

## Abstract

A second technological revolution is imminent. As with the first major technological revolution, which was driven by the decentralisation of computing power from that of mainframe networks to that of the “smart” personal-workstation computer, this second revolution is also characterised and driven by decentralised online-communication technology. In terms of this second revolution, however, modern computer users are poised to free themselves of centralised Internet servers for communication and/or file-sharing by connecting to other users directly.

This second revolution has even been given a name, that is, “People-to-People” or “Peer-to-Peer computing”, called “P2P computing” for short.

Although the basic concept of P2P computing cannot lay claim to being entirely new, it is associated by many a user with file-sharing, particularly with MP3 music files, whilst P2P computing, in fact, constitutes much more, such as instant messaging, group collaboration and distributed computing. Closer to the Information Studies and library environments would be the P2P data replication of a digital-library collection.

The extensive use of P2P computing will have a significant impact on the bandwidths available within organisations, to the extent that an unprepared network may very well be overwhelmed by the increased traffic generated by the use of P2P file-sharing.

A well-founded understanding of the usage patterns of the individuals using P2P applications will ensure better management and control by the system administrators of and over the expensive information technology (IT) infrastructure. This will also ensure better response to and usage of other systems and applications, such as e-mail and the general World Wide Web utilising the same infrastructure.

The principal aim of this study, which was undertaken at the Rand Afrikaans University, is to identify such user behaviour and utilisation patterns in students interacting with P2P systems available on the Internet.

The selected respondents, in their capacity of Information Science students, all completed a questionnaire aimed at examining, among other things, aspects such as

- information technology (IT) utilisation
- the frequency and duration of user sessions
- the type of content downloaded and shared
- the social aspects associated with P2P file-sharing.

P2P computing can be seen as a radically new mindset, setting the stage for a decentralised Internet future. Although distributed computing is dogged by various inherent problems, for example copyright and privacy issues, the most troublesome problem still is that of bandwidth usage. It is to be hoped, therefore, that the application of the research results of this study will ensure the better management of existing infrastructure and resources, as well as the development and rendering of new value-added services within organisations.



## Opsomming

'n Tweede tegnologiese revolusie is ophande. Net soos in die geval van die eerste groot tegnologiese revolusie, wat deur die desentralisasie van rekenaarkrag vanaf dié van hoofraamnetwerke na dié van die “slim” persoonlikewerkstasie-rekenaar aangedryf is, word hierdie tweede revolusie ook gekenmerk aan en aangedryf deur die desentralisasie van intydsekommunikasie-tegnologie. Met dié tweede revolusie is moderne rekenaargebruikers egter spronggereed om hulself te bevry van gesentraliseerde Internetbedieners vir kommunikasie en/of die deel van inligting deur regstreeks met ander gebruikers te skakel.

Hierdie tweede revolusie het selfs al 'n naam gekry, te wete die “Portuur-na-Portuur-” of “Persoon-tot-Persoon-rekenarisering”, kortweg in Engels “P2P computing” genoem.

Hoewel die basiese konsep van “P2P”-rekenarisering niks nuuts is nie, assosieer die meeste rekenaargebruikers dit met lêerdeling (“file-sharing” in Engels), en dan spesifiek met MP3-musieklêers, terwyl “P2P”-rekenarisering in werklikheid veel meer behels, byvoorbeeld kitsboodskapversending (“instant messaging” in Engels), samewerking binne groepsverband (“group collaboration” in Engels) en verspreide rekenaarverwerkings (“distributed computing” in Engels). Nader aan die Inligtingstudie- en biblioteek-omgewing sou “P2P”-rekenarisering vir die dataherhaling (“data replication” in Engels) van 'n digitalebiblioteek-versameling toegepas kon word.

Die gebruik op groot skaal van “P2P”-rekenarisering sal 'n groot impak hê op die bandwydtes wat organisasies tot hul beskikking het, in dié mate dat 'n onvoorbereide netwerk heel moontlik oorweldig sou kon word deur die verhoogde netwerkverkeer wat weens die gebruik van “P2P”-lêerdeling gegeneer word.

'n Grondige begrip van die gebruikspatrone van die individue wat “P2P”-toepassings benut, sal daarom verseker dat die stelseladministrateurs beter beheer en kontrole oor die duur intligtingstegnologie-infrastruktuur kan uitoefen, wat weer 'n beter reaksie op en gebruik van ander stelsels en toepassings sal verseker, soos e-pos en die algemene “World Wide Web”, wat dieselfde infrastruktuur gebruik.

