



Schweizerische Eidgenossenschaft
Confédération suisse
Confederazione Svizzera
Confederaziun svizra

Eidgenössisches Departement des Innern EDI
Bundesamt für Meteorologie und Klimatologie MeteoSchweiz

What Can a Small Country Do? The MeteoSwiss Implementation of the COSMO Suite on the Cray XT4

Petra Baumann

13th ECMWF Workshop on the Use of HPC in Meteorology
3.-7. November 2008





MeteoSwiss Model Setup

ECMWF IFS (global)

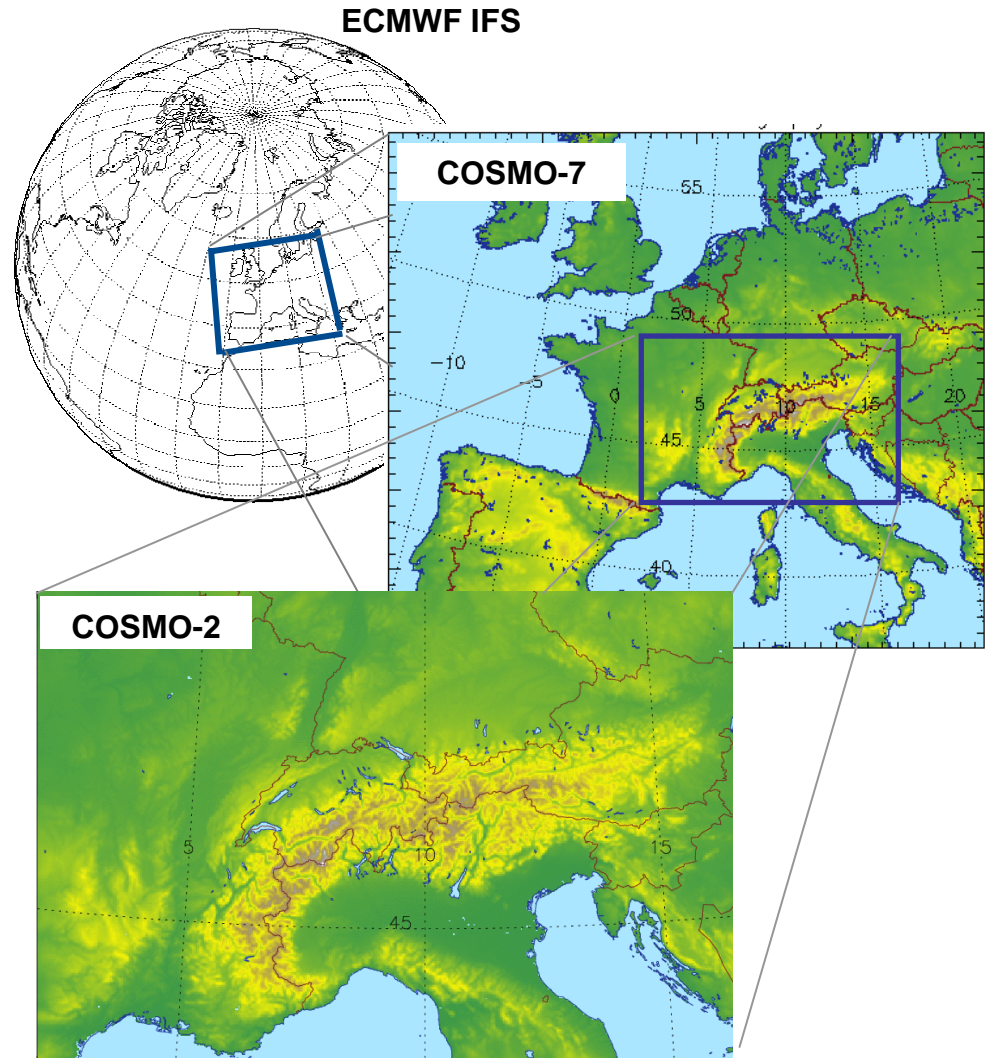
- 25km, 91 layers
- 2 x 240h per day
+ 2 x 78h per day

COSMO-7 (regional)

- 6.6km, 60 layers,
393 x 338 grid points
- 2 x 72h per day

COSMO-2 (local)

- 2.2km, 60 layers,
520 x 350 grid points
- 8 x 24h per day





Computing Facilities

Swiss National Supercomputing Centre (CSCS)
in Manno, Ticino





HPC Platforms



Production system: Cray XT4 'buin'

- 16 service nodes (AMD Opteron dual core, 2.6 GHz)
- 448 compute nodes (AMD Opteron dual core, 2.6 GHz)
- Special purpose purchase of CSCS to guarantee the high level of availability needed by MeteoSwiss
- Reserved usage by MeteoSwiss during operational time slots



Failover system: Cray XT3 'palu'

- 32 service nodes (AMD Opteron dual core, 2.6 GHz)
- 1664 compute nodes (AMD Opteron dual core, 2.6 GHz)
- Shared machine, also used for development

