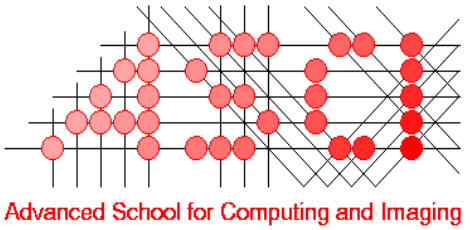


25 Years of Parallel and Distributed Computing Research in ASCI

Henri Bal

Vrije Universiteit Amsterdam

High Performance Distributed Computing

https://www.vuhpdc.net/

Introduction

- Overview of my research
- Importance of
  - ASCI
  - Infrastructure
  - Computer Science collaborations
  - Applications (like Imaging)

VU DAS-1
cluster (1997)

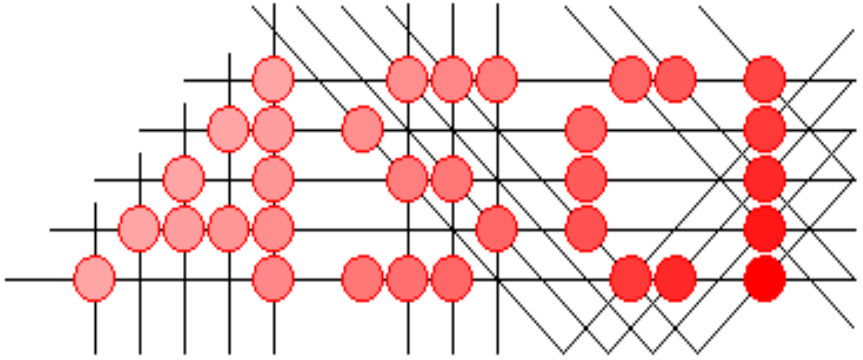
Topics

- ASCI
- DAS (Distributed ASCI Supercomputer)
- Albatross: wide-area algorithms
- Ibis: distributed computing
- MCL: Many-core (GPU) programming
- Swan: Smartphone computing & IoT
- Clownfish: edge/cloud computing

Research schools

- **Dutch initiative from 1990s**
- **Aims:**
  - **Stimulate top research & collaboration**
  - **Provide PhD education (courses)**
- **Dutch CS landscape became organized into:**
  - **Programming and Algorithmics (IPA)**
  - **Information and Knowledge Systems (SIKS)**
  - **Computing and Imaging (ASCI)**

ASCI: Advanced School for Computing and Imaging (1995)



Andy Tanenbaum

Arnold Smeulders

Henri Bal

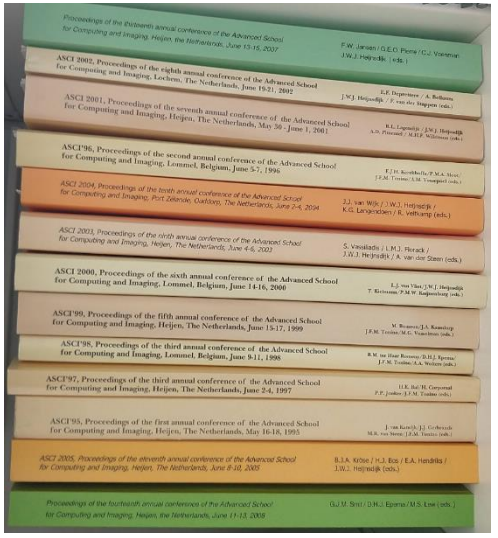
Remco Veltkamp

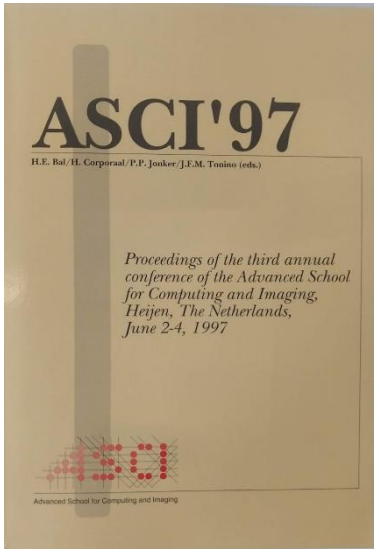
Scientific directors

vrije Universiteit

ASCI activities

- PhD-level courses
- DAS infrastructure (1997-now)
- Yearly “Centerparcs” conference 1995-2008
  - Highly successful, 100-150 participants per year
  - Created a “computing and imaging community”





- Now: ICTopen + NCCV + CompSys

vrije Universiteit

Looking back: impact of ASCI

- 6 NWO/M DAS grants
- Major incentive for VL-e (20 M€ funding)
  - Virtual Laboratory for e-Science (2003-2009)
  - Ultimately led to Netherlands eScience Center
- Strong participation in COMMIT/ (2011-2017)
- Efficient Deep Learning
  - NWO Perspective program led by Henk Corporaal
- Real-Time Video Surveillance Search
  - NWO-TTW project Snoek + Bal + Schiphol consortium

Infrastructure: DAS (pre-)history

- **Andy Tanenbaum already built a cluster in 1984**
  - **Pronet token ring network**
  - **8086 CPUs**
- **He built several Amoeba processor pools**
  - **MC68000, MC68020, MicroSparc**
  - **VME bus, 10 Mb/s Ethernet, Myrinet, ATM**

DAS NWO Proposal (29 April 1996)

DAS — THE DISTRIBUTED ASCI SUPERCOMPUTER

PART A

1. PROPOSER

This proposal is being made by ASCI (Advanced School for Computing and Imaging).

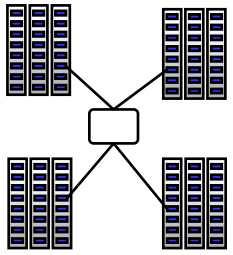
The contact person is the Scientific Director of ASCI:

Prof. dr. A.S. Tanenbaum
Dept. of Mathematics and Computer Science
Vrije Universiteit
De Boelelaan 1081a
1081 HV Amsterdam

Tel: 020 444-7780
Fax: 020 444-7653
Email: ast@cs.vu.nl

The research itself will be carried out by a number of co-operating research groups within ASCI. The primary groups are headed by the following people: Dr. H.E. Bal, VU; Prof. dr. L.O. Hertzberger, UvA; Dr. P. Jonker, TUD; Prof. dr. F.J. Peters; Prof. dr. ir. H.J. Sips, TUD; Dr. P.M.A. Sloot, UvA; Prof. dr. A.S. Tanenbaum, VU; Prof. dr. H.A.G. Wijshoff, RUL; and Prof. dr. I.T. Young, TUD. Other groups within ASCI will also contribute. Most of the groups within the Computer Systems Section of ASCI are participating, as well as several groups from the Imaging Section.

What is DAS?



- **Distributed common infrastructure for Dutch Computer Science**
  - **Distributed: 4-6 clusters at different locations**
  - **Common: single owner (ASCI), single design team**
    - **Users have access to entire system**
  - **Dedicated to CS exper**
    - **Interactive (distributed)**
    - **Able to modify/break th**
  - **Dutch: small scale**

A Medium-Scale Distributed System for Computer Science Research: Infrastructure for the Long Term

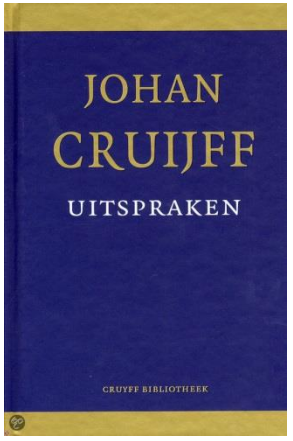
Henri Bal, VU University Amsterdam
Dick Epema, Delft University of Technology
Cees de Laat, University of Amsterdam
Rob van Nieuwpoort, Netherlands eScience Center
John Romein, ASTRON
Frank Seinstra, Netherlands eScience Center
Cees Snoek, University of Amsterdam
Harry Wijshoff, Leiden University

Like any science that heavily uses experimentation, computer science depends on

vrije Universiteit

About SIZE

- Only ~200 nodes in total per DAS generation
- Less than 1.5 M€ total funding per generation
- Johan Cruyff:
  - "Ieder nadeel heb zijn voordeel"
  - Every disadvantage has its advantage



