

# SPEC Research Group

Sam Kounev

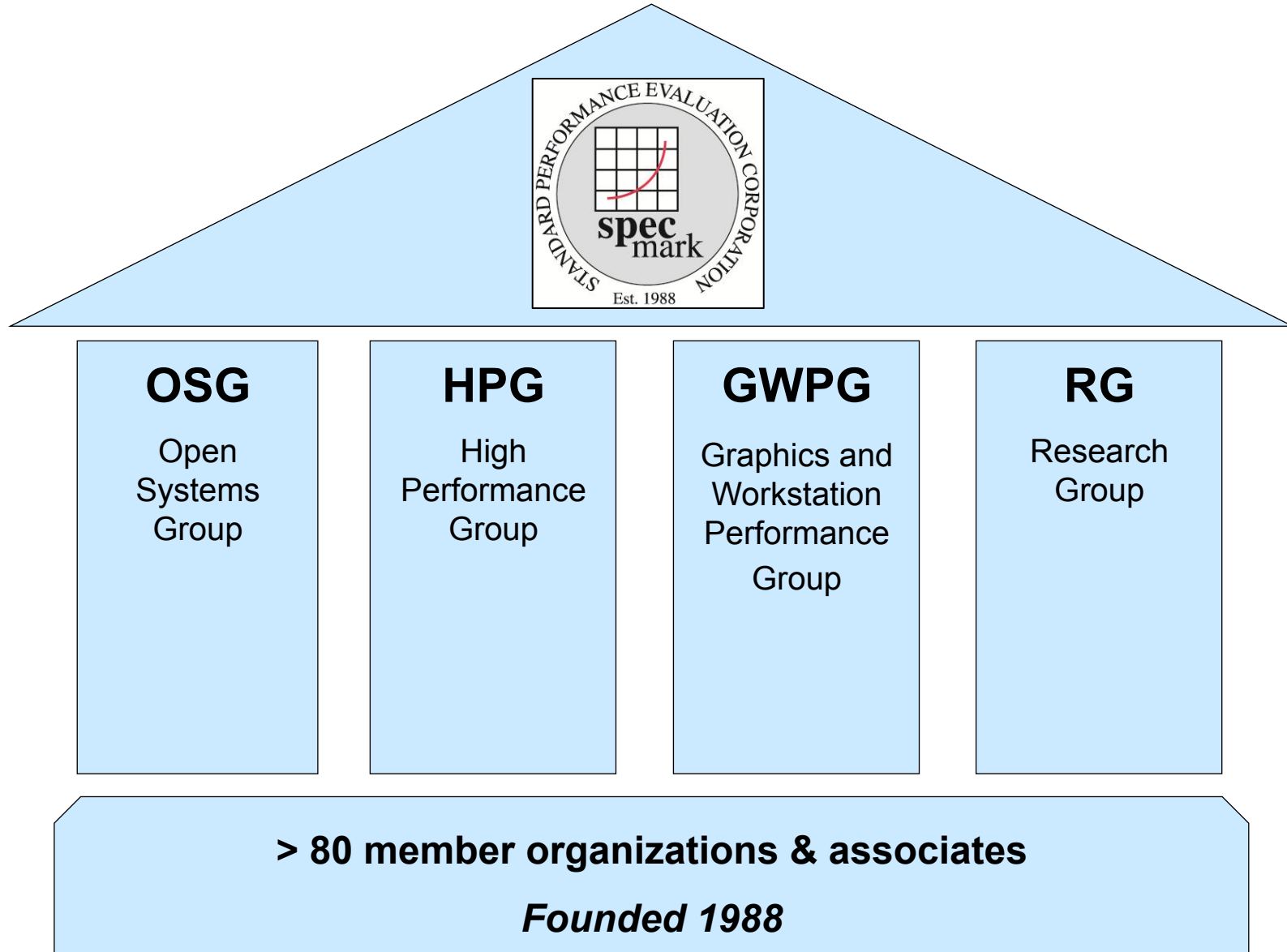
SPEC 2015 Annual Meeting