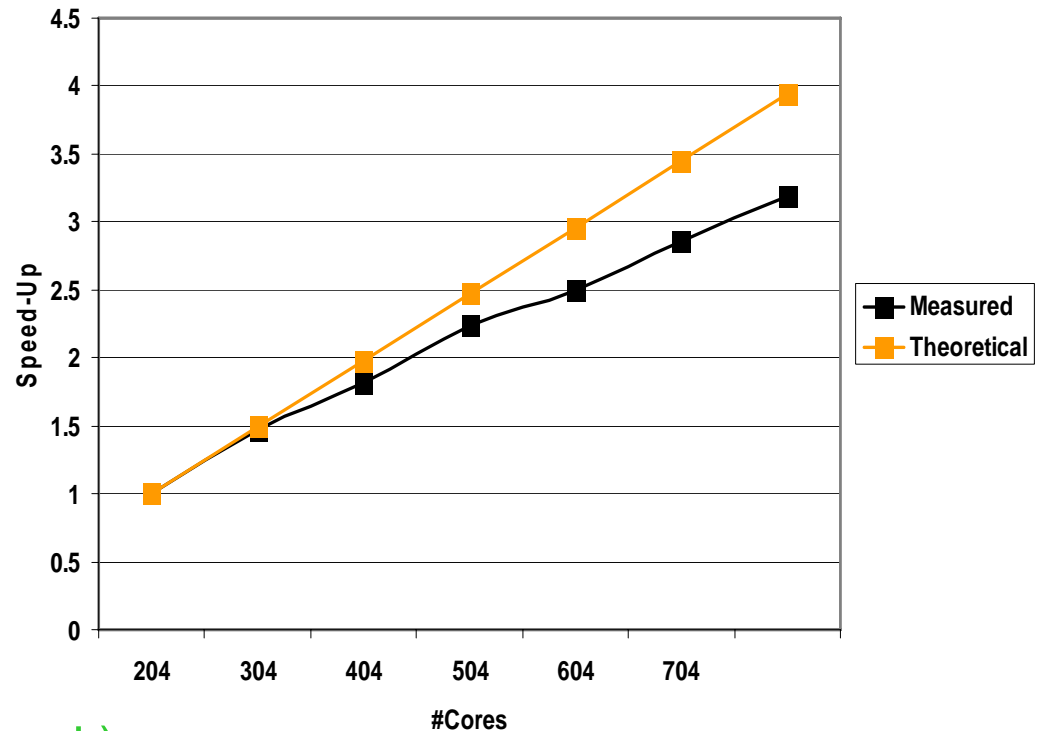
- UNICOS/lc operating system (Linux on service nodes, Catamount on compute nodes)
- Lustre parallel file system



COSMO Scalability on the CRAY XT4

- Code: FORTRAN90, MPI-only parallelization
- Compiler: Portland Group
- 24h COSMO-2 forecast

NX	NY	NIO	#Cores	t in s
10	20	4	204	3760
15	20	4	304	2566
20	20	4	404	2067
20	25	4	504	1682
24	25	4	604	1503
25	28	4	704	1318
25	32	4	804	1177



Operational setting:
390 Gflops sustained (9 % of peak)

(Data: J.-G. Piccinali, CSCS)



