# Accelerating Transport System Micro-Simulations using Cuda

Peter Heywood, Paul Richmond, Steve Maddock and Matthew Jung

The University of Sheffield, United Kingdom


Caution
40
40
40
40
A46
Carna
TRANSPORT IRELAND
Van Steenbergen
TRANS ORION
www.transorion.pl

# Demand on Transport Networks

- **Road** travel projections for 2010 to 2040 (UK) [1]
  - Up to 42% increase of car ownership
  - 19% to 55% growth in UK road traffic

# Demand on Transport Networks

- **Road** travel projections for 2010 to 2040 (UK) [1]
  - Up to 42% increase of car ownership
  - 19% to 55% growth in UK road traffic
- **Rail** travel projections for 2010 to 2037 (UK) [2]
  - Edinburgh - Central London: 116%
  - Birmingham - Central London: 95%

# Demand on Transport Networks

- **Road** travel projections for 2010 to 2040 (UK) [1]
  - Up to 42% increase of car ownership
  - 19% to 55% growth in UK road traffic
- **Rail** travel projections for 2010 to 2037 (UK) [2]
  - Edinburgh - Central London: 116%
  - Birmingham - Central London: 95%
- **Air** travel projections 2014 to 2034 (IATA) [3]
  - 121% increase - 3.3 billion to 7.3 billion passengers

# Demand on Transport Networks

- **Road** travel projections for 2010 to 2040 (UK) [1]
  - Up to 42% increase of car ownership
  - 19% to 55% growth in UK road traffic
- **Rail** travel projections for 2010 to 2037 (UK) [2]
  - Edinburgh - Central London: 116%
  - Birmingham - Central London: 95%
- **Air** travel projections 2014 to 2034 (IATA) [3]
  - 121% increase - 3.3 billion to 7.3 billion passengers
- All result in an increase of **Pedestrian** traffic

# Cost & Disruption

- Real world changes are **expensive** & **disruptive**
- £709 billion spent maintaining the UK strategic Motorway and A road network in 2010/2011 [4]
- Need for a **cheaper** & **less disruptive** solution

# How can GPUs Help?

- Improved tools for **planning** transport networks
  - Design of new infrastructure e.g. building layouts
  - Impact of new Highways, Railways etc.

# How can GPUs Help?

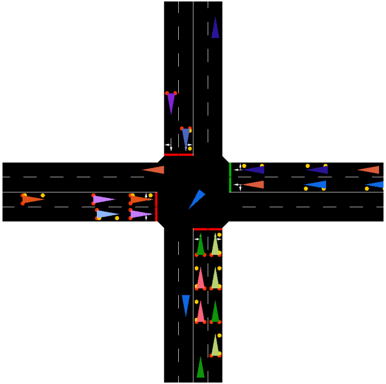
- Improved tools for **planning** transport networks
  - Design of new infrastructure e.g. building layouts
  - Impact of new Highways, Railways etc.
- Improved tools for **managing** transport networks
  - Intelligent crowd management & evacuation
  - Smart Motorways, Green-waving, ...

# How can GPUs Help?

- Improved tools for **planning** transport networks
  - Design of new infrastructure e.g. building layouts
  - Impact of new Highways, Railways etc.
- Improved tools for **managing** transport networks
  - Intelligent crowd management & evacuation
  - Smart Motorways, Green-waving, ...
- Achieved through high-performance simulations & Interactive Visualisation

## Predictive Simulation

- Simulate *many* scenarios *many* times
- Aggregate & analyse results to find the optimal solution
- High performance is critical



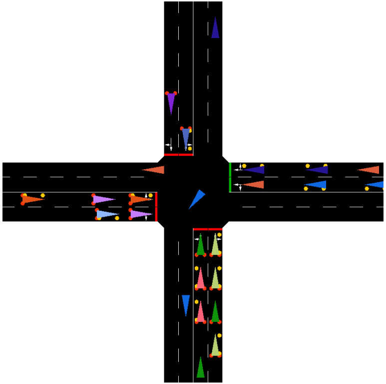
An example of traffic microsimulation visualisation (`sumo-gui`)

