Massivizing Online Games: Distributed Computing Challenges and High Quality Time

Alexandru Iosup

Parallel and Distributed Systems Group
Delft University of Technology

Our team: Undergrad Adrian Lascateu, Alexandru Dimitriu (UPB, Romania), …
Grad Siqi Shen, Yong Guo, Ruud vd Bovenkamp (TU Delft, the Netherlands), …
Staff Otto Visser, Dick Epema, Henk Sips, Fernando Kuipers (TU Delft), Thomas Fahringer, Radu Prodan (U. Innsbruck), Nicolae Tapus, Vlad Posea (UPB), …
Why Social Gaming?
What is This Talk About?

Massivizing Social Games: Distributed Computing Challenges and High Quality Time – A. Iosup
Online Gaming used to be art, may now be computing

Online Gaming used to be multimedia, is now DC

Online Gaming used to be networking, is now all DC

Online Gaming used to be v-worlds, is now many apps
(TU) Delft - the Netherlands - Europe

- Founded 13th century
  - Population: 100,000

- Founded 1842
  - Population: 13,000

Population: 16.5 M
The Parallel and Distributed Systems Group at TU Delft

Johan Pouwelse  P2P systems
Henk Sips      multicore P2P systems
Dick Epema     grids/clouds e-Science P2P systems
Alexandru Iosup online gaming grids/clouds P2P systems

Home page
- www.pds.ewi.tudelft.nl

Publications
- see PDS publication database at publications.st.ewi.tudelft.nl
What’s in a name? MSG, MMOG, MMO, …

Over 250,000,000 active players

Massively Social Gaming  =  (online) games with massive numbers of players (100K+), for which social interaction helps the gaming experience

1. Virtual World Sim
   Explore, do, learn, socialize, compete
   +

2. Game Data
   Player stats and relationships, others
   +

3. Game Content
   Graphics, maps, puzzles, quests, culture
MSGs are a Popular, Growing Market

- 25,000,000+ subscribed players (from 250,000,000+ active)
- Over 10,000 MSGs in operation
- Subscription market size $7.5B+/year, Zynga $600M+/year
Bungie, Computing then Serving 1.4PB/yr.

- Halo 3 is one of the many successful games
- Halo 3 players get, in 1.4PB
  - Detailed player profiles
  - Detailed usage stats
  - Ranking

- CERN produces ~15PB/year (10x larger)
  - (Not) faster than the speed of light, the Higgs boson (?)
World of Warcraft, a Traditional HPC User (since 2003)

- 10 data centers
- 13,250 server blades, 75,000+ cores
- 1.3PB storage
- 68 sysadmins (1/1,000 cores)

Zynga, an Amazon WS User

**ZYNGA GAME**

**FarmVille**

**AVERAGE NUMBER OF ACTIVE PLAYERS**

27 million daily / 75 million monthly

**THAT'S A LOT**

FarmVille boasts 118 million total installs. It has more monthly active users than the population of France.

**AVERAGE SESSION**

33 minutes

**PLAYER PROFILE**

N/A average age

60% female, 40% male

**MOST POPULAR TIME TO PLAY (EST)**

FarmVille: 8-9 AM

Selling in-game virtual goods:

“Zynga made est. $270M in 2009 from.”

http://techcrunch.com/2010/05/03/zynga-revenue/

“Zynga made more than $600M in 2010 from selling in-game virtual goods.”

S. Greengard, CACM, Apr 2011
Agenda

1. What’s in a Name?

2. Three Current Challenges
   1. Platform Scalability Challenge
   2. Gaming Analytics Challenge
   3. Content Generation Challenge

3. The Next Five Years

4. Conclusion
@large Research Challenge: V-World Platform for MMOGs

Scaling quickly to millions of players
- 1M in 4 days, 10M in 2 months
- Up-front and operational costs
- Performance, Scalability, & Cost
Impact on Game Experience

Responsive game vs. Unresponsive game

September 7, 2012

[Source: Nae, Iosup, and Prodan, ACM SC 2008 and IEEE TPDS 2011]
Computational Model for the Server