Small is beautiful

- We have superior wide-area latencies
  - "The Netherlands is a 2×3 msec country" (Cees de Laat, Univ. of Amsterdam)
- Able to build each DAS generation from scratch
  - Coherent distributed system with clear vision
- Despite small scale: 3 CCGrid SCALE awards
  - 2008: Ibis: ``Grids as promised" (Ibis team, VU)
  - 2010: WebPIE: web-scale reasoning (Urbani, VU)
  - 2014: BitTorrent analysis (Iosup, TUD)

vrije Universiteit

TU Delft Delft University of Technology

MultimediaN

netherlands eScience center

Universiteit van Amsterdam

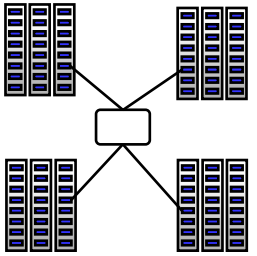
Universiteit Leiden

ASTRON

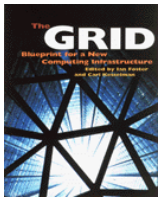
SURF SARA

vrije Universiteit

DAS generations: visions



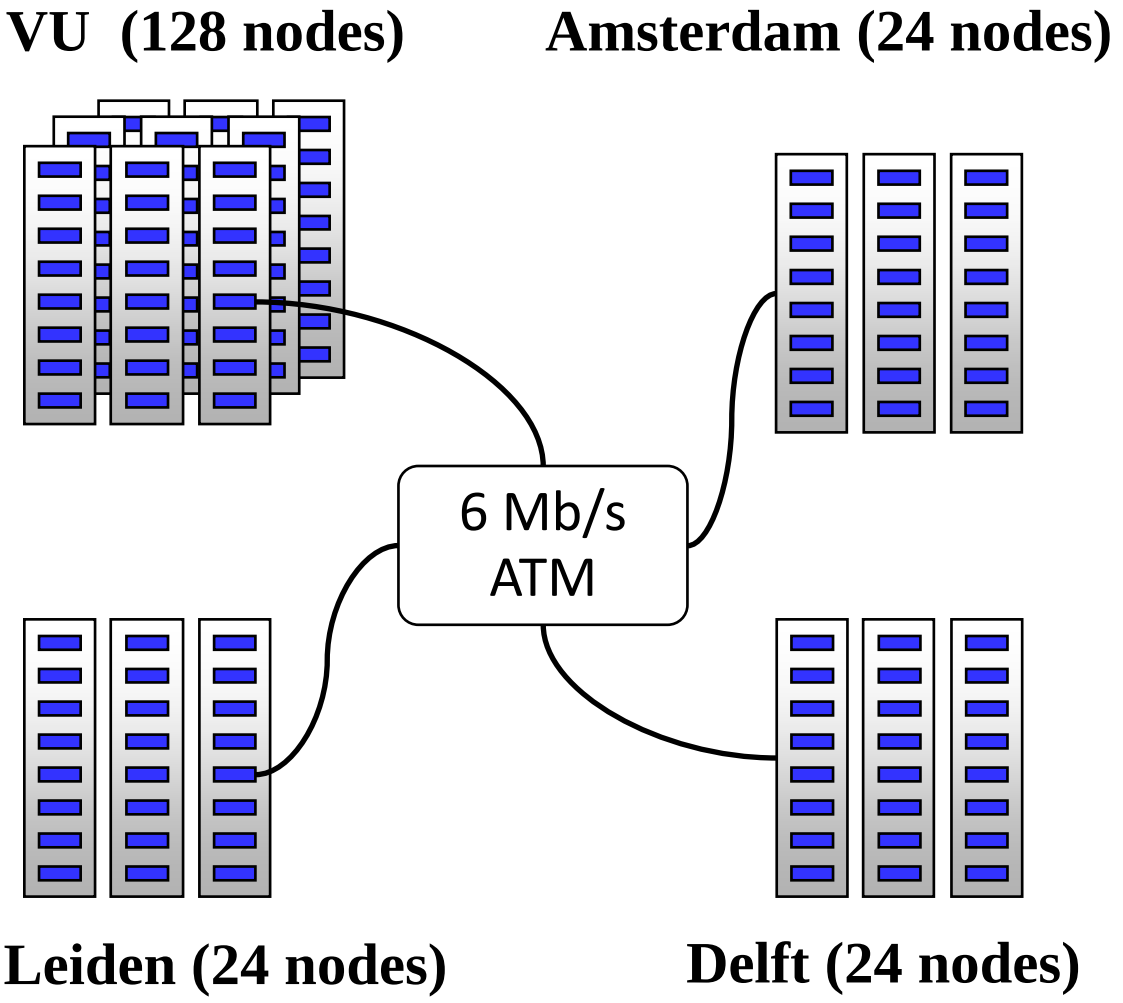
- DAS-1: Wide-area computing (1997)
- DAS-2: Grid computing (2002)



- DAS-3: Optical Grids (2006) StarPlane
- DAS-4: Clouds, diversity, green IT (2010) DAS 4
- DAS-5: Harnessing diversity, data-explosion (2015)
- DAS-6: Distributed research ecosystem (2020)

DAS-1 (1997-2002)
A homogeneous wide-area system

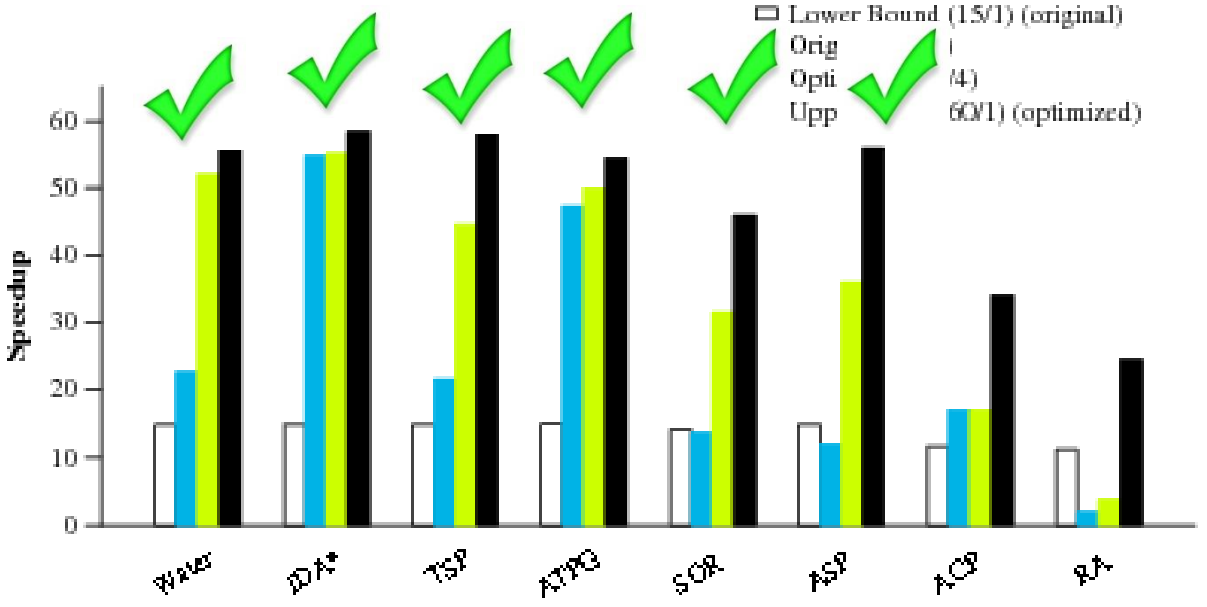
200 MHz Pentium Pro
Myrinet interconnect
BSDI → Redhat Linux
Built by Parsytec



vrije Universiteit

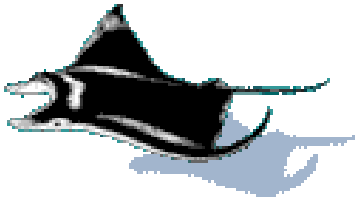
Albatross project (Aske Plaat)

- Optimize algorithms for wide-area systems
  - Exploit hierarchical structure ➔ locality optimizations
- Compare:
  - 1 small cluster (15 nodes)
  - 1 big cluster (60 nodes)
  - wide-area system (4×15 nodes)



