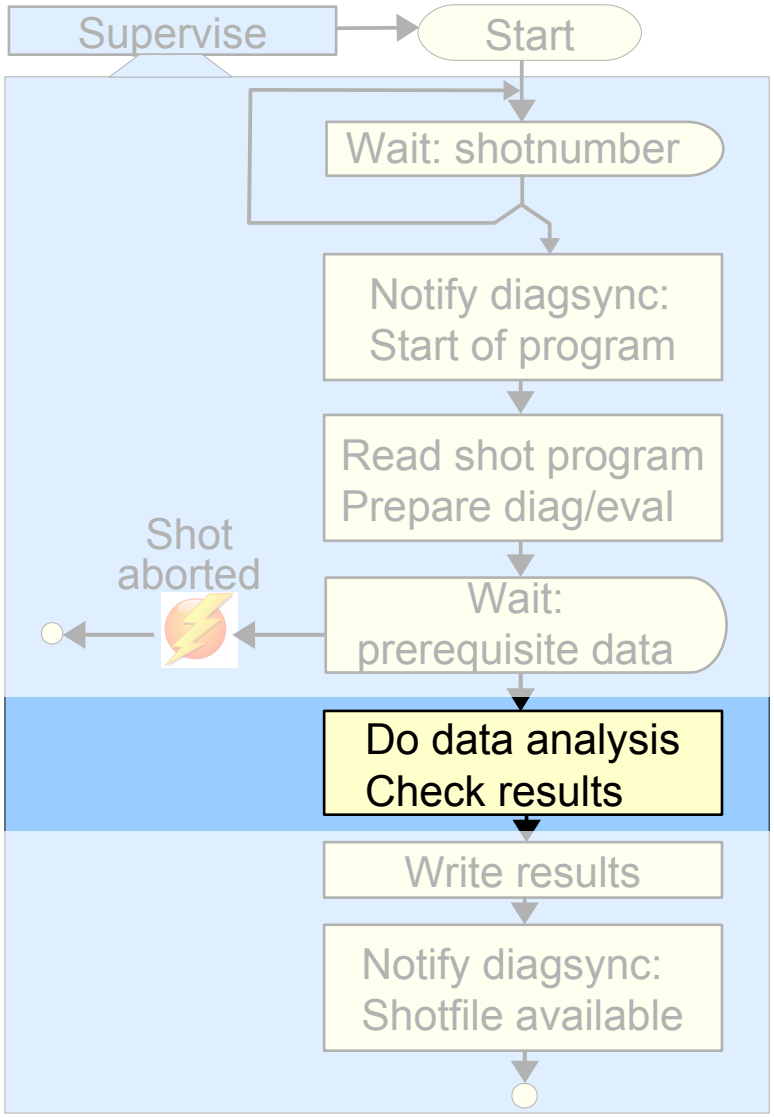
- Wait until all prerequisite data are available:
  - Old method: Wait until shot file is available in file system
  - Current method:
    - Attempt to read shotfile which does not yet exist blocks the program
    - Request for notification is send to a “diagnostic synchronisation server” (diagsync)
    - Diagsync sends notification as soon as it knows that shotfile exists
      - RPC callback
      - Shortcut:
        On central computer:
        local sockets
- Same method (shotfile structure) for both raw data and evaluated data

Flow chart of automatic data evaluation



- If other programs need our results:
  Notify diagsync of
  - Start of program
  - End of program
    Shotfile available

Flow chart of automatic data evaluation



- Do data analysis:
  - No human interaction allowed
- Check results:
  - Strongly depends on particular algorithm, e.g. compare measured with reconstructed quantities