Die hoofdoel van hierdie studie, wat aan die Randse Afrikaanse Universiteit onderneem is, is om sodanige gebruiksgedrag en -patrone te identifiseer by studente wat interaktief verkeer met “P2P”-stelsels wat op die Internet beskikbaar is.

Die gekose respondente, in hul hoedanigheid van Inligtingkunde-studente, het almal vir dié doel 'n vraelys ingevul waarin daar onder meer na die volgende aspekte ondersoek ingestel is:

- Inligtingtegnologie-gebruik
- Die duur en frekwensie van gebruikersessies
- Die tipe inligting wat afgelaai en gedeel word
- Die sosiale aspekte wat met “P2P”-lêerdeling gepaard gaan.

“P2P”-rekenaarkunde kan as 'n radikaal nuwe ingesteldheid beskou word, wat die weg vir 'n degesentraliseerde Internet-toekoms baan. Hoewel verspreide rekenaarkunde (“distributed computing” in Engels) deur verskeie inherente probleme geteister word, byvoorbeeld kopiereg- en privaatheidskwessies, bly die grootste vraagstuk steeds dié rondom bandwydtegebruik. Die toepassing van die navorsingsresultate van hierdie studie sal daarom hopelik verseker dat die bestaande infrastruktuur en hulpbronne beter bestuur word, asook dat nuwe toegevoegdewaarde-dienste binne organisasies ontwikkel en gelewer word.



## List of figures

<b>Figure 2.1:</b> Evolution of the Internet (Mowery, 2002) .....	<b>7</b>
<b>Figure 4.1:</b> A taxonomy of computer systems (Milojicic, 2002) .....	<b>48</b>
<b>Figure 4.2:</b> A taxonomy of the distributed systems (Milojicic, 2002) .....	<b>48</b>
<b>Figure 4.3:</b> P2P application types (Miller, 2001: 30) .....	<b>52</b>
<b>Figure 5.1:</b> Quadrant Model of Scientific Research (Stokes, 1997: 73) ...	<b>74</b>



## List of charts

<b>Chart 1:</b> Gender composition .....	<b>81</b>
<b>Chart 2:</b> Age distribution .....	<b>81</b>
<b>Chart 3:</b> Year of study .....	<b>82</b>
<b>Chart 4:</b> Degree registration .....	<b>83</b>
<b>Chart 5:</b> Ratings in terms of computer usage .....	<b>83</b>
<b>Chart 6:</b> ICT hardware owned or accessed .....	<b>84</b>
<b>Chart 7:</b> Relative age of ICT hardware .....	<b>85</b>
<b>Chart 8:</b> Connection to the Internet .....	<b>86</b>
<b>Chart 9:</b> Frequency of use of each application .....	<b>87</b>
<b>Chart 9a:</b> Duration of a single session of SMS/MMS, e-mail and WWW .....	<b>88</b>
<b>Chart 9b:</b> Duration of a single session of online chatting and P2P file-sharing .....	<b>88</b>
<b>Chart 9c:</b> Duration of a single session of online banking and online purchasing .....	<b>89</b>
<b>Chart 10:</b> Use of Peer-to-Peer file-sharing application .....	<b>90</b>
<b>Chart 11:</b> Number of files downloaded in a single session .....	<b>91</b>
<b>Chart 12:</b> Frequency of downloads .....	<b>91</b>
<b>Chart 13:</b> Type of content downloaded .....	<b>92</b>
<b>Chart 14:</b> Sharing of files on the Peer-to-Peer file-sharing application ...	<b>93</b>
<b>Chart 15:</b> Number of files uploaded in a single session .....	<b>93</b>
<b>Chart 16:</b> Frequency of uploads .....	<b>94</b>
<b>Chart 17:</b> Type of content uploaded .....	<b>94</b>
<b>Chart 18:</b> Number of files shared at any given point in time .....	<b>95</b>

## List of tables

<b>Table 2.1:</b> Some of the basic top-level DNS domains (Obiztek, 2004) ..	<b>11</b>
<b>Table 2.2:</b> Examples of popular Web search engines (Shannon, 2000)	<b>13</b>
<b>Table 2.3:</b> Major categories of newsgroups (Laquey, 1995: 68) .....	<b>15</b>
<b>Table 2.4:</b> Internet vs next-generation Internet (Stresing, 2003) .....	<b>21</b>
<b>Table 4.1:</b> Criteria for identifying P2P activity (Miller, 2001: 22) .....	<b>44</b>
<b>Table 4.2:</b> The “Ten Commandments” of computer ethics (Hellriegel, 1996: 631) .....	<b>58</b>
<b>Table 4.3:</b> Comparison between client/server and P2P collaboration (Milojicic, 2002) .....	<b>60</b>
<b>Table 5.1:</b> Differences between research design and research methodology .....	<b>75</b>
<b>Table 5.2:</b> Differences between quantitative and qualitative research methodologies .....	<b>76</b>



