



43 YEARS OF HAVING “FUN” WITH HPC SYSTEMS

... AND BUILDING TOOLS TO OPTIMIZE THE PERFORMANCE OF HPC APPLICATIONS

27 MAR 2025 | BERND MOHR



THE EARLY YEARS

1981 TO 1992

/

FRIEDRICH ALEXANDER UNIVERSITÄT ERLANGEN-NÜRNBERG

Member of the Helmholtz Association



TELEFUNKEN TR440 @ FAU

My first computer: used in programming class in 1st semester computer science

- Fastest computer developed in Europe at this time (end of the 1970ies)
- **Dual** processor! ⇒ BUT, FAU had the only **3-processor** version!
- Amazing huge 1.5 Mbyte main memory
- Programming via punch cards (Algol60, Fortran IV, etc)



- Info

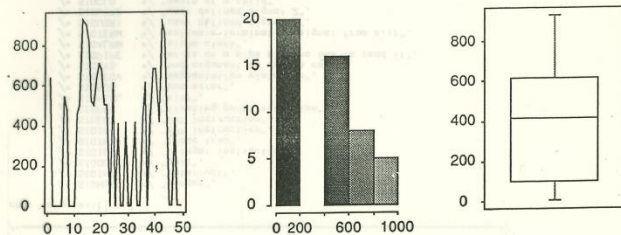
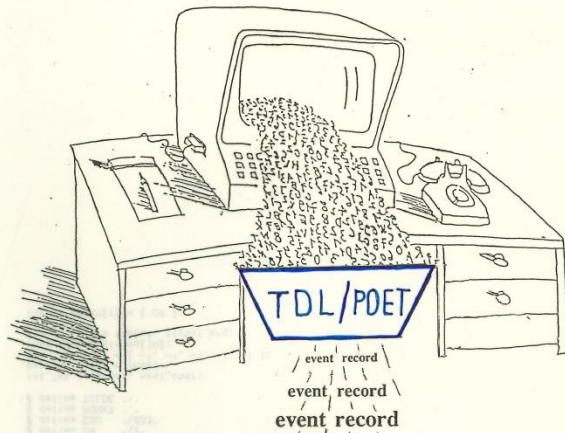
- English: https://www.chessprogramming.org/TR_440
- German: https://de.wikipedia.org/wiki/TR_440