Requirements for automated programs

No human interaction allowed

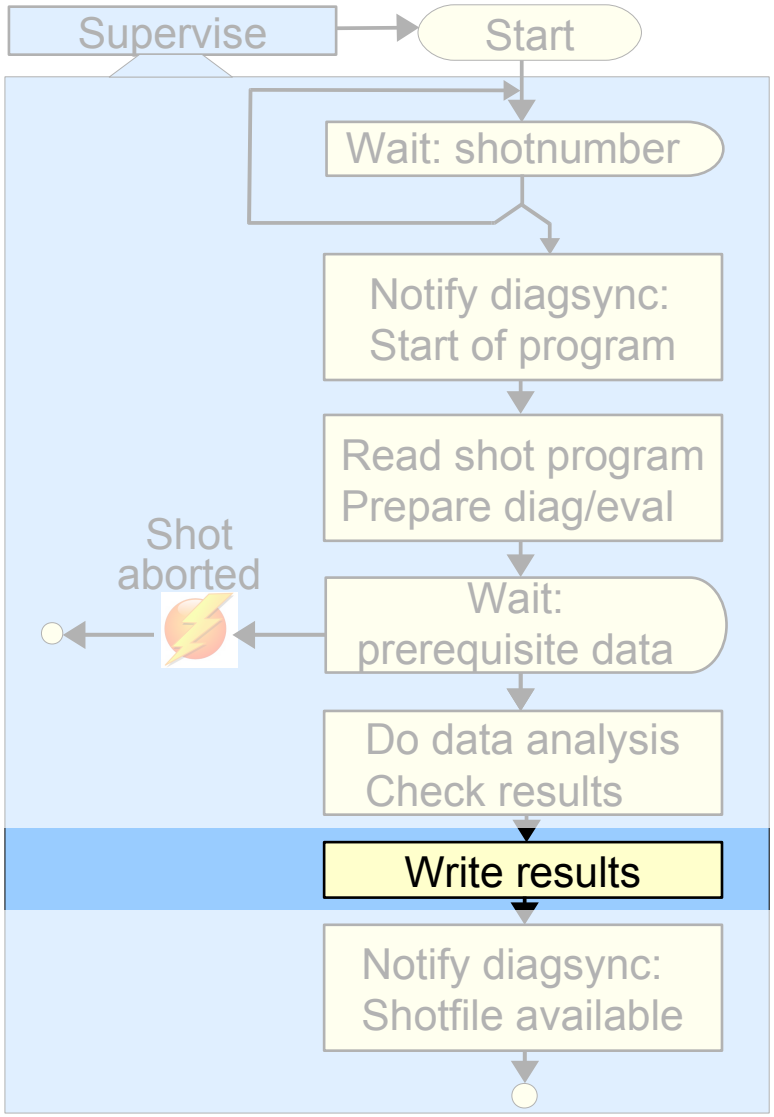
- Only input parameter: shotnumber
- All other parameters must be set by
  - *Either:* default values which are known to work in most cases
  - *Or*: “intelligent algorithms” based on available plasma parameters and measurements
- Illegal measurements must be detected automatically (how ?)
- Quality of results must be checked automatically (how ?)

- All actions should be logged for debugging purposes
- All dependencies must be documented for reproducibility

Automated data validation

- Check results:
  - Strongly depends on particular algorithm,
    e.g. compare measured with reconstructed quantities
  - Equlibrium reconstruction by FP:
    - One suspect probe:
      Replace it by using “replacement coefficients” from other probes
    - Several suspect probes: return error
  - Equilibrium reconstruction by cliste:
    - Iteratively switch off suspect probes (up to 3)
    - If one probe is switched off for many time points:
      Notify session leader (probe seems to be broken)
    - Return error if no convergence in a given number of iterations is reached (convergence criteria are relaxed dynamically)
  - Integrated data analysis:
    - Return error if residua exceed given limit

