Scalability of IFS on massively parallel computers with special focus on Data Assimilation

Mats Hamrud

Acknowledgement to Colleagues at ECMWF

Talk overview

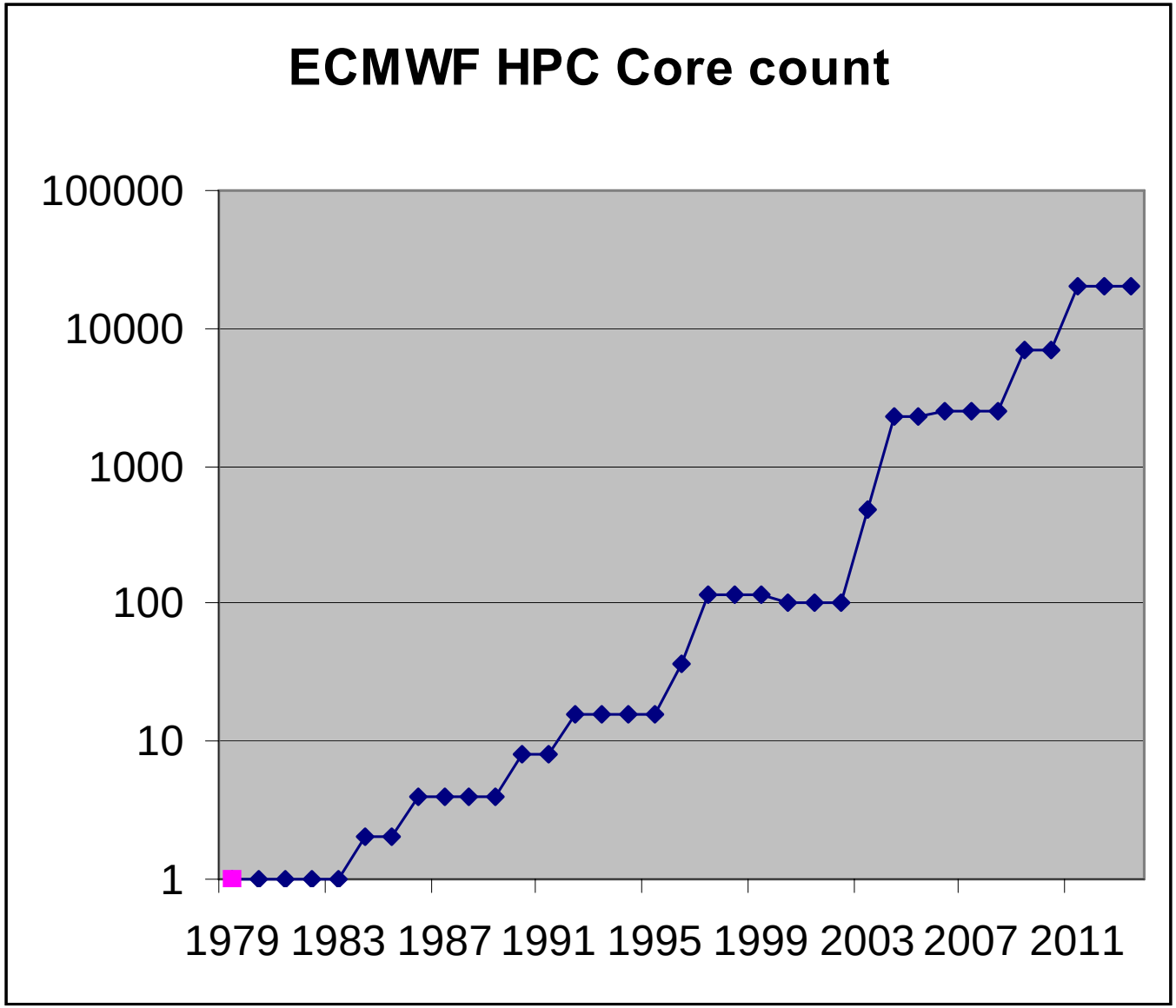
- **Project plan**
- **Future High Performance computers**
- **Current scaling properties of ECMWF Data Assimilation System**
- **Conclusion**

Project Plan

- **Project to initially run for two years**
- **Prepare for the possibility on running ECMWF's operational codes on massively parallel computers**
- **Main focus on Data Assimilation where the greatest challenge is foreseen**
- **Collaborative effort with other groups**