## Table of contents

<b>Acknowledgements</b> .....	<b>i</b>
<b>Abstract</b> .....	<b>ii</b>
<b>Opsomming</b> .....	<b>iv</b>
<b>List of figures</b> .....	<b>vi</b>
<b>List of charts</b> .....	<b>vii</b>
<b>List of tables</b> .....	<b>viii</b>





# Chapter 1

## Introduction

<b>1.1</b>	<b>Background to the research project</b> .....	<b>1</b>
<b>1.2</b>	<b>Rationale of the study</b> .....	<b>1</b>
1.2.1	What is “P2P computing”? .....	1
<b>1.3</b>	<b>Research problem and sub-problems</b> .....	<b>2</b>
<b>1.4</b>	<b>Research methodology</b> .....	<b>3</b>
1.4.1	Literature review .....	3
1.4.2	Empirical research .....	5
<b>1.5</b>	<b>Conclusion and recommendations</b> .....	<b>5</b>

# Chapter 2

## The Internet as information infrastructure

<b>2.1</b>	<b>Introduction</b> .....	<b>6</b>
<b>2.2</b>	<b>Defining the Internet</b> .....	<b>6</b>
<b>2.3</b>	<b>History of the Internet</b> .....	<b>6</b>
2.3.1	Phase 1 (1960 - 1985): Early computer networks .....	7
2.3.2	Phase 2 (1986 - 1995): Infrastructure development and growth ...	10
2.3.3	Phase 3 (1996 - the present): Creating commercial content and applications .....	12
<b>2.4</b>	<b>Applications utilising the Internet</b> .....	<b>12</b>
2.4.1	Electronic mail (e-mail) .....	12
2.4.2	The World Wide Web (WWW) .....	13
2.4.2.1	Difference between the Internet and the World Wide Web .....	14
2.4.3	Electronic commerce (E-commerce) .....	14
2.4.4	File Transfer Protocol (FTP) .....	15
2.4.5	Newsgroups .....	15
2.4.6	List serve .....	16
2.4.7	Peer-to-Peer computing .....	16
<b>2.5</b>	<b>Main features of the Internet</b> .....	<b>17</b>

2.5.1	Seamlessness .....	17
2.5.2	Currency .....	17
2.5.3	Global reach .....	18
2.5.4	Comprehensiveness .....	19
2.5.5	Interactivity .....	19
<b>2.6</b>	<b>The future of the Internet .....</b>	<b>20</b>
<b>2.7</b>	<b>Summary .....</b>	<b>22</b>

## Chapter 3

### The online community and the global village

<b>3.1</b>	<b>Introduction .....</b>	<b>23</b>
<b>3.2</b>	<b>Social communities revisited .....</b>	<b>23</b>
3.2.1	Impact of technology on a traditional community .....	25
<b>3.3</b>	<b>Online communities .....</b>	<b>25</b>
3.3.1	Defining online communities .....	27
3.3.2	Characteristics of online communities .....	28
3.3.2.1	Online communities are user-based .....	28
3.3.2.2	Online communities are specific .....	28
3.3.2.3	Online communities are organised .....	28
3.3.2.4	Online communities are integrated .....	29
3.3.2.5	Online communities are interactive .....	29
3.3.3	Online community activities .....	29
3.3.3.1	E-mail .....	29
3.3.3.2	Newsgroups .....	31
3.3.3.3	Chat .....	32
3.3.3.4	Message boards/Conferencing .....	33
<b>3.4</b>	<b>Benefits of an online community .....</b>	<b>34</b>
<b>3.5</b>	<b>Driving forces behind the online community .....</b>	<b>35</b>
3.5.1	Content attractiveness .....	35
3.5.2	Member loyalty .....	35
3.5.3	Member profiles .....	35
3.5.4	Transaction offerings .....	35
<b>3.6</b>	<b>The global village .....</b>	<b>36</b>

<b>3.7</b>	<b>Summary .....</b>	<b>37</b>
------------	----------------------	-----------