- Single sequential loop
- 3 steps in each loop:
  1. Game-world state update
  2. Entity interaction computation (dominant for MSGs)
  3. Entity state updates
- Load generated by (2) non-deterministic $\leftarrow$ human factor

(Source: Nae, Iosup, and Prodan, ACM SC 2008)
Model for Entity Interaction Computation

- Player to Player/Player to Environment interaction
  - **Low** interaction: $O(n)$, e.g. RTS
  - **Medium** interaction: $O(n \cdot \log(n))$ – RPG quest maps
  - **High** interaction: $O(n^2)$ – RPG war maps, FPS
  - **Very High** interaction: $O(n^2 \cdot \log(n))/ O(n^3)$, unit-target matching, team path-finding, maxflow alg., …
Online games hosting model

• Generic Online Games (non-MM)
  • **Static:** dedicated isolated single servers

• MMOGs
  • **Static:** dedicated clusters - using parallelization techniques

• Problems with these approaches
  1. Large amount of over-provisioning
  2. Non-efficient coverage of the world for the provided service

[Source: Nae, Iosup, and Prodan, ACM SC 2008]
Cloud Computing

VENI - @larGe: Massivizing Online Games using Cloud Computing
Proposed hosting model: dynamic

- Using data centers for dynamic resource allocation

  - Main advantages:
    1. Significantly lower over-provisioning
    2. Efficient coverage of the world is possible

[Source: Nae, Iosup, and Prodan, ACM SC 2008]
MSG Ecosystem: Model

- **Game operators**
  - Past player activity/business model $\rightarrow$ Predicted load $\rightarrow$ requests

- **Data centers**
  - Local time-space **renting policy** $\rightarrow$ offers
  - Time-Space renting policy, e.g., 1 node-hour
  - Resource allocation: central request-offer matching
  - Rules for ranking request-offer match:
    1. The offer size and type vs. the request
    2. The geographical proximity offer-request
    3. The finer grained resources (quantity & time)

[Source: Nae, Iosup, and Prodan, ACM SC 2008]
Experimental Setup [1/3]

Discrete-Event Simulator

• Input
  • Traces from RuneScape, a real top-5 MMOG
    • 7 countries, 3 continents
    • More than 130 game worlds
  • Consisting of
    • Geographical location
    • Number of clients
    • Over 10,000 samples at 2 min. interval, 2 weeks

• Output (for every time-step)
  • Resource allocation decisions
  • Resource allocation
  • Performance metrics

[Source: Nae, Iosup, and Prodan, ACM SC 2008]
Experimental Setup [2/3]

Environment

• 1 game operator
• 17 data centers
• 11 data center time-space renting policies

<table>
<thead>
<tr>
<th>Location</th>
<th>Data Centers</th>
<th>Machines (total)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Europe</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Finland</td>
<td>2</td>
<td>8 machines</td>
</tr>
<tr>
<td>Sweden</td>
<td>2</td>
<td>8 machines</td>
</tr>
<tr>
<td>U.K.</td>
<td>2</td>
<td>20 machines</td>
</tr>
<tr>
<td>Netherlands</td>
<td>2</td>
<td>15 machines</td>
</tr>
<tr>
<td><strong>North America</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>U.S. (West)</td>
<td>2</td>
<td>35 machines</td>
</tr>
<tr>
<td>Canada (West)</td>
<td>1</td>
<td>15 machines</td>
</tr>
<tr>
<td><strong>Australia</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Australia</td>
<td>2</td>
<td>8 machines</td>
</tr>
</tbody>
</table>

[Source: Nae, Iosup, and Prodan, ACM SC 2008]
Experimental Setup [3/3]

Performance Metrics

• Resource over-allocation [%]
  - The wasted resources vs. optimal allocation at each simulation time step for all utilized machines (cumulative).

• Resource under-allocation [%]
  - The amount of resources needed but not allocated, for all machines (computed individually).