Austin, TX, February 5, 2015





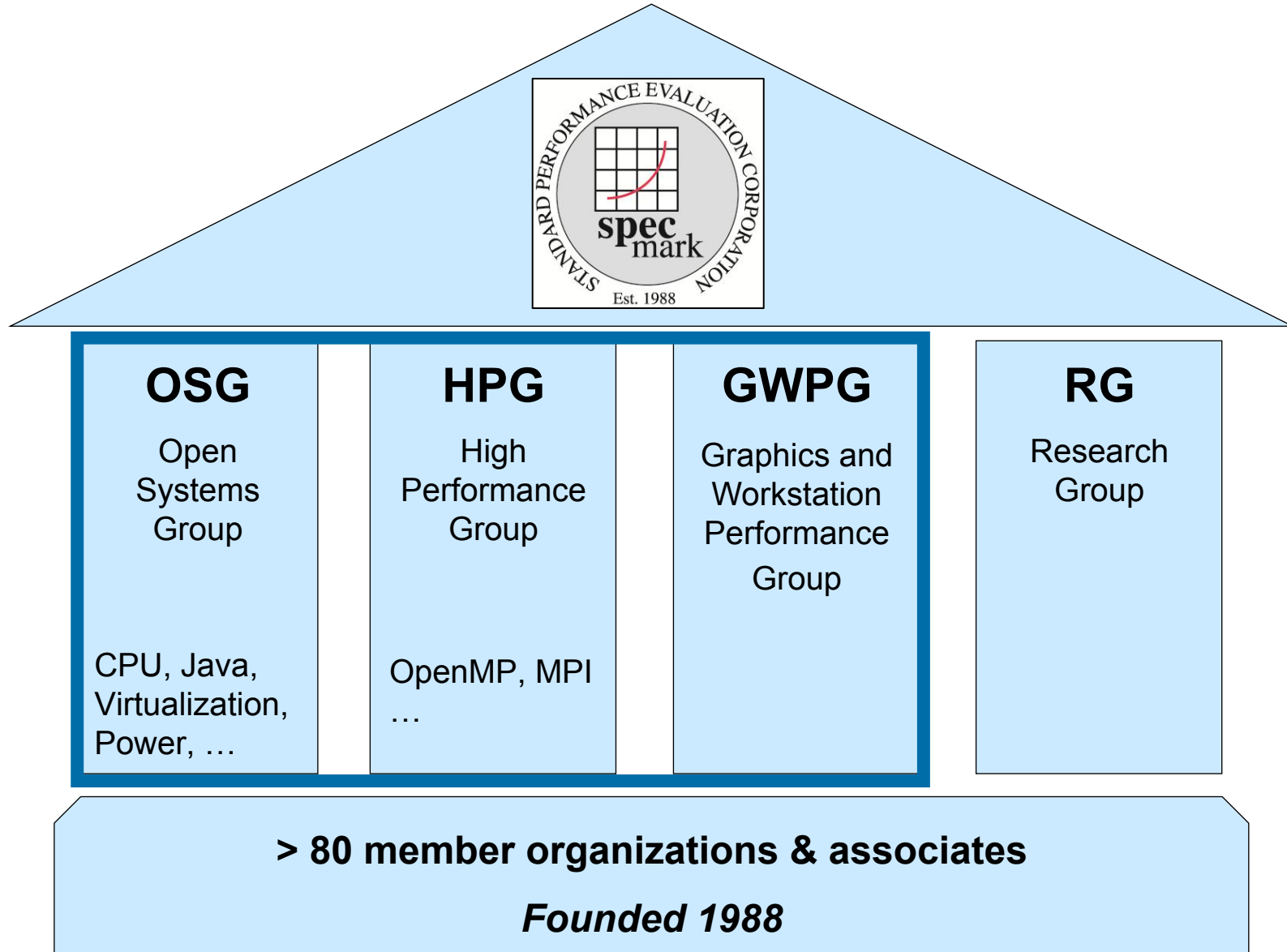
# Standard Performance Evaluation Corporation