ECMWF HPC through time

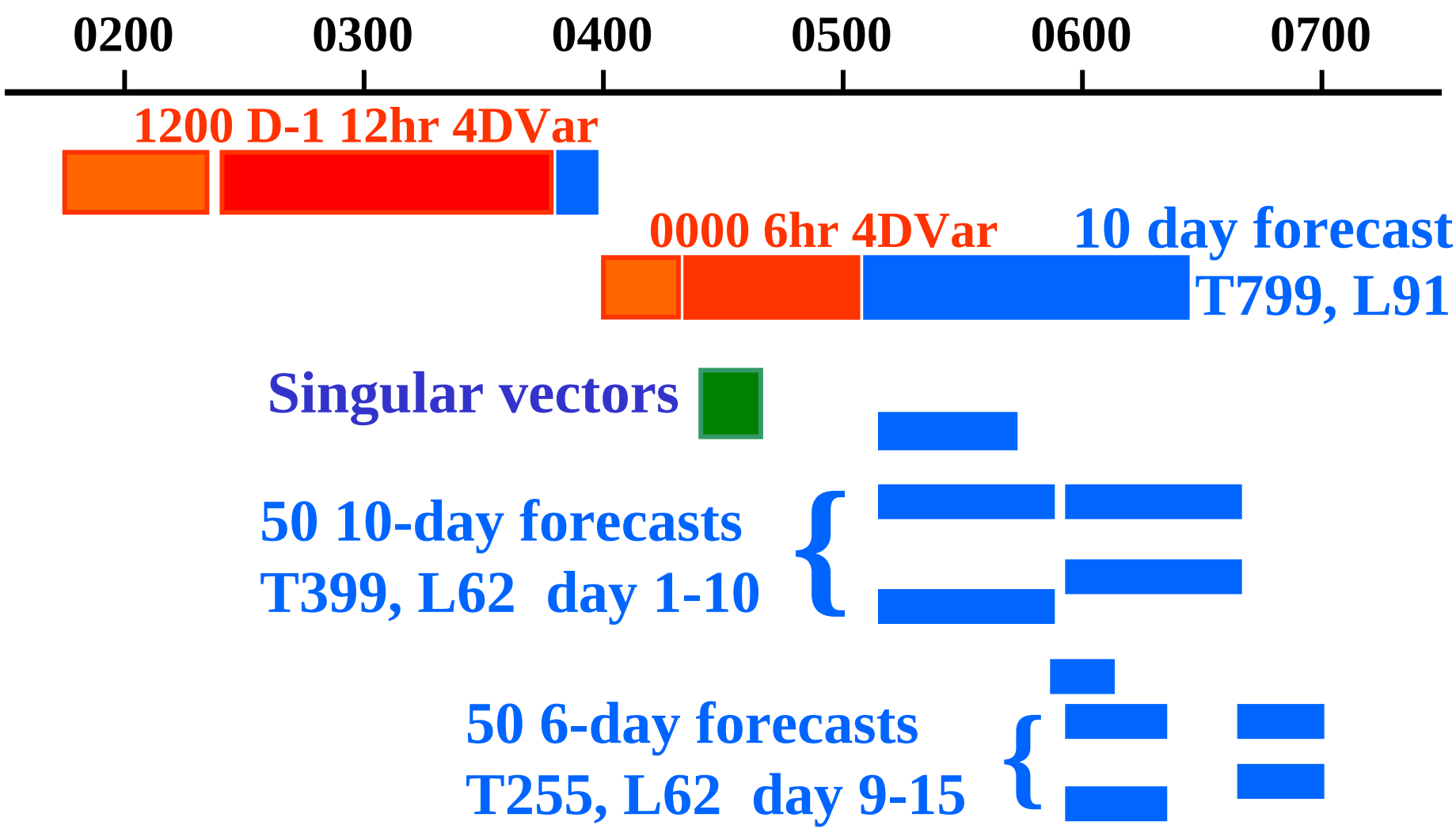
Cray 1A	1979-1984	1	1
Cray XMP-22	1984-1986	1	2
Cray XMP-48	1986-1990	1	4
Cray Y-MP	1990-1992	1	8
Cray C90	1992-1996	1	16
Fujitsu VPP700	1996-1997	36	36
Fujitsu VPP700	1997-2000	116	116
Fujitsu VPP5000	2000-2003	100	100
IBM P690	2003-2004	30*2	480*2
IBM P690+	2004-2006	70*2	2240*2
IBM P5-575+	2006-2009	155*2	2480*2
IBM P6	2009-2011	~240*2	~8000*2


ECMWF HPC Core count
100000
10000
1000
100
10
1
1979 1983 1987 1991 1995 1999 2003 2007 2011