MASTER / PH.D CONTEXT

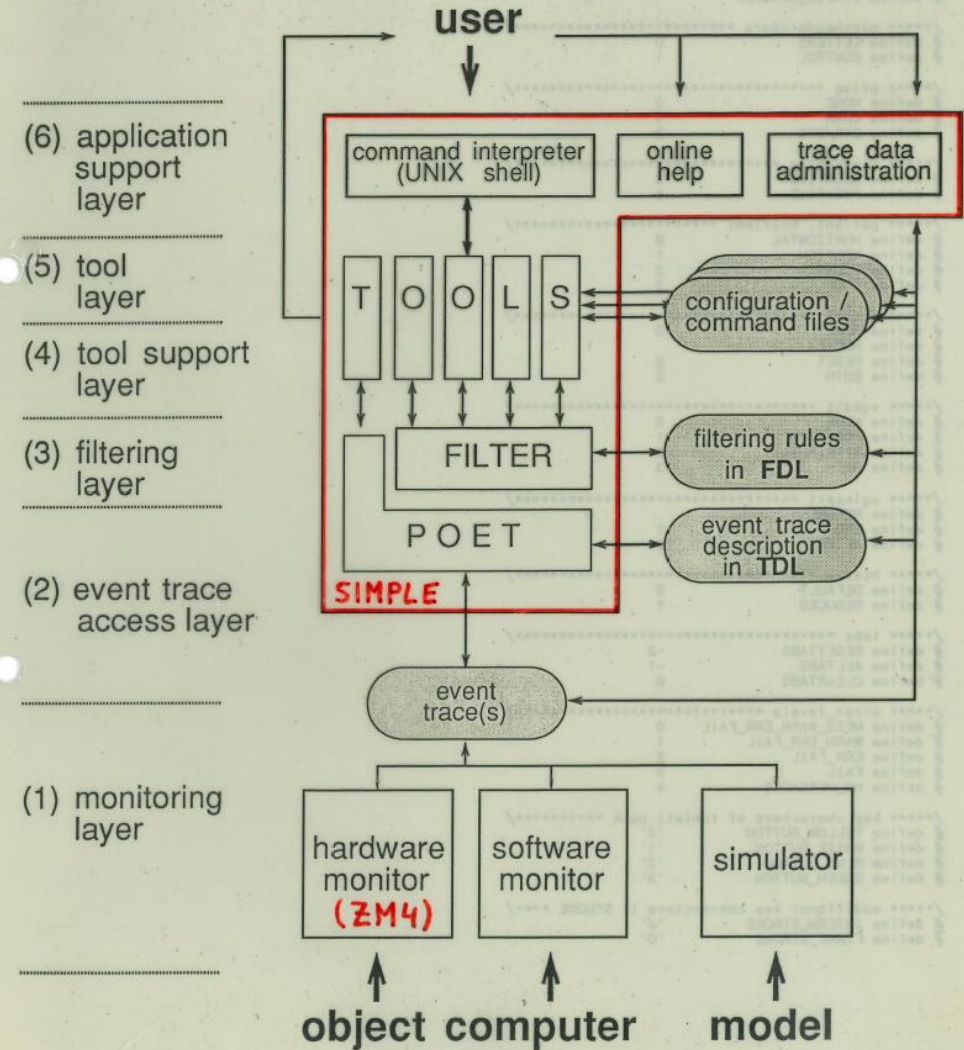
- Der SFB 182
 - „Multiprozessor- und Netzwerkkonfigurationen“
 - Four 3-year phases (1987 – 1998)
- Work Package C1
 - „Messung, Modellierung und Bewertung von Multiprozessoren und Rechnernetzen“
- Parallel system development @ FAU IMMD
 - EGPA (Erlangen General Purpose Array)
 - DIRMU (Distributed Reconfigurable Multiprocesser kit)
 - MEMSY (Modular Expandable Multiprocessor System)



Source related and
Integrated
Multiprocessor and -computer
Performance evaluation,
Logging, and visualization
Environment



□ general structure of SIMPLE:



object system / operating system	monitor / interface	application
DIRMU / DIRMOS	logic analyzer ZM4 / parallel port	numerical application simulation program
Transputer	ZM4 / link adapter ZM4 / bus adapter	communication system TRACOS
SUPRENUM / PEACE	ZM4 / 7 segment display	ray tracing
IBM-PC / OS/2, MSDOS	ZM4 / Centronics	protocol software B-ISDN, FDDI
IBM-PC / XENIX	ZM4 / Centronics	protocol software
SUN4 / SunOS	ZM4 / VME bus	X-Windows
SIEMENS robot control	ZM4 / SMP bus	robot control software
CCC3280 / XELOS	software monitor	multiprocessor UNIX
IBM-PC network	ZM4 / Centronics	Electrical Load Supervision Control System

IBM Zurich
Research Lab

IBM ENC
Heidelberg

Siemens
Erlangen

Siemens
Munich

Fudan Univ.
Shanghai
Page 6

SUPRENUM

- German: **SUP**er**RE**chner für **NUM**erische Anwendungen
- English: super-computer for numerical applications

- German research project to develop a parallel computer from 1985 through 1990
- Major effort aimed at developing a national expertise in HPC both at HW and SW level
- Developed 256-node Suprenum-1
 - Fastest massively parallel MIMD computer in the world during a period in 1992
 - Still, project was canceled and considered a commercial failure