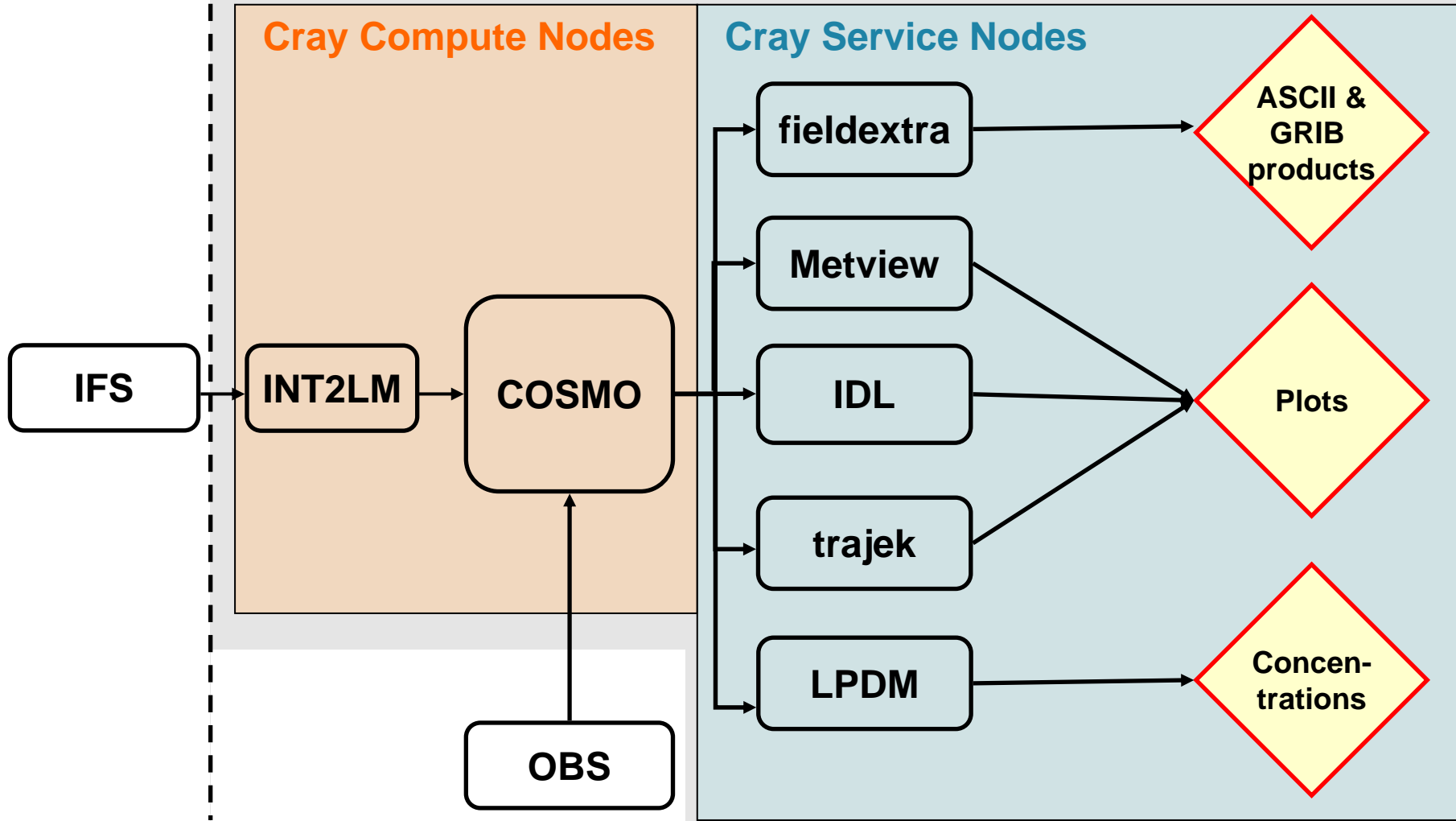
Data Flow

ECMWF

CSCS

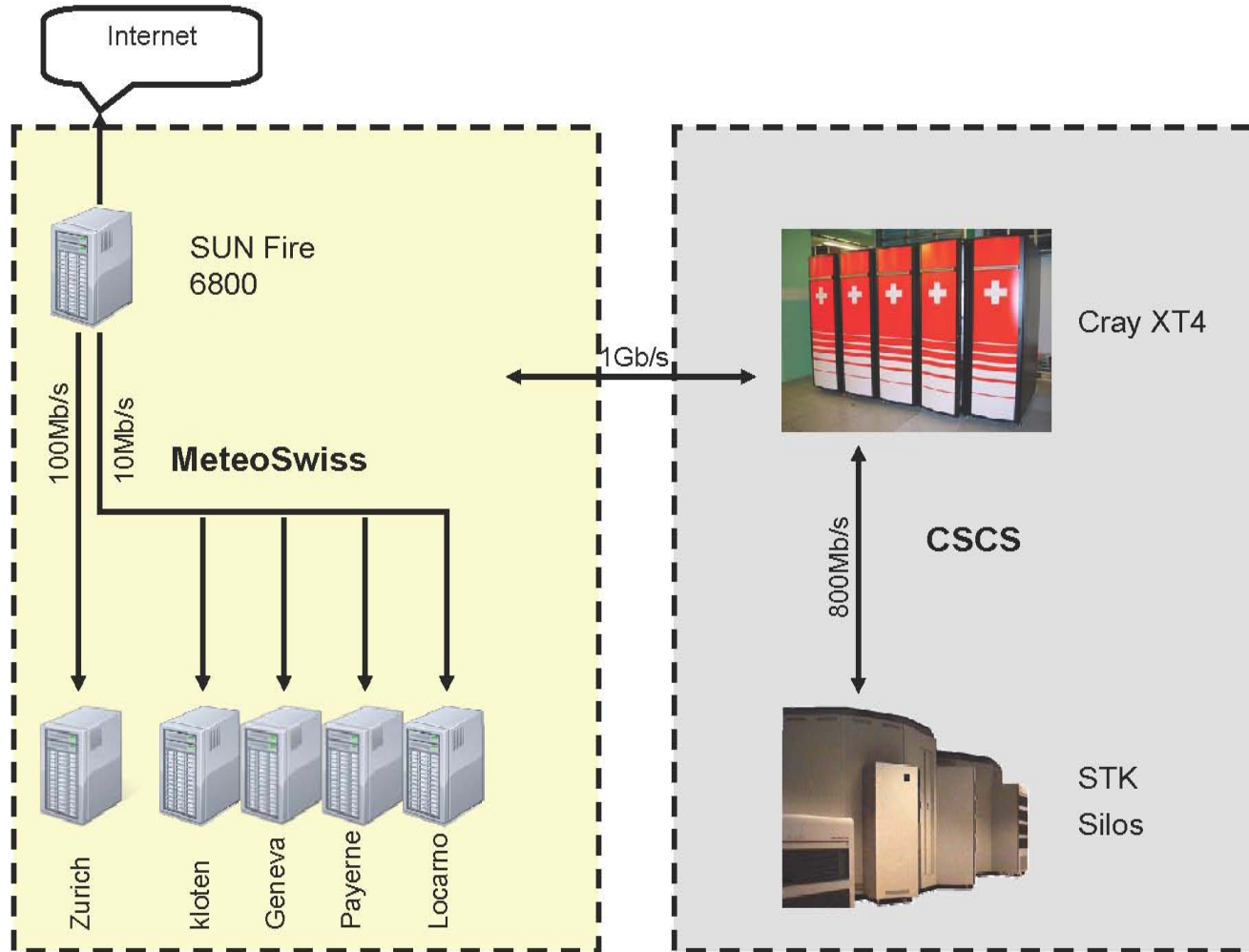
Cray Compute Nodes

Cray Service Nodes





Hardware and Communications





Queuing

- **PBS Pro** workload management system
- **YOD** parallel execution library under Catamount