Wide-area programming systems

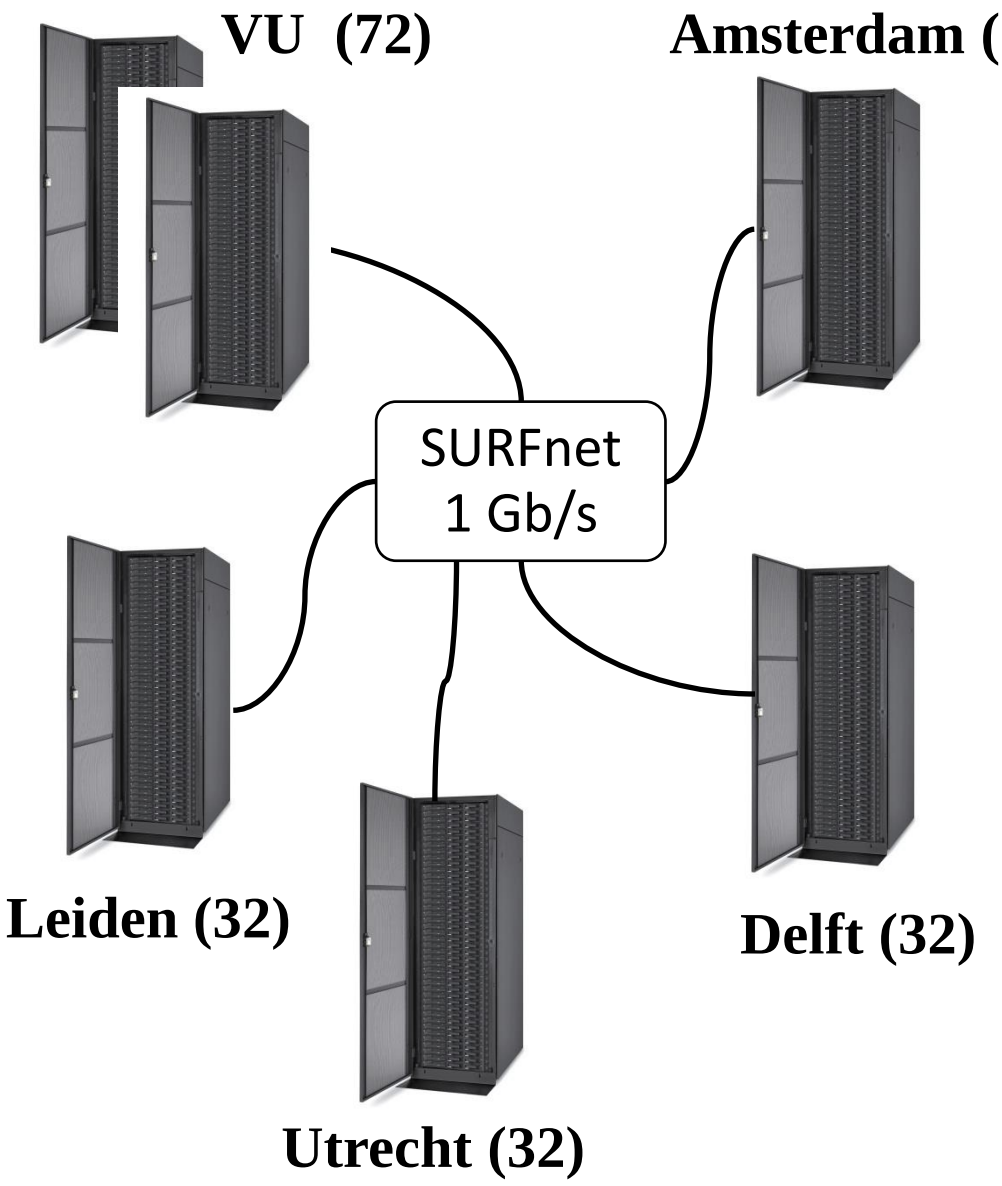
- **Manta (Jason Maassen, Ronald Veldema, Rob van Nieuwpoort)**
  - **High-performance Java [TOPLAS 2001]**



- **MagPIe (Thilo Kielmann):**
  - **MPI's collective operations optimized for hierarchical wide-area systems [PPoPP'99]**

- **KOALA (Dick Epema, TU Delft):**
  - **Multi-cluster scheduler with support for co-allocation**

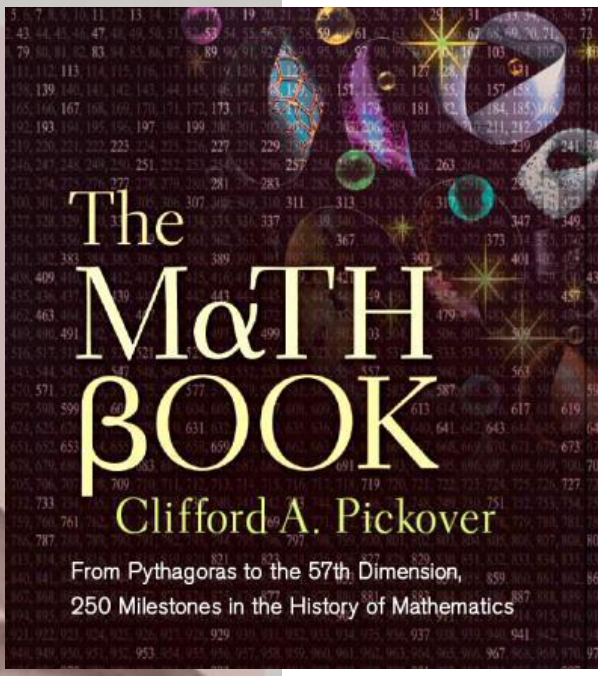
DAS-2 (2002-2006)
a Computer Science Grid



889.063.398.406 bordposities

AMSTERDAMSE SUPERCOMPUTER LOST BORDSPEL AWARI OP

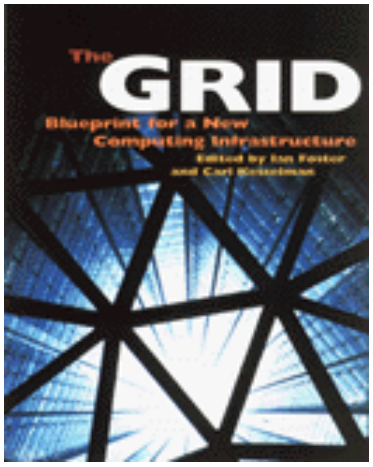
Awari spelen tegen de computer is niet leuk meer. Tegen een database met alle bordposities, achterwaarts bepaald, kan geen mens op en remise is het hoogst haalbare.



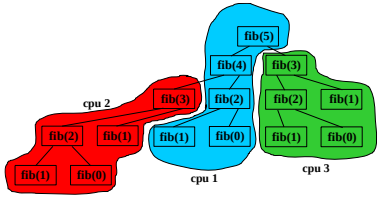
John Romein fully solves Awari on DAS-2 (NRC, 14 Sept. 2002)

Grid experiments

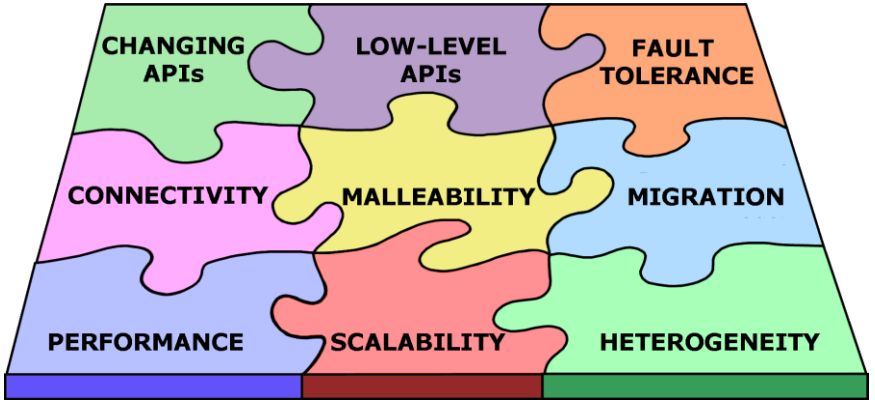
- Do clean performance measurements on DAS
- Combine DAS with EU grids to test heterogeneity
  - Show the software ``also works'' on real grids

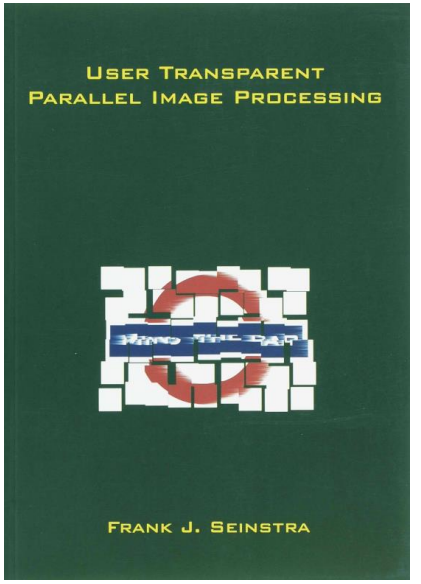


Grid programming systems



- **Satin (Rob van Nieuwpoort):**
  - **Transparent divide-and-conquer parallelism for grids**
- **JavaGAT:**
  - **Middleware-independent API for grid applications**
- **Ibis: Java-centric grid computing**
  - **Funded by NWO, VL-e, EU**

