

New HPCS at Slovak Hydrometeorological Institute

Oldřich Španiel

oldrich.spaniel@shmu.sk

Operational suit at SHMI from 2003

- **Agreement on co-operation with ZAMG**
- **unfailing operational suit**
 - ◆ **over 1000 integration**
 - ◆ **over 100 GB of output data**

Invitation To Tender for High Performance Computing System

ITT – 50 mandatory requirements

- ◆ benchmark test, reliability, acceptance test
 - ◆ requirements for HW
 - ◆ requirements for SW, OS
 - ◆ requirements for delivering and upgrade
 - ◆ service, maintenance
 - ◆ training and documentation
 - ◆ technical information for installation and operation
- compilation and linking
 - Lancelot - execution of EE927
 - Morgane – execution of 001
 - Switchover test
-
- elapsed time limit for Morgane - 28 min

Invitation To Tender for High Performance Computing System

Benchmark domain

- 320x288 points
 - horizontal resolution 8.3km
 - 51 vertical levels
 - time step 337.5s
 - 48h integration => 512 steps
 - quadratic spectral truncation
- **July 2003**
declaration of tender
 - **October 2003**
evaluation committee
 - **December 2003**
contract signed
 - **January 2004**
delivery of HW

Story of Furniture van & Removal crew & Happy users



The hardware fashion



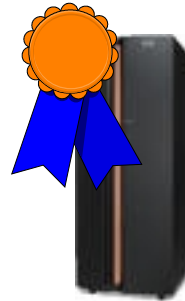
SGI Altrix 3700
32 CPUs Itanium 2 - 1.3 GHz