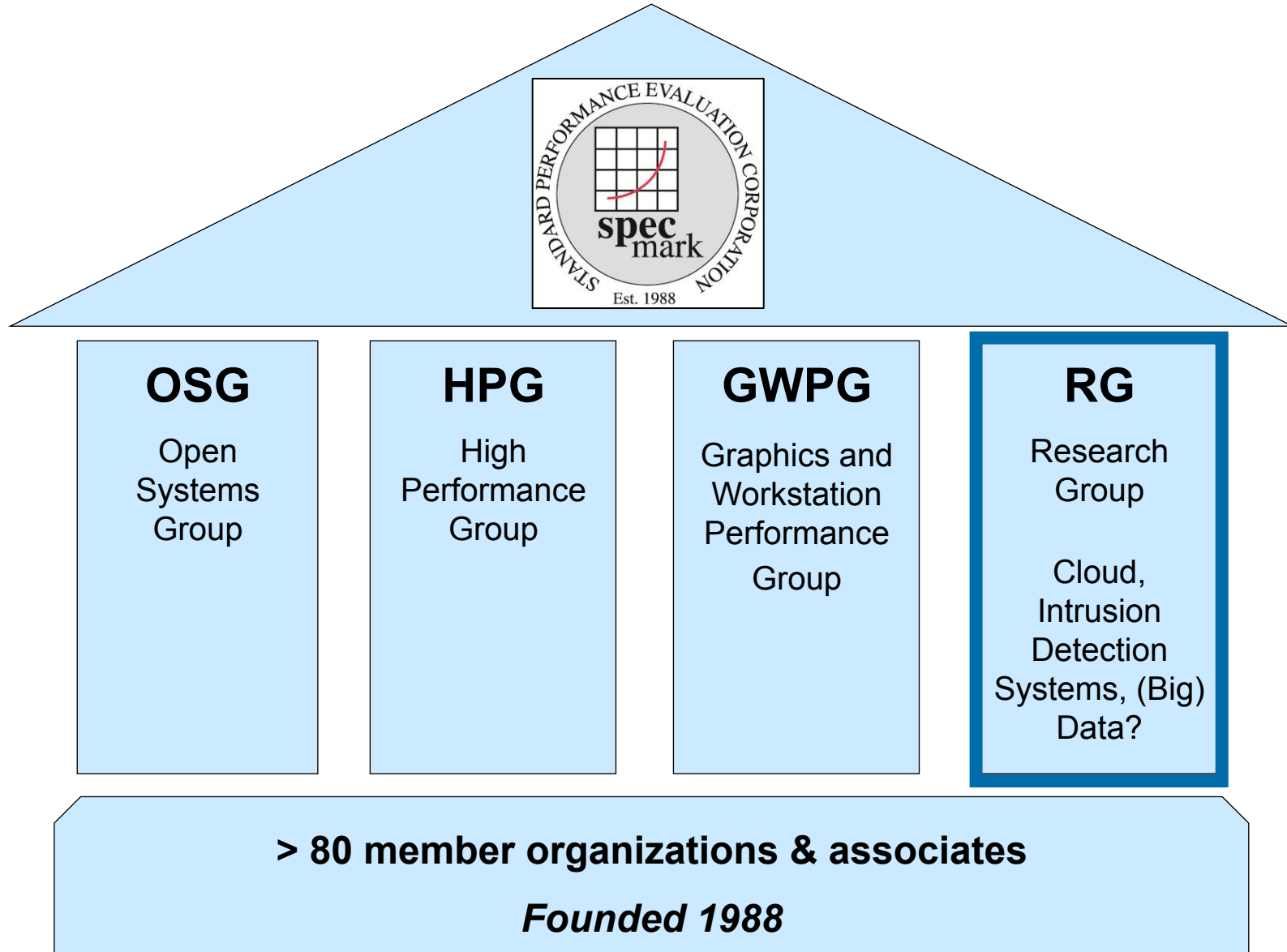
# Standard Performance Evaluation Corporation

## Development of Industry Standard Benchmarks



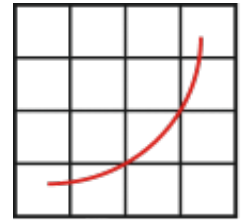


# Standard Performance Evaluation Corporation Platform for Collaborative Research



# SPEC Research Group (RG)

*The Research Group of the  
Standard Performance Evaluation Corporation*



**spec**<sup>®</sup>

## Mission Statement



- ▶ Provide a **platform for collaborative research efforts** in the area of quantitative system evaluation and analysis
- ▶ Foster interactions and **collaborations** between industry and academia
- ▶ **Scope:** computer benchmarking, performance evaluation, and experimental system analysis in general
- ▶ **Focus on** standard scenarios, metrics, benchmarks, analysis methodologies and tools

Find more information on: <http://research.spec.org>

# Members (Dec 2014)



# Scope

- Standard scenarios, metrics, and (research) benchmarks
- Methodologies, techniques and tools for quantitative analysis
  - Measurement, load testing, profiling, workload characterization, dependability & efficiency evaluation of computing systems, etc.
- **Performance in a broad sense**
  - **Classical performance metrics**  
Response time, throughput, scalability and efficiency including energy efficiency, etc.
  - **Non-functional system properties** *under the term dependability* (availability, reliability, and security)

