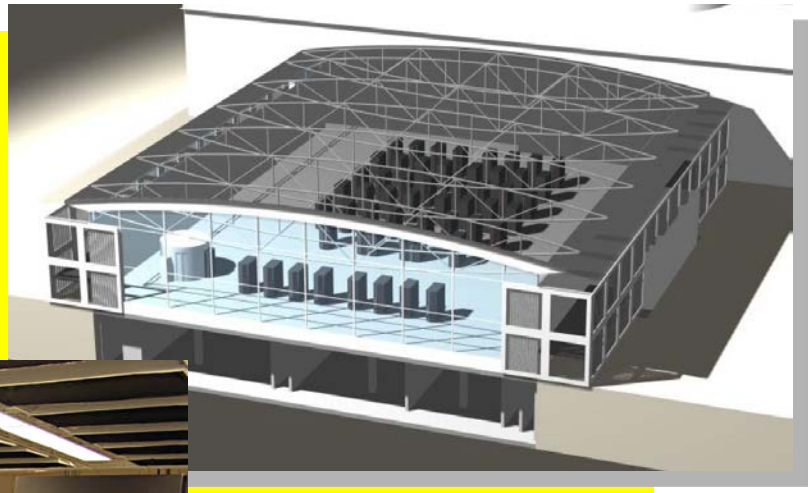
Beyond Automatic Performance Analysis

Prof. Dr. Michael Gerndt
Technische Univeristät München
gerndt@in.tum.de

Performance Analysis and Tuning is Essential

AUTOTUNE

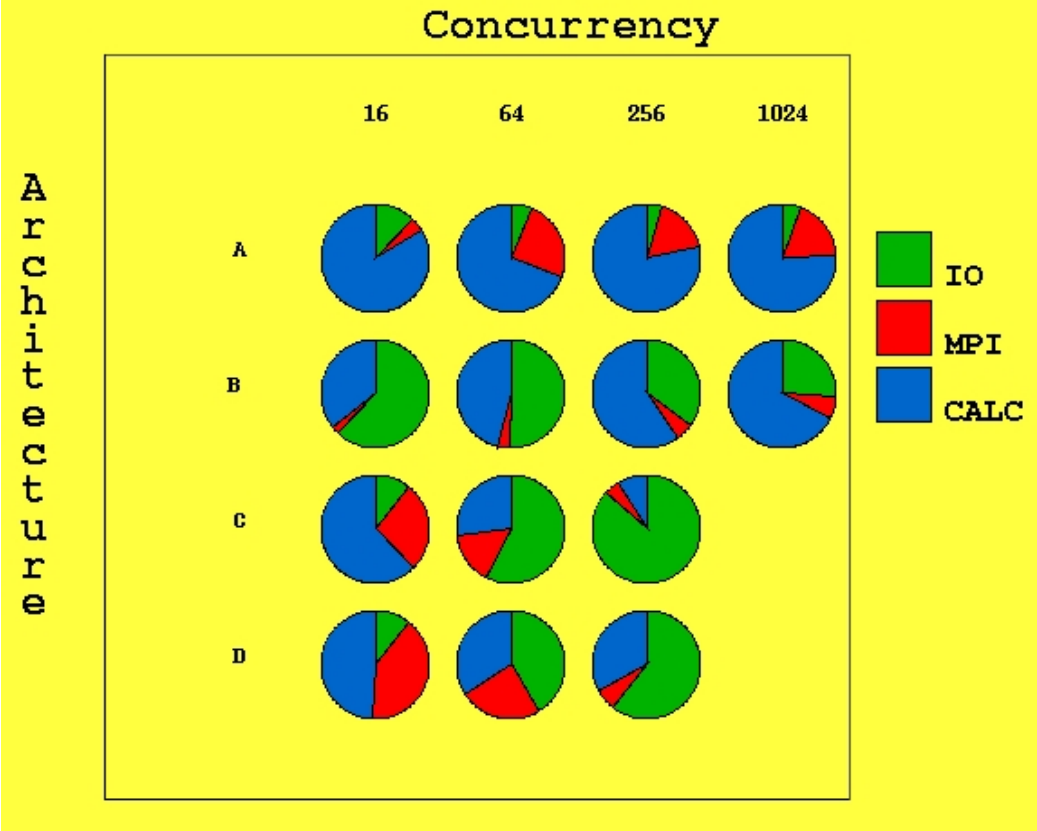
www.autotune-project.eu



Performance Analysis at Scale

www.autotune-project.eu

A high level description of the performance of cosmology code MADCAP.