NEC SX6-4B
4 CPUs



SUN – Sun Fire 15K
72 CPUs UltraSPARC III – 1.2 GHz



IBM Regatta p690
32 CPUs Power 4 T+ - 1.7GHz



HP AlphaServer GS1280
32 CPUs EV7 - 1.15 GHz

HW specification

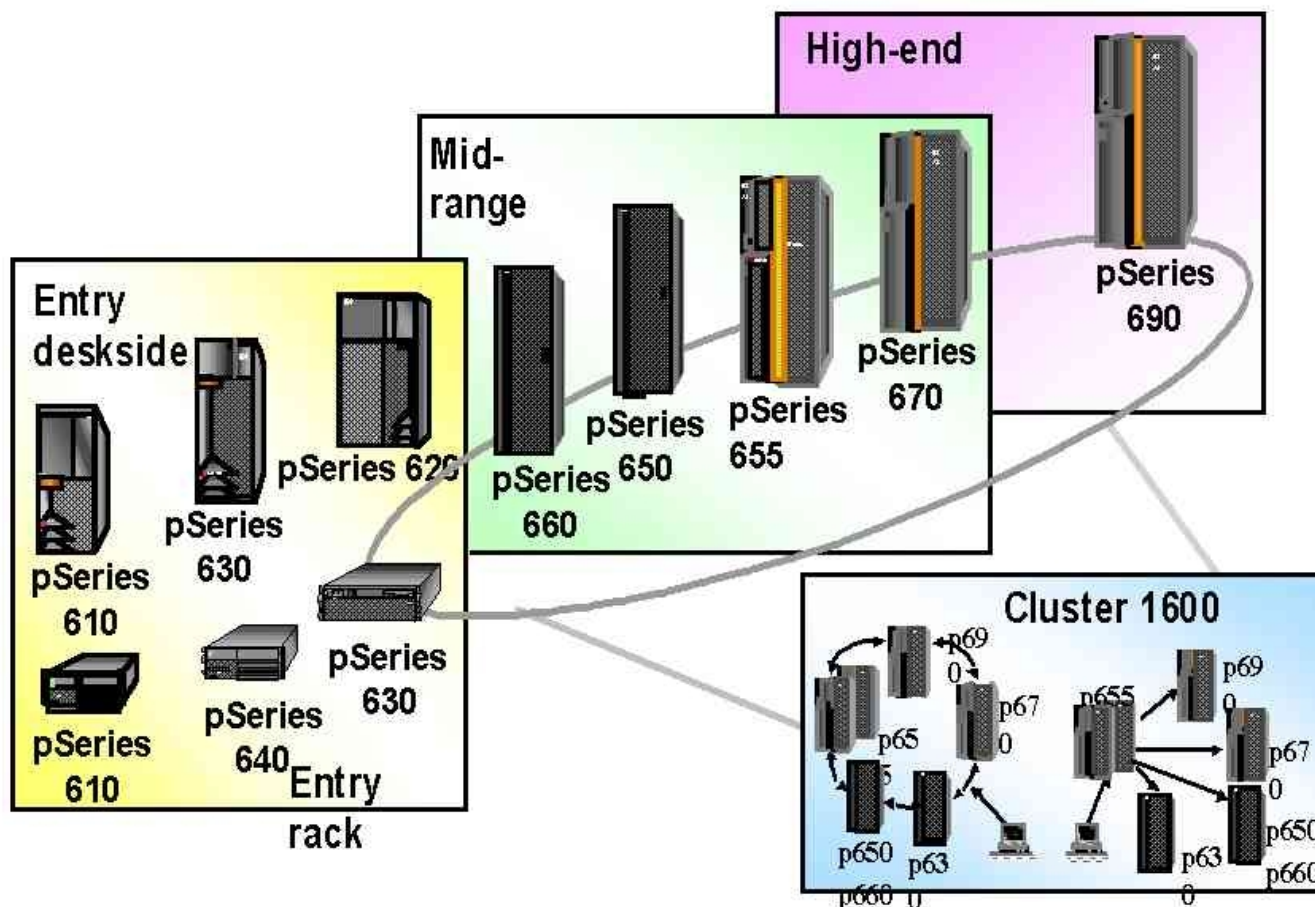
- IBM @server pSeries 690
- 32 CPUs POWER 4 Turbo+ 1,7 GHz
- 32 GB RAM memory
- IBM FAST T600 Storage Server + EXP700, 1.5 TB

SW specification

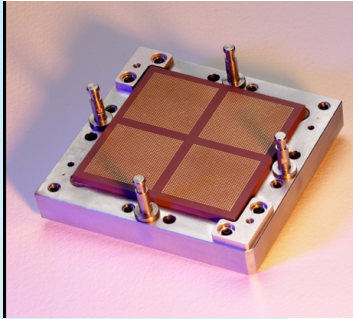
- AIX 5.2
- Fortran compiler XLF 8.1.1.0
- C,C++ compiler
- Engineering and Scientific Library ESSL
- Mathematics Library MASS 3.0
- Parallel Environment (MPI):PE 3.2.0.16
- LoadLeveler 3.2
- TotalView