# Selected SPEC RG Activities

## Working Groups

**Cloud Computing** (Chair: Alexandru Iosup, TU Delft)

**Intrusion Detection Systems** (Chair: Marco Vieira, Uni-Coimbra)

**BigData Benchmarking** (Chair: Tilmann Rabl, bankmark)

**DevOps Performance** (Chair: André van Hoorn, Uni-Stuttgart)

Repository of peer-reviewed tools, experimental data & traces

Maintain a portal for all kinds of performance-related resources

Organization of conferences and workshops





# Selected Peer – Reviewed Tools



**Faban** – automates the running of multi-tier server performance tests/workloads

**Kieker** – appl. performance monitoring & dynamic software analysis



Benchmarking of event processing systems

**DiSL** – Java bytecode instrumentation



Performance evaluation of storage systems

# Example Tool: LIMBO

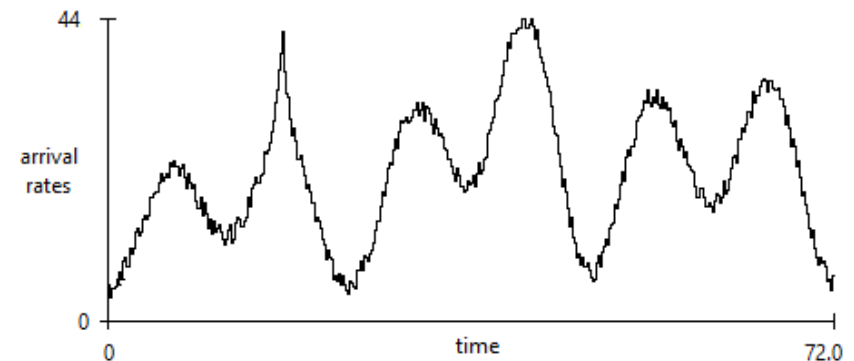
Problem: How to capture the load intensity variations (e.g. requests per sec) in a compact mathematical model?

## Load Intensity Modeling Tool

Automated model extraction from recorded traces

Creation and composition of custom models

Emulation of job arrivals for load generation



<http://descartes.tools/limbo>

# Research Benchmark

**Research Benchmarks are not** intended for direct comparison and marketing of existing products

Provide a basis for in-depth quantitative analysis and evaluation

Representative application scenarios and workloads

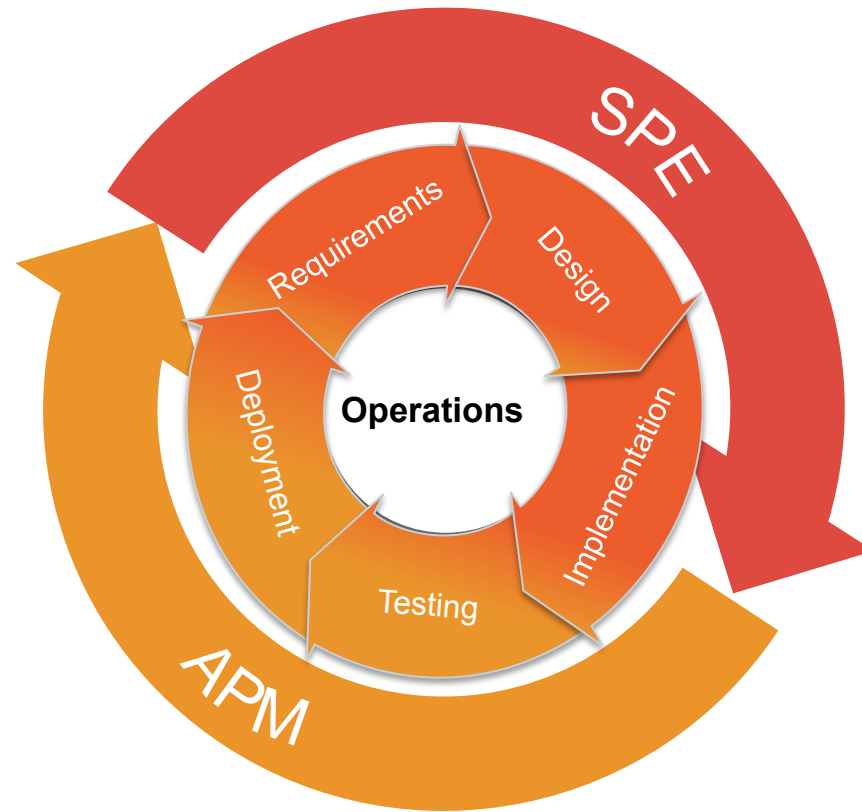
Flexible and customizable to different usage scenarios

Provide a range of possible metrics

# Highlights from SPEC RG's 4<sup>th</sup> Year

- Roughly **45 member organizations**
- Two new **working groups** (RG BigData and RG DevOps)
- Technical co-sponsorship of QEST 2014 and ICAC 2015
- Keynote presentation at QEST 2014
- Two new tools **DynamicSpotter** and **LIMBO** accepted
- A **technical report** on hypercall vulnerabilities published
- **ISSRE paper** nominated for Best Paper Award
- RG Newsletter now established as a regular publication