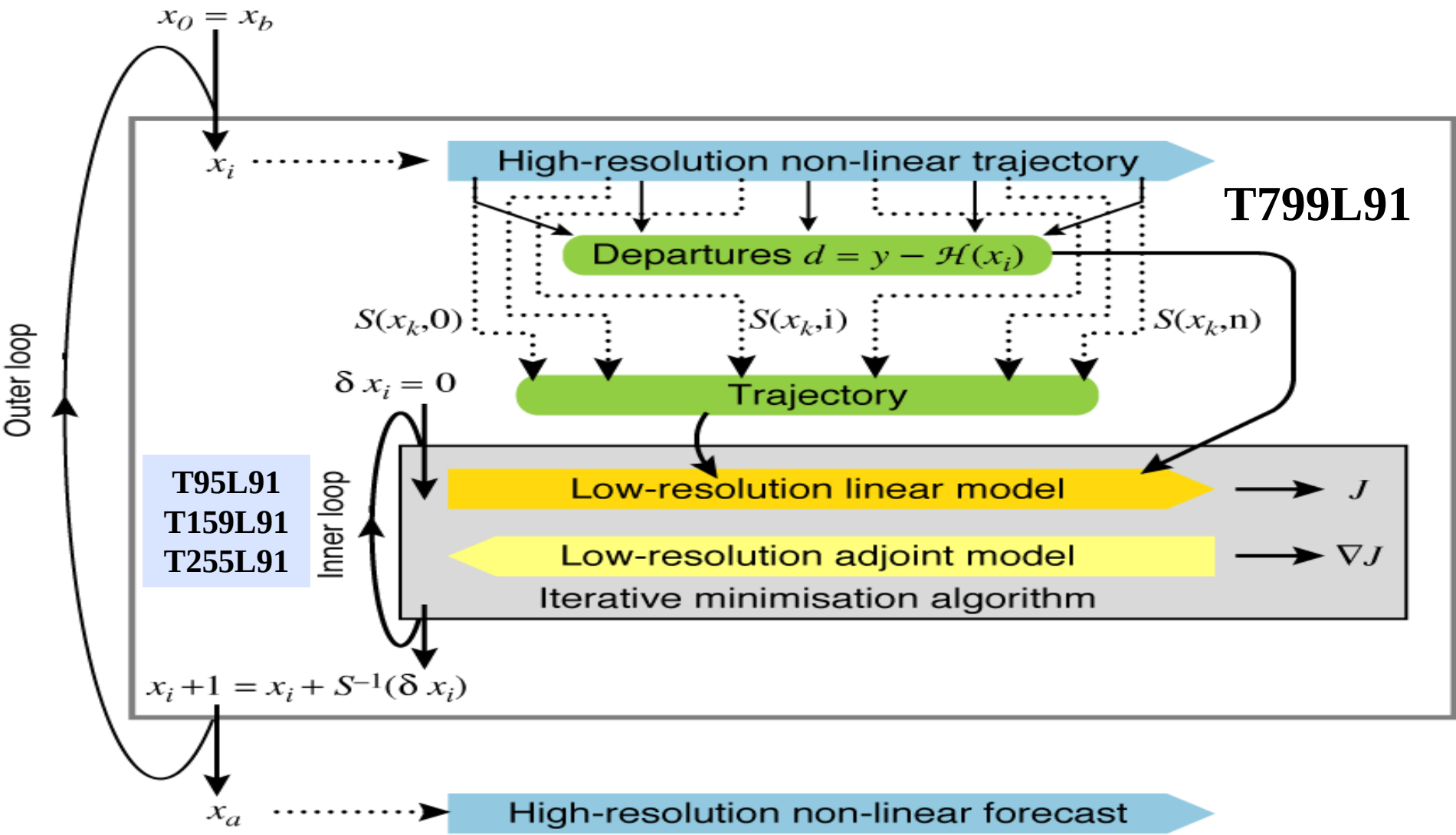
Next HPC at ECMWF

- **May have in the order of 100,000 cores based on current budget and industry trends**
- **Contract with IBM for Power6/7 expires June 2013**
- **Benchmark needs to be ready by the end of 2010**
- **All components of benchmark should scale reasonably to > 25,000 cores**
- **Operational codes need to be ready for making efficient use of ~100,000 cores early in 2013**