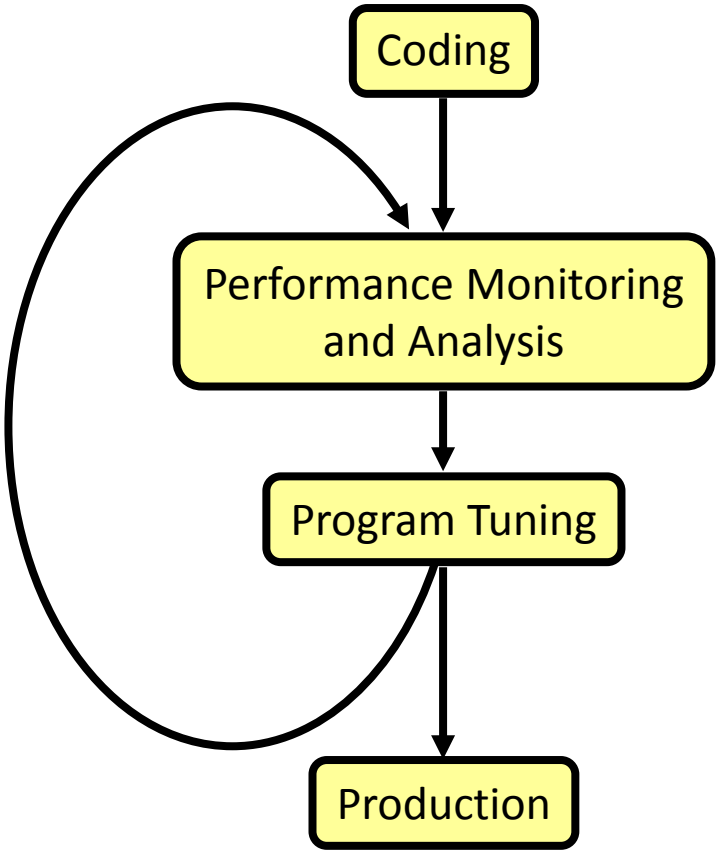


Source: David Skinner, NERSC

Performance Analysis for Parallel Systems

AUTOTUNE
www.autotune-project.eu

- Development cycle
  - Assumption: Reproducibility
- Instrumentation
  - Static vs Dynamic
  - Source-level vs binary-level
- Monitoring
  - Software vs Hardware
  - Statistical profiles vs event traces
- Analysis
  - Source-based tools
  - Visualization tools
  - Automatic analysis tools