- <https://en.wikipedia.org/wiki/SUPRENUM>





THE POSTDOC YEARS

1993 TO 1995

/

UNIVERSITY OF OREGON, EUGENE

Member of the Helmholtz Association



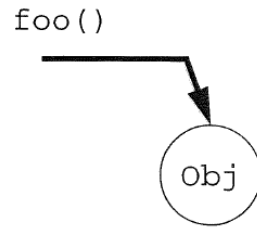
CONTEXT

- ARPA funded project “pC++”
 - **Programming Environments, Compiler Technology and Runtime Systems for Object-Oriented Parallel Processing**
 - Dennis Gannon, Indiana University
 - Postdocs: Pete Beckman, Francois Bodin
 - **pC++** compiler and runtime system, **Sage++** toolkit
 - **Languages, Libraries and Performance Evaluation Tools for Scalable Parallel Systems**
 - A. Malony, J. Cuny, University of Oregon
 - Postdoc: Bernd Mohr
 - **TAU** program analysis tools

pC++ – The Programming Language Ideas

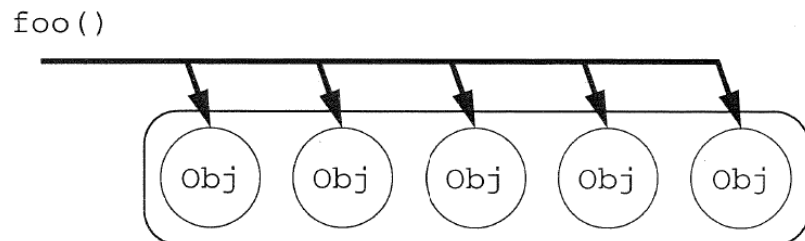
- **regular C++:** programmers apply operators and functions to objects as “messages”

```
class Obj {  
    int x;  
    void foo();  
};  
Obj myObj;  
myObj.foo();
```



- **pC++:** this concept is extended so that an operator or function can be applied to a large set, grid, array (:= **collection**) of objects (:= **elements**) in parallel

```
Collection Vector { ... };  
Vector<Obj> paraObj(AlignObj, DistrObj);  
paraObj.foo();
```



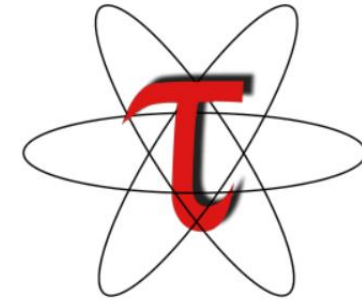
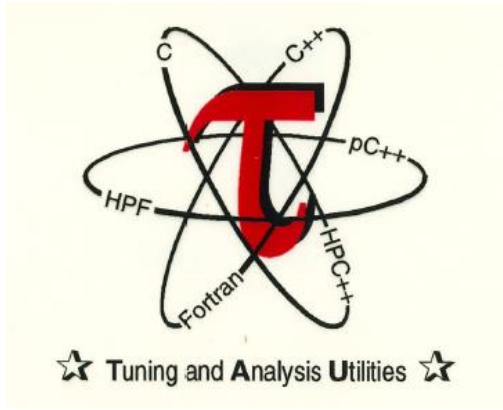
pC++ – Supported Systems

- Shared memory systems
 - Kendall Square KSR-1 / KSR-2
 - Sequent Symmetry (under Dynix + PTX)
 - SGI (Power) Challenge + Onyx
 - Convex SPP-1
 - (BBN TC2000)
 - Distributed memory systems
 - TMC CM-5
 - Intel Paragon
 - IBM SP-1 / SP-2
 - Cray T3D / **T3E**
 - Meiko-CS2
 - Workstation Clusters with PVM + MPI (homogeneous)
 - UNIX Workstations (SUN, HP, DEC, IBM, SGI, ...)
 - serialized
 - thread-based (Pthreads, LWP, AT&T tasks, Awesime)
- ➡ The same pC++ program will run without modification on all platforms