# RG DevOps Performance WG



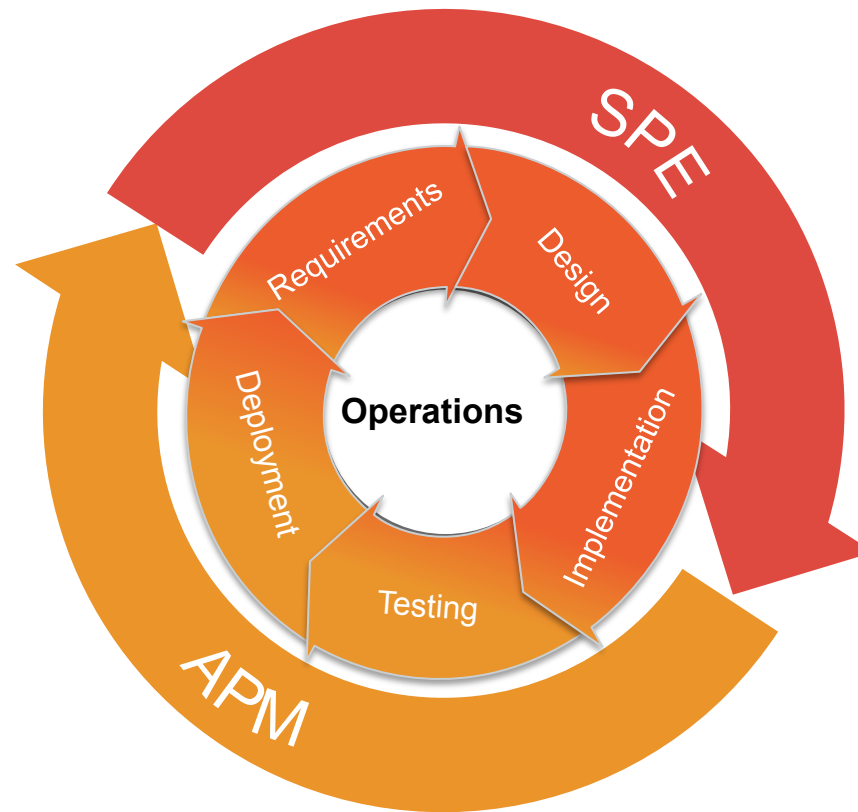
## ■ Contact

- Chair: André van Hoorn (University of Stuttgart)
- Vice chair: Andreas Brunnert (fortiss GmbH)

## ■ Web site

- <http://research.spec.org/devopswg/>

# RG DevOps Performance WG



**Mission:** foster and facilitate research in combining **measurement-based** application performance management (APM) and **model-based** software performance engineering (SPE) activities for **business-critical application systems**.

<http://research.spec.org/devopswg/>

# RG DevOps Performance WG

- Current member organizations

The logo for fortiss, featuring the word "fortiss" in a blue, lowercase, sans-serif font.The logo for KIT (Karlsruher Institut für Technologie), featuring a green fan-like graphic to the left of the letters "KIT" in a bold, black, sans-serif font, with the full name "Karlsruher Institut für Technologie" in a smaller font below.The logo for NOVATEC, featuring a blue graphic of three slanted parallel lines to the left of the word "NOVATEC" in a bold, blue, sans-serif font.The logo for Julius-Maximilians-Universität Würzburg, featuring a blue square with a white border containing the text "UNIVERSITÄT WÜRZBURG" in blue, with "Julius-Maximilians-" in a smaller font above.The logo for the University of Stuttgart, featuring a circular pattern of dots to the left of the text "University of Stuttgart" in a bold, black, sans-serif font, with "Germany" in a smaller font below.

University of Stuttgart  
Germany

The logo for Christian-Albrechts-Universität zu Kiel, featuring a purple box with the letters "C | A | U" in white, and a white box below containing the text "Christian-Albrechts-Universität zu Kiel" in a black, sans-serif font.

- Contact

- Chair: André van Hoorn (University of Stuttgart)
  - Vice chair: Andreas Brunnert (fortiss GmbH)
  - Secretary: Felix Willnecker (fortiss GmbH)
  - Release manager: Nils Ehmke (Kiel University)
- 
- Bi-weekly calls, currently @ Fri - 4:00-5:00 (Berlin/CET)
  - The next F2F meeting in February 2015 (Würzburg, Germany)

# RG DevOps: Topics of Interest

- Performance and workload model extraction
  - Dynamic/adaptive instrumentation techniques
  - Combination of static and dynamic analysis
- Feedback between development (Dev) and operations (Ops)
  - Interchange formats for metrics and models
  - Process models to better integrate SPE and APM activities
- Runtime performance management techniques
  - Automatic problem detection, diagnosis, and resolution
- Load testing
- Benchmarks for SPE and APM methods, techniques, and tools



# RG DevOps: Selected Activities

- **Report on DevOps performance (10+ authors)**
- **Using industry common APM solutions for model generation**  
(cooperating with dynaTrace) – see Poster @ ICPE 2015
- **Expert-guided automatic diagnosis of performance problems in enterprise applications**
- **Landscaping model extraction approaches**