## Predictive Simulation

- Simulate *many* scenarios *many* times
- Aggregate & analyse results to find the optimal solution
- High performance is critical

## Interactive Visualisation

- Decision makers are often not modelling specialists [5]
- Interactive visualisation increases accessibility of simulations
- Aids decision making process



An example of traffic microsimulation visualisation (`sumo-gui`)

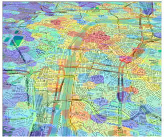
# Simulation Resolution

Transport network simulations can typically be classified as:

# Simulation Resolution

Transport network simulations can typically be classified as:

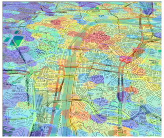
- **Macroscopic** (*Top-Down*)
  Aggregates characteristics of environment



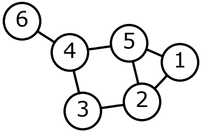
# Simulation Resolution

Transport network simulations can typically be classified as:

- **Macroscopic** (*Top-Down*)
  Aggregates characteristics of environment
- **Mesoscopic** (*Middle-Out*)
  Model groups (platoons) of individuals as a single unit



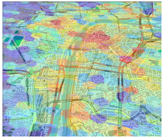




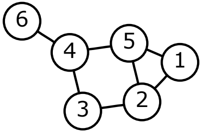
# Simulation Resolution

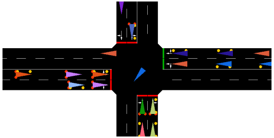
Transport network simulations can typically be classified as:

- **Macroscopic** (*Top-Down*)
  Aggregates characteristics of environment
- **Mesoscopic** (*Middle-Out*)
  Model groups (platoons) of individuals as a single unit
- **Microscopic** (*Bottom-Up*)
  Model individuals within the system



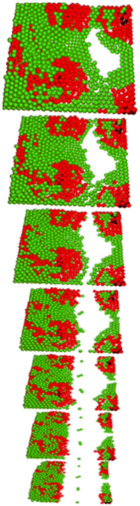




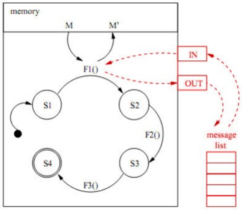
# Agent Based Modelling (ABM)

- Method for describing model behaviour at an individual level
- Complex behaviours emerge from simple rules and local interaction
- Computationally Expensive
- Not *embarrassingly parallel* but it is well suited to GPU acceleration



# FLAME GPU

- Flexible Large-scale Agent Modelling Environment for the GPU
- Template-based simulation environment for generation of high performance simulations
- Agents represented using a form of state machine
    - Provides high level abstraction
- www.flamegpu.com




State machine agent with message based communication

http://on-demand.gputechconf.com/gtc/2015/presentation/S5133-Paul-Richmond.pdf

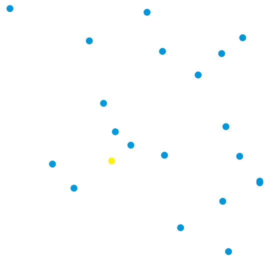
# FLAME GPU Agent Communication

- Agent communication specialised through partitioning techniques [6]

# FLAME GPU Agent Communication

- Agent communication specialised through partitioning techniques [6]

- **Non Partitioned Messaging**
  - All to All Communication - $O(n^2)$ message iteration loop

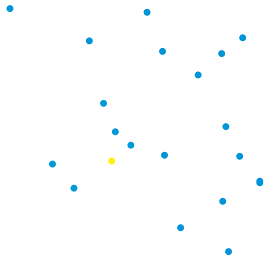


Non-Partitioned Messaging

# FLAME GPU Agent Communication

- Agent communication specialised through partitioning techniques [6]