Cray Compute Nodes

advance reservation

- Permitted exclusively for MeteoSwiss operators
- Used for operational computations

high-priority queue

- Permitted solely for MeteoSwiss operators and further special MeteoSwiss users
- Higher priority than normal queue
- Used to run model test chains
- Backup in case of advance reservation failure

normal queue

- Permitted for all users

Cray Service Nodes

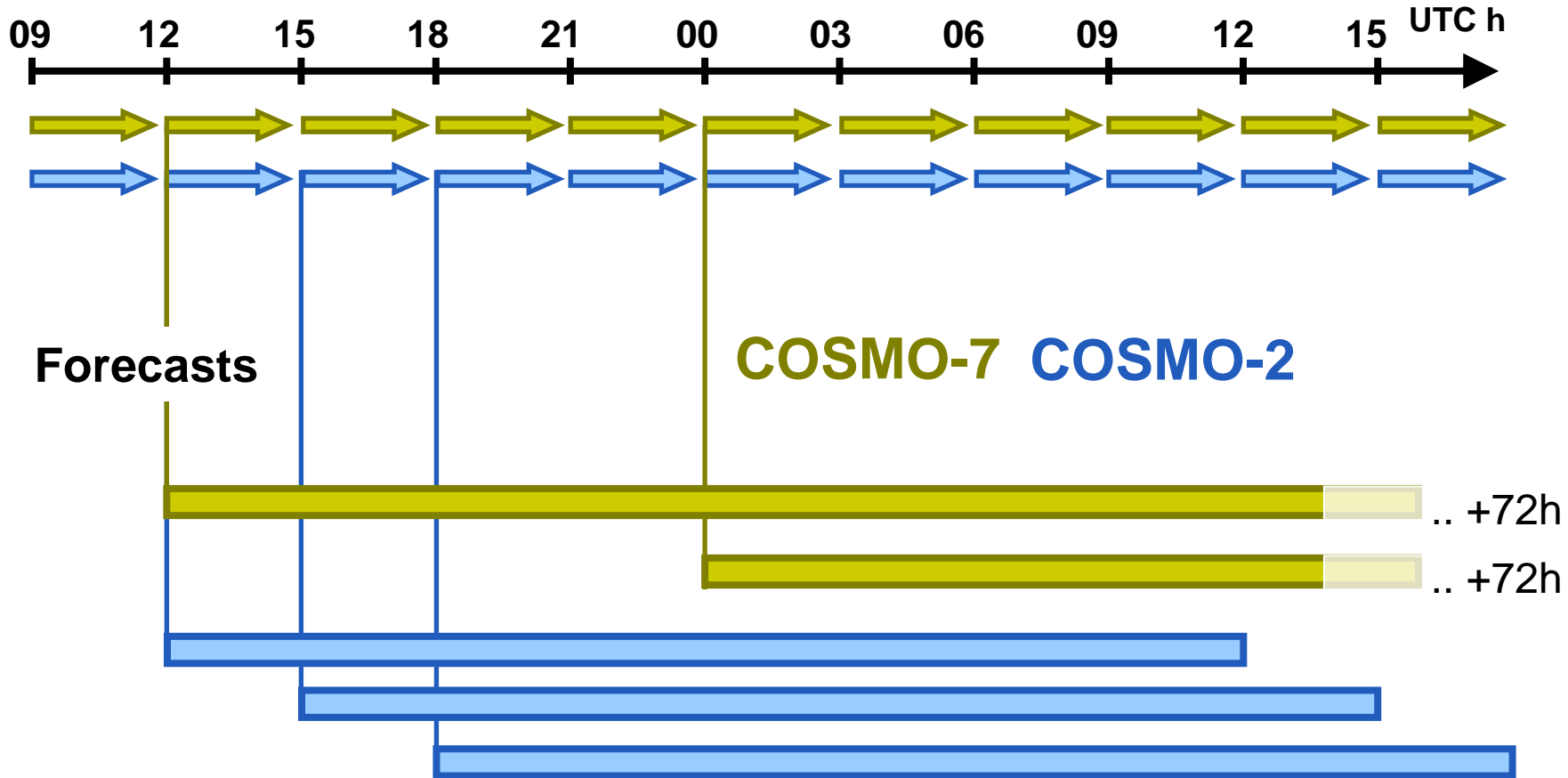
post-processing queue

- Permitted solely for MeteoSwiss operators
- Used for operational post-processing



