

# Generating Shifting Workloads to Benchmark Adaptability in Relational Database Systems

---

Tilmann Rabl, Andreas Lang, Thomas Hackl,  
Bernhard Sick, Harald Kosch

Faculty of Computer Science  
University of Passau



---

TPC Technology Conference on Performance Evaluation & Benchmarking  
August 24, 2009

# Introduction

---

Hot database research topics:

- Adaptability
- Automatic and autonomic tuning

*How can they be benchmarked?*

# Introduction

---

Hot database research topics:

- Adaptability
- Automatic and autonomic tuning

*How can they be benchmarked?*

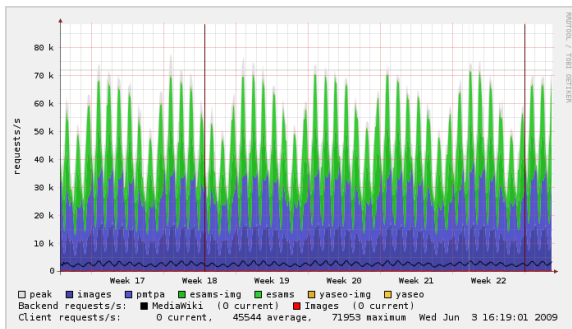
Today's solution

- Starting from scratch
- Changing the workload completely

# Motivation

*How does real-world workload look like?*

Homogeneous workload:

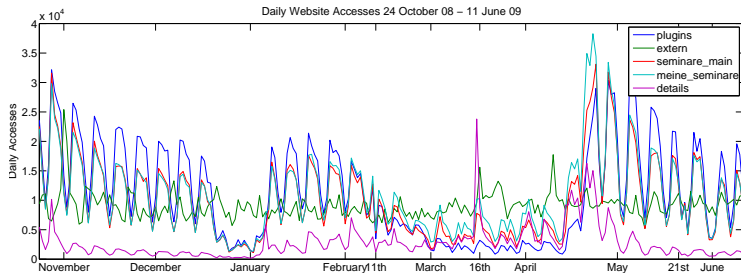


- Daily and weekly patterns

# Motivation

*How does real-world workload look like?*

Special purpose workload:



- Daily and weekly patterns
- Workload classes: working days, holidays, outliers
- Trends

# Benchmark Design Challenges

---

## Design Challenges

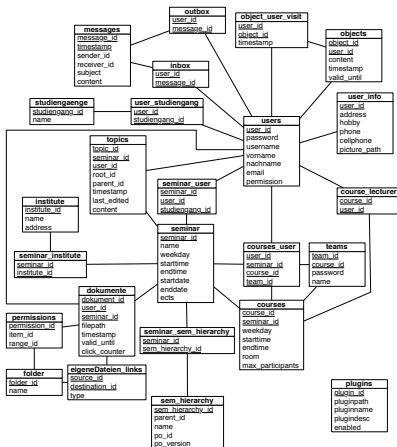
- Realistic data & workload patterns
- Scalable
- Configurable
- Random generators

## Basis

- Stud.IP eLearning management system
- Online application
- 1 year web server log data
- Complete database dump

# Benchmark Design

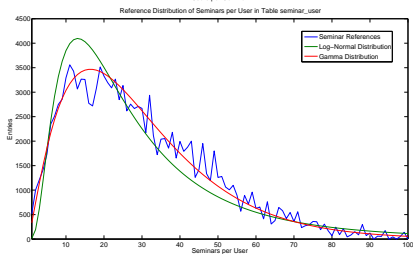
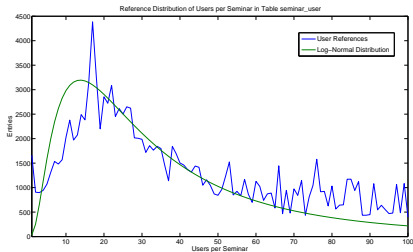
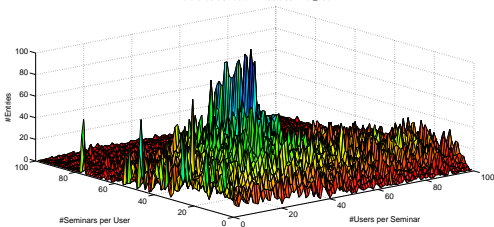
- 25 tables (original 200)
- 30 queries
- Only course management
- Most important:  
seminar - seminar\_user - user