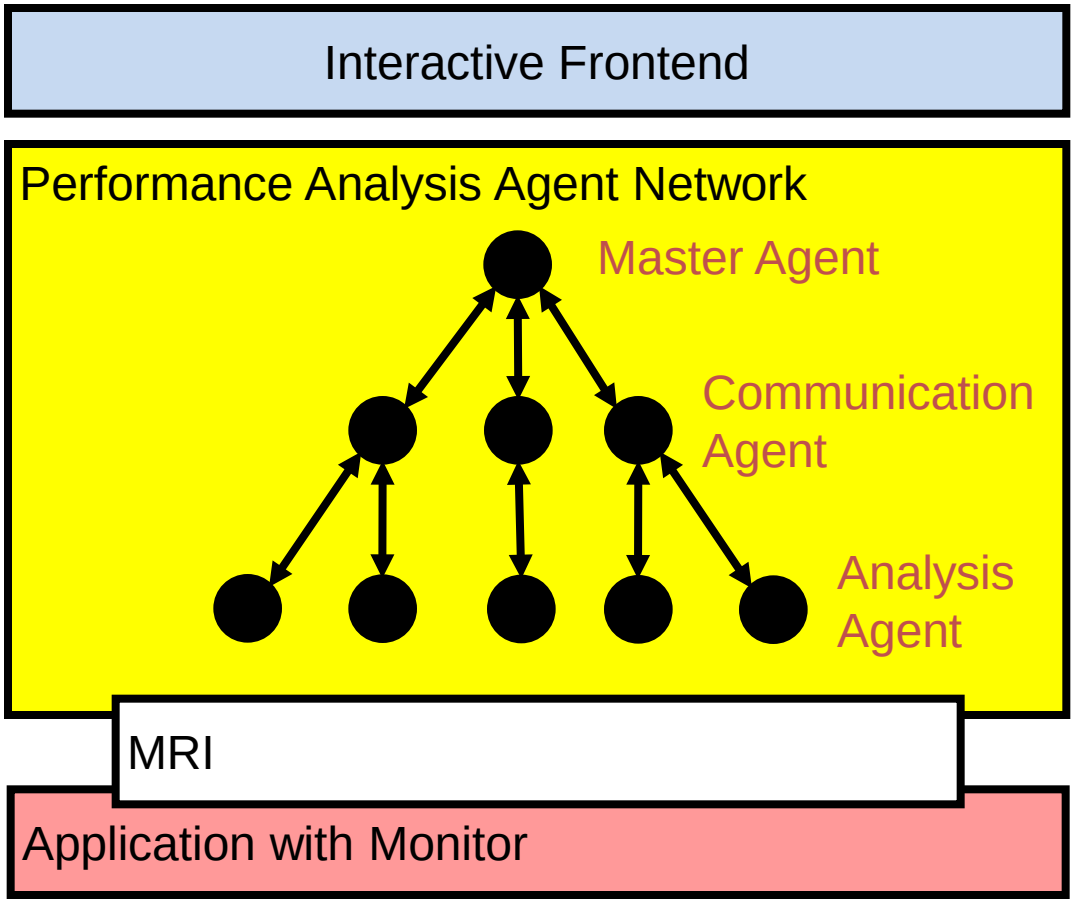
Periscope

- Automated search
  - Based on formalized performance properties
- Online analysis
  - Search performed while application is executing
- Distributed search
  - User specified number of analysis agents
  - Additional cores for agents
- Profile data only
  - even for MPI Waittime analysis

Properties

- StallCycles(Region, Rank, Thread, Metric, Phase)
  - Condition: Percentage of lost cycles >30%
  - Confidence: 1.0
  - Severity: Percentage of lost cycles
- StallCyclesIntegerLoads
  - Requires access to two counters
- L3MissesDominatingMemoryAccess
  - Condition: Importance of L3 misses (theoretical latencies)
  - Severity: Importance multiplied by actual stall cycles

Periscope Design



Agent Search Strategies

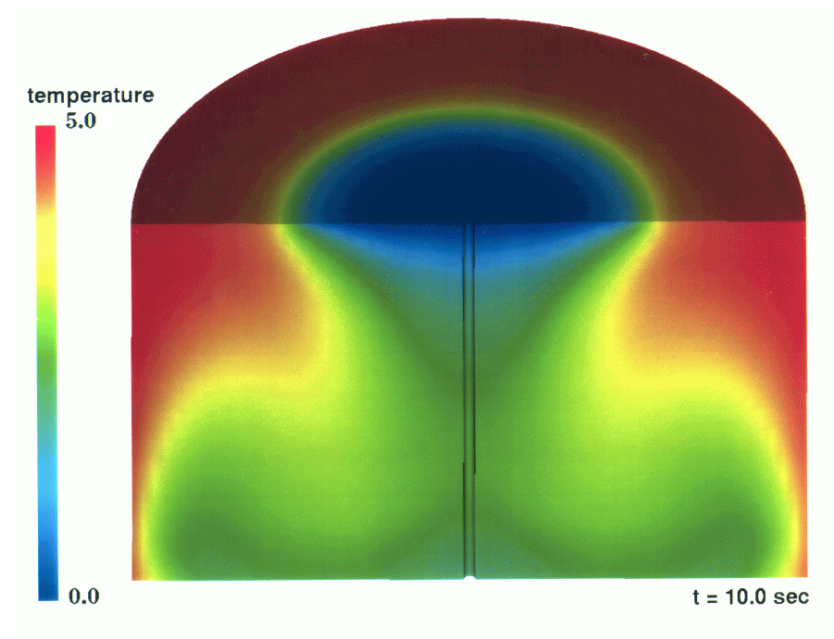
- Application phase is a period of program's execution
  - Phase regions
    - Full program
    - Single user region assumed to be repetitive
  - Phase boundaries have to be global (SPMD programs)
- Search strategies
  - Determine hypothesis refinement
    - Region nesting
    - Property hierarchy-based refinement