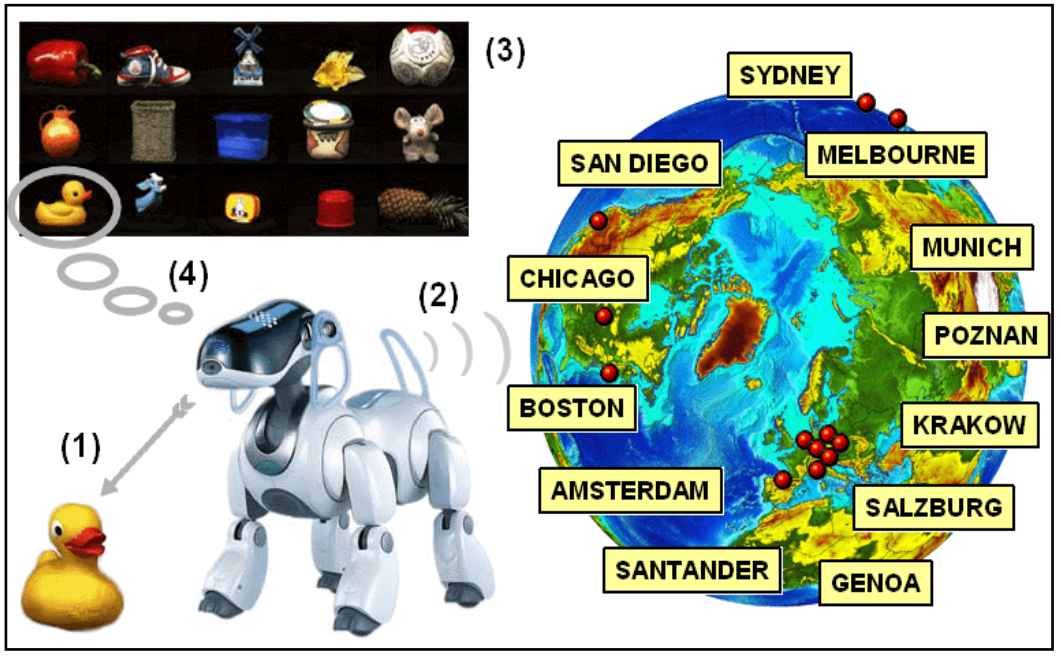




PhD UvA (2003)

vrije Universiteit

Color-based Object Recognition by a Grid-connected Robot Dog



Problems:

Scaling to large systems

Platform dependence

Middleware dependence

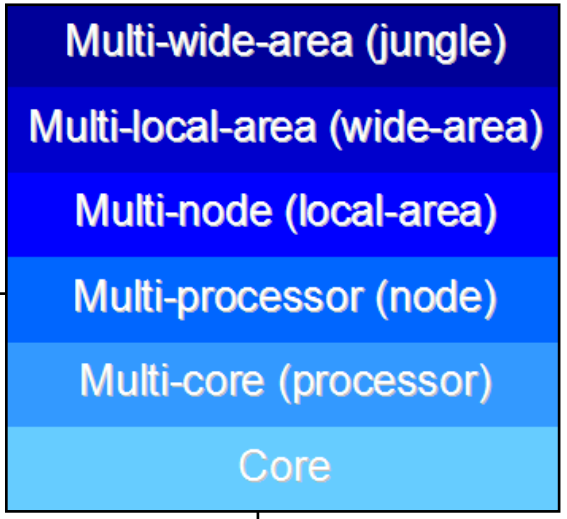
Connectivity (firewalls)

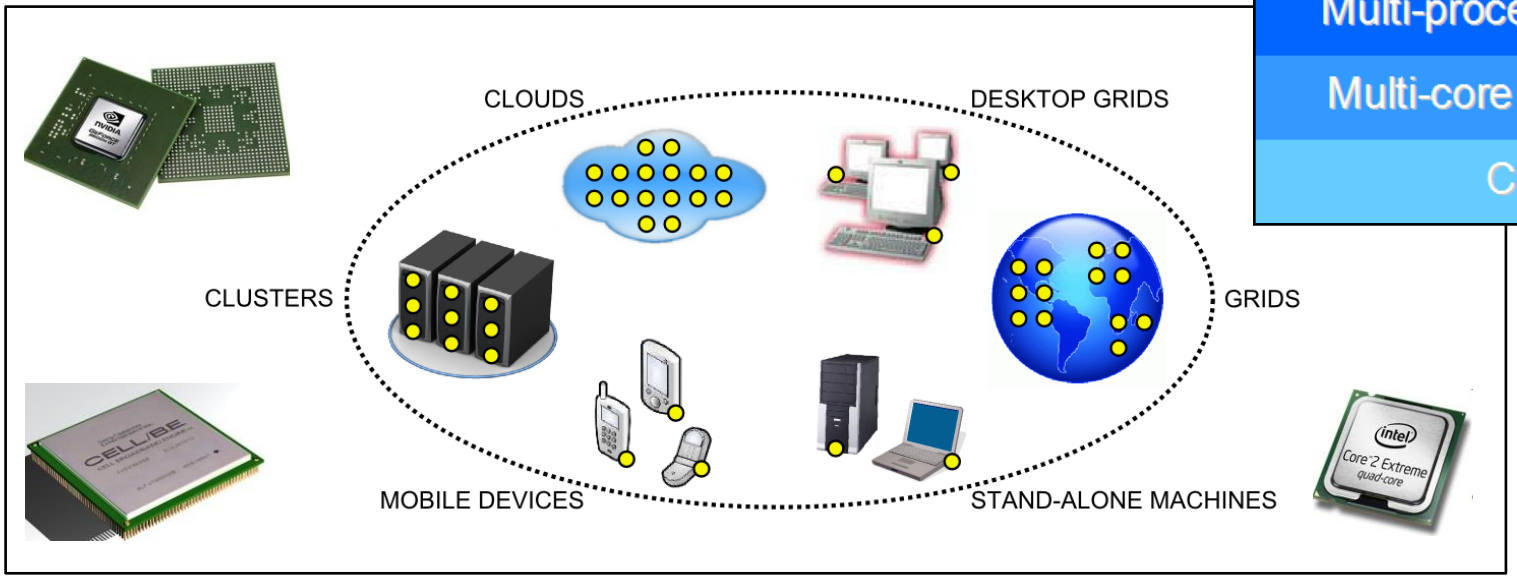
Fault-tolerance

Seinstra et al (IEEE Multimedia, 2007) + AAAI'07: Most Visionary Research Award)

Jungle Computing with Ibis

- ‘Worst case’ computing as required by end-users
  - Distributed
  - Heterogeneous
  - Hierarchical (incl. multi-/many-cores)



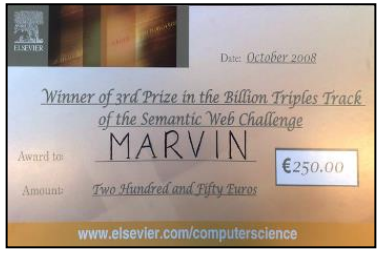


Ibis Results: Awards & Prizes

AAAI-VC 2007
Most Visionary Research Award

1st Prize: DACH 2008 - BS

1st Prize: DACH 2008 - FT

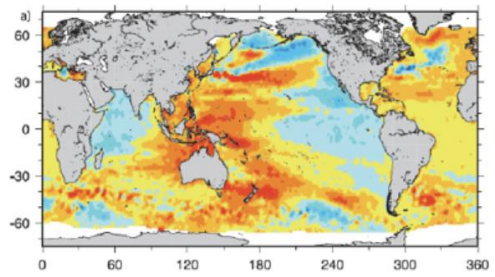


3rd Prize: ISWC 2008

1st Prize: SCALE 2008

1st Prize: SCALE 2010

https://www.research-software.nl/software/xenon (NLeSC)