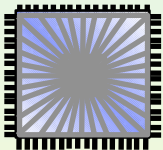
IBM UNIX system pSeries



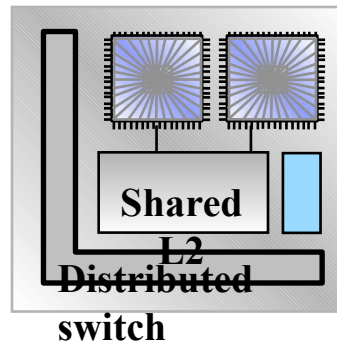
GigaHz Chip / Module Structure Buildup



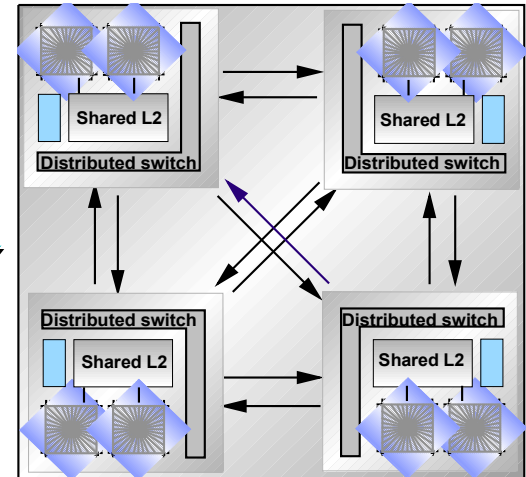
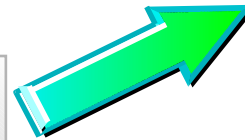
**4.5" square glass-ceramic
8-way SMP in 20 sq.ins !**



**POWER4+ 1.7 GHz
microprocessor**



**2-way POWER4+ SMP
system on a single chip!
(184 million transistors)**



**8-way (4
chip)POWER4+
SMP system on a
Multichip Module
pSeries 690
basic building block**

Four POWER4+ chip slots

The diagram illustrates the architecture of a POWER4+ chip slot. It features a central vertical bar labeled 'MEMORY SLOT'. To the left of this bar, two horizontal lines represent the 'GX Bus'. Each 'GX Bus' line has a green arrow pointing towards the 'MEMORY SLOT' bar. The 'MEMORY SLOT' bar is divided into two main sections: 'MEMORY' (top) and 'SLOT' (bottom). Each section is connected to a 'Mem Ctrl' (Memory Controller) block, which is in turn connected to an 'L3' (Level 3) cache block. The connections are as follows: the top 'MEMORY' section is connected to the top 'Mem Ctrl' block, which is connected to the top 'L3' block; the bottom 'SLOT' section is connected to the bottom 'Mem Ctrl' block, which is connected to the bottom 'L3' block. The 'Mem Ctrl' blocks are connected to the 'L3' blocks via bidirectional arrows. The 'L3' blocks are also connected to the 'Mem Ctrl' blocks via bidirectional arrows. The 'GX Bus' lines are connected to the 'MEMORY' and 'SLOT' sections via bidirectional arrows.

processors (8 processors) on an MCM, with 1

The diagram illustrates a 4-processor MCM architecture. It consists of four identical processing blocks arranged in a 2x2 grid. Each block contains a 'Shared L2' cache and a 'Distributed switch'. The 'Shared L2' cache is connected to the 'Distributed switch' via a vertical line. The 'Distributed switch' is connected to the external 'GX Bus' via a horizontal line. The 'GX Bus' is represented by a central horizontal line with four vertical lines extending to each block. The 'GX Bus' is labeled '4 GX Bus links for external connections'. The 'Shared L2' cache is labeled 'L3 cache shared across all processors'. The 'Distributed switch' is labeled 'Distributed switch'.