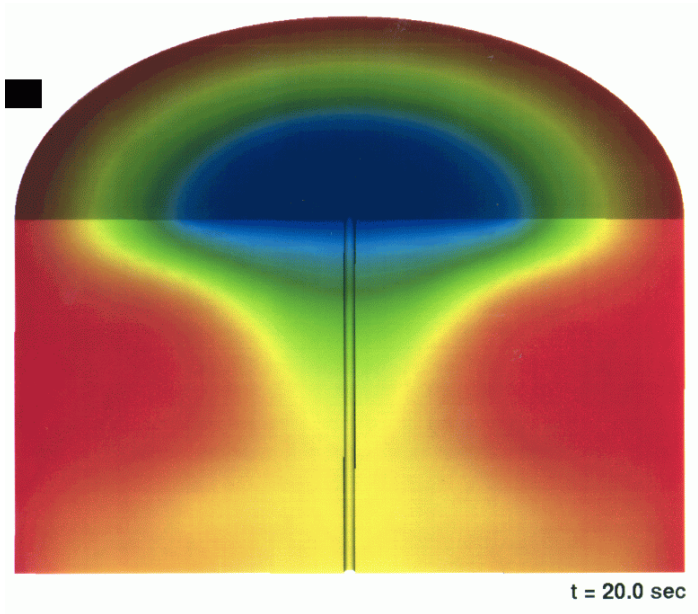
Agent Search Strategies

- Stall cycle analysis
- Stall cycle analysis with breadth first search
- MPI strategy
- OpenMP strategy
- OMP Scalability strategy
- Benchmarking strategy
- Default strategy

Crystal Growth Simulation

www.autotune-project.eu





Results

www.autotune-project.eu

```
USER_REGION; cx.f:61;    78.361;  IA64 Pipeline Stall Cycles
46.250;       Stalls due to L1D TLB misses ...
36.018;       L3 misses dominate memory access
30.294;       Stalls due to waiting for FP register

CALL_REGION; cx.f:70;  52.324; IA64 Pipeline Stall Cycles
29.468;       L3 misses dominate memory access
26.858;       Stalls due to L1D TLB misses ...
23.972;       Stalls due to waiting for FP register

velo;         cx.f:622; 52.316; IA64 Pipeline Stall Cycles
29.465;       L3 misses dominate memory access
26.854;       Stalls due to L1D TLB misses ...
23.972;       Stalls due to waiting for FP register

LOOP_REGION;  cx.f:731; 32.512; IA64 Pipeline Stall Cycles
24.098;       L3 misses dominate memory access
20.035;       Stalls due to waiting for FP register
```

Integration in Eclipse (PTP)

AutoTune

www.autotune-project.eu

Where is the problem?

Filter problems for region

What is the most severe problem?

lel_Jugene/velo.f - Eclipse

Navigate Search Project Run Window Help

properties256.psc | velo.f | crecvxs.f

```
334     *                 nh,600,
335     *                 MPI_COMM_WORLD, error)
336      endif
337
338      IF (NV.GE.0) then
339          call mpi_recv(wn(1,1,0),LX*LDIM2, MPI_DOUBLE_PRECISION,
340     *                  MPI_ANY_SOURCE,
341     *                  600, MPI_COMM_WORLD,status,error)
342      endif
343
344C     NACH RECHTS SENDEN(VN)
345
346      IF (NR.GE.0)
347     *  CALL CSENDXS(77,LDIM3,VN(1,LDIM2,1),8*LX,8*LX*(LDIM2+1),NR,0)
348      IF (NL.GE.0)
349     *  CALL CRECVXS(77,LDIM3,VN(1,0,1),8*LX,8*LX*(LDIM2+1),NL,0)
350
```

Project Explorer | SIR Outline View

- SIR File: /gpfs/automountdir/hom
  - subroutine: CRECVXS (0/4)
    - call: MPI_RECV (4/4)
  - subroutine: VELO (0/6)
    - call: MPI_RECV (4/4)
    - call: MPI_ALLREDUCE (2/2)