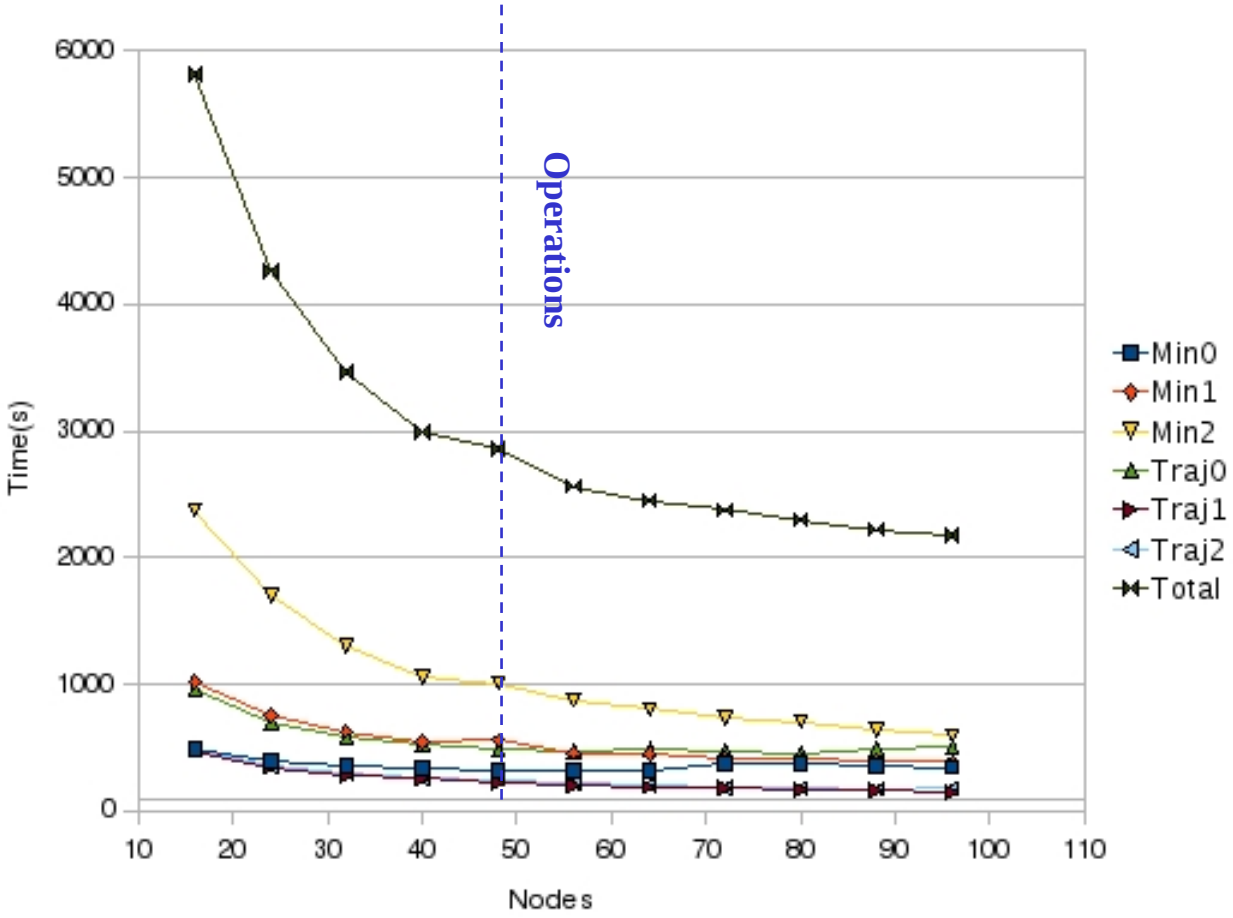
The 0000 UTC Operational Suite



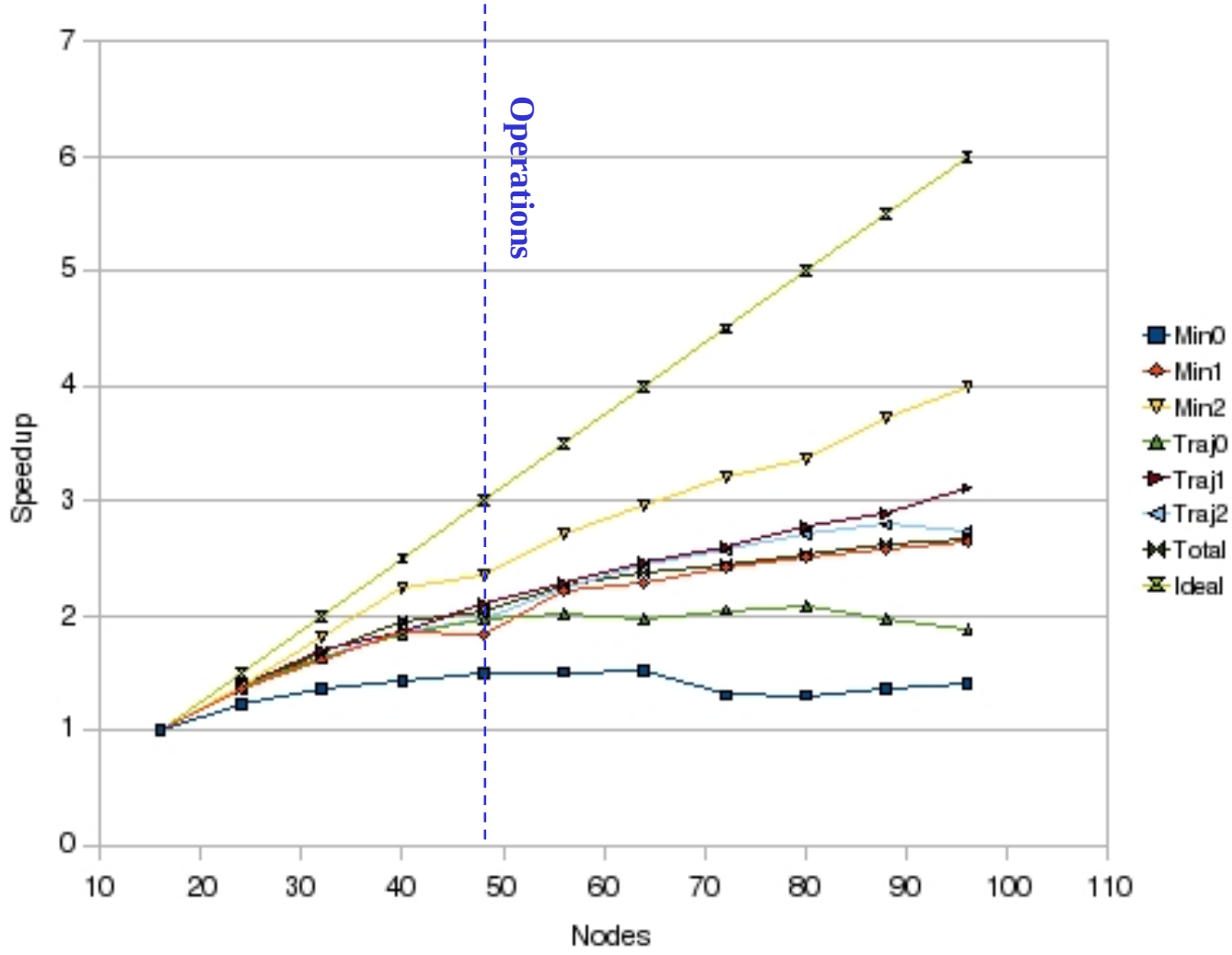