• Significant under-allocation events (count)
  - Number of times the resource under-allocation is >1%, for a period of 2 minutes → people leave.
Resource Provisioning and Allocation

**Static vs. Dynamic Provisioning**

[Source: Nae, Iosup, and Prodan, ACM SC 2008]
Impact of Load Prediction Accuracy

Q: How does the prediction accuracy impact resource provisioning? A: Good prediction matters.

[Source: Nae, Iosup, and Prodan, ACM SC 2008]
Impact of Interaction Compute-Intensiveness

Q: How are different MMOG types handled under dynamic resource provisioning?

(Interaction models Low ~ $O(n)$, Medium, High, Very High ~ $O(n^3)$)

A: Over-provisioning, Under-provisioning worse with increase in interaction compute-intensiveness

[Source: Nae, Iosup, and Prodan, ACM SC 2008]
Latency Tolerance: From None to High

Q: What is the impact of latency tolerance on hosting?

A: (left) very sensitive  (mid) sensitive  (right) non-sensitive
    very costly          costly          cheap
Also Studied

- Via real game measurements
  - Interactivity model (short-term msmt.)
  - Effects of underperforming platform (long-term msmt.)

- Via prototype implementation
  - Match model-reality [TPDS’11]

- Via simulation
  - Impact of virtualization [NetGames’11][IJ AMC’11]
  - Economic models [under submission]
@large Research Challenge: Continuous Analytics for MMOGs

Analyzing the behavior of millions of players, on-time
- Data mining, data access rights, cost v. accuracy, …
- Reduce upfront costs
- Low response time & Scalable
- Large-scale Graph Processing
@large: Social Everything!

- **Social Network** = undirected graph, **relationship** = edge
- **Community** = sub-graph, density of edges between its nodes higher than density of edges outside sub-graph

(Analytics Challenge)

**Improve gaming experience**

- Ranking / Rating
- Matchmaking / Recommendations
- Play Style/Tutoring

**Self-Organizing Gaming Communities**

- Player Behavior
The Game Trace Archive (upcoming)

- Share gaming traces and best-practices on using them
- Support simulations and real-world experiments

<table>
<thead>
<tr>
<th>Name</th>
<th>Period</th>
<th>Size (GB)</th>
<th>Node (M)</th>
<th>Edge (M)</th>
<th>Category</th>
</tr>
</thead>
<tbody>
<tr>
<td>KGS</td>
<td>2002/02-2009/03</td>
<td>2</td>
<td>0.8</td>
<td>27.4</td>
<td>Chess Game</td>
</tr>
<tr>
<td>FICS</td>
<td>1997/11-2011/09</td>
<td>168</td>
<td>0.4</td>
<td>144.2</td>
<td>Chess Game</td>
</tr>
<tr>
<td>BBO</td>
<td>2009/11-2009/12</td>
<td>10</td>
<td>3.9</td>
<td>12.9</td>
<td>Card Game</td>
</tr>
<tr>
<td>XFire</td>
<td>2008/05-2011/12</td>
<td>58</td>
<td>7.7</td>
<td>34.7</td>
<td>OMGN</td>
</tr>
<tr>
<td>Dota League</td>
<td>2006/07-2011/03</td>
<td>23</td>
<td>0.1</td>
<td>3.0</td>
<td>RTS</td>
</tr>
<tr>
<td>DotaLicious</td>
<td>2010/04-2012/02</td>
<td>1</td>
<td>0.1</td>
<td>0.6</td>
<td>RTS</td>
</tr>
<tr>
<td>Dota Garena</td>
<td>2009/09-2010/05</td>
<td>1</td>
<td>0.3</td>
<td>0.1</td>
<td>RTS</td>
</tr>
<tr>
<td>WoWAH</td>
<td>2006/01-2009/10</td>
<td>3</td>
<td>0.1</td>
<td>N/A</td>
<td>MMORPG</td>
</tr>
</tbody>
</table>
The CAMEO Framework

1. **Address community needs**
   - Can analyze skill level, experience points, rank
   - Can assess community size dynamically