# Data Generation

- Based on original data
- Maximum likelihood estimation
- Standard probability distributions
- Scalable

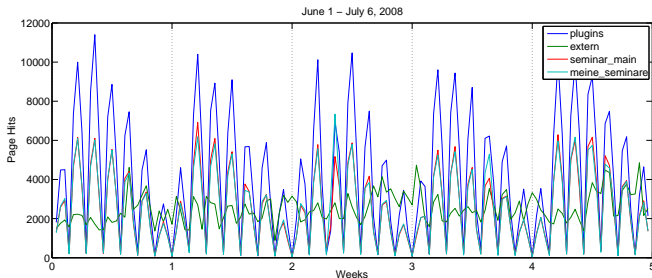
Reference distribution in table seminar\_user





# Workload Modeling

- Workload interpreted as time series
- Classes of queries (*meine\_seminare*)
- Classes of days (*Monday during lecture period*)
- Daily approximation by polynomials



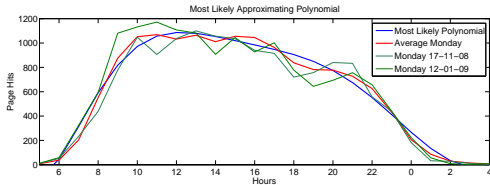
# Polynomial Approximation

Optimal approximating polynomial of degree  $K$ :

$$p_a(x) = \sum_{k=0}^K a_k p_k(x)$$

Properties of  $p_k$ :

- different ascending degrees
- leading coefficient one
- pair wise orthogonal (inner product)



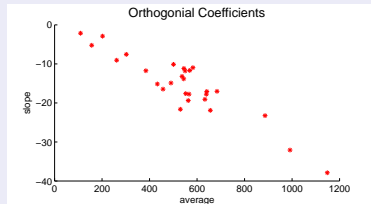
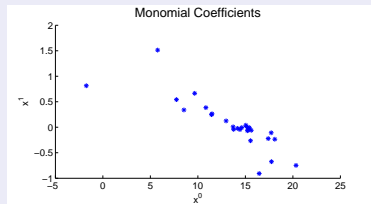
# Workload Generation

## Unchanging Behavior

Evaluate polynomial at certain points in time

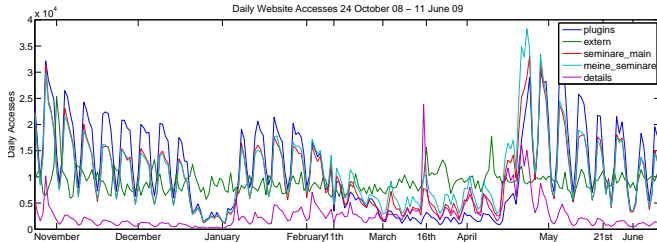
## Random Generation

- Time series as point in *shape space*
- Weighting factors  $a_k$  are normally distributed
- Multivariate Gaussian to generate weighting vector  $\mathbf{a}$



# Benchmarking Objectives

- Basic Performance
  - Shifting workloads
  - Measure peak performance
- Adaptivity
  - Different daily patterns
  - Measure adaptability
- Robustness
  - Introduction of outliers
  - Measure over adaptation
- Energy and Space Efficiency
  - Shifting workloads
  - Measure energy / space efficiency



# Conclusion

---

## Conclusion

- Online eLearning management benchmark
- New generator model for query workloads
- More realistic benchmarking for automatic / autonomic tuning

# Conclusion

---

## Conclusion

- Online eLearning management benchmark
- New generator model for query workloads
- More realistic benchmarking for automatic / autonomic tuning

## Future Work

- Tune and test
- Apply techniques to standard benchmarks
- Trends in workloads
- Schema evolution

# Questions?

---

Thank you.