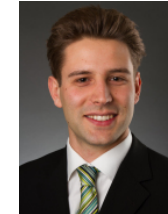
# RG Cloud WG



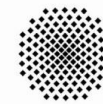
Alexandru Iosup  
Chair



Aleks Milenkoski  
Secretary



Nikolas Herbst  
Release Manager



<http://research.spec.org/working-groups/rg-cloud-working-group.html>

# RG Cloud WG: Vision

- “a broad approach, **relevant for both academia and industry**, to cloud benchmarking, quantitative evaluation, and experimental analysis.”
- “To develop new **methodological elements** for gaining deeper understanding not only of **cloud performance**, but also of **cloud operation and behavior**”
- “... through diverse quantitative evaluation tools, including **benchmarks, metrics, and workload generators.**”

# RG Cloud WG: Topics of Interest

- New methodological elements
  - Online workload generation
  - Workload characterization and modeling
  - Techniques for performance measurement and evaluation, etc.
- New quantitative evaluation tools
  - Benchmarks, Metrics, Workload generators
- Targeted topics, usually triggered by a question
  - How to present cloud usage scenarios in a uniform, formal way?
  - How to measure elasticity? (And other metric-related questions)
  - How to benchmark a PaaS cloud (for specific application domains)?

# A Finished Product

Textual and visual formalism for describing cloud usage scenarios

Value chains, value chains with mediators, hybrid service provisioning, ...

A. Milenkoski, A. Iosup, S. Kounev, K. Sachs, P. Rygielski, J. Ding, W. Cirne & F. Rosenberg.  
Cloud Usage Patterns: A Formalism for Description of Cloud Usage Scenarios.  
Technical Report SPEC-RG-2013-001 v.1.0.1, SPEC Research Group - Cloud Working  
Group, Standard Performance Evaluation Corporation (SPEC), April 2013.



SPEC RG Cloud Working Group  
<http://research.spec.org/working-groups>



# Defining Elasticity

Def: The degree to which a system is able to **adapt** to **workload changes** by **provisioning and deprovisioning** resources in an **autonomic manner**, such that at each point in time the **available resources match** the **current demand** as closely as possible.

*N. Herbst, S. Kounev and R. Reussner*

***Elasticity in Cloud Computing: What it is, and What it is Not.***

*in Proceedings of the 10th International Conference on Autonomic Computing (ICAC 2013), San Jose, CA, June 24-28, 2013.*

[ [slides](#) | [http](#) | [.pdf](#) ]

*[http://en.wikipedia.org/wiki/Elasticity\\_\(cloud\\_computing\)](http://en.wikipedia.org/wiki/Elasticity_(cloud_computing))*

# Work-In-Progress: BUNGEE

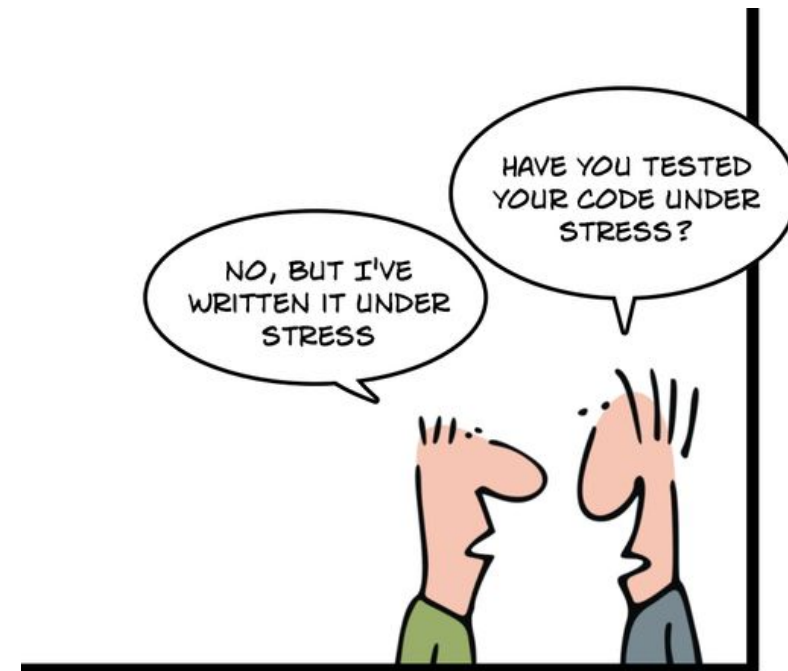
Framework for benchmarking elasticity

Current focus: IaaS cloud platforms

Planned collaboration with OSG Cloud



<http://descartes.tools/bungee>



# RG Cloud: Why and How to Join?

<http://research.spec.org/working-groups/rg-cloud-working-group.html>

## Why join?

Active, growing group for impactful work on cloud computing

Participate in the future of standardization work at SPEC

Discuss how performance can be measured and engineered

Find out about novel methods and current trends

Get in contact with leading organizations in the field

Find potential employees and/or interns

Find more information on: <http://research.spec.org>



## How to join?

Join SPEC (low-cost yearly membership fee, once per organization)