Global Climate Modeling

- **Understand future local sea level changes**
- **Needs high-resolution simulations**
- **Distributed computing (multiple resources)**

netherlands eScience center

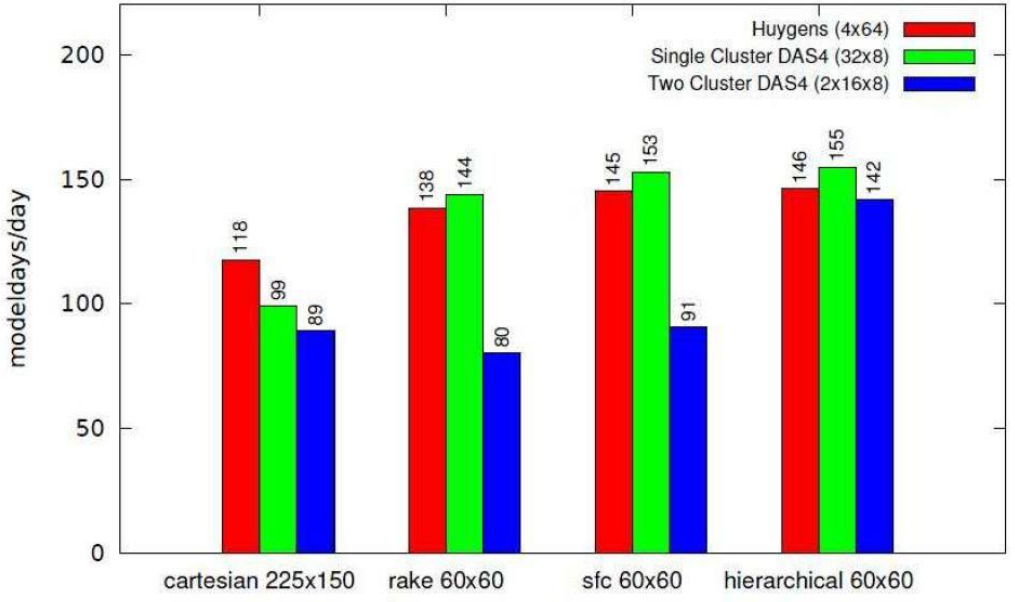
IMAU Institute for Marine and Atmospheric research Utrecht

COMMIT/

vrije Universiteit

Distributed Climate Modeling (Jason Maassen)

- Use Ibis to couple different simulation models
  - Land, ice, ocean, atmosphere
- Wide-area optimizations similar to Albatross project (16 years ago), like hierarchical load balancing



vrije Universiteit

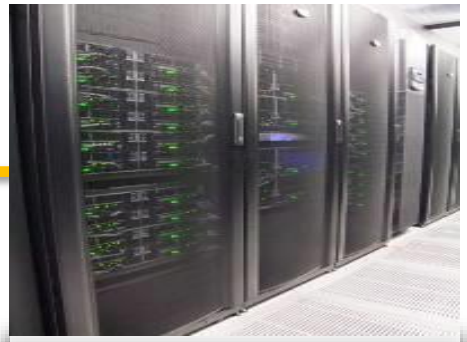
Enlighten Your Research Global award

#7 STAMPEDE (USA)

10G

CARTESIUS (NLD)

10G



EMERALD (UK)

KRAKEN (USA)

10G

SUPERMUC (GER) #10

IMAU Institute for Marine and Atmospheric research Utrecht

LMU LUDWIG-MAXIMILIANS-UNIVERSITÄT MÜNCHEN

OXFORD e-Research CENTRE

UNIVERSITY OF OXFORD

NCAR NATIONAL CENTER FOR ATMOSPHERIC RESEARCH

RUTGERS

ROSENSTIEL SCHOOL OF MARINE AND ATMOSPHERIC SCIENCE RSMAS UNIVERSITY OF MIAMI

VU UNIVERSITY AMSTERDAM

UNIVERSITY OF MIAMI

Universiteit Utrecht

netherlands eScience center

vrije Universiteit

CONCURRENCY AND COMPUTATION: PRACTICE AND EXPERIENCE
Concurrency Computat.: Pract. Exper. 2017; 29: e3853
Published online 19 May 2016 in Wiley Online Library (wileyonlinelibrary.com). DOI: 10.1002/cpe.3853

Report on the first workshop on negative and null results in eScience

Ketan Maheshwari, Daniel Katz, Silvia D. Olabarriaga, Justin Wozniak, Douglas Thain

e3908 | First Published: 02 August 2016

First Page | Full text | PDF | References | Request permissions

Special Issue Papers

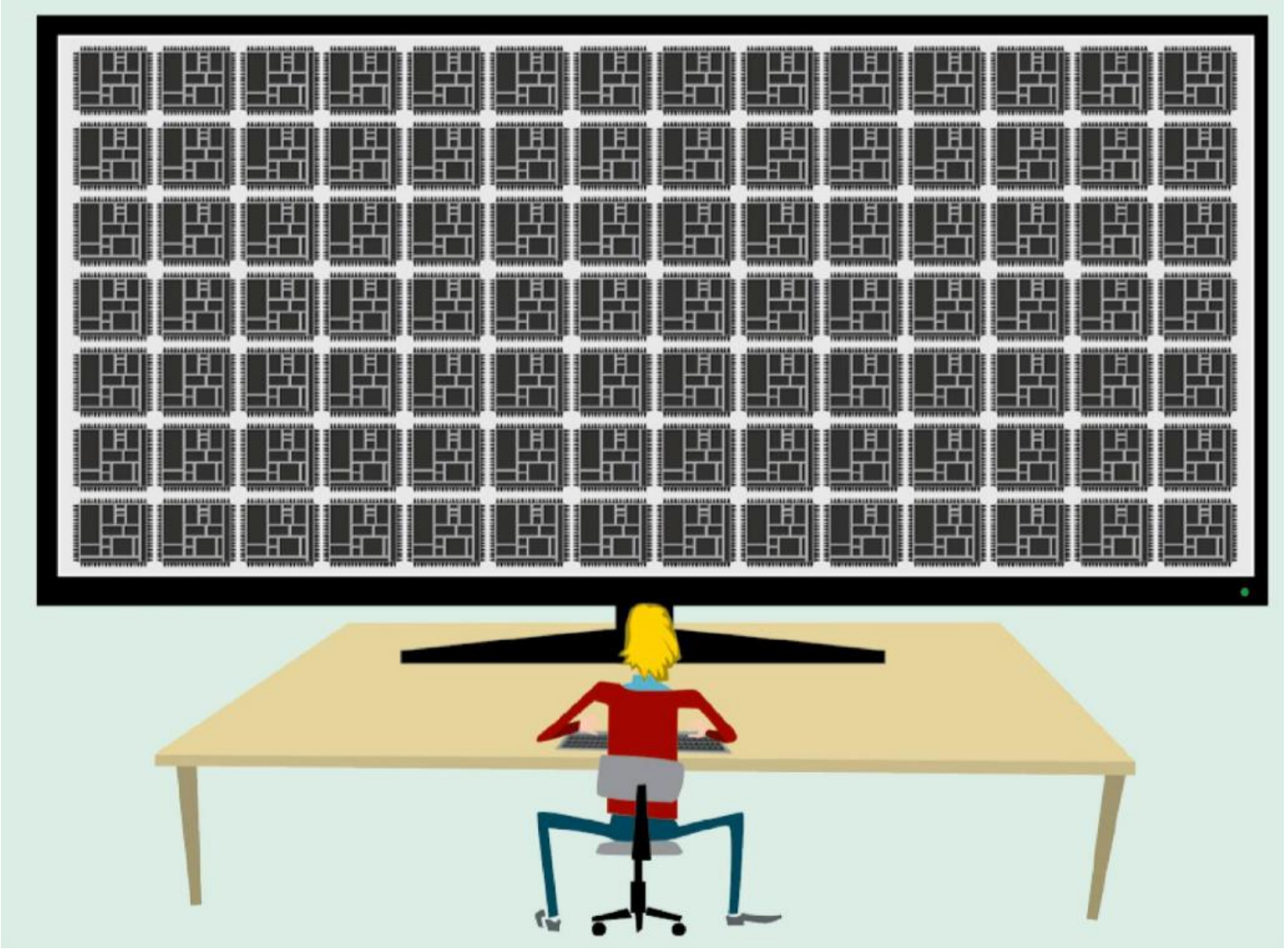
Full Access

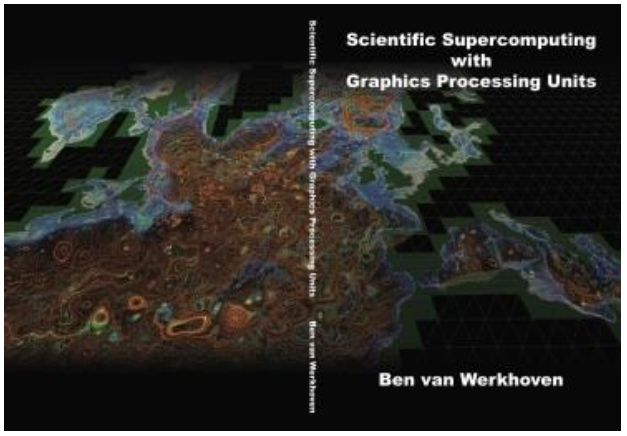
On the complexities of utilizing large-scale lightpath-connected distributed cyberinfrastructure