2. **Using on-demand technology: Cloud Comp.**
   - Dynamic cloud resource allocation, Elastic IP

3. **Data management and storage: Cloud Comp.**
   - Crawl + Store data in the cloud (best performance)

4. **Performance, scalability, robustness: Cloud Comp.**

---

A. Iosup, CAMEO: Continuous Analytics for Massively Multiplayer Online Games on Cloud Resources. ROIA, Euro-Par 2009 Workshops, LNCS 6043, (2010)
@large: Sample Analytics Results
Activity and Social Network

- **Bridge Base Online (BBO):** 1M+ players, top free site
- **Dataset:** 100K players
  - 9K group
  - Social relationships from bridge pairing
- **Large (~10K) online social groups can coordinate**
- **Identified player behavior**
  - community builder,
  - community member,
  - random player,
  - faithful player

M. Balint, V. Posea, A. Dimitriu, and A. Iosup, An Analysis of Social Gaming Networks in Online and Face to Face Bridge Communities, LSAP 2011.
@large: Sample Analytics Results

Analysis of Meta-Gaming Network

- “When you play a number of games, not as ends unto themselves but as parts of a larger game, you are participating in a metagame.” (Dr. Richard Garfield, 2000)
- XFire: since 2008 (3+ years), 500K of 20M players

![Graph showing distribution of time spent in games](image)
@large Research Challenge: Content Generation for MMOGs

Generating content on time for millions of players
- Player-customized: Balanced, Diverse, Fresh
- Up-front and operational costs
- Response time, Scalability, & Cost
Player-Customized Content
Skill Level Distribution in RuneScape

- **RuneScape**: 135M+ open accounts (world record)
- **Dataset**: 3M players (largest measurement, to date)
  - 1,817,211 over level 100
  - Max skill 2,280
- **Number of mid- and high-level players is significant**
- **New Content Generation Challenge**
@large: Sample Analytics Results
Skill Level Distribution in RuneScape

- Runescape: 135M active accounts, 7M active (2008)
- (largest MMOG msmt.)

- Player skill: distribution changes over time

(Procedural) Game Content (Generation)

Derived Content
NewsGen, Storification

Hendricks, Meijer, vd Velden, Iosup,
Procedural Content Generation for
Games: A Survey,
ACM TOMCCAP, 2012

Game Design
Rules, Mechanics, ...

Game Scenarios
Puzzle, Quest/Story, ...

Game Systems
Eco, Road Nets, Urban Envs, ...

Game Space
Height Maps, Bodies of Water, Placement Maps, ...

Game Bits
Texture, Sound, Vegetation, Buildings, Behavior,
Fire/Water/Stone/Clouds
The POGGI Content Generation Framework

Only the puzzle concept, and the instance generation and solving algorithms, are produced at development time.

* A. Iosup, POGGI: Puzzle-Based Online Games on Grid Infrastructures, EuroPar 2009 (Best Paper Award)
Puzzle-Specific Considerations
Generating Player-Customized Content

Puzzle difficulty

• Solution size
• Solution alternatives
• Variation of moves
• Skill moves

Player ability

• Keep population statistics and generate enough content for most likely cases
• Match player ability with puzzle difficulty
• Take into account puzzle freshness

Full Move Set:
- A: Right
- B: Up
- C: Down
- D: Left

Solution:
- Target: [ ]
- Pins: X A B C D E

Solution details:
- X: Right
- A: Right
- B: Up
- X: Up
(Best solution: 4 moves)

Solution:
- Target: [ ]
- Pins: X A B C D E

Solution details:
- B: Up
- X: Up
- B: Left
- C: Down
- C: Left
- B: Down
- B: Right
- B: Down
- E: Right
- E: Down
- E: Right
- B: Up
- A: Up
- B: Left
- C: Down
- C: Right
- E: Down
- X: Left
- E: Left
- X: Down
- X: Left
(Best solution: 21 moves)
Agenda