Details: <http://research.spec.org/faq.html#MembershipAndMeetings>



# RG Big Data Working Group

- Differentiating from traditional large database applications, e.g., OLTP, OLAP
- Clear goals for the aspects important to measure
- Sound rules and metrics to measure performance
- Tools to help instantiate sample big data systems

# RG Big Data Working Group

- Big Data Benchmark Community (BDBC)
  - Original founders: Chaitan Baru (SDSC), Tilmann Rabl (University of Toronto), Milind Bhandarkar (Pivotal/Greenplum), Raghu Nambiar (Cisco), Meikel Poess (Oracle)
  - [clds.ucsd.edu/bdbc/community](http://clds.ucsd.edu/bdbc/community)
- Refining ideas from BigBench for a **SPEC BigData benchmark**
- **Published papers** in SIGMOD, TPCTC, Big Data Journal
- **5 WBDB workshops** on Big Data Benchmarking (2012-2014)
- 2 Springer Publications

# RG BigData: Targets for 2015

## Organizing two workshops

6<sup>th</sup> WBDB, June 16-17 2015, Toronto, CA

7<sup>th</sup> WBDB, September 10-11, 2015 Delhi, India

## Publications/ Conference Participation

Present SPEC WG on Big Data, XLDB

Position paper, targeted for SIGMOD Record

Big Data Survey style paper including SPEC RG vision on Big Data, targeted for CACM

## Refining ideas introduced in BigBench for a SPEC Big Data benchmark

# RG IDS Working Group

Mission statement:

*“dissemination of techniques and tools for quantitative evaluation of host and network intrusion detection systems, with a focus on intrusion detection systems in dynamic virtualized environments”*

Topics of interest:

- Workload generation using attack injection
- IDS evaluation metrics

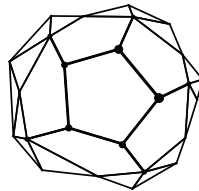


• U



C

• SIEMENS



nebula®

# Attack Injection



# RG IDS: Example Results

Aleksandar M., Bryan D. P., Nuno A., Marco V., and Samuel K. Experience Report: An Analysis of Hypercall Handler Vulnerabilities @ The 25th IEEE International Symposium on Software Reliability Engineering (ISSRE 2014) - Research Track, Italy, 2014.



Aleksandar M., Bryan D. P., Nuno A., Marco V., and Samuel K. hInjector: Injecting Hypercall Attacks for Evaluating VMI-based Intrusion Detection Systems (Poster) @ The 2013 Annual Computer Security Applications Conference (ACSAC 2013), USA, 2013.

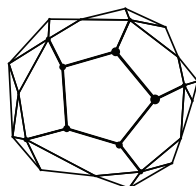


**hInjector available @ <https://github.com/hinj/hInjector>**

Aleksandar M., Marco V., Samuel K., Alberto A., Bryan D. P. Evaluating Computer Intrusion Detection Systems: A Survey of Common Practices @ *ACM Computing Surveys*. Under Revision.