Jason Maassen, Ben van Werkhoven, Maarten van Meersbergen, Henri E. Bal, Michael Kliphuis, Sandra E. Brunnabend, Henk A. Dijkstra, Gerben van Malenstein, Migiel de Vos, Sylvia Kuijpers, Sander Boele, Jules Wolfrat, Nick Hill, David Wallom, Christian Grimm, Dieter Kranzlmüller, Dinesh Ganpathi, Shantenu Jha, Yaakoub El Khamra, Frank O. Bryan, Benjamin Kirtman, Frank J. Seinstra

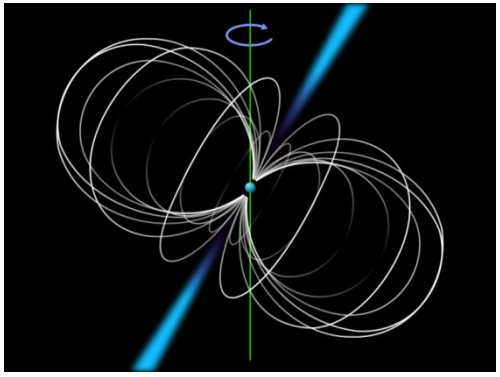
e3853 | First Published: 19 May 2016

Many-core (GPU) programming





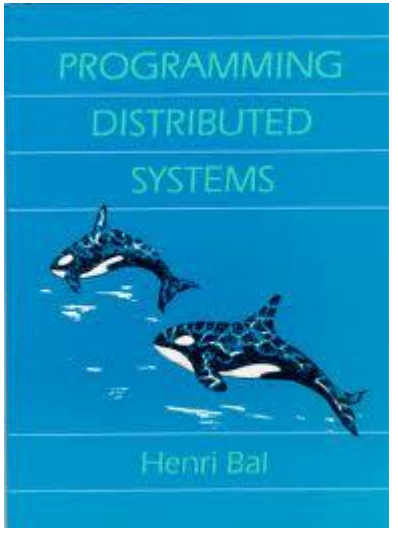
Case studies



- Convolution algorithms - Ben van Werkhoven
- Radio astronomy pulsar search - Alessio Sclocco
- Mapping algorithm to architecture is difficult, especially as the architecture is difficult:
  - Many levels of parallelism
  - Limited resources (registers, shared memory)
  - Everything must be in balance to obtain performance
  - Importance of autotuning
    - https://www.research-software.nl/software/kernel-tuner

Why is GPU programming hard?

- **Bottom line: tension between**
  - **control over hardware to achieve performance**
  - **higher abstraction level to ease programming**
- **Programmers need understandable performance**
- **Old problem in Computer Science, but now in extreme form**



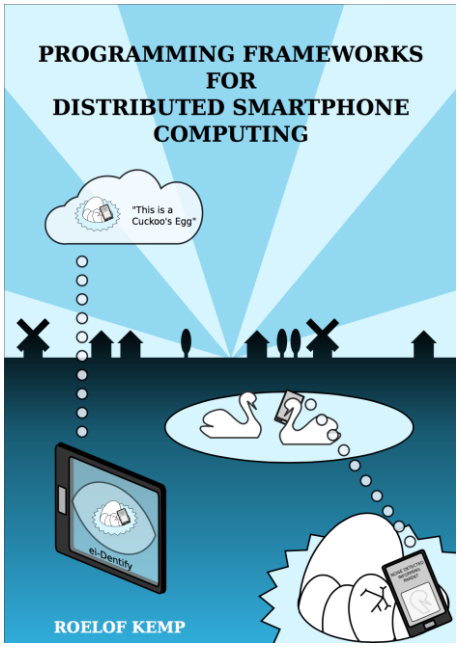
(1989)

Stepwise refinement for performance

- **MCL Methodology:**
  - **Programmers work on *multiple levels of abstraction***
  - **Integrate *hardware descriptions* into programming model**
  - ***Performance feedback* from compiler, based on hardware description and kernel**
  - **Cooperation between compiler and programmer**
- **Cashmere: Satin (divide & conquer) + MCL**
  - **Load balancing for heterogeneous GPU systems**
- **Applied to image source identification (NFI)**

Pieter Hijma - Programming Many-Cores on Multiple Levels of Abstraction (PhD, 2015)
Extensive survey paper on Optimization Techniques for GPUs

Side hobby: smartphones & IoT



A Smartphone-based Infrastructure for Decentralized Partnership Formation

Nicolae Vladimir Bozdog
Vrije Universiteit

2012 2014 2019

Energy-efficient Stream Processing for a Smart Device Ecosystem

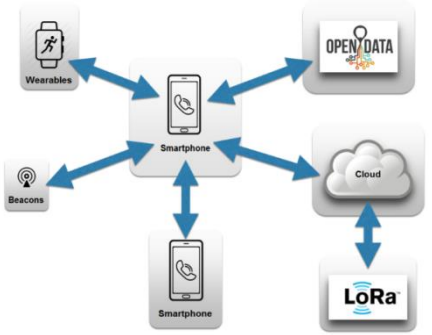
Roshan Das
Vrije Universiteit

2021

- **Cloud offloading (Cuckoo)**
- **Sensing With Android Nodes (SWAN)**
- **Distributed ad-hoc sensing**
- **Stream processing**

SWAN

- Programming environment for smartphones & IoT
- Common low-level interface (data collection and aggregation) for many sensor types
  - Smartphones
  - Smartwatches
  - Remote phones
  - Real-time open data
  - IoT data

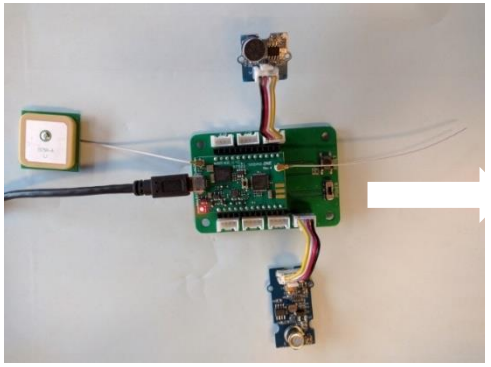


- Processing on phones, watches, clouds
- Programming in SWAN-song DSL
- App: weather-based bike-routing

IoT pilot testbed Alkmaar

- **LoRaWAN based testbed covering whole city**
  - **Two gateways & 20 Sodaq sensor nodes**
  - **Set up mostly by Kees Verstoep**