Multi-incremental 4D-Var at ECMWF



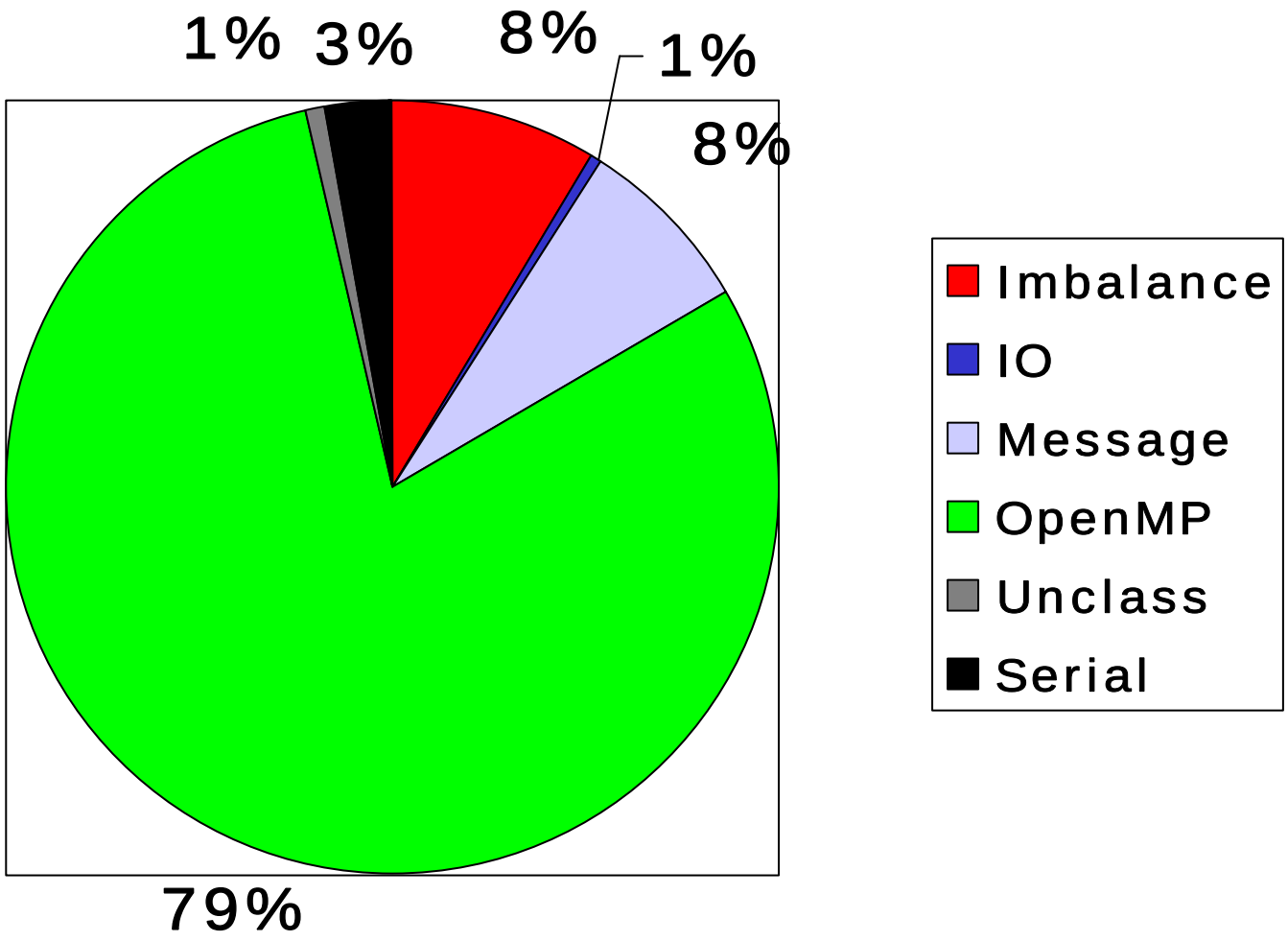
4D Var run-time, 32 user threads per node