Flow chart of automatic data evaluation



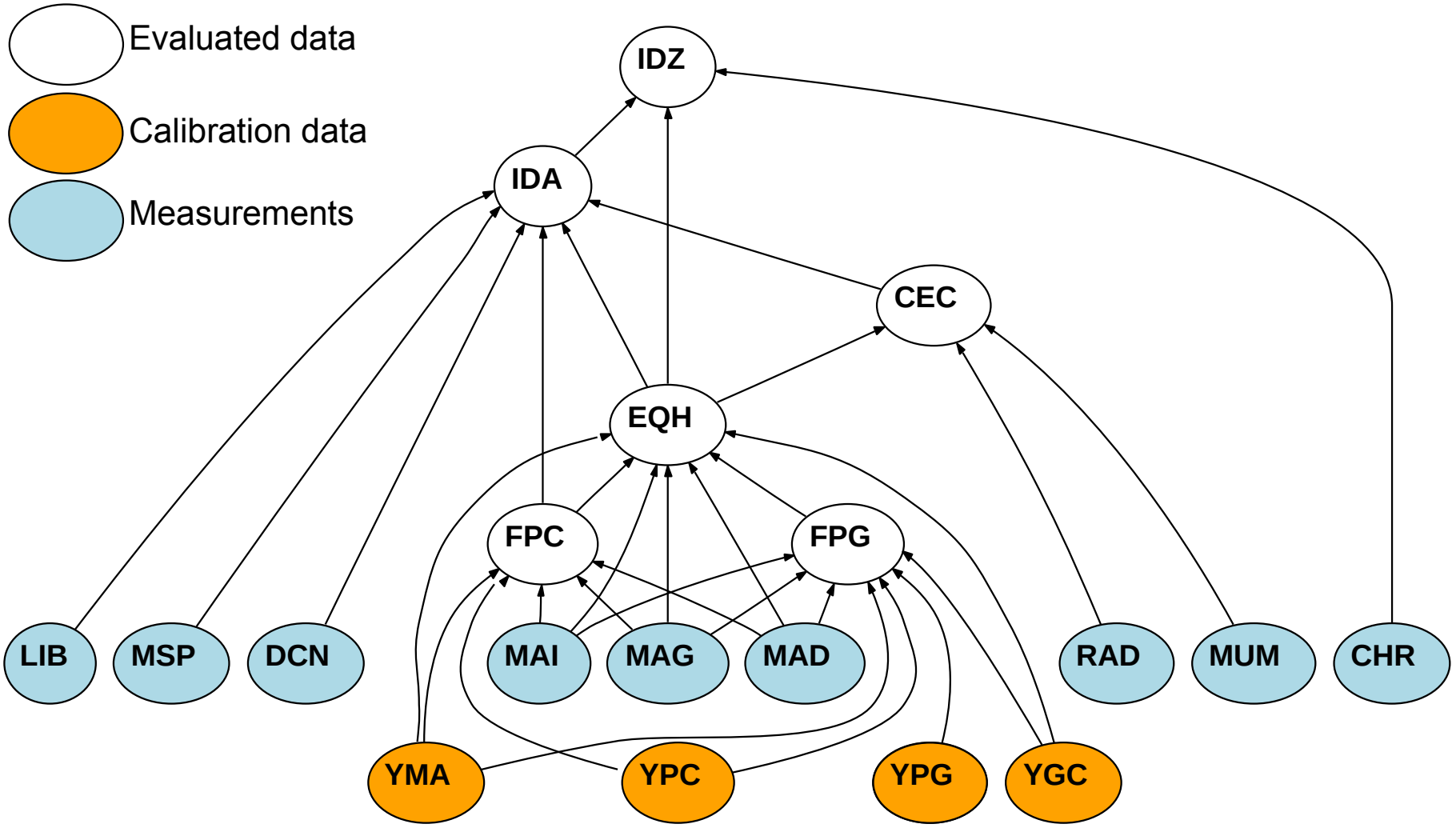
- Write results to shot file:
  - How to deal with invalid or suspect results ?
  - Document all dependencies

Invalid or suspect results

- Do not write invalid values at all
  - ⊕ Nobody can use these values by accident
  - ⊖ Not possible if several signals share a common base vector (e.g. time vector)
- Denote invalid quantities with a special value:
  IEEE (quiet) NaN
  - ⊕ Any try to use this value in further calculations results in NaN again
  - ⊕ NaN are treated as missing values by many applications (e.g. plotting with IDL)
  - ⊖ Value is lost forever, even if it might hold some limited information
- Link an object of type “Qualifier” to the signal in the shotfile, which describes the (physical) quality of the signal
  - ⊕ Can mark any parts of a signal as invalid, without erasing the signal
  - ⊖ Must be checked before using the signal
  - ○ Furthermore: Global qualifier for the whole shotfile
- Proposal: Use continuous qualifier (0.0→1.0) and error bars

Data analysis chain: dependencies

Example: Dependency chain for IDZ (Z_{eff} from integrated data analysis)