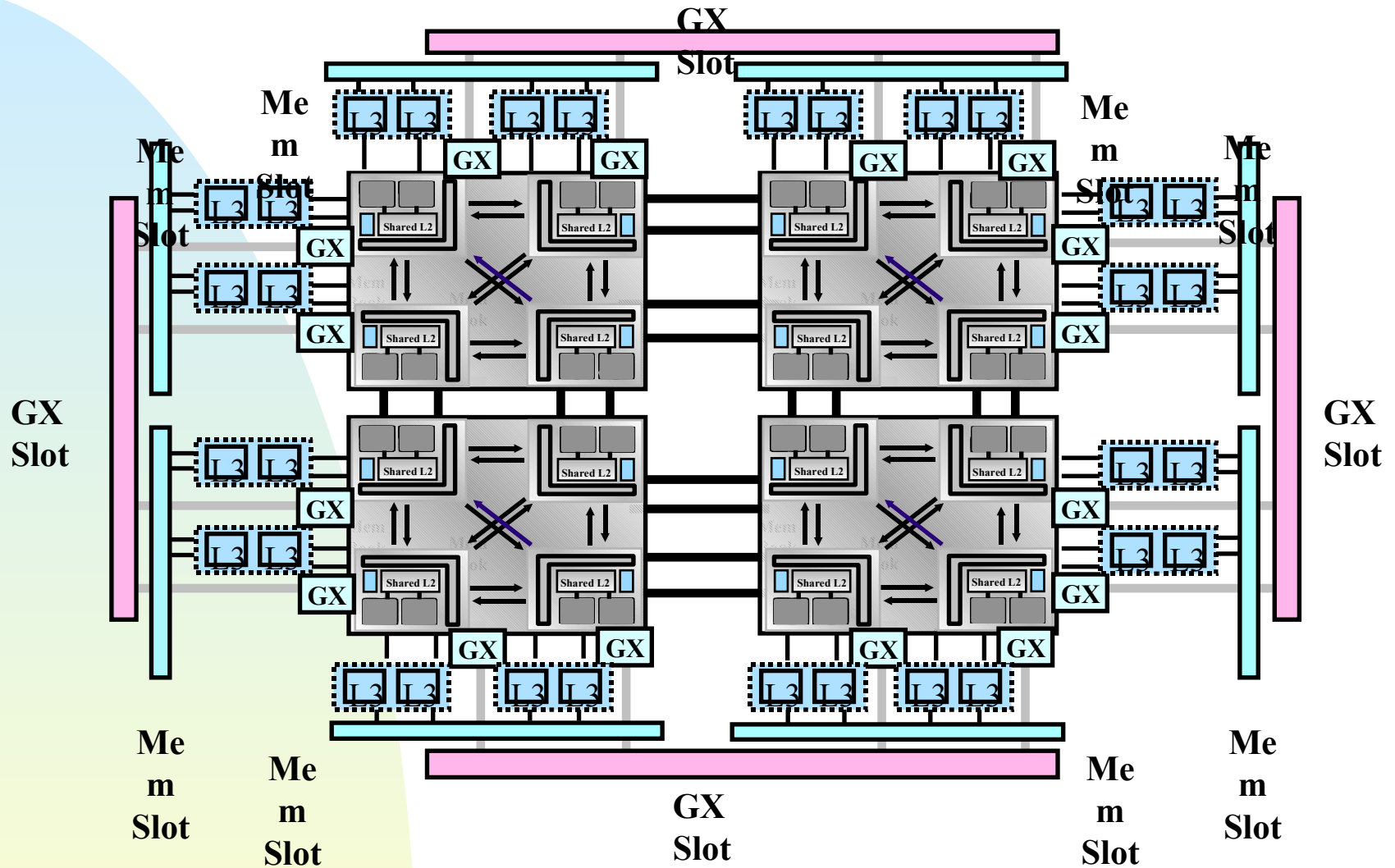
L3 cache shared across all processors

4 GX Bus links for external connections



4 GX Bus links for external connections

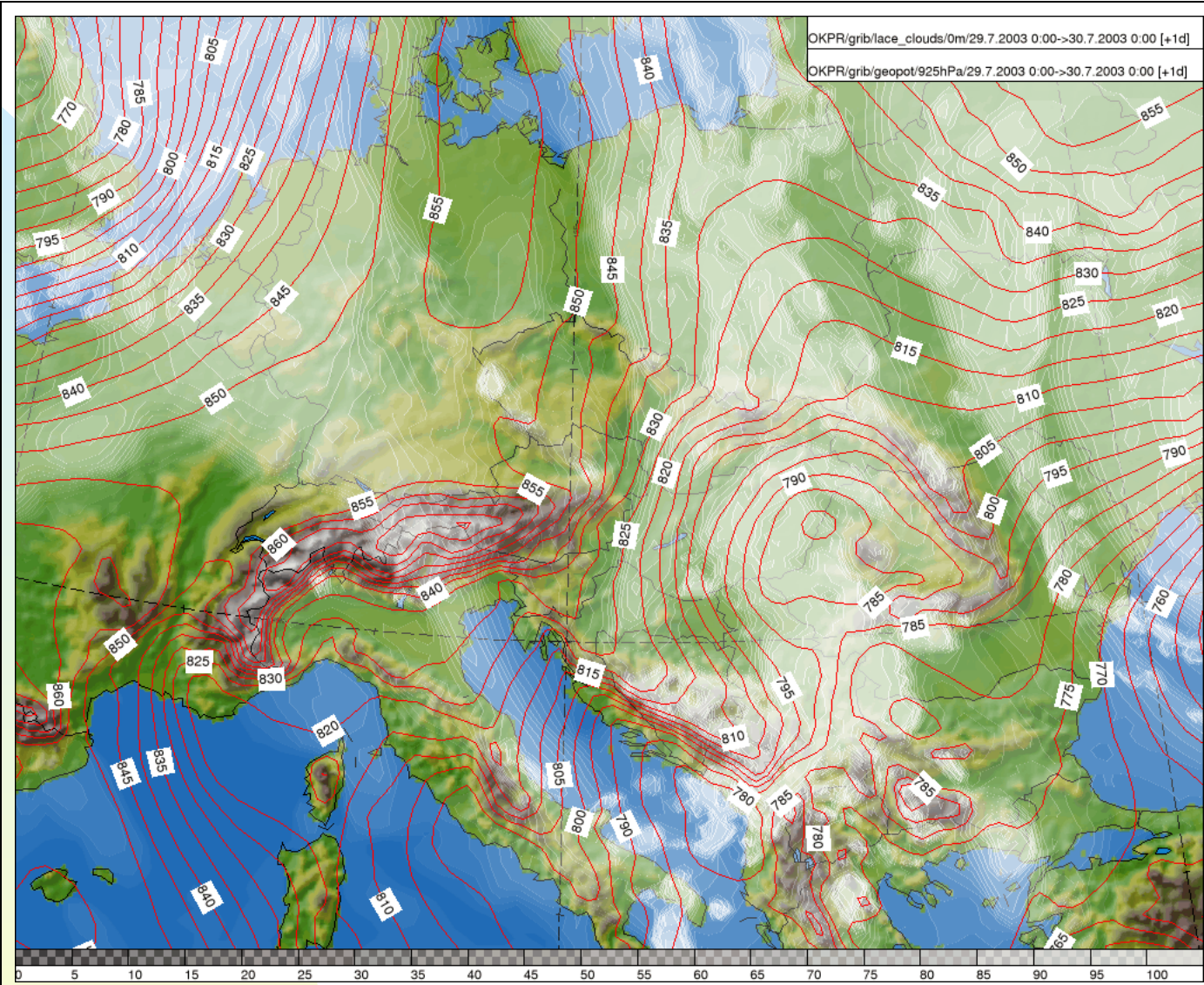
Full 32-way p690 Backplane Structure



the operational suite - basic characteristics

- domain size 2882x2594 km
(320x288 points in quadratic grid)
- domain corners [2.19 ; 33.99 SW][39.06 ; 55.63 NE]
- horizontal resolution 9.0 km
- vertical levels 37
- time step 400 s
- forecast length 48h / twice per day (00 and 12 UTC)
- mode dynamical adaptation
- code version AL25T1
- LBC - RETIM&BDPE, coupling frequency - 3 h
- 1.7.2004

- **The Products & Solutions Support Center (PSSC) IBM Montpellier**
 - ◆ **porting of ALADIN code (not so complex)**
 - ◆ **optimisation of code**
 - ◆ **optimisation of memory manager and I/O**
 - ◆ **provision of reliability of operational suite (AIX Workload Manager (WLM) & LoadLeveler)**
- **ECMWF, Marocco, Tunisia, Hungary and Slovakia**



Conclusion