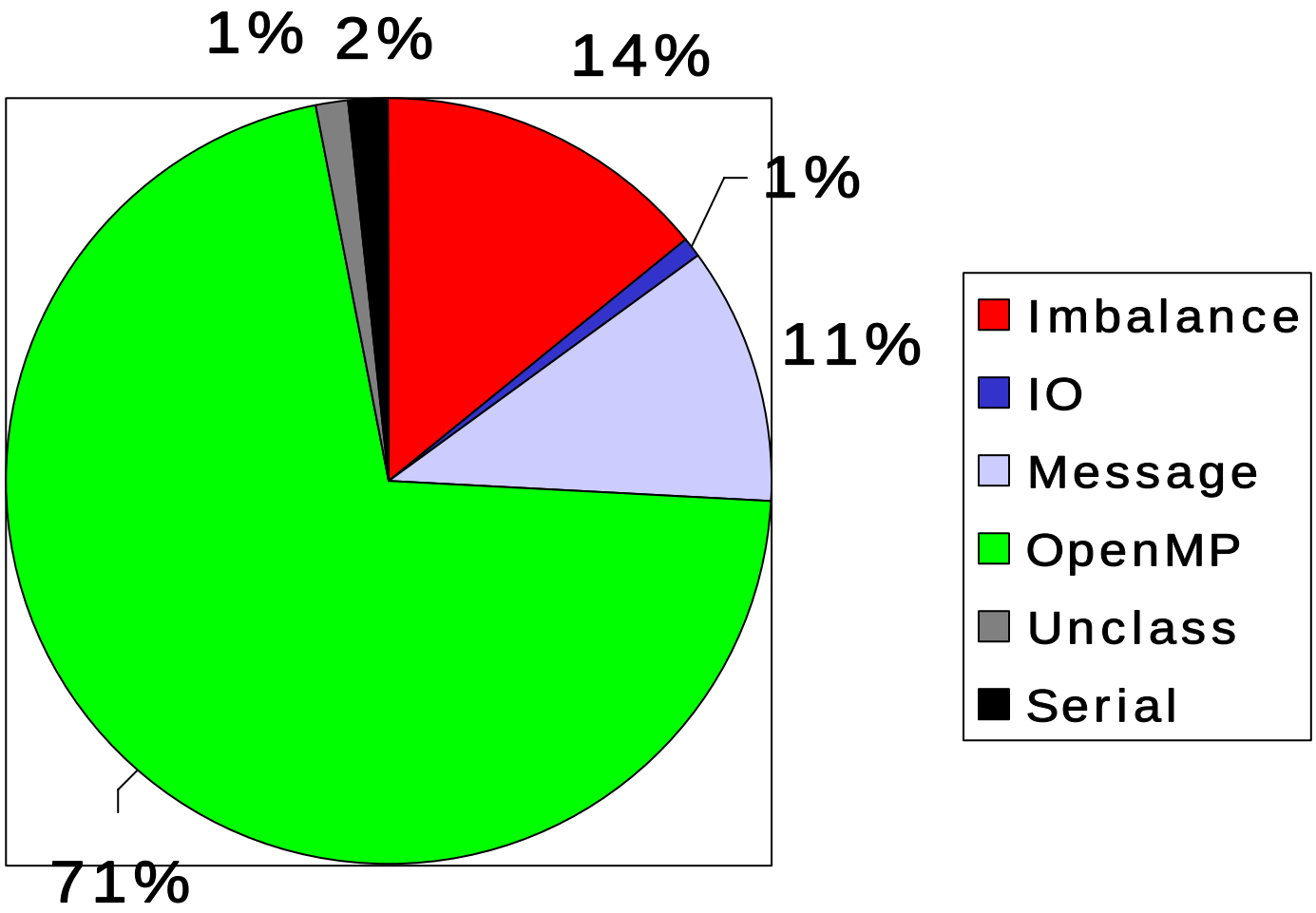
Incremental 4D Var speedup