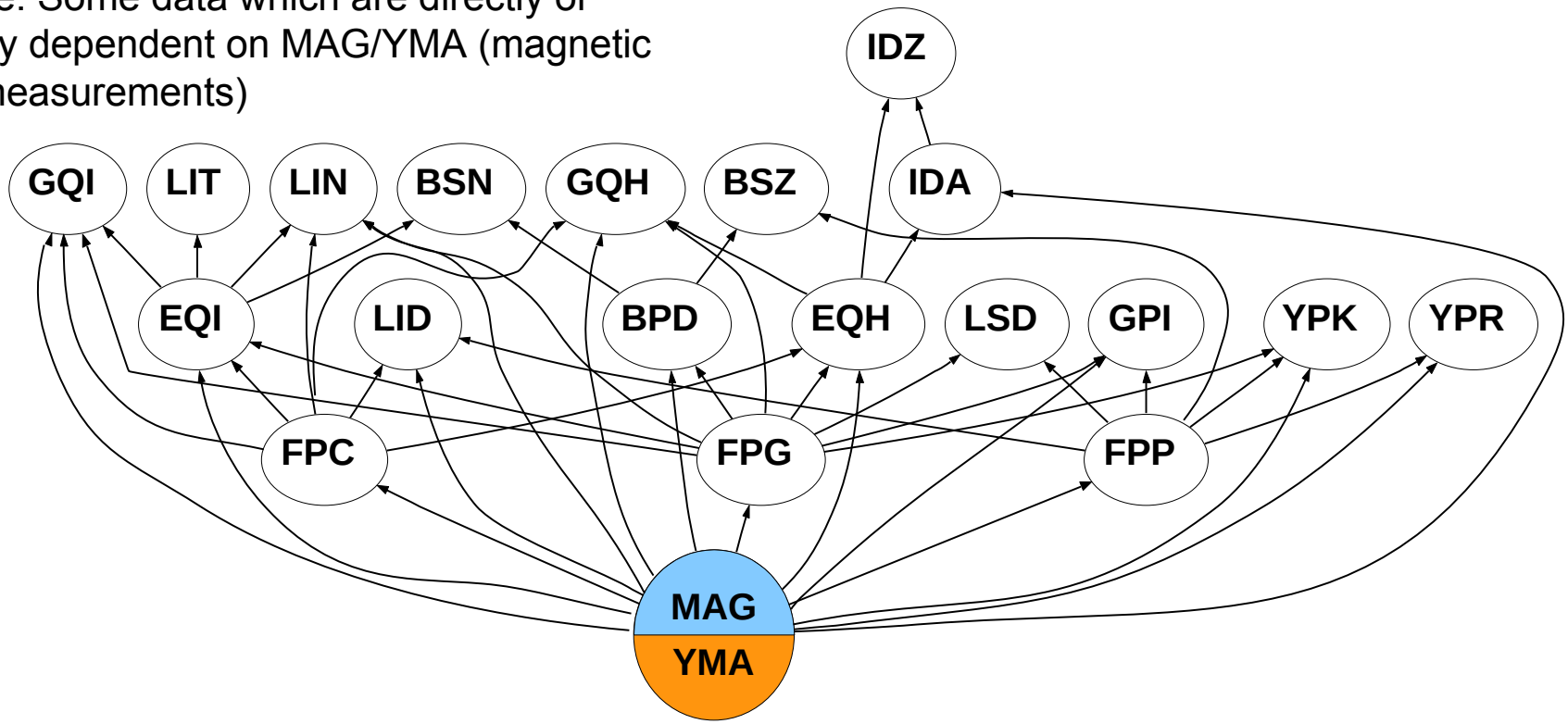


Keeping evaluated data up to date

Aim:
Re-calculate any level-*n* shotfile when it becomes inconsistent, i.e. when the data it depends on have been modified

Prerequisite:
Document dependencies of shotfiles and analysis programs

Example: Some data which are directly or indirectly dependent on MAG/YMA (magnetic probe measurements)



Dependence Data Base (DDB)

Status:

- About 20 active diagnostics in DDB
- Shotfile is marked as “invalid” when a predecessor becomes invalid, but without any further consequences
- E-mail when shotfile becomes invalid
  Human expertise: Is it necessary to re-calculate shotfile ?
- Dependencies can also easily be logged in the shotfile itself (even if diagnostic is not recorded in DDB)

Future work:
Redesign of DDB (Oracle Database)

“Ideal” implementation:

- Include *all* level-n diagnostics in DDB
- Warn the user if he tries to use data from an invalid shotfile
- Automatically re-calculate any invalid shotfiles if possible
  Really ? (Huge chain to re-calculate; computer load ?)

Summary

- Despite the need of flexibility automated data analysis is routinely used at ASDEX Upgrade
- User friendly function calls:
  - Start of discharge cycle (shotnumber server)
  - Shotfile available (diagsync server)
- Invalid or suspect results may be flagged in shotfiles
  - NaN
  - Physical Qualifier (discrete)
    Continuous qualifier ?
- Data Depence Base helps to keep evaluated data up to date
  - Warn when trying to read out-of-date data
  - Automatically re-calculate shotfiles ?