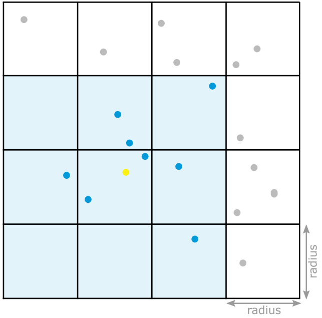
- **Non Partitioned Messaging**
  - All to All Communication - $O(n^2)$ message iteration loop
- **Discrete Partitioned Messages**
  - Message from Non-mobile discrete agents.
  - Receive messages from agents with a specified `radius`



Non-Partitioned Messaging

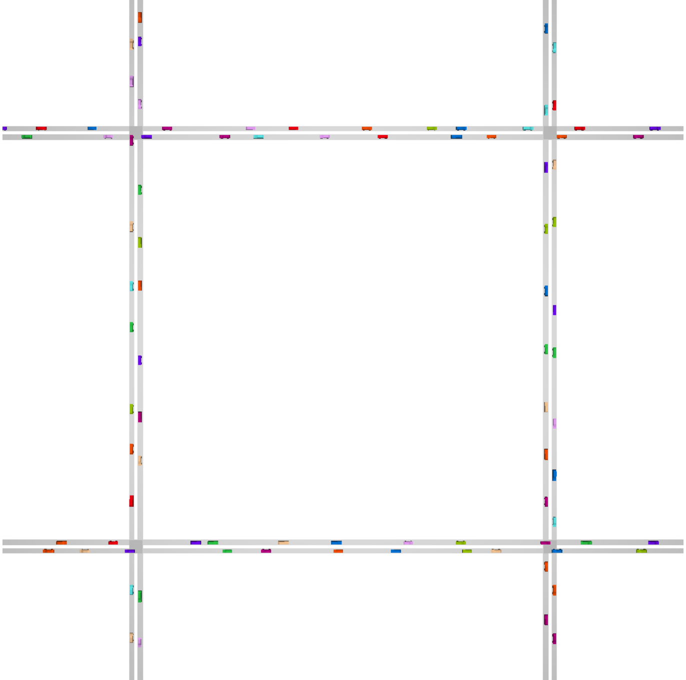
# FLAME GPU Agent Communication

- Agent communication specialised through partitioning techniques [6]
- **Non Partitioned Messaging**
  - All to All Communication - $O(n^2)$ message iteration loop
- **Discrete Partitioned Messages**
  - Message from Non-mobile discrete agents.
  - Receive messages from agents with a specified `radius`
- **Spatially Partitioned Messaging**
  - Messages from continuous space agents in 2D or 3D environment
  - Receive messages from agents with a specified `radius`



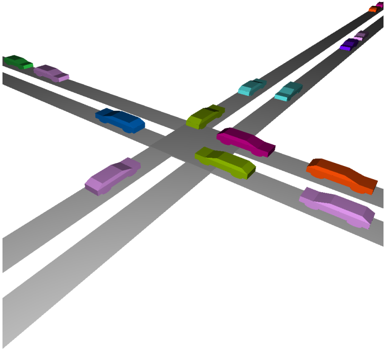

Spatially Partitioned Messaging

# Road Network Simulation using FLAME GPU



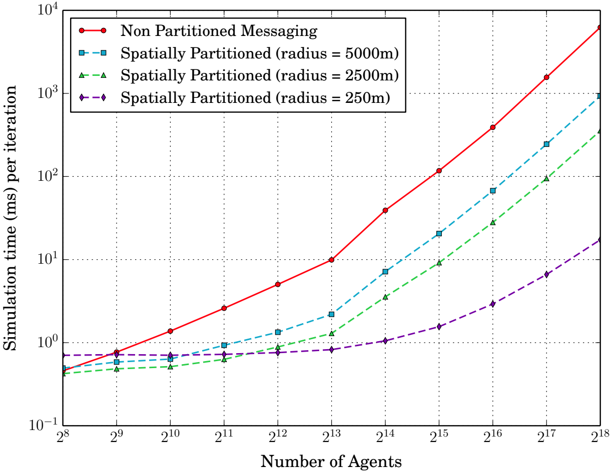
# Motivation & Implementation

- Evaluate the suitability of FLAME GPU for Road Network Simulation
  - Implemented Gipps' car following model [7]
  - Safety-distance model considers driver and vehicle characteristics
  - Artificial Grid road network
- Real-time rendering enabled by geometry instancing & Cuda OpenGL Interoperability