## Chapter 4

### Peer-to-Peer (P2P) computing

<b>4.1</b>	<b>Introduction .....</b>	<b>39</b>
<b>4.2</b>	<b>P2P: A historical perspective .....</b>	<b>39</b>
4.2.1	Direct exchange .....	40
4.2.2	Distributed processing .....	41
4.2.3	Online collaboration .....	41
<b>4.3</b>	<b>P2P: A modern-day perspective .....</b>	<b>42</b>
4.3.1	Defining P2P .....	42
4.3.1.1	A simple definition .....	42
4.3.1.2	A complex definition .....	42
4.3.1.3	P2P computing vs P2P network vs P2P communications .....	44
<b>4.4</b>	<b>Objectives of P2P technology .....</b>	<b>45</b>
4.4.1	Cost-sharing/Cost reduction .....	45
4.4.2	Improved scalability/reliability .....	46
4.4.3	Resource aggregation and interoperability .....	46
4.4.4	Increased autonomy .....	46
4.4.5	Anonymity/Privacy .....	47
4.4.6	Dynamism .....	47
4.4.7	Enabling <i>ad hoc</i> communication and collaboration .....	47
<b>4.5</b>	<b>P2P architecture .....</b>	<b>48</b>
<b>4.6</b>	<b>P2P models .....</b>	<b>48</b>
4.6.1	Pure P2P .....	49
4.6.2	P2P with a simple discovery server .....	49
4.6.3	P2P with discovery and lookup servers .....	50
4.6.4	P2P with discovery, lookup and content servers .....	50
<b>4.7</b>	<b>Applications utilising P2P computing .....</b>	<b>51</b>
4.7.1	P2P communications (instant messaging) .....	52
4.7.2	P2P file-sharing .....	53
4.7.2.1	P2P file-sharing in a nutshell .....	53
4.7.2.2	Issues surrounding P2P file-sharing .....	55

4.7.3	P2P group collaboration .....	59
4.7.4	P2P distributed computing .....	61
<b>4.8</b>	<b>Benefits of Peer-to-Peer computing .....</b>	<b>62</b>
<b>4.9</b>	<b>Drawbacks of Peer-to-Peer computing .....</b>	<b>64</b>
<b>4.10</b>	<b>P2P environments .....</b>	<b>64</b>
4.10.1	Implementation in tertiary-education institutions .....	65
4.10.1.1	P2P learning .....	65
4.10.1.2	Features of P2P computing contributing to P2P learning .....	66
4.10.2	P2P and libraries .....	67
<b>4.11</b>	<b>Implementation within a business .....</b>	<b>68</b>
<b>4.12</b>	<b>Summary .....</b>	<b>70</b>

## Chapter 5

### Empirical research

<b>5.1</b>	<b>Introduction .....</b>	<b>71</b>
<b>5.2</b>	<b>What is research? .....</b>	<b>71</b>
<b>5.3</b>	<b>Stokes' research paradigm .....</b>	<b>72</b>
5.3.1	Where this study fits into Stokes' quadrant .....	74
<b>5.4</b>	<b>Research methodology .....</b>	<b>74</b>
5.4.1	Research objectives .....	75
5.4.2	Differences between quantitative and qualitative research methodologies .....	76
5.4.3	A quantitative-research methodology .....	76
5.4.4	A qualitative-research methodology .....	77
<b>5.5</b>	<b>Research design .....</b>	<b>78</b>
5.5.1	Rationale of this study .....	78
5.5.2	The research problem .....	79
<b>5.6</b>	<b>Test sample, data collection and data .....</b>	<b>80</b>
<b>5.7</b>	<b>Analysis and interpretation of questionnaire results .....</b>	<b>81</b>
5.7.1	Section 1: General information .....	81
5.7.2	Section 2: Information and communication technology (ICT) hardware .....	83

5.7.3	Section 3: Information technology (IT) utilisation .....	86
5.7.4	Section 4: Peer-to-Peer file-sharing .....	89
5.7.5	Section 5: Social aspects of Peer-to-Peer file-sharing .....	95
5.8	<b>Summary</b> .....	96

## Chapter 6

### Conclusion and recommendations

6.1	<b>Summary</b> .....	97
6.2	<b>Conclusions</b> .....	97
6.3	<b>Recommendations</b> .....	99
6.4	<b>Future research</b> .....	99
<b>References</b> .....		101
<b>Appendix A</b> .....		113
<b>Appendix B</b> .....		126