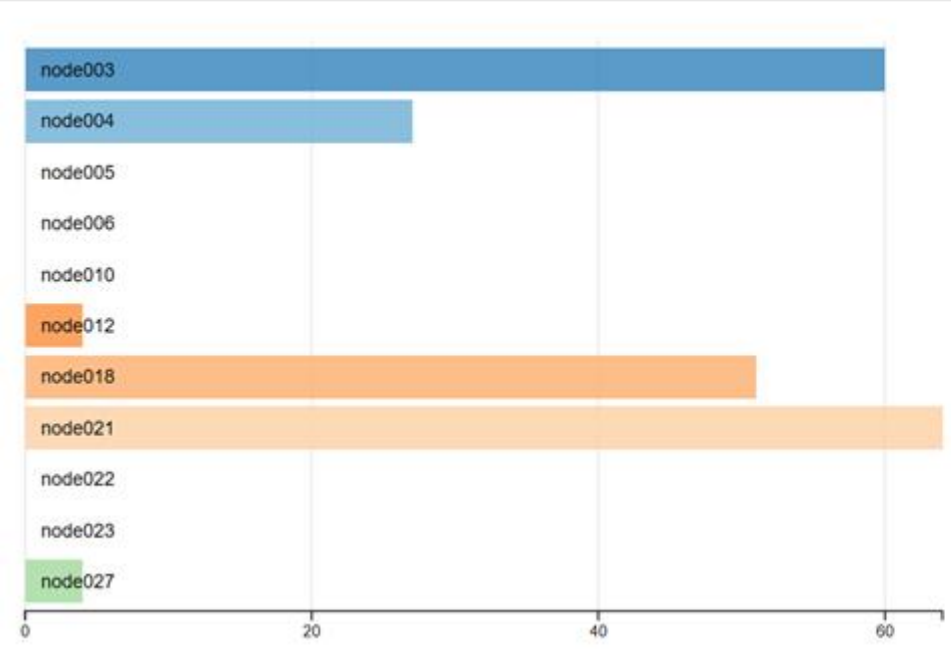
Gateway

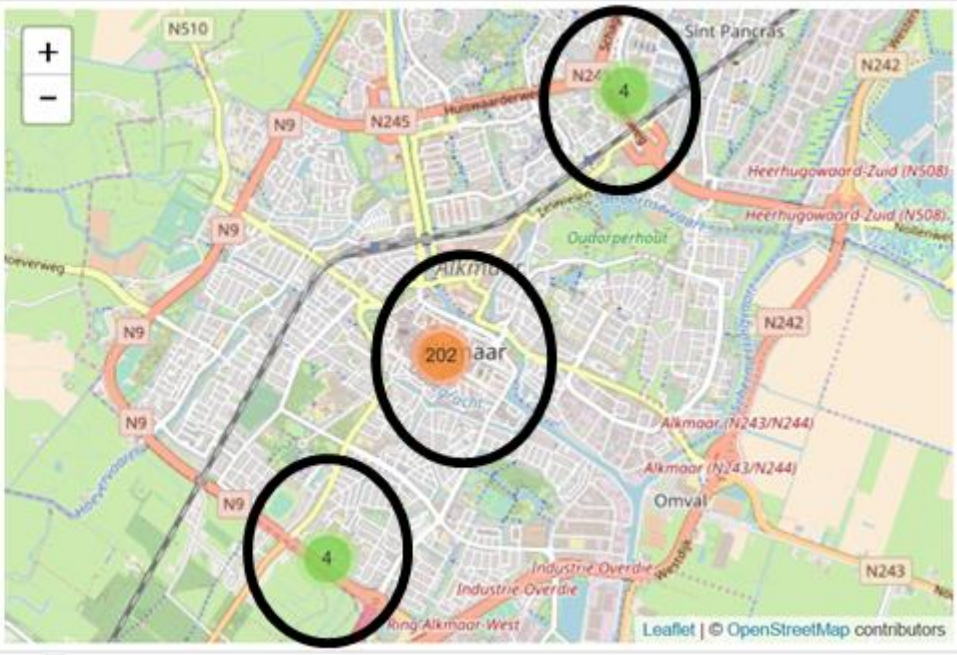


Sensor nodes

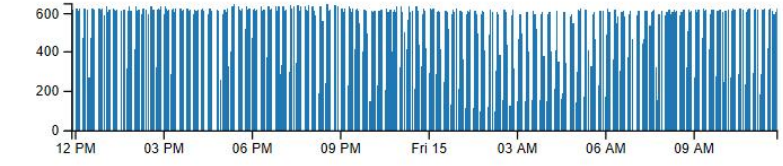
vrije Universiteit

datasciencealkmaar

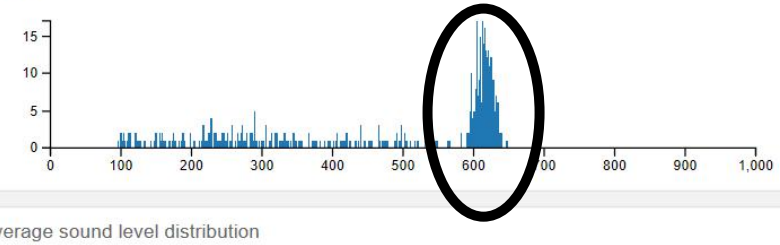




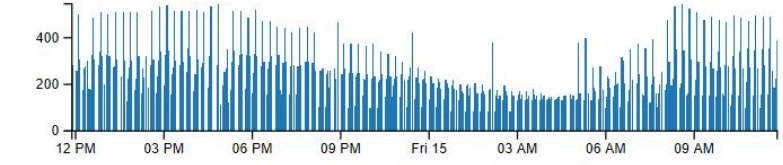
Peak sound level



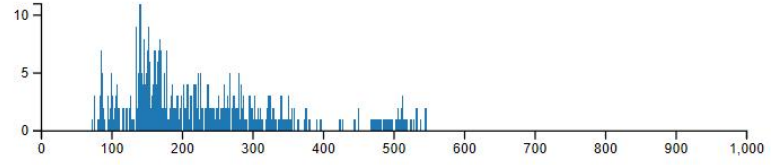
Peak sound level distribution



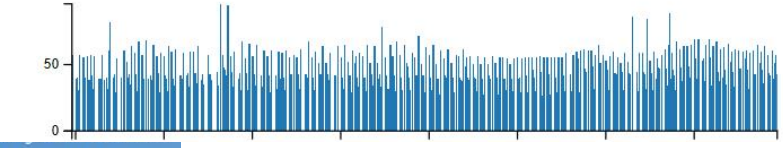
Average sound level



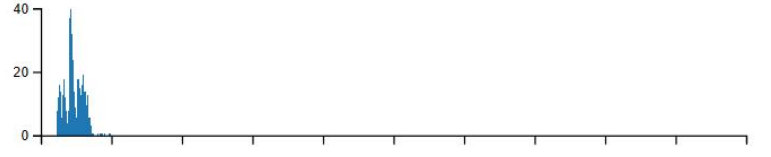
Average sound level distribution



Air quality



Air quality distribution

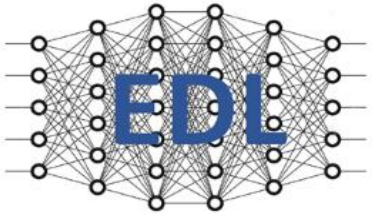


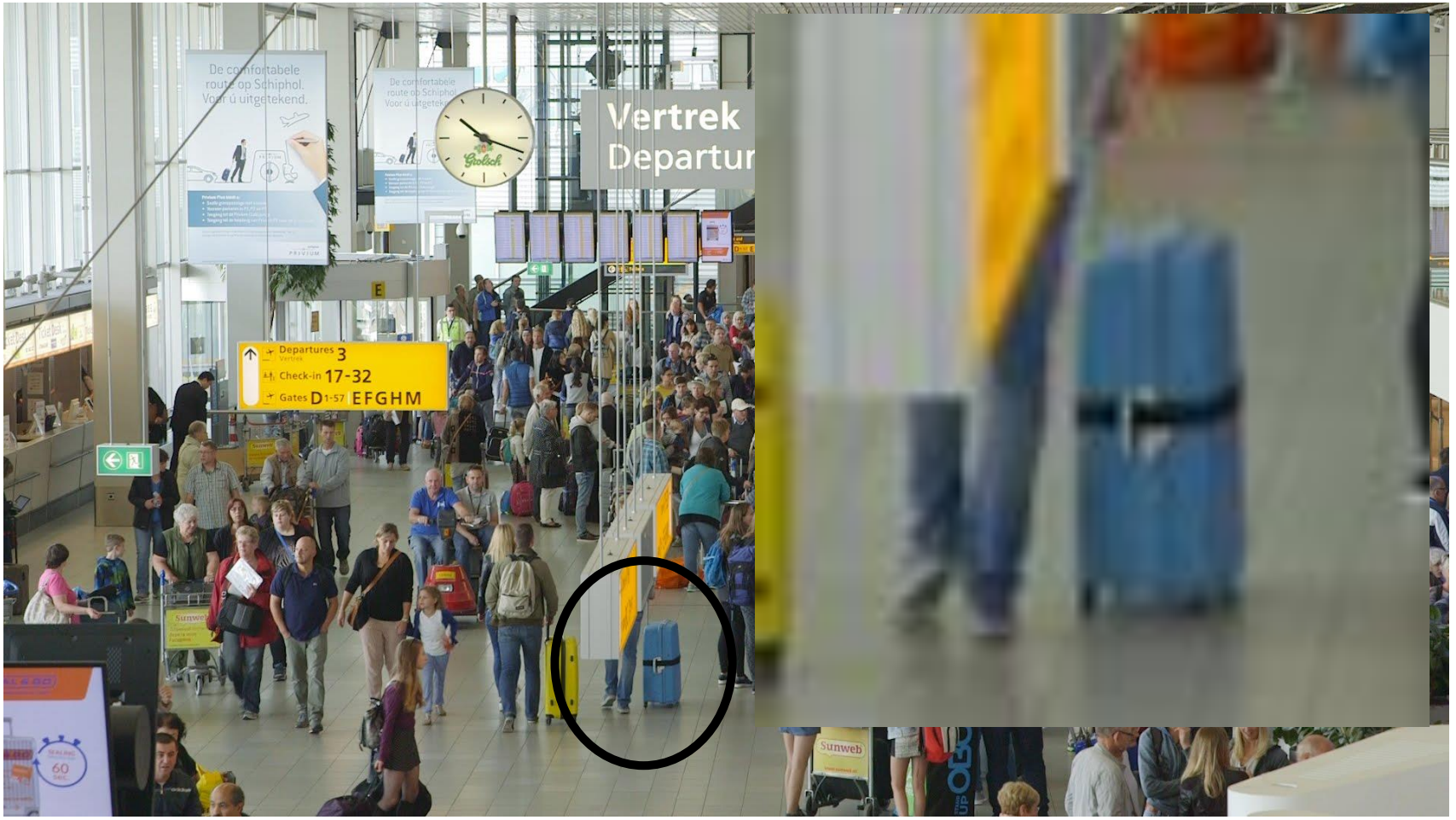
Next: combining IoT sensors & HPC

- Data streams become larger
- Computing demands on sensor data increase
  - Deep learning (inferencing)
  - Activity recognition in video streams
  - Reasoning on sensor streams (stream reasoning)
- Often need real-time response => edge computing
- NWO project: OffSense (2020-2025)
  - Generalized offloading model
  - With Lin Wang, Animesh Trivedi, Alexandru Iosup; PhDs: George Karlos, Matthijs Jansen

Real-Time Video Surveillance Search

- With Cees Snoek (UvA) & Schiphol consortium
- Part of NWO-TTW program Efficient Deep Learning + separate NWO-TTW-OTP grant