Clustering Results View | Periscope Properties View

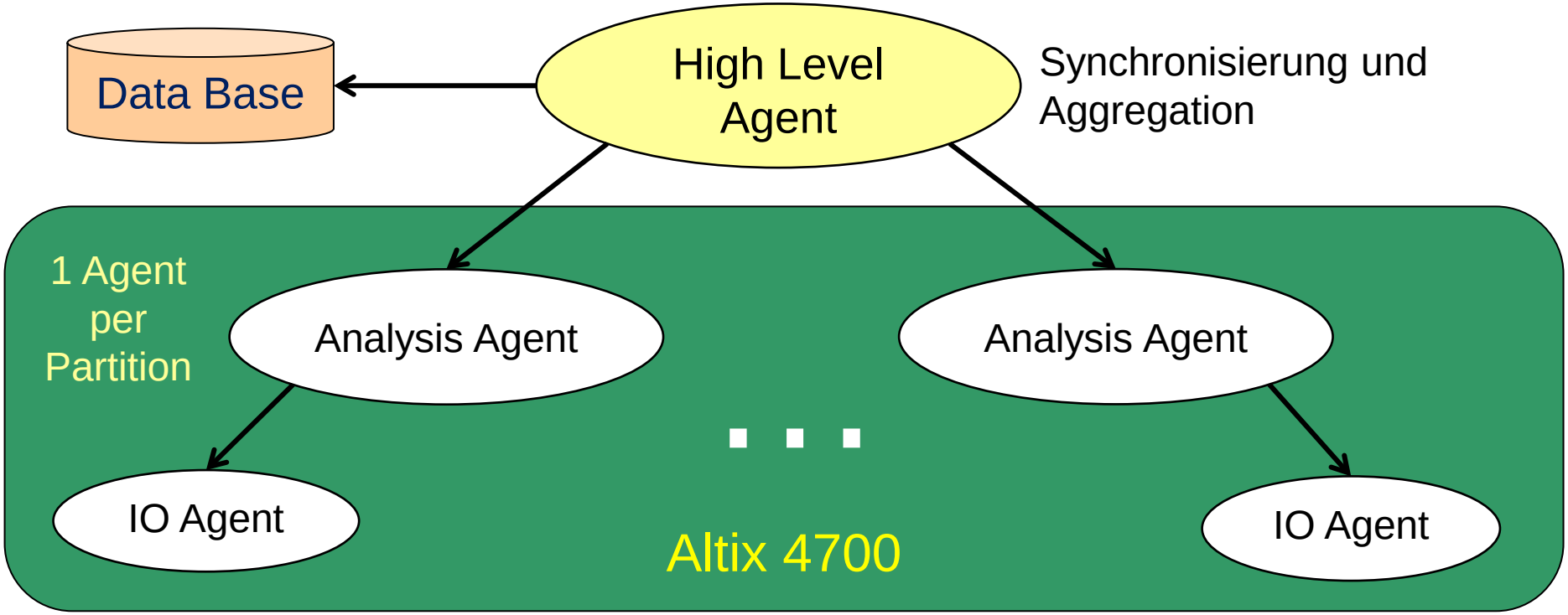
Name	Filename	RFL	Sever...	Region	Process
Excessive MPI time due to late process in allre			5,77	Types Group	
Excessive MPI time due to late process in a	velo.f	528	5,77	CALL_REGION	255
Excessive MPI time in receive due to late senc			34,81	Types Group	
Excessive MPI time in receive due to late se	crecvxs.f	12	27,24	CALL_REGION	15, 31, 47, 63, 79, 95, 111, 127, 143, 159, ...
Excessive MPI time in receive due to late se	velo.f	339	50,02	CALL_REGION	240, 241, 242, 243, 244, 245, 246, 247, 24...
Excessive MPI time in receive due to late se	velo.f	339	33,72	CALL_REGION	255
Excessive MPI time in receive due to late se	crecvxs.f	12	28,27	CALL_REGION	255
Excessive MPI communication time			29,09	Types Group	
Excessive MPI communication time	velo.f	339	50,05	CALL_REGION	240, 241, 242, 243, 244, 245, 246, 247, 24...
Excessive MPI communication time	crecvxs.f	12	28,45	CALL_REGION	255
Excessive MPI communication time	crecvxs.f	12	27,43	CALL_REGION	15, 31, 47, 63, 79, 95, 111, 127, 143, 159, ...
Excessive MPI communication time	velo.f	528	5,77	CALL_REGION	255
Excessive MPI communication time	velo.f	339	33,73	CALL_REGION	255