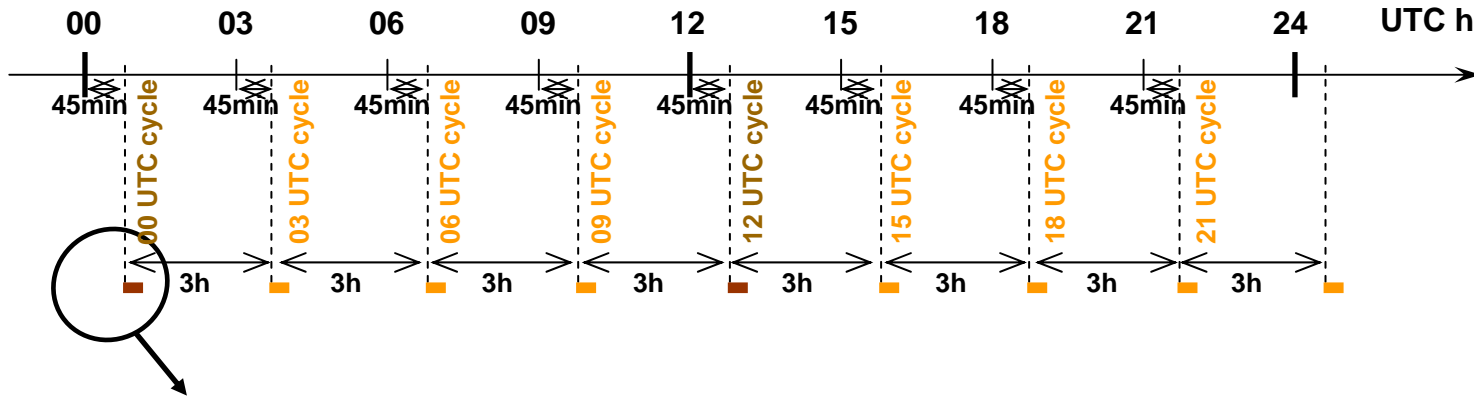
Production Scheme

Assimilation cycle

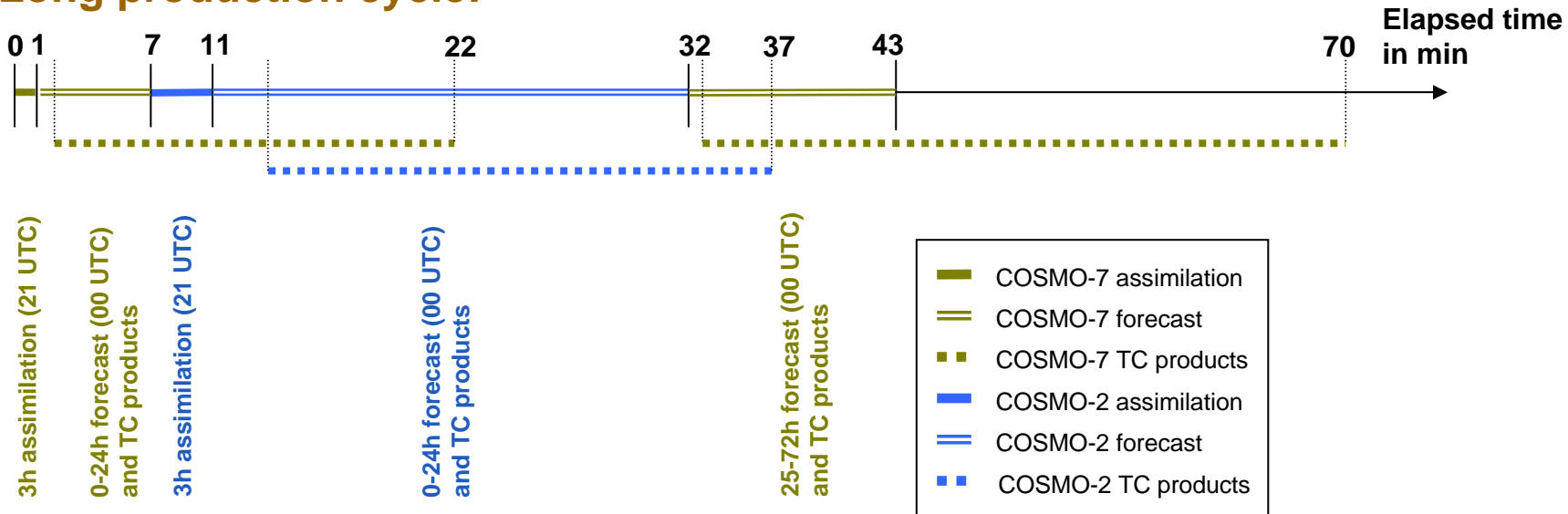




Production Scheme

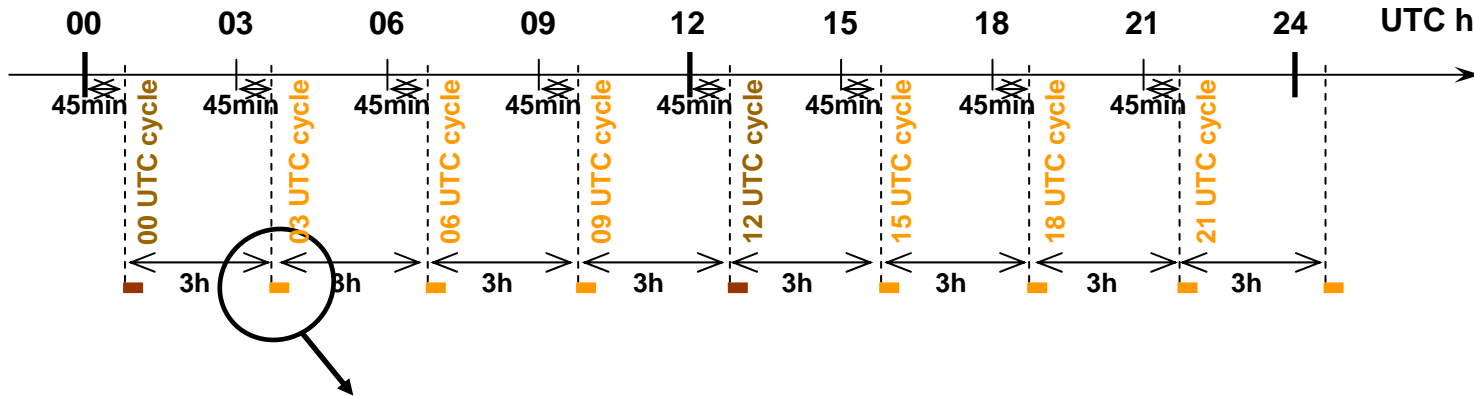


Long production cycle:

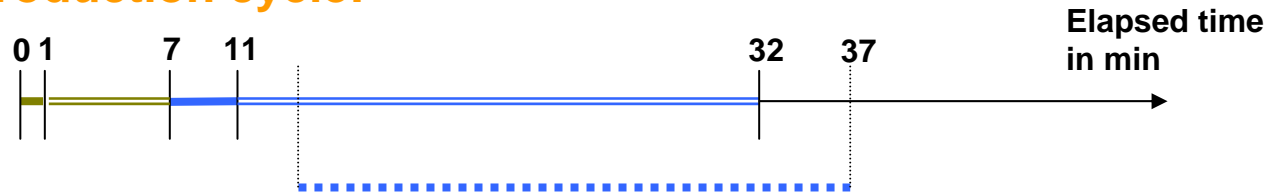




Production Scheme



Short production cycle:



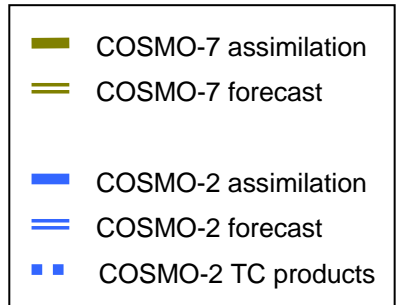
3h assimilation (00 UTC)

0-24h forecast (03 UTC)

3h assimilation (00 UTC)

0-24h forecast (03 UTC)
and TC products

25-54h forecast (03 UTC)



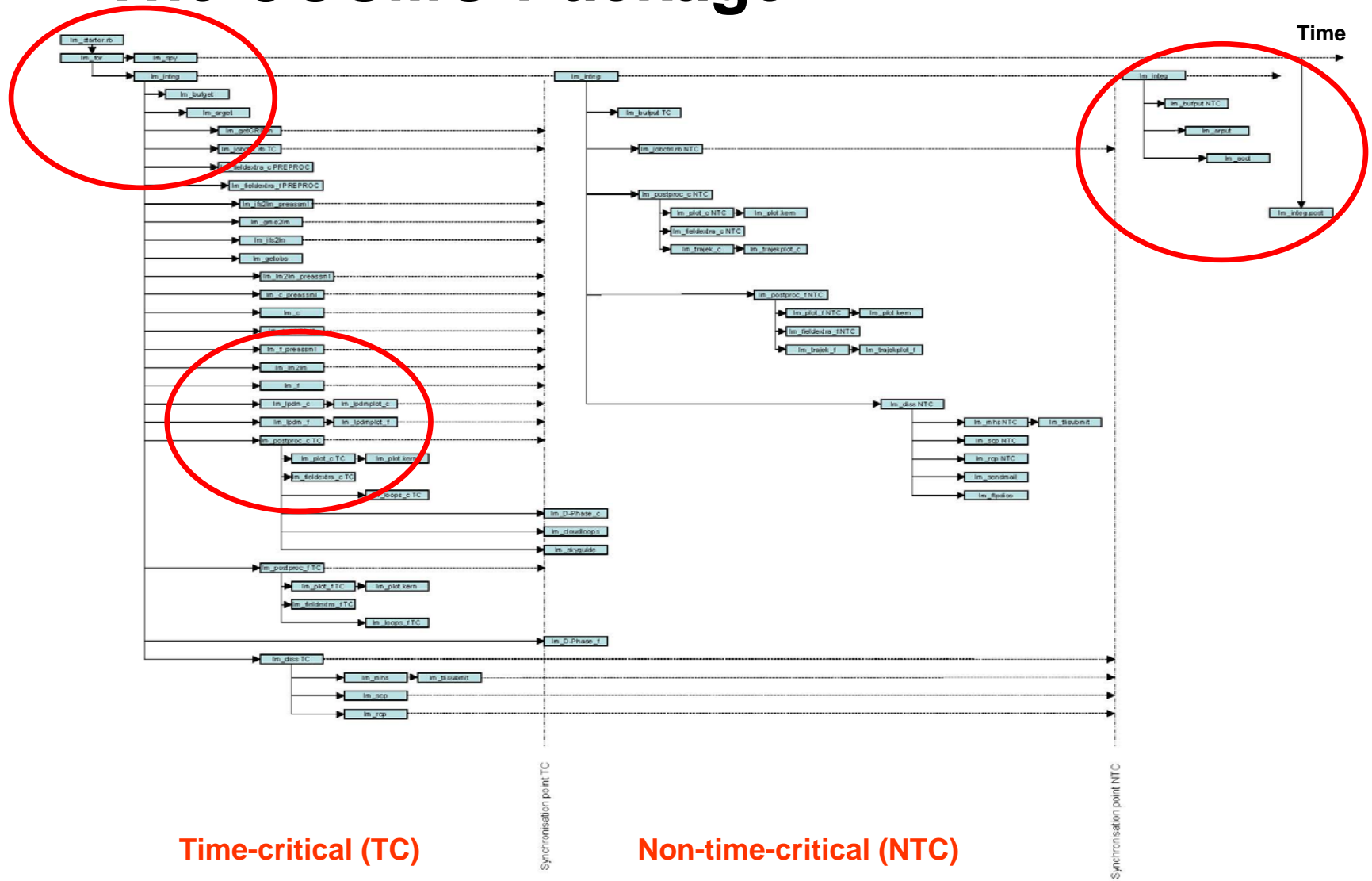


The COSMO Package

- Steering software, comprising ~50 shell and Ruby scripts
- Running on Cray service node
- Key features
 - Starter with task schedule list input, running as cron job
 - Concurrent run of supervising spy and task processing
 - Time-critical and non-time-critical part of task execution
 - Concurrent model integration, post-processing, and dissemination in time-critical part
 - Inter-process-communication via files