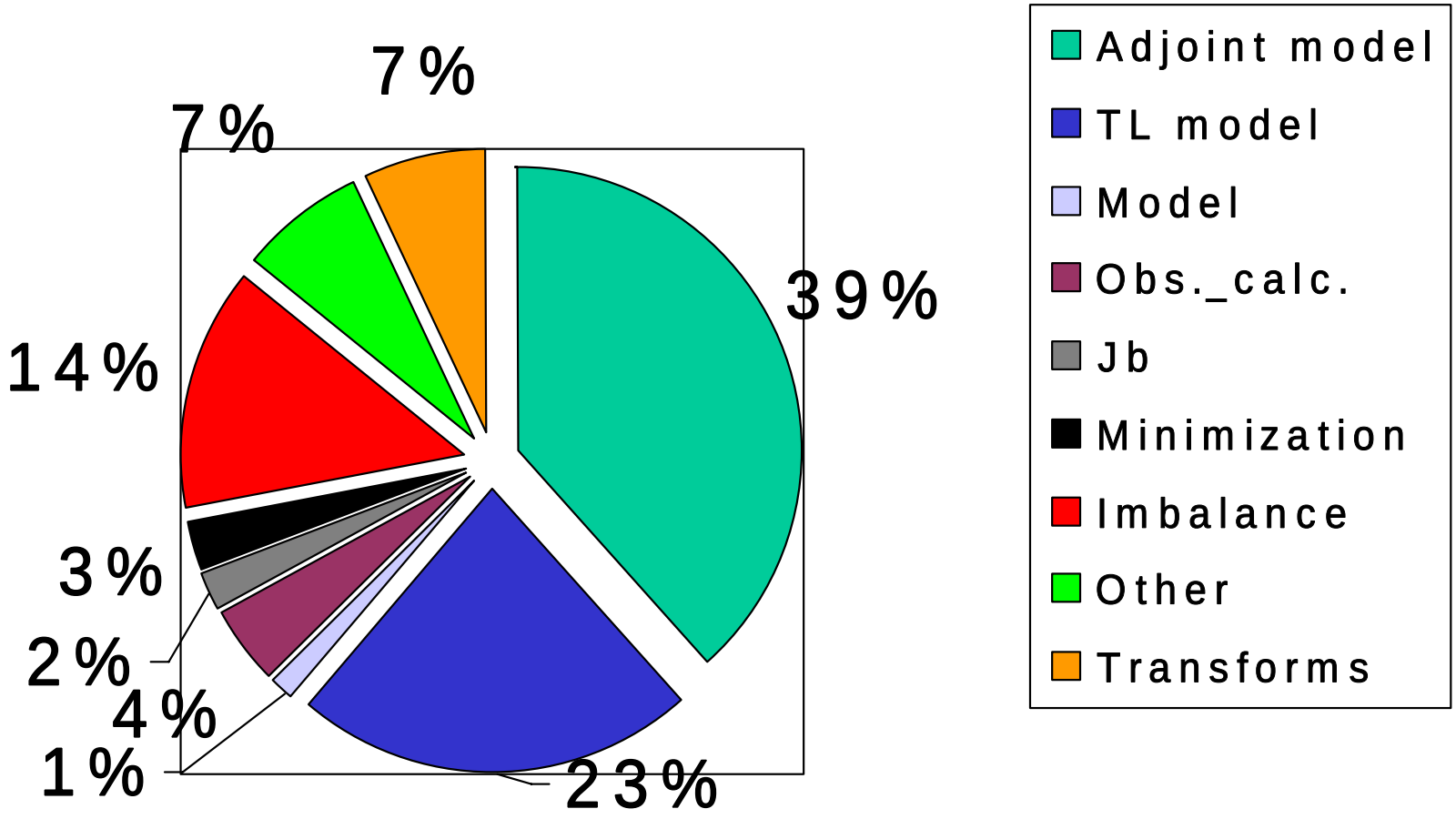
Final minimization, 24 Nodes (96x8)