1. What’s in a Name?
2. Three Current Challenges
3. The Next Five Years
   1. Cloudification
   2. Mobile Social Gaming
   3. Social Everything!
   4. Content, Content, Content
4. Conclusion
Cloudification: PaaS for MSGs

(Platform Challenge)
Build MSG platform that uses (mostly) cloud resources

- Close to players
- No upfront costs, no maintenance
- Compute platforms: multi-cores, GPUs, clusters, all-in-one!
- Performance guarantees
- Code for various compute platforms—platform profiling
- Load prediction miscalculation costs real money
- What are the services?
- Vendor lock-in?
- My data
Mobile Social Gaming and the SuperServer

(Platform Challenge)

Support MSGs on mobile devices

- Mobiles everywhere (2bn+ users)
- Gaming industry for mobiles is new Growing Market
- SuperServer to generate content for low-capability devices?
- Battery for 3D/Networked games?
- Where is my server? (Ad-hoc mobile gaming networks?)
- Security, cheat-prevention

![US Mobile Gaming Revenues, by Segment, 2009-2014](chart.png)
Social Everything!

- **Social Network** = undirected graph, **relationship** = edge
- **Community** = sub-graph, density of edges between its nodes higher than density of edges outside sub-graph

(Analytics Challenge)

Improve gaming experience

- Ranking / Rating
- Matchmaking / Recommendations
- Play Style/Tutoring

Self-Organizing Gaming Communities

- Player Behavior
Content, Content, Content

(Content Challenge)
Produce and distribute content for 1BN people

• Game Analytics → Game statistic
• Crowdsourcing
• Storification
• Auto-generated game content
• Adaptive game content
• Content distribution/
  Streaming content
Massivizing Online Gaming

- Million-user, multi-bn market
- V-World, Content, Analytics

Current Technology

- Upfront payment
- Cost and scalability problems
- Makes players unhappy

Summary

@large: Our Vision

- HPC has to help
- Economy of scale with clouds

@large: Ongoing Work

- Content: POGGI Framework
- Platform: edutain@grid
- Analytics: CAMEO Framework

@large: The Future

- Happy players
- Happy cloud operators

Publications Gaming and Clouds

2008: ACM SC
2009: ROI A, CCGrid, NetGames, EuroPar (Best Paper Award), ...
2010: IEEE TPDS, Elsevier CCPE
2011: Book Chapter CAMEO, IEEE TPDS, IJ AMC
2012: IPDPS, CCGrid, ...

Graduation (Forecast)
2012—14: 3PhD, 6Msc, 6BSc
Thank you for your attention!
Questions? Suggestions? Observations?

More Info:
- http://www.st.ewi.tudelft.nl/~iosup/research.html
- http://www.st.ewi.tudelft.nl/~iosup/research_gaming.html
- http://www.st.ewi.tudelft.nl/~iosup/research_cloud.html

Alexandru Iosup

A.Iosup@tudelft.nl
http://www.pds.ewi.tudelft.nl/~iosup/ (or google “iosup”)
Parallel and Distributed Systems Group
Delft University of Technology
Resource Provisioning and Allocation
Compound Metrics

- Trade-off Utility-Cost still needs investigation
- Performance and Cost are not both improved by the policies we have studied

(Variable) Blackbox Performance Engineering

- Performance Evaluation of Four Commercial Clouds
  - Amazon EC2, GoGrid, Elastic Hosts, Mosso
  - Resource acquisition
  - Single- and Multi-Instance benchmarking

- Low compute and networking performance\(^1\)

- Performance variability over time\(^2\)


2- Iosup et al., On the Performance Variability of Production Cloud Services, CCGrid 2011, pds.twi.tudelft.nl/reports/2010/PDS-2010-002.pdf
Continuous Analytics for MMOGs

MMOG Data = raw and derivative information from the virtual world (millions of users)

Continuous Analytics for MMOGs = Analysis of MMOG data s.t. important events are not lost

- Data collection
- Data storage
- Data analysis
- Data presentation
- ... at MMOG rate and scale