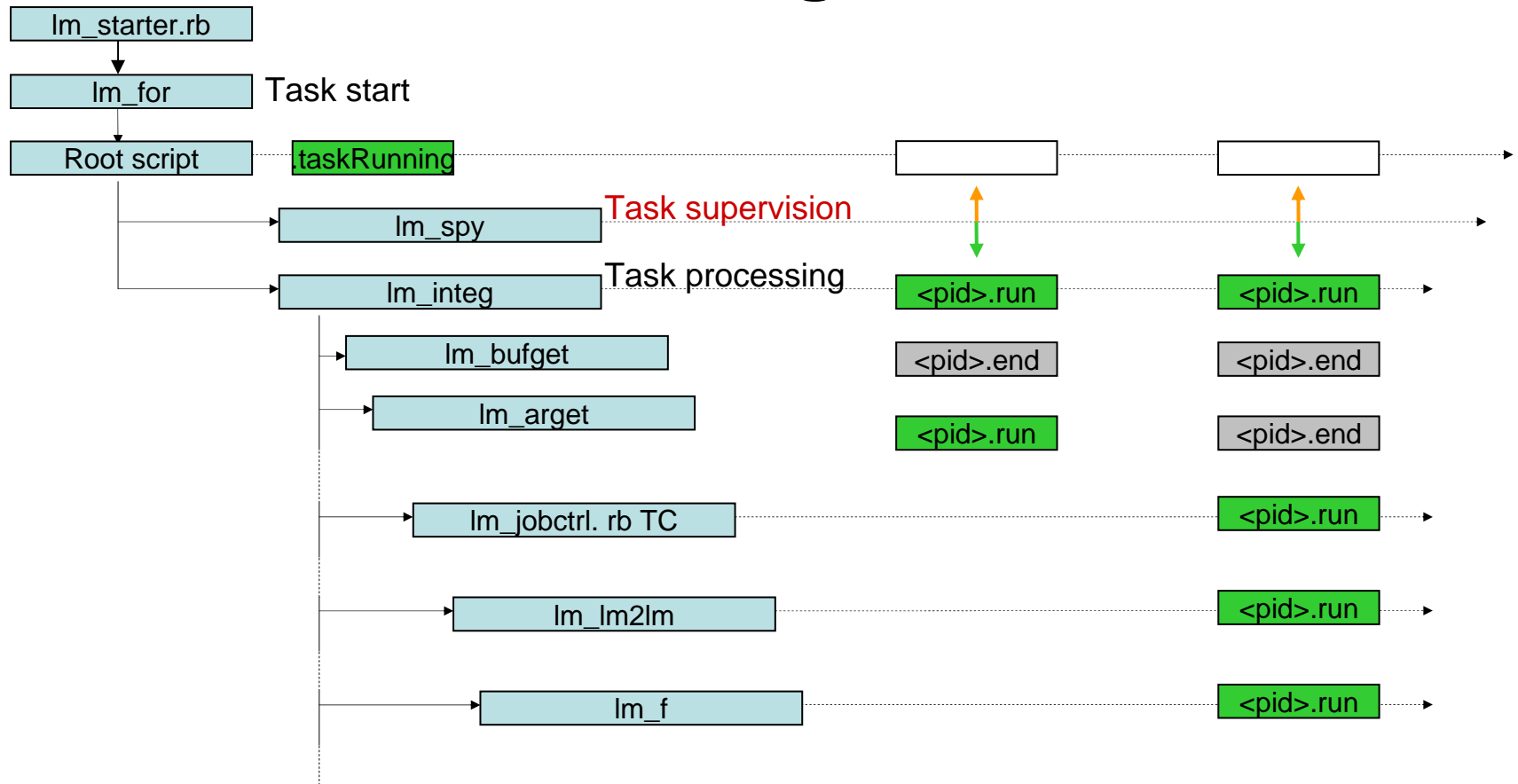


The COSMO Package





The COSMO Package

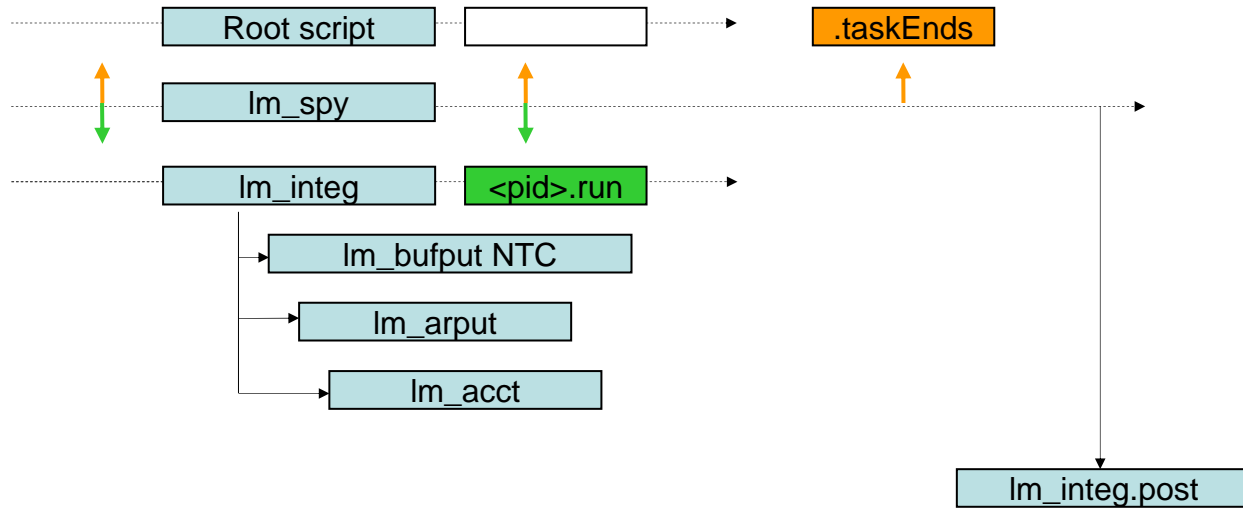


Task supervision

- Information: periodically collect all messages for log file and mail messages
- Control: periodically check disk and abnormal exit of package modules, using `<pid>.run/<pid>.end` file mechanism, check for `.taskEnd`
- Final actions: task status, post-mortem actions, and housekeeping