TAU LOGO EVALUATION



★ Tools Are Us ★

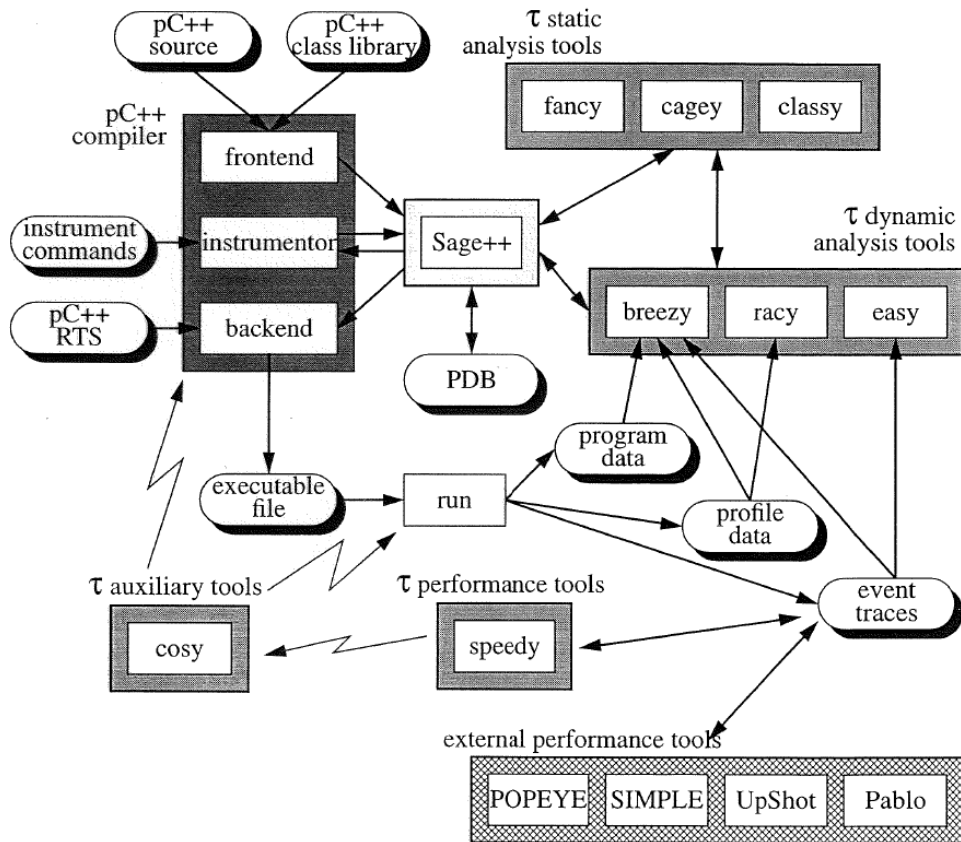


TAU Performance System ®



<https://www.cs.uoregon.edu/research/paraducks/>

The pC++ Programming Environment



☆ Tools Are Us ☆

- Currently available τ tools:
 - **cosy** (**C**ompile manager **S**tatus displa**Y**)
 - **fancy** (**F**ile **A**ND **C**lass displa**Y**)
 - **cagey** (**C**ALL **G**raph **E**xtended displa**Y**)
 - **classy** (**C**LASS hierarch**Y** browser)
 - **racy** (**R**outine and data **A**Ccess profile displa**Y**)
 - **speedy** (**S**peedup and **P**arallel **E**xecution **E**xtrapolation **D**ispla**Y**)
 - **breezy** (**B**reakpoint **E**xecutive **E**nvironment for visuali**Z**ation and data displa**Y**)

- Prototypes:
 - **easy** (**E**vent **A**ND **S**tate displa**Y**)
 - **dandy** (**D**istributed **A**rray **N**avigator **D**ispla**Y**)
 - **crafty** (**C**ont**R**ol flow **A**ND **F**unc**T**ion displa**Y**)
 - **geeky** (**G**Eeky **E**ditng and symbol loo**K**up displa**Y**)
 - **POPEYE, DAQV** (data and performance visualization)

- τ can work with a local or remote pC++ language system
- τ originally designed for C++/pC++ programs