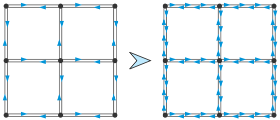


Close up view of instanced vehicles

# Benchmarks: Fixed Network, Variable Population

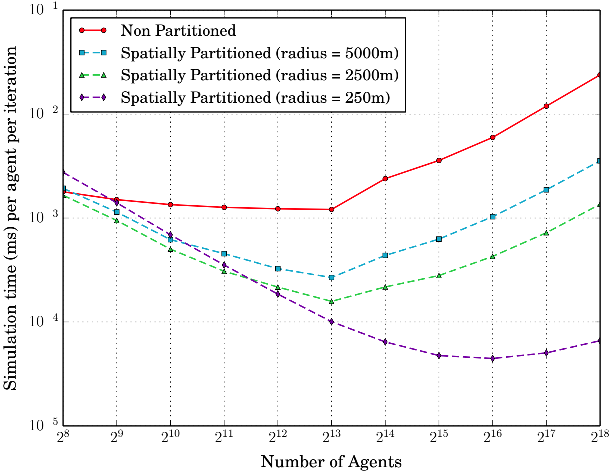



Average Simulation time for increasing agent population on a fixed size road network



- Spatially partitioned messaging outperforms non-partitioned messaging
- Smaller radii outperforms larger radii beyond partitioning scheme overhead cost
- Tesla K20c
- More details see “Road Network Simulation using FLAME GPU” [8]

# Benchmarks: Fixed Grid, Variable Population - Per Agent



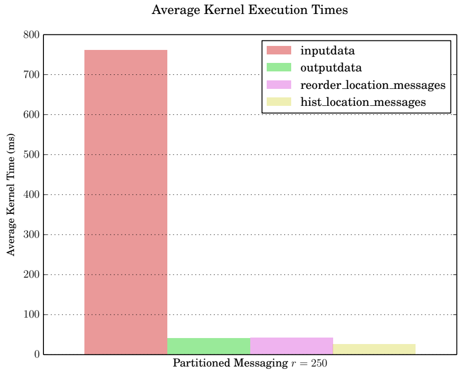

Average Simulation time per agent for increasing agent population on a fixed size road network

- Performance per iteration, divided by population size
- Distinct gradient change at $2^{13}$ agents - **hardware utilisation** vs **larger message lists**
- Maximum message count

| | |
|---|---|
| Non-partitioned | 262144 |
| Partitioned $r = 5000$ | 19662 |
| Partitioned $r = 2500$ | 9720 |
| Partitioned $r = 250$ | 309 |

# Results: Fixed Grid, Variable Population - Kernel Profiling

- Kernel times averaged over 10 iterations
- $2^{15}$ (32768) Agents, $r = 250$
- `inputdata` kernel is dominant
    - Message list iteration