Filter: | Search: | RE | 0 Loaded - 13 Shown - 1 Selected -

PerSyst: Periscope System Monitoring

AUTOTUNE
www.autotune-project.eu

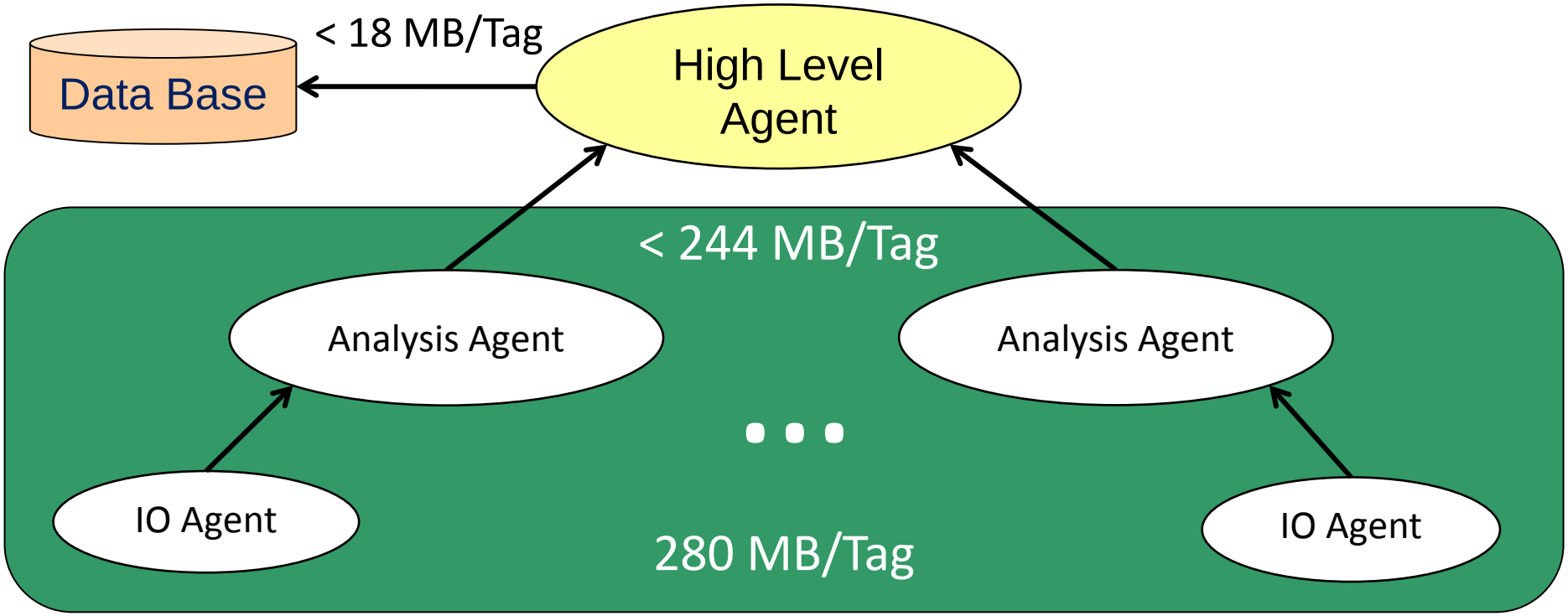
- Distributed fault tolerant architecture
- Incremental analysis



PerSyst: Data Reduction

www.autotune-project.eu

- Aggregation of data in properties
- Aggregation per job



AutoTune

AUTOTUNE

www.autotune-project.eu

- Automatic application tuning
  - Performance and energy
- Parallel architectures
  - HPC and parallel servers
  - Homogeneous and heterogeneous
  - Multicore and GPU accelerated systems
  - Reproducable execution capabilities
- Variety of parallel pardigms
  - MPI, HMPP, parallel patterns

Partners

AUTOTUNE

www.autotune-project.eu

- Technische Universität München
- Universität Wien
- CAPS Entreprises
- Universitat Autònoma de Barcelona
- Leibniz Computing Centre
- National University of Galaway, ICHEC