- Recognize anomalous activities in dense crowds
- Huge AI problem (e.g., lack of labelled data)
- Huge compute problem: many cameras, need real-time response
  - Live video analytics as a service (Apostolo et al.)

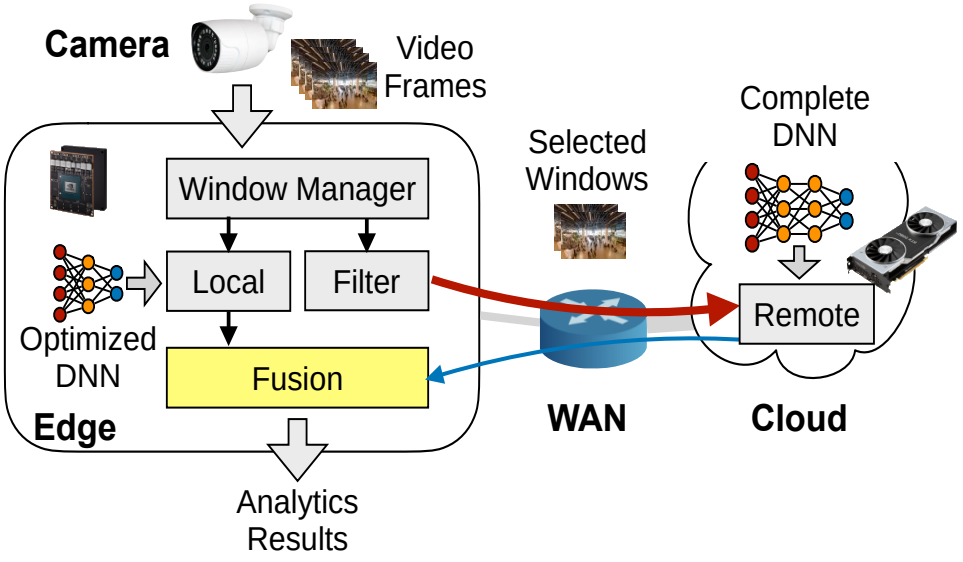




vrije Universiteit

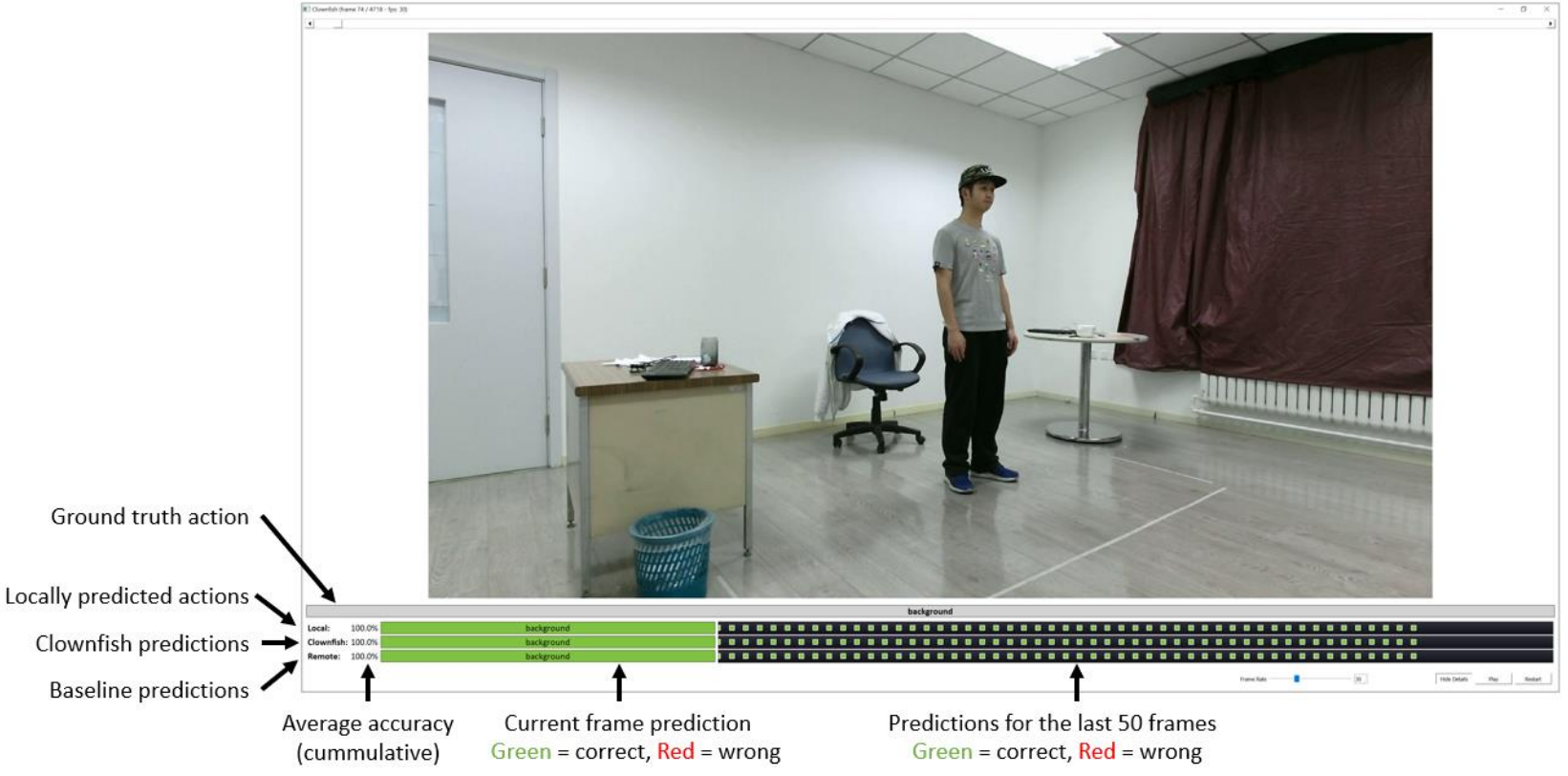
Clownfish (Vinod Nigade, Lin Wang)

- **Cloud can run full DL models - best accuracy**
- **Edge can run small model - fast response**
- **Clownfish: achieve edge-cloud symbiosis**



Clownfish (Vinod Nigade, Lin Wang)

- **Cloud can run full DL models - best accuracy**
- **Edge can run small model - fast response**
- **Clownfish: achieve edge-cloud symbiosis**



Conclusions

- ASCI collaborations had huge impact on career
  - Computing: VL-e, COMMIT/, many NWO projects
  - Imaging: object recognition, NFI, Schiphol
- HPC bridges technology and applications, collaborations with many applications domains
  - AI (Awari, WebPIE, stream reasoning, DL); Astronomy, Climate, Physics, Bioinformatics, …
- Unique infrastructures like DAS are crucial
  - Next: https://slices-ri.eu/
- Community building is key
  - E.g. Manifesto Future Computer Systems and Networking Research in the Netherlands