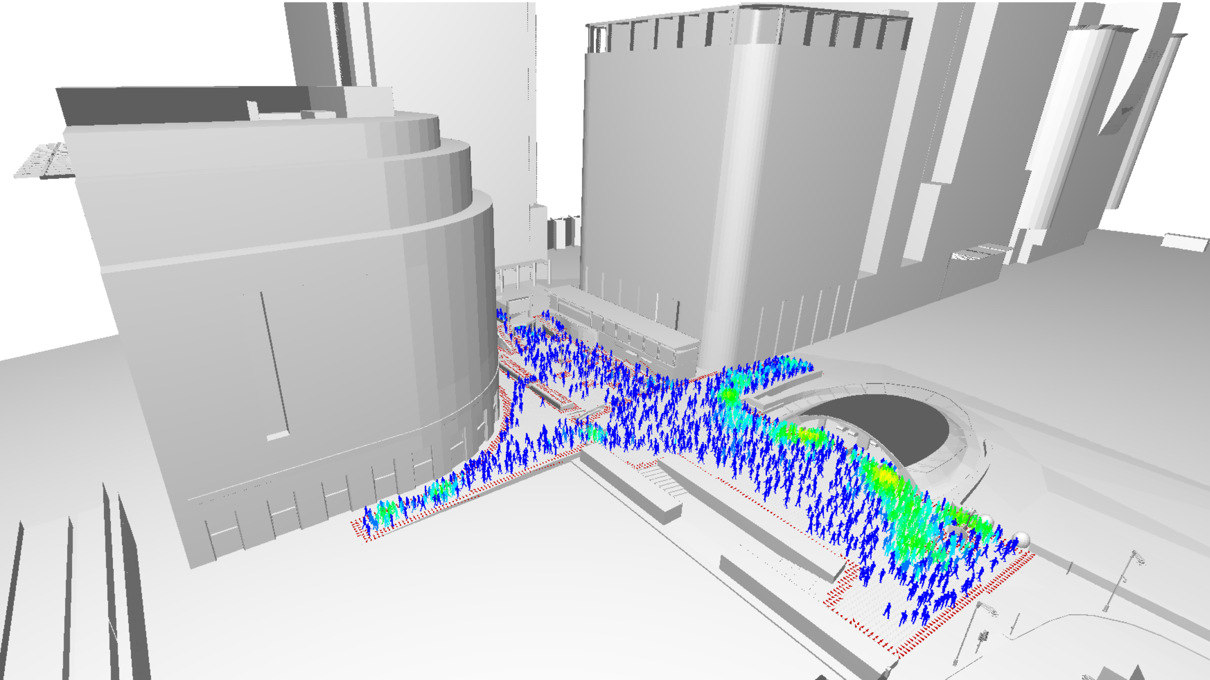



Average Kernel execution times for spatially partitioned messaging with $r = 250$

# Agent Communication

- Motivated to demonstrate FLAME GPU suitability for Road Network Simulation
- Demonstrated good performance
- Highlighted the limiting factor: Large message lists
- Need for a specialised communication strategy for network constrained agents

# Pedestrian Crowd Simulation using FLAME GPU



# Pedestrian Simulation

- Crowd simulations provide insight into how an environment will be used
- Pedestrians move towards target exit while interacting in a realistic fashion
- Spatially Partitioned messaging offers significant performance improvements
- Cheap Visualisation via instanced rendering

# Virtual Reality Pedestrian Simulation using Omnideck 6


CATAPULT
Intelligent Mobility Table
CATAPULT
OMNIFINITY™

# Virtual Reality Crowd Simulation

- Accessibility increased by immersive visualisation
- Requires immersive user input & realistically populated environment
- Omnifinity Omnideck 6
  - 6m Diameter treadmill (4m active) [9]
  - 16 triangular sections of rollers
  - Tracks user location in virtual environment
  - The Transport Systems Catapult in Milton Keynes UK have the first non-military Omnideck 6

Omnideck Omnifinity 6 at ITEC2015 [10]
©MSE Omnifinity AB

# What Next?

## Microscopic Simulation

- Develop message partitioning scheme for network based communication
  - Applicable to non-transport simulations
  - Specialised For Road Networks
- Multi-modal simulation of vehicles and pedestrians

## Macroscopic Transport Simulation

- Working with an Industrial Partner
- Accelerate Macroscopic assignment and simulation using GPUs for large scale models

# References I


[1] Department for Transport, "Road traffic forecasts 2015." https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/260700/road-transport-forecasts-2013-extended-version.pdf, Mar. 2015.

[2] Atkins, "HS2 Baseline Forecasting Report ." https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/365502/HS2_Baseline_Forecasting_Report_August_2012v4_1_TRACKED_V0.1_final.pdf, 2013.