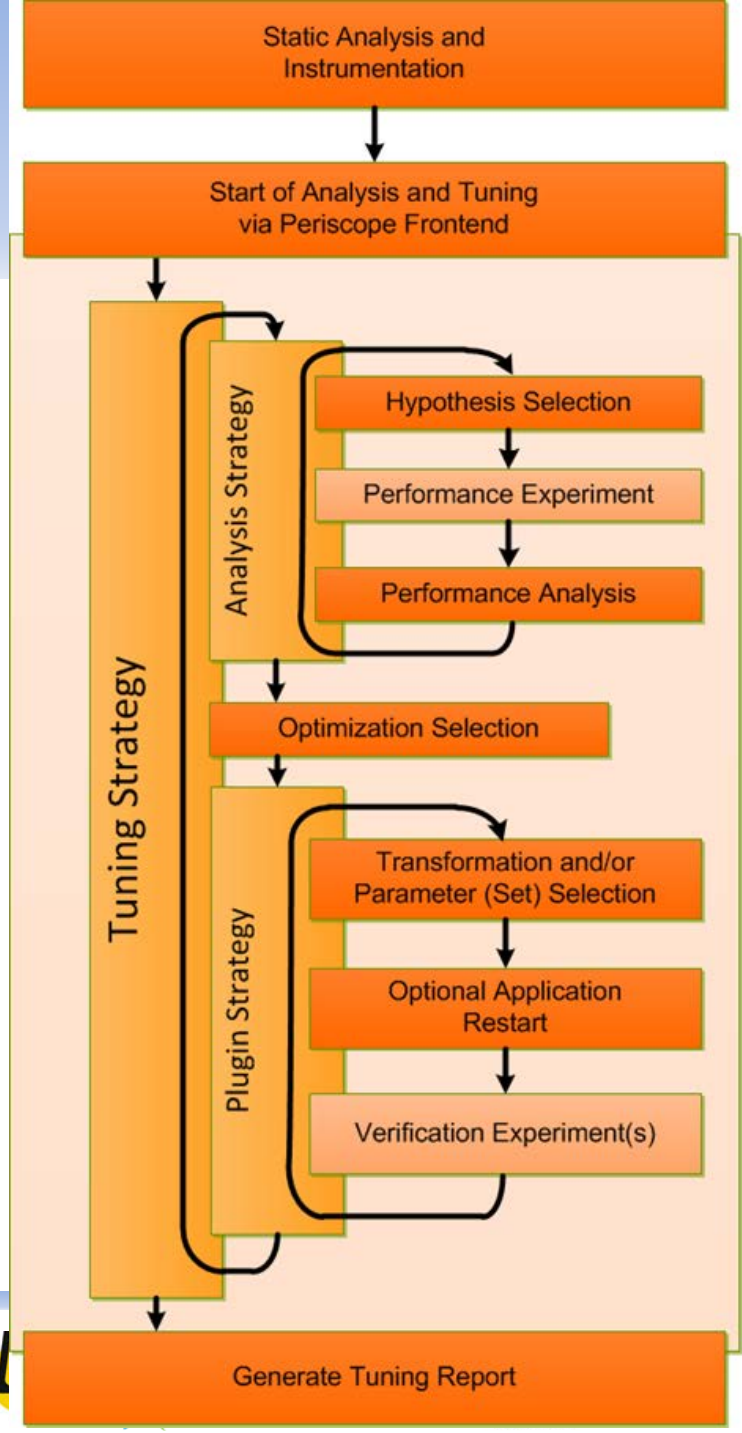
Autotune Approach

AUTOTUNE
www.autotune-project.eu

- Predefined tuning strategies combining performance analysis and tuning
- Plugins
  - Compiler based optimization
  - HMPP tuning for GPUs
  - Parallel pattern tuning
  - MPI tuning
  - Energy efficiency tuning

Periscope Tuning Framework

- Online
  - Analysis and evaluation of tuned version in single application run
  - Multiple versions in single step due to parallelism in application
- Result
  - Tuning recommendation
  - Adaptation of source code and /or execution environment
  - Impact on production runs



Autotuning Extension in HMPP

AUTOTUNE
www.autotune-project.eu

- Directives to provide optimization space to explore
  - Parameterized loop transformations
  - Alternative/specialized code declaration to specify various implementations

```
#pragma hmppcg(CUDA) unroll(RANGE), jam
 for( i = 0 ; i < n; i++ ) {
  for( j = 0 ; j < n; j++ ) {
    . . .
    VC(j,i) = alpha*prod+beta * VC(j,i);
  }
 }
```