THE LATER YEARS

1996 TO NOW

/

JÜLICH SUPERCOMPUTING CENTRE

Member of the Helmholtz Association



CONTEXT: 25 YEARS OF AUTOMATIC TRACE ANALYSIS

- 1999 – 2004

APART

- EU ESPRIT + IST Working Group
- <http://www.fz-juelich.de/apart/>*



- Sequential analyzer EXPERT
- <http://www.fz-juelich.de/zam/kojak/>*

*HINT: <https://web.archive.org/>

- 2006 – now

VI-HPS

- Helmholtz Virtual Institute
- <http://www.vi-hps.org/>

scalasca

- Parallel analyzer SCOUT
- <http://www.scalasca.org>



- Trace Visualizer
- <http://www.vampir.eu/>



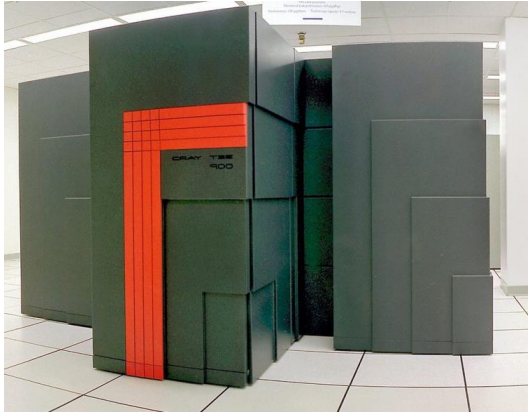
- Instrumentation and measurement framework
- <http://www.score-p.org>



SCORE-P/SCALASCA: SUPPORTED ARCHITECTURES OVER THE YEARS

- **Instrumentation and measurement only (visual analysis on front-end or workstation)**
 - Cray XT, XE, XK, XC, EX
 - IBM BlueGene/L, BlueGene/P, BlueGene/Q
 - K Machine, Fujitsu FX10 and FX100, Fugaku
 - Tianhe 1A and 2
 - Intel MIC (KNC, KNL)
- **Full support (instrumentation, measurement, and automatic analysis)**
 - Linux IA32, IA64, x86_64, PPC, ARM, and ARM64 based clusters
 - IBM AIX Power3/4/5/6/7/8/9 based clusters

JÜLICH HPC PRODUCTION SYSTEMS (LATE 90ies)



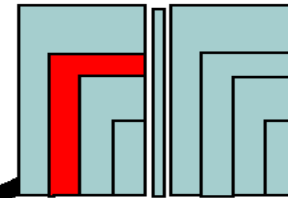
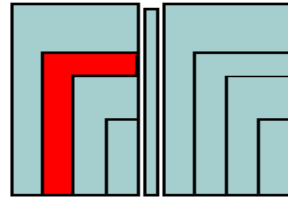
CRAY Supercomputer Complex

CRAY T3E-600

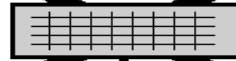
Massively-parallel Supercomputer
 * 512 Processing Elements
 * Peak Performance 307 GFLOPS
 * Main Memory 64 GByte

CRAY T3E-1200

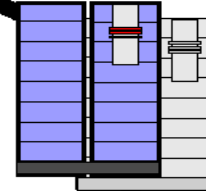
Massively-parallel Supercomputer
 * 512 Processing Elements
 * Peak Performance 614 GFLOPS
 * Main Memory 262 GByte