[3] International Air Transport Association (IATA), "Press Release: New IATA Passenger Forecast Reveals Fast-Growing Markets of the Future." http://www.iata.org/pressroom/pr/Pages/2014-10-16-01.aspx, 2014.

[4] UK Department for Transport, "Cost of maintaining the Highways Agency's motorway and A road network per lane mile." https://www.gov.uk/government/publications/cost-of-maintaining-the-highways-agency-s-motorway-and-a-road-network-per-lane-mile, 2011.

[5] H. Neffendorf, G. Fletcher, R. North, T. Worsley, and R. Bradley, "Modelling for intelligent mobility." https://ts.catapult.org.uk/documents/10631/169582/Modelling+Intelligent+Mobility,+Feb+2015/73b7c9f9-d05a-4fca-ad9f-0e226e48d6b7, Feb. 2015.

[6] P. Richmond, "Flame gpu technical report and user guide," tech. rep., technical report CS-11-03. Technical report, University of Sheffield, Department of Computer Science, 2011.

[7] P. G. Gipps, "A model for the structure of lane-changing decisions," *Transportation Research Part B: Methodological*, vol. 20, no. 5, pp. 403–414, 1986.

# References II


[8] P. Heywood, P. Richmond, and S. Maddock, "Road network simulation using flame gpu," in *Euro-Par 2015: Parallel Processing Workshops*, pp. 430–441, Springer, 2015.

[9] Omnifinity AB, "Omnideck 6 - Technical Product sheet." `http://www.omnifinity.se/media/`, 2015.

[10] Omnifinity AB, "Omnideck Media Pack." `http://www.omnifinity.se/media/`, 2015.

Described the challenge of increasing demand

Highlighted current performance limitations

Demonstrated immersive virtual reality for transport system simulation

@ptheywood
ptheywood.uk
p.heywood@sheffield.ac.uk

www.flamegpu.com
www.sheffield.ac.uk/dcs/research/groups/graphics

The University Of Sheffield.

sheffield.ac.uk

CATAPULT Transport Systems

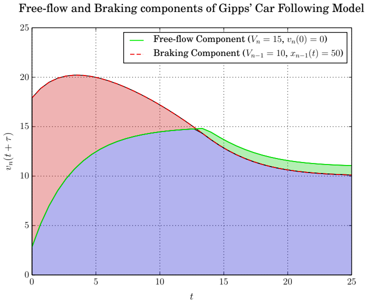
ts.catapult.org.uk

# Additional Slides

# Gipps' Car Following Model Equation

$$v_n(t+\tau) = \min\Big\{ v_n(t) + 2.5 a_n \tau (1 - v_n(t)/V_n)(0.025 + v_n(t)/V_n)^{\frac{1}{2}},$$
$$b_n\tau + \sqrt{{b_n}^2\tau^2 - b_n[2[x_{n-1}(t) - s_{n-1} - x_n(t)] - v_n(t)\tau - v_{n-1}(t)^2/\hat{b}]}\Big\}$$

| | |
|---|---|
| $a_n$ | the maximum acceleration of vehicle $n$ |
| $b_n$ | the most severe braking that the vehicle $n$ will undertake |
| $s_n$ | the effective size of vehicle $n$, including a margin |
| $V_n$ | the target speed of vehicle $n$ |
| $x_n(t)$ | the location of the front of vehicle $n$ at time $t$ |
| $v_n(t)$ | the speed of vehicle $n$ at time $t$ |
| $\tau$ | constant reaction time for all vehicles |
| $\hat{b}$ | estimate of leading vehicles most severe braking |



# FLAME GPU & Omnideck Integration

- Simulator listens for UDP packets
- Updates user agent and camera positions
- Camera height set based on floor height-map
- Simulated Pedestrians respond to the user as a pedestrian agent
- Visualisation uses GLFW and the Oculus Runtime

```
typedef struct UdpData
{
    unsigned int commandID:4;
    unsigned int unused:4;
    double lon;
    double lat;
    double altitude;
    double yaw;
    double pitch;
    double roll;
}UdpData;
```