Collaboration with industry

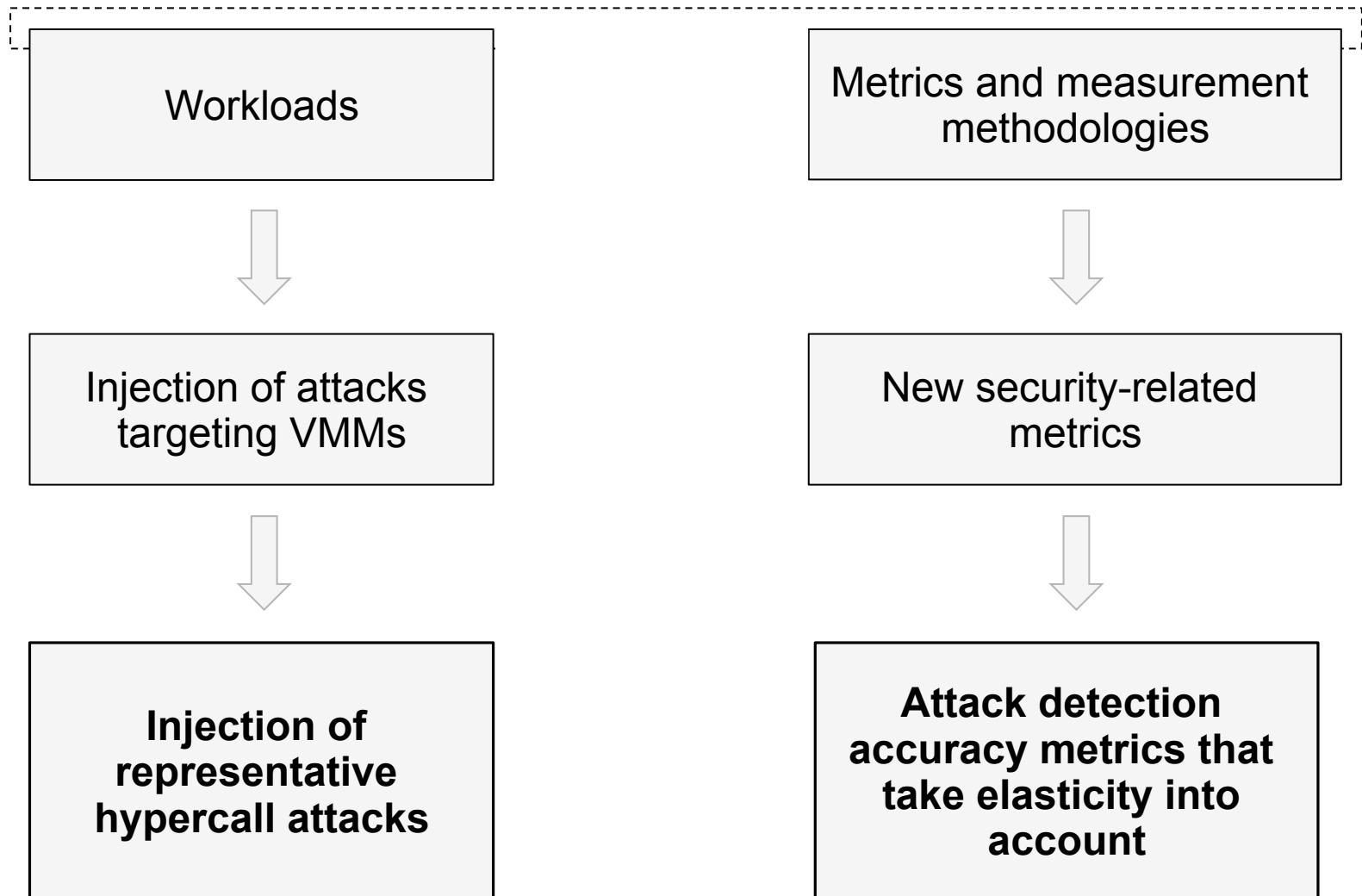


nebula®



# hInjector Tool

## IDS evaluation in virtualized environments



# Collaboration with SPEC OSG Power

- University of Würzburg (Jóakim v. Kistowski, Sam Kounev)
  - Accelerator development
  - Workload balancing for multi-node systems prototyping
  - Calibrating workloads by core loading
  - Plans to provide a SERT-based tool for broader testing
- ICPE 2015
  - Industry paper: “Analysis of the influences on server power consumption and energy efficiency for CPU-intensive workloads”
  - Tutorial on “How to build a benchmark” in collaboration with Power, CPU & TPC
- Paper planning for MASCOT 2015 & ICPE 2016, Journal paper



# Collaboration with SPEC OSG Power

- University of Würzburg (Jóakim v. Kistowski, Sam Kounev)
  - Accelerator development
  - Workload balancing for multi-node systems prototyping
  - Calibrating workloads by core loading, so integrating load balancing into the core architecture of worklet design
- Plans to provide a SERT-based tool for broader testing
- ICPE 2015 paper: Analysis of the Influences on Server Power Consumption and Energy Efficiency for CPU-Intensive Workloads
- Tutorial on “How to Build a Benchmark” at ICPE 2015 in collaboration Power, CPU & TPC
- Paper planning for MASCOT 2015 & ICPE 2016, Journal paper
- Will jointly add credibility to the SPEC power work and to the interactions with academia
- ON HOLD: the ideas around Jóakim and his team managing the open sourcing of worklets (based on the Chauffeur WDK that SPEC is now shipping) as possible future candidates for the SERT or for SPECpower vNext
  - o Everyone agrees it’s a great idea but the effort required may be too much for Jóakim and his team to allocate

# Example Collaborations with OSG

**University of Alberta** is a supporting contributor to SPEC OSG

- Developed a set of inputs to selected benchmarks to be released soon after the release of SPEC CPU v6

## Additional Contributions

- Sanity check on procedures to generate inputs
- Characterization of benchmark behavior variations
- Suggestions to improve a few benchmarks

# Benefits for SPEC

- **Involvement** of researchers and students in SPEC
- **Exchange** of ideas and experiences
- **Feedback** from an academic perspective
- Participation in **benchmark development**
  
- **Visibility** and **credibility** in the academic community
- Attraction of **new members** and **benchmark projects**

# Questions?

<http://research.spec.org/>