- Runtime API
  - Optimization space description
  - Static code information collect
  - Dynamic information collect (i.e. timing, parameter values)

Energy Efficiency Plugin (LRZ)

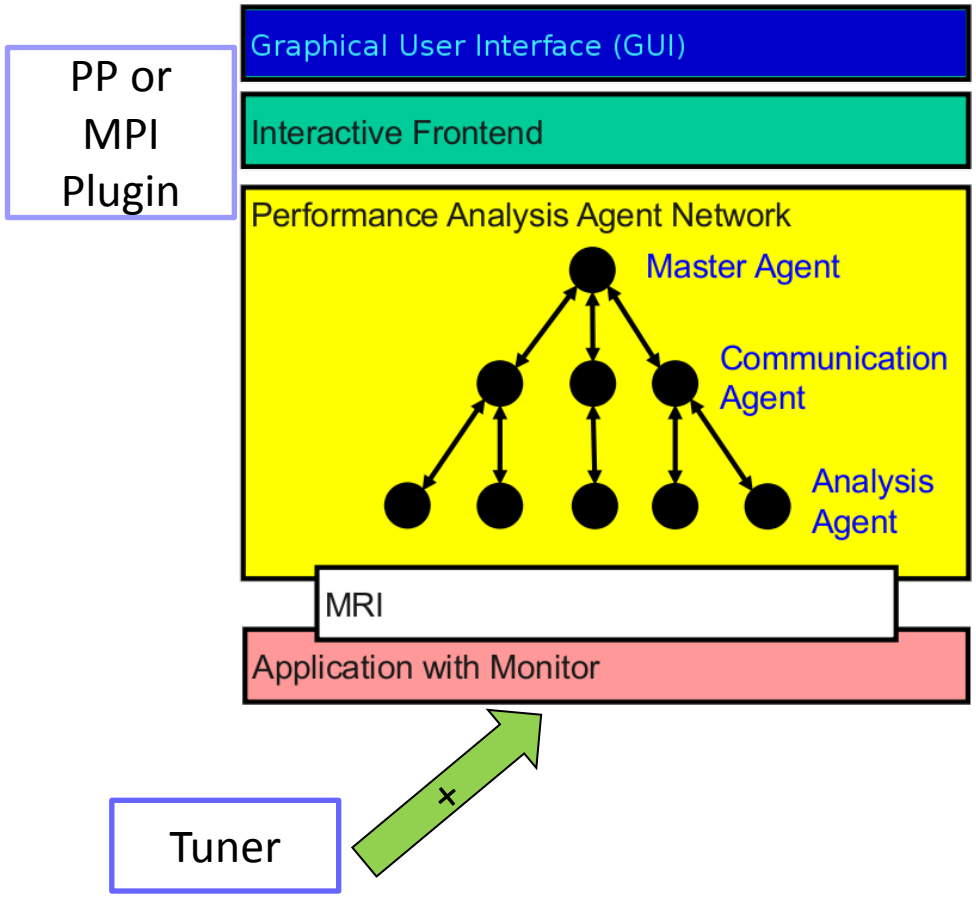
AUTOTUNE
www.autotune-project.eu

1. Preparation phase:
   - Selection of possible core frequencies
   - Selection of regions for code instrumentation
2. "While"-loop (until region refinement):
   "For"-loop (all frequencies) :
   a) Set new frequency of tuned region
   b) Periscope analysis (instrumentation, (re-)compiling , start and stop experiment)
   c) Measure execution time + energy of tuned region
   d) Evaluate experiment
   "End for"
3. Evaluate results of all experiments in refinement loop
4. Store best frequencies-region combination

Parallel Pattern/MPI Tuning

AUTOTUNE
www.autotune-project.eu

- PP or MPI Plugin encloses a performance model (ex. M/W) based on Periscope-provided performance data
- Automatically generated tuning decisions are sent to Tuner
- Tuner dynamically modifies the application before next experiment



Expected Impact

- Improved performance of applications
- Reduced power consumption of parallel systems
- Facilitated program development and porting
- Reduced time for application tuning
- Leadership of European performance tools groups
- Strengthened European HPC industry

THANK YOU