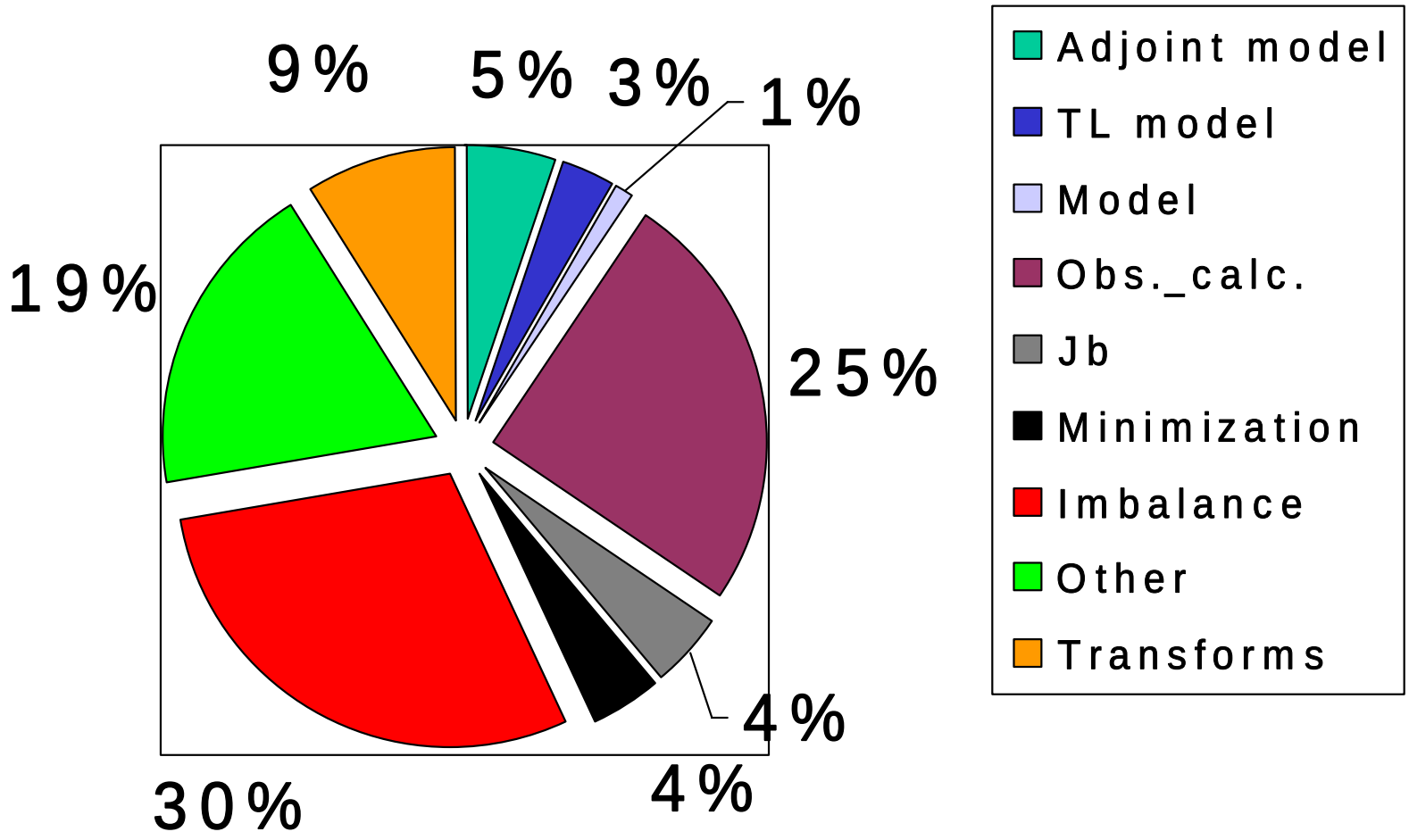
Final minimization, 48 Nodes (192x8)



Final minimization, 48 Nodes



First minimization, 48 Nodes



Questions we need to answer

- **What is inhibiting scaling of the current 4D Var?**
- **What is the impact on scaling of the planned scientific developments for the IFS ?**
- **Should the continued use of Incremental 4D Var within the foreseeable future be assumed?**

Items from ECMWF 10-year strategy

- Non-hydrostatic model
- Long window 4-D var
- Error correlations of observations
- Ensemble data assimilation
- Additional variables (rain ,CO2,aerosols…)
- Modularization of IFS
- Increased resolution (10km in 2015)

Conclusion

- **The ECMWF Data Assimilation System presently does not scale as we would wish it to do**
- **We need to resolve these problems or look at alternative algorithms**