Gigabit Backend Network
 * GigaRing, HiPPI, ATM

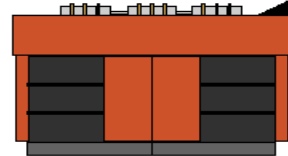


CRAYJ90/16
 Vectorsupercomputer
 for Interactive Work



CRAY T90/12

Vectorsupercomputer
 * Peak Performance 20 GFLOPS
 * Main Memory 8 GByte



CRAYJ90/12
 File Server

Mass Storage:
 * 3 TByte Disks
 * 72 TByte Tape Silo

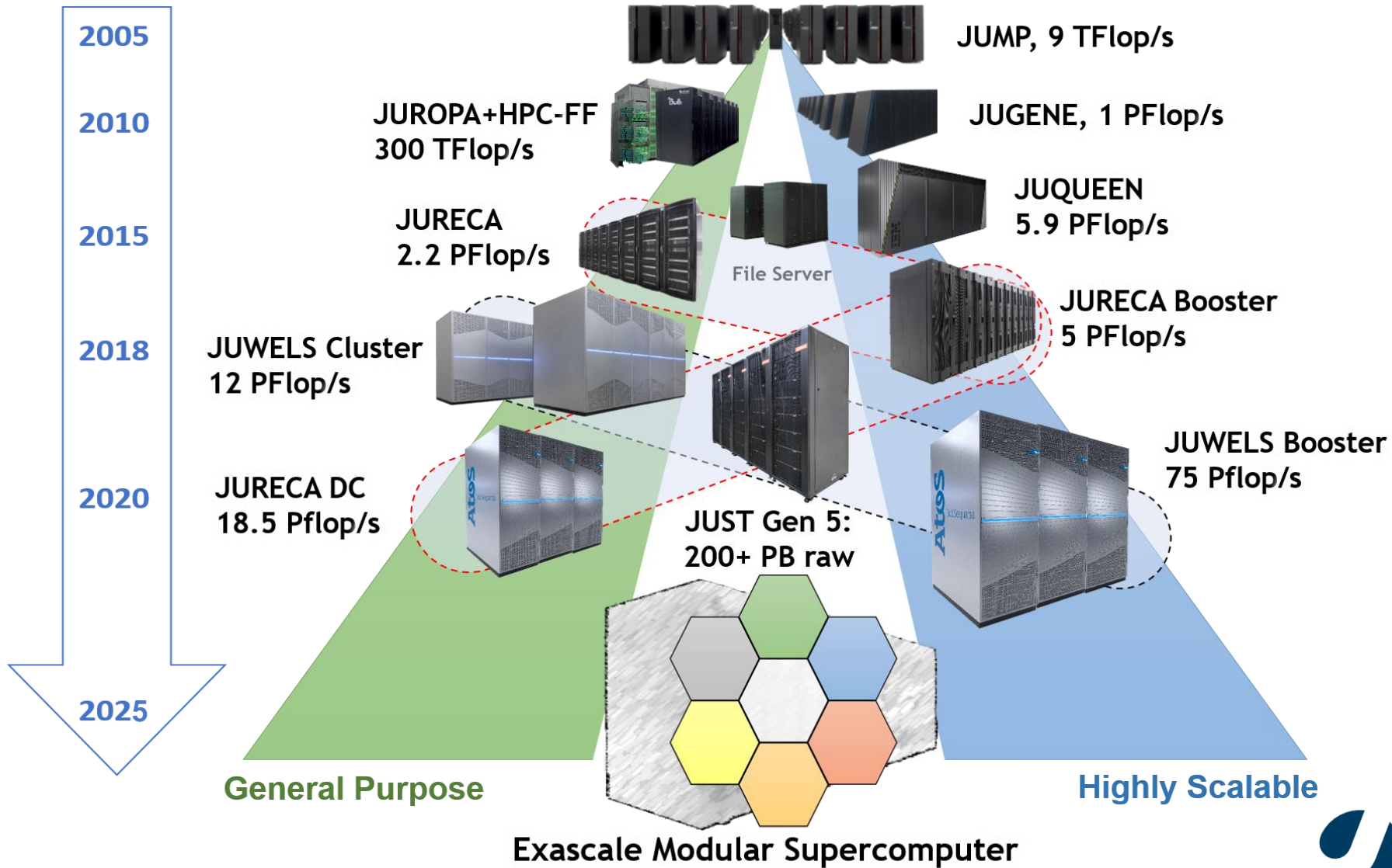


ZAM, 17.09.1999



Also:
 140-node
 Intel Paragon XP/S
 10 GFLOPS

JÜLICH HPC PRODUCTION SYSTEMS (2000 ⇒)





37 YEARS OF PERFORMANCE TOOLS DEVELOPMENT: SUCCESS STORIES AND FAILURES

<https://www.youtube.com/watch?v=IhIrigOFHwo>

QUESTIONS ?