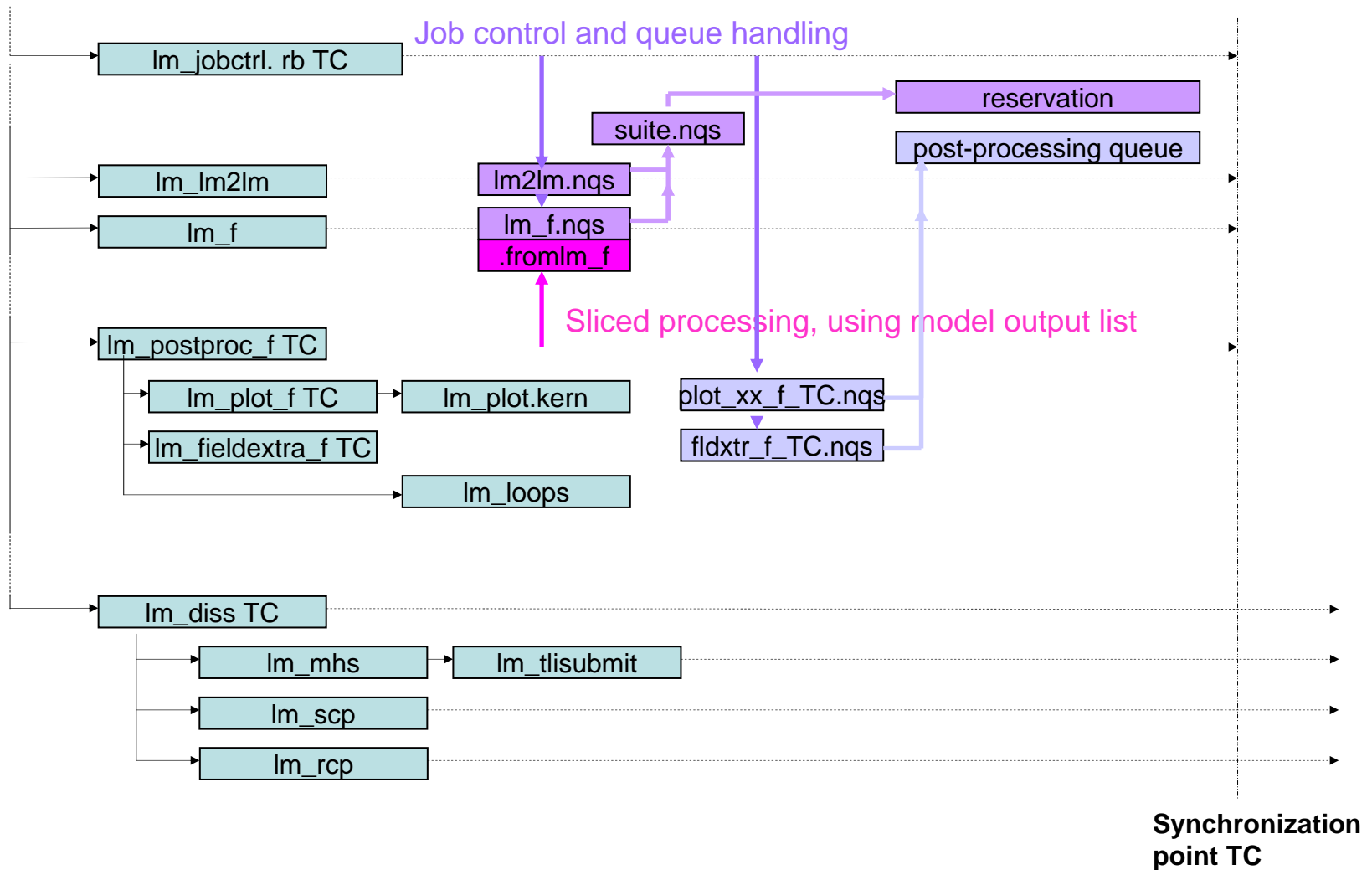


The COSMO Package





The COSMO Package





Reliability Statistics

6 months period Apr-Sep 2008:

366 times +72h COSMO-7 production: (96% reliability)

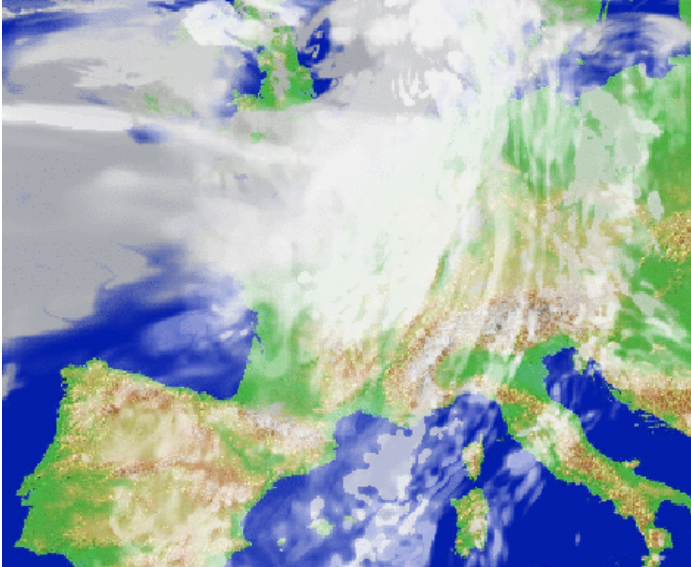
	#Total	#CSCS problems	#COSMO package
delay TC end > 12h	0	0	0
5h < delay TC end <12h	2	2	0
1.1h < delayTC end < 5h	5	4	0
TC end OK, but missing products	6	3	2

1464 times +24h COSMO-2 production: (97% reliability)

	#Total	#CSCS problems	#COSMO package
delay TC end > 3h	14	7	2
0.5h < delay TC end <3h	14	7	3
TC end OK, but missing products	9	3	3



Many thanks for your attention!



Questions?